1912. Vol. XVI.

Annales de l'Institut Commercial de Kiew.

ИЗВѢСТІЯ Кіевскаго Коммерческаго ИНСТИТУТА.

1912.

Книга XVI.

КІЕВЪ—1912.

ВІД ПРОФЕСОРА КНЕУ
В.М. ФЕЩЕНКО

1912. Vol. XVI.

Annales
de l'Institut Commercial de Kiew.

ИЗВѢСТІЯ
Кіевскаго Коммерческаго
ИНСТИТУТА.

1912.

Книга XVI.

КІЕВЪ—1912.

ОГЛАВЛЕНІЕ.

TABLE DES MATIÈRES.

1912. Vol. XVI.

Annales de l'Institut Commercial de Kiew.

ИЗВѢСТІЯ Кіевскаго Коммерческаго ИНСТИТУТА.

1912.

Книга XVI.

КІЕВЪ—1912.

Печатано по опредѣленію Учебнаго Комитета Кіев. Коммерч. Института.
Директоръ **М. Довнаръ-Запольскій.**

Типографія И. И. Чоколова, Кіевъ, Б.-Житомірская № 20, соб. д.

Оглавленіе.

Е. Е. Слуцкій.

ТЕОРІЯ КОРРЕЛЯЦІИ

и

ЭЛЕМЕНТЫ УЧЕНІЯ О КРИВЫХЪ РАСПРЕДѢЛЕНІЯ.

○ ○ ○

(Пособіе къ изученію нѣкоторыхъ важнѣйшихъ методовъ современной статистики).

КІЕВЪ.
1912.

E. E. Slutsky.

Théorie de la Corrélation

et Traité abrégé de Courves de Fréquence.

Manuel pour servir à l'étude de quelques méthodes principales de la statistique moderne.

Kieff.
1912.

ОГЛАВЛЕНІЕ.

Замѣченныя опечатки.

Страница	*Строка*	*Напечатано:*	*Должно быть:*
1	Примѣч. 1-е	Докладъ, читанный и т. д. 26 апр. 1912 г.	Докладъ, читанный и т. д. 5 апр. 1912 г.
5	3 св.	или	, или
63	15 „	многихъ	мнимыхъ
136	На чер. 27 по винѣ цинкографіи оказались двѣ лишнія линіи: пунктирная и тонкая непрерывная.		
160	4—3 сн.	детерминантовъ	детерминантовъ.
„	1 „	почти достаточны.	совершенно достаточны.

Теорія корреляціи и элементы ученія о кривыхъ распредѣленія. [1]

Е. Слуцкаго.

ПРЕДИСЛОВІЕ.

За два послѣднія десятилѣтія теоретическая статистика сдѣлала колоссальные успѣхи: усовершенствованіе старыхъ методовъ, открытіе и разработка новыхъ, появленіе прекрасныхъ работъ по біологіи и обществовѣдѣнію, иллюстрирующихъ методы и доказывающихъ ихъ безспорную научную цѣнность, наконецъ, созданіе—хотя и небольшого пока еще—кадра научныхъ дѣятелей, систематически примѣняющихъ и разрабатывающихъ дальше новые методы,—все это, взятое вмѣстѣ, позволяетъ говорить о наступленіи новой эры въ статистикѣ.

Это движеніе зародилось и развилось въ Англіи, въ другія страны оно только начинаетъ проникать. Выросло оно изъ потребностей современной біологіи по иниціативѣ недавно умершаго знаменитаго Francis'a Galton'a. Но Galton не былъ математикомъ, и заслуга теоретической разработки новыхъ идей и созданія школы почти безраздѣльно должна составить славу Karl'a Pearson'a, имя котораго въ исторіи нашей науки будетъ стоять рядомъ съ именами Лапласа, Гаусса и Пуассона. Новая школа должна поэтому по всей справедливости называться школой Гальтона—Пирсона.

Общее оживленіе интереса къ теоретической статистикѣ позволяетъ надѣяться, что распространеніе идей новой школы на всѣ страны и на всѣ области возможнаго ихъ примѣненія—дѣло не особенно далекаго будущаго. Скромная задача автора—содѣйствовать этому естественному и неизбѣжному процессу.

---

[1] Докладъ, читанный въ засѣданіи „Общества Экономистовъ" 26 апрѣля 1912 г.

Примѣненіе новыхъ методовъ сравнительно просто, и обучиться ему нетрудно. Для того, чтобы примѣнять формулы, достаточно понимать ихъ смыслъ и умѣть производить указываемыя ими вычисленія, трудъ которыхъ упрощается употребленіемъ спеціальныхъ таблицъ, вычисленныхъ также по иниціативѣ Пирсона [1]). Однако, одной рутиной обойтись нельзя. Во всякомъ вопросѣ могутъ встрѣтиться неожиданныя новыя детали, могутъ возникнуть недоумѣнія относительно границъ приложимости метода и значенія результатовъ. А это требуетъ уже не одного только знанія рецептовъ для вычисленій, но и пониманія духа теорій и ихъ математическаго обоснованія.

Мы приходимъ такимъ образомъ къ кардинальному требованію, которое жизнь ставитъ дѣятелямъ статистики: *статистикъ долженъ быть математикомъ, ибо его наука есть наука математическая.*

Вотъ почему авторъ удѣлилъ такъ много вниманія формуламъ и математическимъ доказательствамъ. При этомъ играло роль еще одно соображеніе. Готовые рецепты хороши только въ областяхъ старыхъ, прочно установившихся. Пересадить новые методы на новую почву, не давая имъ обоснованія, авторъ считаетъ предпріятіемъ, не имѣющимъ за собой никакихъ гарантій успѣха.

Объемъ математическихъ знаній, требуемыхъ для пониманія большей части приводимыхъ авторомъ выводовъ и доказательствъ, сравнительно невеликъ. Самыхъ элементарныхъ свѣдѣній по аналитической геометріи и дифференціальному исчисленію (свѣдѣній, которыя могутъ быть пріобрѣтены въ нѣсколько дней) достаточно для усвоенія элементовъ теоріи корреляціи. Дальнѣйшія обобщенія въ области этой послѣдней, равно какъ и первая часть, трактующая о кривыхъ распредѣленія, требуютъ нѣсколько большихъ математическихъ знаній.

Авторъ пытался своимъ изложеніемъ удовлетворить различныя категоріи возможныхъ читателей. Доказательства поэтому упрощены, насколько это допускалось требованіемъ строгости

---

[1]) У насъ имѣются въ изданіи А. Леонтовича: „Элементарное пособіе къ примѣненію методовъ Gauss'а и Pearson'а при оцѣнкѣ ошибокъ, въ статистикѣ и біологіи". Часть III. Вспомогательныя таблицы. Кіевъ, 1911 г.

изложенія. Тѣ изъ математическихъ выкладокъ, которыя авторъ считалъ доступными для наименѣе подготовленныхъ читателей, представлены съ большими подробностями, чѣмъ въ томъ нуждается настоящій математикъ.

Наконецъ, изложеніе авторъ стремился развивать такимъ образомъ, чтобы читатель, пропустивъ трудное для него мѣсто, могъ все-таки поднять дальше упущенную имъ нить и понять смыслъ формулъ и порядокъ ихъ практическаго приложенія. Авторъ не льститъ себя однако надеждой, что послѣдняя задача разрѣшена имъ вполнѣ удовлетворительно.

Главная тема настоящей работы—это теорія корреляціи. Обойти теорію кривыхъ распредѣленія авторъ не считалъ, однако, возможнымъ, и ей посвящено, впрочемъ, далеко не полное и слишкомъ, можетъ быть, сжатое изложеніе въ I части.

Читателю, слабо знакомому съ математикой и интересующемуся главнымъ образомъ лишь методомъ корреляціи, можно посовѣтовать послѣ ознакомленія съ первыми четырьмя параграфами первой части сразу-же перейти къ чтенію второй части.

Часть I.

Элементы ученія о кривыхъ распредѣленія.

§ 1. *Общее понятіе о кривой распредѣленія или кривой частоты.*

Разсматривая любую совокупность индивидуумовъ, имѣющихъ общій имъ всѣмъ измѣримый признакъ, мы замѣчаемъ, что признакъ этотъ не у всѣхъ индивидуумовъ имѣетъ одну и ту-же величину. Было время, когда статистики пренебрегали этими различіями, концентрируя все свое вниманіе на ариѳметической средней признака. Въ настоящее время нѣтъ надобности уже бороться съ этимъ устарѣлымъ самоограниченіемъ. Можетъ считаться почти общимъ достояніемъ та мысль, что ариѳметическая средняя еще слишкомъ мало говоритъ намъ о характерѣ всей статистической группы и, что задача статистики сводится къ тому, чтобы по возможности полно и по возможности просто описать весь составъ подлежащей разсмотрѣнію совокупности.

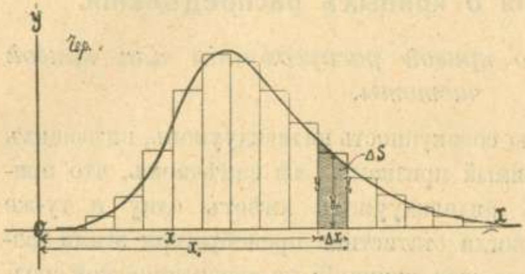
Первое рѣшеніе этой задачи состоитъ въ полномъ и детальномъ описаніи распредѣленія признака въ совокупности. Элементарная, но всегда необходимая форма такого описанія извѣстна хорошо всякому статистику. Это—таблица, въ которой величина признака подраздѣлена на интервалы, и для каждаго интервала указана численность соотвѣтствующей подгруппы [1]).

---

[1]) Къ сожалѣнію, часто эти подгруппы слишкомъ велики, чтобы можно было составить о распредѣленіи надлежащее представленіе. Даже при наличности довольно точной группировки низшія и высшія (особенно послѣднія) группы оказываются слишкомъ широкими. Напр., группировка крестьянъ по размѣрамъ посѣвной площади: отъ 0—5, 5—10, 10—15, 15—25, 25—50 и свыше 50 десятинъ. Такая группировка почти неподдается раціональной разработкѣ. Необходимо настаивать на томъ, чтобы группировка была возможно детальнѣе, чтобы признакъ, лежащій въ основаніи классификаціи, былъ раздѣленъ на равные интервалы, и особенно важно детальное подраздѣленіе на низшемъ и на высшемъ предѣлѣ группы

Дальнѣйшимъ шагомъ является изображеніе совокупности при помощи кривой распредѣленія. Съ чисто формально-математической стороны дѣло обстоитъ при этомъ весьма просто.

По оси x (черт. 1) откладываемъ величины признака, подраздѣляя его на возможно малые интервалы. Численность соотвѣтствующихъ подгруппъ изображаемъ площадью прямоугольника, построеннаго на каждомъ интервалѣ, какъ на основаніи. Если бы было извѣстно распредѣленіе во всѣхъ деталяхъ, то для безконечно большой группы мы могли бы сдѣлать наши интервалы безконечно малыми и въ предѣлѣ мы получили бы, вмѣсто ступенчатой фигуры, фигуру, ограниченную непрерывной кривой. Это и будетъ кривая распредѣленія.



Отрѣзокъ площади, заключенный между двумя ординатами y и y_1, отвѣчающими абсциссамъ x и x_1, изображаетъ число индивидуумовъ, обладающихъ признакомъ, величина котораго лежитъ между x и x_1. Средняя ордината (y_m) этого отрѣзка есть высота того прямоугольника, площадь котораго равна площади отрѣзка ($\triangle S$). Т. обр. $y_m \times \triangle x = \triangle S$. Отсюда

$$y_m = \triangle S / \triangle x, \text{ т. е.}$$

средняя ордината отрѣзка кривой представляетъ изъ себя число индивидуумовъ, приходящееся въ данномъ интервалѣ въ среднемъ на единицу величины интервала. Эту величину мы назовемъ средней частотой въ данномъ интервалѣ. Зная, напр., что въ нѣкоторой совокупности людей имѣется 2000 лицъ, ростъ которыхъ заключается между 2 метрами 40 см. и 2 м. 45 см., мы находимъ среднюю частоту—2000/5 = 400 чел. на 1 см. разницы въ ростѣ.

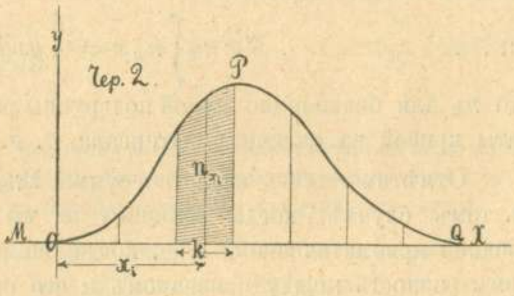
Очевидно, однако, что въ разныхъ частяхъ нашего интервала частота будетъ различная. Однако, чѣмъ ближе мы возьмемъ двѣ ординаты кривой, тѣмъ меньше онѣ будутъ отличаться другъ отъ друга, и тѣмъ меньше будетъ отличаться отъ каждой изъ нихъ средняя ордината, изображающая частоту. Взявъ безконечно малый интерваль, мы найдемъ, что въ предѣлѣ *средняя* ордината сольется съ *ординатой кривой* въ данной точкѣ. Символически:

$$y = \frac{ds}{dx}.$$

Такимъ образомъ, въ то время, какъ площадь кривой распредѣленія показываетъ абсолютное число индивидуумовъ, обладающихъ признакомъ, варіирующимся въ извѣстныхъ границахъ, ордината этой кривой даетъ частоту индивидуумовъ съ данной величиной признака, т. е. число такихъ индивидуумовъ на единицу разницы въ величинѣ признака. Поэтому кривыя распредѣленія называются также кривыми частоты (Frequency Curves).

§ 2. *Моменты распредѣленія.*

Кривая распредѣленія, если мы построимъ ее эмпирически по данному матеріалу, даетъ намъ только то, что въ этомъ матеріалѣ уже имѣется, но даетъ въ болѣе наглядной формѣ. Для цѣлей болѣе глубокаго изученія одной наглядности, однако, недостаточно. Намъ нужно имѣть численныя характеристики различныхъ свойствъ статистической группы. Одну изъ нихъ мы знаемъ уже. Это — средняя величина признака. Чтобы ее найти, нужно величину признака помножить на число индивидуумовъ, обладающихъ признакомъ въ такомъ размѣрѣ, и раздѣлить на общее число индивидуумовъ. Въ тѣхъ, однако, случаяхъ, когда число индивидуумовъ очень велико, этотъ способъ дѣлается непримѣнимымъ. Во-первыхъ, потому, что намъ можетъ не быть извѣстной точная величина признака у каждаго индивидуума, во-вторыхъ, даже если она и извѣстна, счетная работа становится непосильной. Поэтому намъ необходимъ способъ приближеннаго вычисленія средней ариѳметической.



Пусть MPQ (чер. 2) будетъ кривой распредѣленія. O—точка, отъ которой мы измѣряемъ величину признака. Раздѣлимъ всю площадь нашей кривой на участки, соотвѣтствующіе равнымъ интерваламъ, и пусть величина интервала равна $k = \triangle x$. Величины признака, соотвѣтствующія серединамъ интерваловъ, т. е. разстояніямъ между точкой O и серединой каждаго интервала, перенумеруемъ по порядку; мы будемъ имѣть тогда $x_1, x_2 \ldots x_i \ldots$ Число индивидуумовъ въ группѣ, предѣлы которой равны $x_i - \frac{1}{2}k$ и $x_i + \frac{1}{2}k$, обозначимъ черезъ n_{x_i} и допустимъ, что всѣ индивидуумы каждой такой группы обладаютъ признакомъ, величина котораго въ точности равна x_i. Тогда, если число индивидуумовъ во всей совокупности равно N, средняя ариѳметическая признака будетъ: $\bar{x} = \frac{1}{N}\Sigma n_{x_i} x_i$, или, опуская для краткости значки при x_i и умножая обѣ части на N:

$$(1) \ldots\ldots\ldots\ldots N\bar{x} = \Sigma n_x x.$$

Чѣмъ меньше будутъ наши интервалы, тѣмъ точнѣе будетъ и эта приближенная формула [1]). Для интерваловъ безконечно малыхъ (т. е. въ предѣлѣ) сумма обращается въ интегралъ, и мы имѣемъ точное выраженіе:

$$(2) \ldots\ldots\ldots N\bar{x} = \int n_x x = \int yx dx,$$ если мы сообразимъ, что n_x для безконечно малой подгруппы равно произведенію ординаты кривой на величину интервала, т. е. на dx.

Отмѣтимъ здѣсь свойство суммы $\Sigma n_x x$ обращаться въ нуль въ томъ случаѣ, когда величину x мы будемъ измѣрять отъ средней ариѳметической, т. е. когда за величину признака примемъ разность между признакомъ и его средней ариѳметической. Въ самомъ дѣлѣ, пусть при этомъ новомъ способѣ измѣренія величина признака будетъ $x' = x - \bar{x}$. Тогда

$$(3) \ldots\ldots \Sigma n_x x' = \Sigma n_x(x - \bar{x}) = \Sigma n_x x - \Sigma n_x \bar{x} =$$
$$= \Sigma n_x x - \bar{x}\Sigma n_x = N\bar{x} - \bar{x}N = 0,$$

что и требовалось доказать.

---

[1]) Статистическая практика показала, что этотъ способъ при неслишкомъ маломъ числѣ индивидуумовъ въ совокупности даетъ достаточно точные результаты даже при небольшомъ числѣ группъ (10—12).

Аналогично средней ариѳметической можно получить рядъ другихъ чиселъ, также характеризующихъ подлежащую изслѣдованію совокупность. Именно, мы можемъ такимъ же способомъ найти средній квадратъ признака, средній кубъ, среднюю четвертую степень и т. д. Эти выраженія Пирсонъ называетъ моментами. Употребляя прежнія обозначенія, мы будемъ имѣть:

		Точно.	Приближенно.
	Нулевой моментъ (средн. нул. степень)	$\mu'_0=\frac{1}{N}\int yx^0dx=1$	$\nu'_0=\frac{1}{N}\Sigma n_x x^0=1$
	Первый моментъ (средн. ариѳметическ.)	$\mu'_1=\bar{x}=\frac{1}{N}\int yxdx$	$\nu'_1=\frac{1}{N}\Sigma n_x x$
(4)	Второй моментъ (средній квадратъ)	$\mu'_2=\frac{1}{N}\int yx^2dx$	$\nu'_2=\frac{1}{N}\Sigma n_x x^2$
	Третій моментъ (средній кубъ)	$\mu'_3=\frac{1}{N}\int yx^3dx$	$\nu'_3=\frac{1}{N}\Sigma n_x x^3$
	Четвертый моментъ (средн. четвер. степ.) и т. д.	$\mu'_4=\frac{1}{N}\int yx^4dx$	$\nu'_4=\frac{1}{N}\Sigma n_x x^4$

Примѣчаніе. Величину средней ариѳметической по соображеніямъ удобства мы будемъ обозначать также буквой h, или $h_x, h_y \ldots$, чтобы обозначить, къ какой перемѣнной величина относится.

Моменты можно находить около любого начала координатъ, но наибольшее значеніе имѣютъ моменты, найденные относительно положенія средней ариѳметической. Именно они то и характеризуютъ распредѣленіе. Если назовемъ точку на оси x, которой соотвѣтствуетъ средняя ариѳметическая величина признака, *центромъ* распредѣленія, то моменты относительно вертикальной оси, проходящей черезъ эту точку, можно будетъ назвать *центральными*. Ихъ принято обозначать тѣми-же буквами, но безъ черточекъ.

Прямое нахожденіе центральныхъ моментовъ неудобно, потому что при этомъ пришлось-бы возводить въ квадратъ, въ кубъ и въ четвертую степень многозначныя числа $(x-\bar{x})$, т. к. средняя ариѳметическая $(\bar{x})$ только случайно можетъ оказаться цѣлымъ числомъ. Поэтому проще всего найти моменты около любого начала, а затѣмъ по этимъ нецентральнымъ моментамъ найти центральные.

Формула для перехода отъ однихъ моментовъ къ другимъ очень проста.

По опредѣленію центральнаго p-аго момента имѣемъ:

(5) . . . . $N\nu_p = \Sigma n_x (x - \bar{x})^p$. Разлагая по биному Ньютона получаемъ:

$$N\nu_p = \Sigma \left\{ n_x \left[x^p - p x^{p-1} \bar{x} + \frac{p(p-1)}{1 \, . \, 2} x^{p-2} \bar{x}^2 - \, . \, . \, . \, . \, . \right.\right.$$
$$\left.\left. . \, . \, . \pm p x \bar{x}^{p-1} \mp \bar{x}^p \right] \right\},$$

гдѣ знакъ $+$ или $-$ зависитъ отъ того, будетъ-ли p четнымъ или нечетнымъ. Суммируя отдѣльные члены и замѣняя для симметріи $\bar{x}$ черезъ ν'_1 находимъ:

$$N\nu_p = \Sigma n_x x^p - p\bar{x} \, \Sigma n_x x^{p-1} + \, . \, . \, . \, . \pm p \bar{x}^{p-1} \Sigma n_x x \mp \bar{x}^p \, \Sigma n_x =$$
$$= N \left\{ \nu'_p - p\nu'_{p-1} \nu'_1 + \frac{p(p-1)}{1 \, . \, 2} \nu'_{p-2} \nu'^2_1 - \, . \, . \, . \, . \pm p \nu'^p_1 \mp \nu'^p_1 \right\}$$

и окончательно, сокращая на N и соединяя два послѣдніе члена вмѣстѣ:

$$(6) \quad . \; . \; . \; . \quad \nu_p = \nu'_p - p\nu'_{p-1} \nu'_1 + \frac{p(p-1)}{1 \, . \, 2} \nu'_{p-2} \nu'^2_1 - \, . \, . \, . \, .$$
$$+ (-1)^{p-1} (p-1) \nu'^p_1 .$$

Въ частности:

$$(7) \quad . \; . \; . \; . \quad \left\{ \begin{array}{l} \nu_1 = 0 \\ \nu_2 = \nu'_2 - \nu'^2_1 \\ \nu_3 = \nu'_3 - 3\nu'_2 \nu'_1 + 2\nu'^3_1 \\ \nu_4 = \nu'_4 - 4\nu'_3 \nu'_1 + 6\nu'_2 \nu'^2_1 - 3\nu'^4_1 \, . \end{array} \right.$$

Вычисленіе располагается обыкновенно такимъ образомъ, что нѣкоторая величина признака, соотвѣтствующая серединѣ интервала, наиболѣе близкаго къ центру распредѣленія, принимается за условный нуль, а величина интервала (k) принимается за единицу. Чтобы потомъ выразить моменты въ прежнихъ единицахъ, нужно помножить p-ый моментъ на k^p.

§ 3. *Среднее отклоненіе и коэффиціентъ измѣнчивости.*

Квадратный корень изъ второго центральнаго момента играетъ особенно крупную роль въ теоретической статистикѣ и ея приложеніяхъ. Обозначая его буквой σ, имѣемъ:

$$(8) \quad \sigma^2 = \nu_2 = \frac{1}{N} \Sigma n_x (x - \bar{x})^2 \quad ^{1}).$$

Эту величину, въ теоріи ошибокъ наблюденія называемую средней квадратичной ошибкой, мы будемъ называть *среднимъ отклоненіемъ* (standard deviation). Если распредѣленіе слѣдуетъ извѣстному закону Гаусса, то около $^2/_3$ всѣхъ индивидуумовъ отклоняется отъ средняго ариѳметическаго въ обѣ стороны не больше, чѣмъ на величину средняго отклоненія. На этомъ основаніи σ можетъ служить мѣрой измѣнчивости, опредѣляющей насколько тѣсны предѣлы, въ которыхъ располагается большинство индивидуумовъ совокупности.

Если мы раздѣлимъ главное отклоненіе на среднюю ариѳметическую, то получимъ такъ наз. коэффиціентъ измѣнчивости

$$(9) \quad V = \frac{\sigma}{h}.$$

Для практическихъ приложеній его удобно выражать въ $^0/_0{}^0/_0$.

Примѣръ. Изучая колебанія среднихъ мѣсячныхъ цѣнъ на рожь за 124 мѣсяца 1893—1903 гг., я нашелъ [2]):

(1) Москва	(„овинная“)	$h_1 = 59{,}40 \pm 0{,}77$ коп.	$\sigma_1 = 12{,}64 \pm 0{,}54$ коп.
(2) Елецъ	(„тяжелая“)	$h_2 = 52{,}64 \pm 0{,}65$ „	$\sigma_2 = 10{,}74 \pm 0{,}46$ „
(3) Самара	(безъ обозн. сорта)	$h_3 = 47{,}04 \pm 0{,}84$ „	$\sigma_3 = 13{,}84 \pm 0{,}59$ „

Мы видимъ здѣсь, что наши три центра отличаются не только средними цѣнами, но и характеромъ ихъ колебаній. По величинѣ колебаній (σ) на первомъ мѣстѣ стоитъ Самара, затѣмъ Москва и наконецъ Елецъ, причемъ разница между Москвой и Самарой меньше, чѣмъ между Москвой и Ельцомъ.

---

[1]) Формула не совсѣмъ точна, т. к. ν_2 есть приближенная величина, которую мы въ большинствѣ случаевъ можемъ, прежде чѣмъ находить σ, исправить однимъ изъ ниже указанныхъ способовъ. (См. § 7).

[2]) См. „Сводъ товарныхъ цѣнъ“ за соотвѣтствующіе годы. 8 мѣсяцевъ пришлось пропустить вслѣдствіе неполноты свѣдѣній.

Если мы, однако, обратимся къ относительному размѣру этихъ колебаній, характеризующему устойчивость цѣнъ и выражаемому при помощи коэффиціента измѣнчивости, то найдемъ слѣдующія числа:

$$V_1 = 100 \frac{\sigma_1}{h_1} = 21{,}28\% \pm 0{,}95$$

$$V_2 = 100 \frac{\sigma_2}{h_2} = 20{,}40\% \pm 0{,}91$$

$$V_3 = 100 \frac{\sigma_3}{h_3} = 29{,}42\% \pm 1{,}36$$

Измѣнчивость Московской и Елецкой цѣны почти одинакова и можетъ быть противупоставлена измѣнчивости самарской цѣны, доходящей почти до 30% средней ариѳметической и составляющей важную экономическую характеристику этого центра хлѣбной торговли.

§ 4. *Вѣроятныя ошибки.*

Рядомъ со средними, средними отклоненіями и коэффиціентами измѣнчивости даны и ихъ вѣроятныя ошибки, вычисленныя по формуламъ:

(10) . . . . . . $E_h = 0{,}67449 \dfrac{\sigma}{\sqrt{N}}$

(11) . . . . . . $E_\sigma = 0{,}67449 \dfrac{\sigma}{\sqrt{2N}}$

(12) . . . . . . $E_V = 0{,}67449 \dfrac{V}{\sqrt{2N}} \sqrt{1 + 2\left(\dfrac{V}{100}\right)^2}$ [1] [2]

Рамки настоящей работы не позволяютъ расширить изложеніе включеніемъ въ него вывода вѣроятныхъ ошибокъ. Наиболѣе необходимое будетъ изложено въ одномъ изъ послѣднихъ параграфовъ, а здѣсь мы ограничимся лишь немногими замѣчаніями.

---

[1] Вычисленіе ихъ сильно облегчается таблицами miss Gibson и Pearl'я & Blakeman'а, перепечатанными у А. Леонтовича. См. op. cit. ч. III, табл. VI и VII.

[2] Первыя двѣ формулы давно извѣстны, послѣдняя дана Пирсономъ въ Proceedings of the Royal Society, Vol. 61, p. 345.

Комплексъ причинъ, складывающійся изъ очень большого числа элементарныхъ вліяній, въ которомъ какъ положительныя, такъ и отрицательныя вліянія одинаково вѣроятны, можетъ быть названъ комплексомъ случайныхъ причинъ. Чѣмъ больше число повтореній явленія, тѣмъ полнѣе случайныя причины компенсируютъ другъ друга, взаимно уничтожаясь въ своихъ вліяніяхъ на данное явленіе. Если законъ Гаусса, о которомъ—ниже, хотя приблизительно осуществляется, то приблизительно около половины случайныхъ отклоненій должно быть меньше „вѣроятной ошибки", около половины больше ея. Чрезвычайно мало вѣроятны отклоненія во много разъ (практически въ 6, въ 5 даже въ 4 раза) превышающія вѣроятную ошибку. Поэтому величина вѣроятной ошибки служитъ критеріемъ, отдѣляющимъ комплексъ случайныхъ вліяній отъ комплекса вліяній основныхъ причинъ, опредѣляющихъ характеръ явленія. Другого критерія въ данномъ вопросѣ не существуетъ. Примѣненіе его таково: Если мы изъ *очень* большого числа наблюденій узнали, напр., среднюю величину явленія, а затѣмъ изъ ограниченной части наблюденій, значительно меньшей по объему, вывели среднюю для этой части, то вѣроятная ошибка послѣдней величины должна служить намъ критеріемъ для отвѣта на вопросъ, существуетъ-ли различіе между всей совокупностью и данной частью ея. Напр., опредѣливъ средній ростъ милліона великороссовъ, а затѣмъ 200 ярославцевъ, мы могли бы сказать, что ростъ великороссовъ вообще и ярославцевъ различны, если бъ *найденная* нами разница превышала въ нѣсколько разъ вѣроятную ошибку второго, значительно менѣе точнаго опредѣленія. Разницу меньшую или немного большую вѣроятной ошибки мы можемъ объяснить вліяніемъ случайныхъ причинъ.

Если мы имѣемъ двѣ группы равныя по численности, напр. знаемъ среднюю изъ мѣсячныхъ цѣнъ въ Москвѣ за 1891—1900 и такую-же среднюю за 1901—1910, то мы вправѣ приписать случайнымъ причинамъ всякую разницу среднихъ, не выходящую, или незначительно выходящую за предѣлы вѣроятной ошибки этой разницы. Вѣроятная ошибка этой разности равна $\sqrt{E^2_1+E^2_2}$, т. е. корню квадратному изъ суммы квадратовъ обѣихъ вѣроятныхъ ошибокъ. Только въ томъ случаѣ, если бъ средняя за 900-е годы превышала среднюю за 90-е годы по крайней мѣрѣ на упятеренную или ушестеренную вѣроятную ошибку разности, мы могли бы утверждать, что уровень московскихъ цѣнъ дѣйствительно измѣнился.

Если мы захотимъ однако приложить этотъ критерій къ рѣшенію вопроса о томъ, можно ли признать существенной разницу между Московской, напр., и Елецкой средней цѣной, или между средними отклоненіями цѣнъ ржи въ Москвѣ и въ Ельцѣ, то мы сдѣлаемъ ошибку, ибо формула $E_{x_1-x_2}=\sqrt{E^2_{x_1}+E^2_{x_2}}$ справедлива только въ томъ случаѣ, когда данныя явленія *независимы другъ отъ друга*, а цѣны въ двухъ сравнительно близкихъ пунктахъ одного рынка таковыми признать нельзя. Только гадательно съ извѣстной долей субъективной увѣренности можно признать разницу существенной, если она *чрезвычайно* сильно превосходитъ вѣроятныя ошибки каждой величины въ отдѣльности. Таково можетъ быть наше сужденіе въ случаѣ предыдущаго примѣра только относительно разницы въ коэффиціентахъ измѣнчивости между Самарой съ одной стороны и Москвой и Ельцемъ съ другой.

Строгій критерій можетъ быть выведенъ лишь при помощи теоріи корреляціи.

§ 5. Законъ Гаусса и обобщеніе его Пирсономъ.

Число индивидуумовъ въ группѣ (N), средняя ариѳметическая ($\bar{x}$) и среднее отклоненіе (σ) въ нѣкоторыхъ случаяхъ совершенно достаточны для того, чтобы дать исчерпывающую характеристику совокупности. Именно, если индивидуумы этой совокупности слѣдуютъ въ своемъ распредѣленіи закону Гаусса, то по этимъ тремъ величинамъ можно найти чисто теоретическимъ путемъ численность любой подгруппы, стоитъ только открыть любую таблицу интеграла вѣроятностей [1]).

---

[1]) Теперь въ изданіи А. В. Леонтовича (цитир. выше) мы имѣемъ превосходныя таблицы Шеппарда, въ которыхъ значенія интеграла вѣроятностей даны въ функціи отношенія перемѣнной x къ ея среднему отклоненію. Если, какъ напр. въ таблицахъ, приложенныхъ къ „Исчисленію вѣроятностей" акад. Маркова, или къ извѣстной книгѣ А. А. Чупрова, аргументомъ служитъ x, дѣленный на величину модуля, то послѣдній легко найдемъ по среднему отклоненію, умноживъ его на $\sqrt{2}$. Иногда (таблицы Enke, см. Леонтовичъ op. cit. ч. I) нужно знать отношеніе x къ вѣроятной ошибкѣ. Послѣдняя равна среднему отклоненію, умноженному на 0,67449.

О значеніи интеграла вѣроятностей и о способахъ пользованія таблицами можно освѣдомиться въ цитированныхъ работахъ.

Законъ Гаусса не обладаетъ, однако, достаточной общностью. Въ самомъ дѣлѣ, каковы тѣ предпосылки, изъ которыхъ онъ выводится.

Каковъ бы ни былъ способъ его вывода, но всегда, явно или неявно, принимаются слѣдующія положенія:

(a) Уклоненія отъ средняго и въ сторону избытка и въ сторону недостатка одинаково вѣроятны.

(b) Присоединеніе новаго, какъ положительнаго, такъ и отрицательнаго отклоненія одинаково вѣроятно, независимо отъ величины суммы уже накопившихся отклоненій [1]).

Элементарный выводъ на основѣ этихъ предпосылокъ можетъ быть сдѣланъ слѣдующимъ образомъ [2]).

Пусть на величину отклоненія отъ средней оказываетъ вліяніе n элементарныхъ причинъ.

Каждая такая причина пусть вызываетъ уклоненіе равное ξ. Въ отдѣльномъ случаѣ мы будемъ имѣть r уклоненій положительныхъ и $n-r$ уклоненій отрицательныхъ. Полное уклоненіе будетъ равно

$$(13)\;.\;.\;.\left\{\begin{array}{ll} x_r \;\;= r\xi-(n-r)\xi=(2r-n)\xi & \text{и аналогично} \\ x_{r+1}=(r+1)\xi-(n-r-1)\xi=(2r+2-n)\xi, & \text{а разность}\end{array}\right.$$

$$(14)\;.\;.\;.\quad \triangle x_r = x_{r+1}-x_r=2\xi.$$

Какова вѣроятность уклоненій x_r и x_{r+1}?

Такъ какъ согласно допущенію положительныя и отрицательныя уклоненія одинаково вѣроятны, то, какъ учитъ теорія вѣроятности, вѣроятность того, что изъ двухъ равновозможныхъ событій одно произойдетъ r, а другое $n-r$ разъ, равна $r+1$-му члену разложенія бинома $(1/2+1/2)^n$, т. е. равна $\left(\frac{1}{2}\right)^n \frac{n!}{r!(n-r)!}$.

Въ предѣлѣ частоты пропорціональны вѣроятностямъ. Слѣдовательно, на N случаевъ уклоненіе равное x_r будетъ встрѣчаться число разъ:

$$(15)\;.\;.\;.\;.\left\{\begin{array}{ll} y_r \;\;= N\cdot\left(\frac{1}{2}\right)^n \frac{n!}{r!\,(n-r)!}, & \text{а уклоненіе равное } x_{r+1} \text{ — число разъ:} \\ y_{r+1}= N\cdot\left(\frac{1}{2}\right)^n \frac{n!}{(r+1)!\,(n-r-1)!} & \end{array}\right.$$

[1]) См. Pearson „Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson". A Rejoinder. Biometrica, Vol. IV, p. 189.

[2]) См. l. cit. примѣч. на стр. 179.

Второе уклоненіе будетъ встрѣчаться чаще на

$$\Delta y_r = y_{r+1} - y_r = N\left(\frac{1}{2}\right)^n \frac{n!}{r!(n-r-1)!}\left(\frac{1}{r+1} - \frac{1}{n-r}\right) =$$
$$= N\left(\frac{1}{2}\right)^n \frac{n!}{r!\,(n-r-1)!} \cdot \frac{n-2r-1}{(r+1)(n-r)}.$$

Ордината середины интервала между y_r и y_{r+1} равна ихъ полусуммѣ:

$$y_{r+\frac{1}{2}} = \frac{1}{2}(y_{r+1} + y_r) = N\left(\frac{1}{2}\right)^n \frac{n!}{r!(n-r-1)!}\ \frac{\frac{1}{2}(n+1)}{(r+1)(n-r)}.$$

Слѣдовательно,

$$\frac{\Delta y_r}{y_{r+\frac{1}{2}}} = \frac{(y_{r+1} - y_r)}{\frac{1}{2}\,(y_{r+1} + y_r)} = \frac{n-2r-1}{\frac{1}{2}(n+1)}.$$

Разность частотъ Δy_r соотвѣтствуетъ разности уклоненій отъ средняго $\Delta x_r = 2\xi$. Раздѣливъ обѣ части на это выраженіе, найдемъ:

$$(16) \quad \ldots\ldots\ldots \frac{1}{y_{r+\frac{1}{2}}}\ \frac{\Delta y}{\Delta x} = \frac{n-2r-1}{\frac{1}{2}(n+1)\,.\,2\xi}.$$

Это выраженіе можно представить въ другомъ видѣ. Именно, абсцисса, соотвѣтствующая ординатѣ $y_{r+\frac{1}{2}}$, равна (см. 13)

$x_{r+\frac{1}{2}} = \frac{1}{2}(x_{r+1} + x_r) = (2r - n + 1)\xi$, откуда $n - 2r - 1 = -\dfrac{x}{\xi}$.

Подставляя это выраженіе въ правую часть ур-ія (16), находимъ:

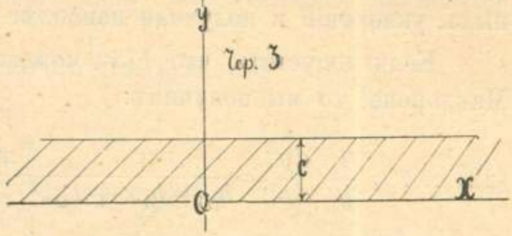
$$(17) \quad \ldots\ldots\ldots \frac{1}{y}\ \frac{\Delta y}{\Delta x} = -\frac{x}{(n+1)\xi^2}.$$

Перейдемъ къ предѣлу. Каково-бы ни было n, $x_r = 0$ при $r = \frac{1}{2}n$, т. е. въ томъ случаѣ, когда ровно половина элементарныхъ уклоненій положительна, а другая половина отрицательна. Если r не равно $\frac{1}{2}n$, то при $n = \infty$, какъ показываетъ ур-іе (13), x можетъ измѣняться отъ $+\infty$ до $-\infty$. Если бы пред. $(n+1)\xi^2 = \infty$, то при конечныхъ значеніяхъ x'а мы имѣли бы:

$\dfrac{\Delta y}{\Delta x} = 0$; $y = Const. = C$, и кривая распредѣленія превратилась бы въ прямую параллельную оси X. Т. к. все число индивидуумовъ равно площади кривой распредѣленія (см. чер. 3), а послѣдняя въ нашемъ случаѣ равна $C \times \infty$, то, для возможности равенства $C \times \infty = N$, необходимо, чтобы

$$C = 0.$$

Итакъ, въ случаѣ, когда пред. $(n+1)\xi^2 = \infty$, мы ни въ одномъ конечномъ интервалѣ не встрѣтимъ индивидуумовъ съ указаннымъ признакомъ. Слѣдов., если мы фактически наблюдаемъ ихъ, то это значитъ, что



$$(18) \; . \; . \; . \; . \; . \; \text{Пред. } (n+1)\xi^2 \Big|_{\substack{n=\infty \\ \xi=0}} = \text{конечному числу} = a^2.$$

Дифференціальное ур-іе кривой распредѣленія будетъ, слѣдоват., имѣть такой видъ:

$$(19) \; . \; . \; . \; . \; . \; . \; . \; . \; . \; \frac{1}{y}\,\frac{dy}{dx} = -\frac{x}{a^2},$$ откуда, интегрируя, безъ труда находимъ:

$$(20) \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; y = y_0 e^{-\frac{x^2}{2a^2}}.$$

Это и есть ур-іе Гауссовой кривой, или, какъ Пирсонъ называетъ ее, *нормальной* кривой распредѣленія.

Если мы отбросимъ сдѣланное выше допущеніе независимости элементарнаго уклоненія отъ суммы уже накопившихся уклоненій, т. е. отъ величины x_r, и наоборотъ допустимъ, что $\xi = f(x_r)$, то мы получимъ наиболѣе общую мыслимую зависимость между частотами и величиною признака. Такова идея Пирсона.

Въ нашемъ ур-іи (17) пред. $(n+1)\xi^2$ уже не будетъ равняться тогда нѣкоторой постоянной a^2, а будетъ какой то функціей x, и дифференціальное ур-іе кривой распредѣленія будетъ имѣть видъ:

$$\frac{1}{y}\,\frac{dy}{dx} = \frac{x}{F(x)},$$

или, если принять за начало координатъ произвольную точку:

$$(21) \quad \frac{1}{y}\,\frac{dy}{dx}=\frac{x-a}{F(x)}\,.$$

Отбросивъ предпосылку независимости элементарной ошибки отъ величины признака, мы тѣмъ самымъ освободились и отъ предпосылки равной вѣроятности положительныхъ и отрицательныхъ уклоненій и получили наиболѣе общую форму зависимости.

Если допустить, что $F(x)$ можетъ быть разложена по ряду Маклорена, то мы получимъ:

$$(22) \quad \frac{1}{y}\,\frac{dy}{dx}=\frac{x-a}{b_0+b_1x+b_2x^2+b_3x^3+b_4x^4+\ldots\ldots}\,.$$

Въ этомъ разложеніи можно взять любое число членовъ, но практически приходится ограничиваться тремя на основаніи слѣдующихъ соображеній. Чтобы найти ур-іе кривой, соотвѣтствующей статистическому матеріалу, нужно по методу Пирсона (см. ниже) вычислить фактическіе моменты и приравнять ихъ теоретическимъ. Для того, чтобы найти кривую типа

$$(23) \quad \frac{1}{y}\,\frac{dy}{dx}=\frac{x-a}{b_0+b_1x+b_2x^2}\,,$$

нужно знать четыре момента фактическаго распредѣленія, а для того, чтобы найти кривую большей общности, необходимы моменты 5-го, 6-го и высшихъ порядковъ. Но, какъ показалъ Пирсонъ [1]), вѣроятныя ошибки моментовъ выше 4-го очень велики и быстро возрастаютъ съ возрастаніемъ порядка момента, поэтому и коэффиціенты кривой, вычисленные при ихъ посредствѣ, также должны быть мало достовѣрны.

Не смотря на указанное ограниченіе, кривыя Пирсона, какъ показалъ большой статистическій опытъ его и его школы, даютъ почти всегда прекрасные результаты, передавая особенности матеріала въ тѣхъ случаяхъ, когда нормальная (Гауссова) кривая отказывается служить статистику.

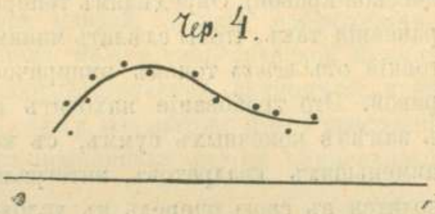
---

[1]) K. Pearson, On the General Theory of Skew Correlation and Non-Linear Regression. Drapers' Company Research Memoirs. Biom. Series II, London, 1905, p. 7—8.

§ 6. Основаніе метода моментовъ.

Какимъ образомъ, однако, статистикъ можетъ воспользоваться теоретической кривой для изображенія своего матеріала? Для этого нужно теоретической кривой придать окончательную форму, вычисливъ коэффиціенты ея по статическимъ даннымъ. Исторически первое рѣшеніе дано было методомъ наименьшихъ квадратовъ.

Идея его заключается въ слѣдующемъ:



Пусть (черт. 4) наблюденіе дало рядъ точекъ, и пусть мы хотимъ найти коэффиціенты ур-ія $y = f(x, a_1, a_2 \ldots a_n)$ [1]) такъ, чтобы получившаяся кривая какъ можно тѣснѣе прилегала къ нашимъ точкамъ. Для этого, по методу наименьшихъ квадратовъ, нужно найти разстоянія между соотвѣтствующими точками и кривой и опредѣлить коэффиціенты ур-ія такъ, чтобы сумма квадратовъ этихъ разстояній была наименьшая.

Недостатокъ этого метода заключается въ томъ, что онъ требуетъ очень большихъ вычисленій даже для параболическихъ

---

[1]) Какое ур-іе взять для этой цѣли зависитъ отъ многихъ обстоятельствъ, и теоретически нельзя дать общаго правила для выбора типа кривой. Въ однихъ случаяхъ парабола n-го порядка ($y = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$) даетъ хорошій результатъ, въ другихъ случаяхъ лучше воспользоваться тригонометрической или показательной кривой. Лучшей кривой нужно считать ту, которая тѣснѣе примыкаетъ къ эмпирической линіи, и для опредѣленія которой нужно вычисленіе меньшаго числа коэффиціентовъ. Не надо думать, что увеличеніе ихъ числа всегда приноситъ значительную пользу, и что, взявъ параболу 6-го, 8-го, 10-го порядка, мы непремѣнно должны получить хорошо подходящую кривую. Важнѣе выборъ типа кривой. Пирсонъ указываетъ, напр., что парабола 6-го порядка съ 7-ю параметрами хуже подходитъ къ эмпирическимъ даннымъ разработаннаго имъ примѣра, чѣмъ его кривая распредѣленія съ 3-мя параметрами. (См. Biometrica Vol. II p. 16—19).

О раціональной мѣрѣ степени соотвѣтствія теоретической кривой съ эмпирической (K. Pearson, On the Criterion that a given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling, Phil. Mag. Vol. 50, 1900, p. 157—175) см. послѣднюю главу настоящей работы.

кривыхъ, а къ приспособленію кривыхъ многихъ другихъ типовъ практически совсѣмъ не приложимъ или приложимъ лишь съ затратой совершенно непосильнаго труда. Нахожденіе вѣроятныхъ ошибокъ найденныхъ коэффиціентовъ во многихъ случаяхъ также или невозможно или очень затруднительно.

Пирсонъ предложилъ слѣдующее видоизмѣненіе этого метода. Вообразимъ себѣ, что намъ даны не отдѣльно расположенныя точки, а непрерывная эмпирическая кривая. (Въ случаѣ отдѣльныхъ точекъ соединимъ ихъ какъ можно болѣе плавной параболической кривой). Опредѣлимъ теперь коэффиціенты теоретическаго уравненія такъ, чтобы сдѣлать минимумомъ сумму квадратовъ разстояній отъ *всѣхъ* точекъ эмпирической кривой до теоретической кривой. Это требованіе находитъ свое аналитическое выраженіе въ замѣнѣ конечныхъ суммъ, съ которыми имѣетъ дѣло методъ наименьшихъ квадратовъ, интегралами и, какъ можно показать, сводится въ свою очередь къ условію равенства моментовъ обѣихъ кривыхъ.

Слѣдующее теоретическое обоснованіе метода дано Пирсономъ (On the Systematic Fitting of Curves, Biometrica Vol. I, p. 267—271). Читатель, не знакомый съ высшей математикой, можетъ пропустить его безъ ущерба для дальнѣйшаго.

Пусть сдѣлана серія измѣреній или наблюденій величинъ перемѣнной y, соотвѣтствующихъ ряду значеній второй перемѣнной x, заключающемуся между $-l$ и $+l$. Требуется найти хорошій методъ приспособленія къ даннымъ наблюденія теоретической или эмпирической кривой $y=\varphi(x, c_1, c_2, c_3, \ldots\ldots c_n)$, гдѣ c_1, c_2, $c_3 \ldots\ldots c_n$ суть произвольныя постоянныя.

Сдѣлаемъ допущенія (1), что функція $\varphi(x)$ можетъ быть разложена по формулѣ Маклорена и (2), что получившійся рядъ болѣе или менѣе быстро сходится.

Пусть разложеніе будетъ:

$$y=\varphi(0)+x\varphi'(0)+\frac{x^2}{1\,.\,2}\varphi''(0)+\ldots\ldots=$$

$$=\alpha_0+\alpha_1 x+\alpha_2\frac{x^2}{1\,.\,2}+\alpha_3\frac{x^3}{1\,.\,2\,.\,3}\ldots\ldots,\text{ гдѣ}$$

α_0, α_1, $\alpha_2 \ldots\ldots$ суть функціи n параметровъ кривой c_1, $c_2 \ldots c_n$. Теоретически возможно поэтому опредѣлить всѣ n величинъ

$c_1, c_2 \ldots c_n$ въ функціи n первыхъ коэффиціентовъ разложенія $\alpha_0, \alpha_1 \ldots \alpha_{n-1}$ и уже черезъ нихъ выразить всѣ остальные коэффиціенты разложенія $\alpha_n, \alpha_{n+1} \ldots\ldots$ Теоретически, слѣд., нашу кривую можно представить въ такомъ видѣ:

$$y = \alpha_0 + \alpha_1 x + \alpha_2 \frac{x^2}{1.2} + \alpha_3 \frac{x^3}{1.2.3} + \ldots\ldots + \alpha_{n-1} \frac{x^{n-1}}{(n-1)!} +$$

$$+ \varphi^{(n)}(\alpha_0, \alpha_1 \ldots \alpha_{n-1}) \frac{x^n}{n!} + \varphi^{(n+1)}(\alpha_0, \alpha_1 \ldots \alpha^{n-1}) \frac{x^{n+1}}{(n+1)!} + \ldots etc.$$

Пусть Y будетъ ордината кривой, данная наблюденіемъ; $y - Y$ будетъ тогда разстояніемъ между теоретической и эмпирической кривыми въ точкѣ, соотвѣтствующей данному x, и нашей задачей является сдѣлать сумму квадратовъ этихъ разстояній возможно меньшей. Согласно съ принципами метода наименьшихъ квадратовъ положимъ:

$$\int (y - Y)^2 dx = \text{минимуму}.$$

Если мы возьмемъ варіацію отъ y, то получимъ разрѣшающее вопросъ ур-іе:

$$(24) \ldots\ldots\ldots \int (y - Y)\, \delta y dx = 0.$$

δy, очевидно, зависитъ отъ варіацій $\alpha_0 \ldots \alpha_{n-1}$, именно:

$$\delta y = \delta\alpha_0 + \delta\alpha_1 x + \delta\alpha_2 \frac{x^2}{1.2} + \delta\alpha_3 \frac{x^3}{1.2.3} + \ldots + \delta\alpha_{n-1} \frac{x^{n-1}}{(n-1)!} +$$

$$+ \left(\frac{d\varphi^{(n)}}{d\alpha_0}\delta\alpha_0 + \frac{d\varphi^{(n)}}{d\alpha_1}\delta\alpha_1 + \ldots + \frac{d\varphi^{(n)}}{d\alpha_{n-1}}\delta\alpha_{n-1}\right)\frac{x^n}{n!} +$$

$$+ \left(\frac{d\varphi^{(n+1)}}{d\alpha_0}\delta\alpha_0 + \frac{d\varphi^{(n+1)}}{d\alpha_1}\delta\alpha_1 + \ldots + \frac{d\varphi^{(n+1)}}{d\alpha_{n-1}}\delta\alpha_{n-1}\right)\frac{x^{n+1}}{(n+1)!} + \ldots etc\ldots$$

$$= \delta\alpha_0\left(1 + \frac{d\varphi^{(n)}}{d\alpha_0}\frac{x^n}{n!} + \frac{d\varphi^{(n+1)}}{d\alpha_0}\frac{x^{n+1}}{(n+1)!} + \ldots\ldots\right) +$$

$$+ \delta\alpha_1\left(x + \frac{d\varphi^{(n)}}{d\alpha_1}\frac{x^n}{n!} + \frac{d\varphi^{(n+1)}}{d\alpha_1}\frac{x^{n+1}}{(n+1)!} + \ldots\ldots\right) +$$

$$+ \delta\alpha_2\left(\frac{x^2}{1.2} + \frac{d\varphi^{(n)}}{d\alpha_2}\cdot\frac{x^n}{n!} + \frac{d\varphi^{(n+1)}}{d\alpha_2}\frac{x^{n+1}}{(n+1)!} + \ldots\right) +$$

$$+ etc.$$

Замѣтимъ, что $\varphi^{(n)}(0, \alpha_0 \ldots . \alpha_{n-1}) \dfrac{x^n}{n!} + \varphi^{(n+1)}(0, \alpha_0 \ldots \alpha_{n-1}) \cdot \dfrac{x^{n+1}}{(n+1)!} + \ldots = R = \dfrac{x^n}{n!} \varphi^{(n)}(\theta x)$, гдѣ θ заключается между 0 и 1, а R есть остаточный членъ ряда Маклорена. Съ этимъ обозначеніемъ мы имѣемъ:

$$\delta y = \delta\alpha_0 \left[1 + \frac{dR}{d\alpha_0}\right] + \delta\alpha_1 \left[x + \frac{dR}{d\alpha_1}\right] + \delta\alpha_2 \left[\frac{x^2}{1.2} + \frac{dR}{d\alpha_2}\right] + \ldots etc.$$

Подставляя въ (24) и замѣчая, что варіація $\delta\alpha$ совершенно произвольна и можетъ быть вынесена поэтому изъ подъ знака интеграла, мы получаемъ:

$$\left[\int (y - Y)\left(1 + \frac{dR}{d\alpha_0}\right) dx\right] \delta\alpha_0 + \left[\int (y - Y)\left(x + \frac{dR}{d\alpha_1}\right) dx\right] \delta\alpha_1 + + etc = 0.$$

Но т. к. $\delta\alpha_0$, $\delta\alpha_1 \ldots$ вполнѣ произвольны, то, чтобы удовлетворить этому ур-ію, коэффиціенты при $\delta\alpha_0$, $\delta\alpha_1 \ldots$ должны независимо обратиться въ нули. Откуда имѣемъ:

$$(25) \ldots \ldots \left\{ \begin{aligned} &\int (y - Y)\left(1 + \frac{dR}{d\alpha_0}\right) dx = 0 \\ &\int (y - Y)\left(x + \frac{dR}{d\alpha_1}\right) dx = 0 \\ &\int (y - Y)\left(\frac{x^2}{1.2} + \frac{dR}{d\alpha_2}\right) dx = 0 \\ &\int (y - Y)\left(\frac{x^3}{1.2.3} + \frac{dR}{d\alpha_3}\right) dx = 0 \\ &etc \ldots\ldots \end{aligned} \right.$$

Пусть $A =$ площади, μ_1, $\mu_2 \ldots$ пусть будутъ первымъ, вторымъ и т. д. моментами теоретической кривой; A', μ'_1, $\mu'_2 \ldots$ — площадью и моментами кривой, данной наблюденіемъ. Моменты взяты около оси Y. Тогда ур-ія (25) могутъ быть написаны такъ:

$$(26) \ldots \begin{cases} A = A' - \int (y-Y)\dfrac{dR}{d\alpha_0}\,dx \\ A\mu_1 = A'\mu'_1 - \int (y-Y)\dfrac{dR}{d\alpha_1}\,dx \\ A\mu_2 = A'\mu'_2 - 2!\int (y-Y)\dfrac{dR}{d\alpha_2}\,dx \\ A\mu_3 = A'\mu'_3 - 3!\int (y-Y)\dfrac{dR}{d\alpha_3}\,dx \\ \ldots\ldots\ldots\ldots\ldots \\ A\mu_{n-1} = A'\mu'_{n-1} - (n-1)!\int (y-Y)\dfrac{dR}{d\alpha_{n-1}}\,dx. \end{cases}$$

Члены съ интегралами должны быть однако малы: (1) такъ какъ они заключаютъ небольшой множитель $(y-Y)$; (2) т. к. $R = \frac{x^n}{n!}\,\varphi^{(n)}(\theta x)$ согласно гипотезѣ при достаточно большомъ n очень незначительно.

Итакъ, пренебрегая членами, заключающими интегралы, получаемъ:

$$(27) \ldots\ldots\ldots \begin{cases} A = A' \\ \mu_1 = \mu'_1 \\ \mu_2 = \mu'_2 \\ \mu_3 = \mu'_3 \\ \ldots\ldots \\ \mu_{n-1} = \mu'_{n-1} \end{cases}$$

A, μ_1 и т. д. выражены при помощи α_0, α_1 .... и т. д., но ничто не препятствуетъ намъ, разъ результатъ полученъ, выразить ихъ опять черезъ параметры кривой $c_1, c_2, \ldots c_n$. Итакъ мы получаемъ слѣдующее правило:

Для того, чтобы подобрать хорошую теоретическую кривую $y = \varphi(x, c_1, c_2 \ldots\ldots c_n)$ къ кривой, данной опытомъ, нужно площадь и моменты первой кривой выразить черезъ ея параметры $(c_1, c_2 \ldots c_n)$ и приравнять ихъ площади и моментамъ эмпирической кривой.

Рѣшеніе (27) является приближеніемъ большей степени, чѣмъ можетъ показаться на первый взглядъ. Ибо, разъ площадь и первые $n-1$ моменты двухъ кривыхъ равны, то и слѣдующіе моменты будутъ приблизительно равны, и въ тѣмъ большей степени, чѣмъ больше n. Но членъ $\int(y-Y)\frac{dR}{d\alpha_s}dx$ исчезаетъ, если высшіе моменты равны. Въ самомъ дѣлѣ, мы можемъ написать:

$$R=\frac{x^n}{n!}\varphi^{(n)}(0)+\frac{x^{n+1}}{(n+1)!}\varphi^{(n+1)}(0)+\ldots\ldots, \text{ а слѣдов.:}$$

$$\int(y-Y)\frac{dR}{d\alpha_s}dx=\frac{d\varphi^{(n)}(0)}{d\alpha_s}\,\frac{1}{n!}(A\mu_n-A'\mu'_n)+$$
$$+\frac{d\varphi^{(n+1)}(0)}{d\alpha_s}\,\frac{1}{(n+1)!}(A\mu_{n+1}-A'\mu'_{n+1})+$$
$$+\ldots\ldots\ldots\ldots\ldots$$

Т. обр. если $A=A'$, то мы имѣемъ множители $\mu_n-\mu'_n$, $\mu_{n+1}-\mu'_{n+1}$ и т. д., величина которыхъ очень мала, и которые умножены на величины сами по себѣ въ силу сходимости ряда Маклорена очень малыя: $\frac{1}{n!}\,\frac{d\varphi^{(n)}(0)}{d\alpha_s}$, $\frac{1}{(n+1)!}\cdot\frac{d\varphi^{(n+1)}(0)}{d\alpha_s}$ и т. д.

Изъ всего этого мы заключаемъ, что равенство моментовъ есть хорошій методъ приспособленія кривыхъ къ даннымъ опыта, и дѣйствительно, какъ показала практика, онъ не хуже метода наименьшихъ квадратовъ. Для параболъ этотъ методъ совпадаетъ съ методомъ наименьшихъ квадратовъ, т. к. даетъ рѣшенія точныя, а не приближенныя, ибо въ этомъ случаѣ разложеніе Маклорена заключаетъ конечное число членовъ. Какъ упоминалось выше методъ моментовъ обладаетъ еще преимуществомъ болѣе легкой приложимости, т. к. онъ оказывается возможнымъ и въ тѣхъ даже случаяхъ, когда первый методъ или абсолютно неприменимъ, или-же требуетъ непомѣрной затраты труда на вычислительную работу. Кромѣ того, во всѣхъ случаяхъ, когда его можно примѣнить, можно найти и вѣроятныя ошибки, полученныхъ коэффиціентовъ.

Для приложенія этого метода необходимо рѣшить слѣдующія задачи:

(1) Нужно умѣть найти моменты любой эмпирической системы наблюденій.

(2) Нужно выразить моменты теоретической кривой въ функціи параметровъ $c_1, c_2, c_3 \ldots . c_n$.

(3) Эти выраженія для моментовъ должны быть таковы, чтобы можно было безъ слишкомъ большого труда разрѣшить систему ур-ій (27).

Покажемъ, какъ рѣшается первая задача.

§ 7. *Нахожденіе эмпирическихъ моментовъ.*

Моменты, найденные по правилу § 2, Пирсонъ называетъ грубыми (raw) и обозначаетъ буквами $\nu_1, \nu_2 \ldots$, если они опредѣлены относительно центра распредѣленія, и $\nu'_1, \nu'_2 \ldots$, если за начало была принята какая нибудь другая точка. Неточность этого способа заключается въ томъ, что мы производимъ вычисленіе такъ, какъ если бы вся сумма индивидуумовъ, приходящихся на интервалъ, была расположена въ серединѣ его; только при безконечно малыхъ интервалахъ этотъ методъ въ предѣлѣ приводитъ къ истиннымъ моментамъ $\mu_1, \mu_2, \mu_3 \ldots$ На практикѣ-же интервалы рѣдко бываютъ настолько малыми, чтобы проистекающей изъ этого способа вычисленія ошибкой можно было бы пренебречь.

При вычисленіи моментовъ мы будемъ различать три случая.

(A) Эмпирическая кривая плавно спускается къ оси X съ обѣихъ сторонъ. Число индивидуумовъ въ крайнихъ группахъ уменьшается настолько постепенно, что кривая, говоря математическимъ языкомъ, имѣетъ съ осью X *на обоихъ* концахъ распредѣленія *касаніе безконечно большого порядка.* Такую кривую Пирсонъ называетъ квази-нормальной (т. к. подобнымъ свойствомъ обладаетъ м. пр. и нормальная кривая Гаусса). Въ этомъ случаѣ истинные моменты легко находятся изъ грубыхъ прибавленіемъ къ нимъ поправокъ Sheppard'а. Выведемъ ихъ [1]).

Пусть ур-іе кривой распредѣленія будетъ

$$y = \varphi(x),$$

гдѣ ydx есть число индивидуумовъ въ подгруппѣ между x и $x + dx$. Пусть k будетъ единица основанія кривой, принятая за интервалъ при группировкѣ сырого матеріала. N — полная численность данной совокупности, а n_r—численность подгруппы r-ой между

---

[1]) On an Elementary Proof of Sheppard's Formulae for correcting Raw Moments and on other allied Points (Editorial), Biometrica, Vol. III p. 308—312). Читатель не математикъ можетъ пропустить выводъ и обратить вниманіе только на результатъ: формулы (34).

$x_r - \frac{1}{2}k$ и $x_r + \frac{1}{2}k$. Пусть y_r' будетъ ордината соотвѣтствующая $x_r + x'$. Тогда:

$$(28) \quad . \ . \ . \ . \ n_r = \int_{-\frac{1}{2}k}^{+\frac{1}{2}k} y_r' dx' = \int_{-\frac{1}{2}k}^{+\frac{1}{2}k} \varphi(x_r + x') dx'.$$

Если $\varphi(x)$ есть непрерывная функція, которая можетъ быть разложена по ряду Тейлора, то

$$n_r = \int_{-\frac{1}{2}k}^{+\frac{1}{2}k} \left[\varphi(x_r) + x'\varphi'(x_r) + \frac{x'^2}{1.2}\varphi''(x_r) + \ . \ . \ . \ . \right] dx'.$$

Интегрируя это выраженіе и замѣчая при этомъ, что $\varphi^{(i)}(x_r)$ есть постоянная, сразу получимъ:

$$n_r = k\varphi(x_r) + \frac{k^3}{24}\varphi''(x_r) + \frac{k^5}{1920}\varphi^{IV}(x_r) + \ . \ . \ . \ . +$$
$$+ \frac{k^{2n+1} \cdot \varphi^{(2n)}(x_r)}{\frac{1}{2}(2n+1)!\, 2^{2n+1}} + etc.$$

Здѣсь x_r есть очевидно любая абсцисса, и это выраженіе даетъ численность подгруппы, соотвѣтствующей любому интервалу, величина котораго k, а средняя абсцисса x_r. Отсюда слѣдуетъ, что, если знакъ суммы распространить на всѣ значенія r, то грубый моментъ s-аго порядка выразится слѣд. образомъ:

$$(29) \ . \ . \ . \ \Sigma(n_r x_r^s) = k\Sigma\left[\varphi(x_r)x_r^s\right] + \frac{k^3}{24}\Sigma\left[\varphi''(x_r)x_r^s\right] +$$
$$+ \frac{k^5}{1920}\Sigma\left[\varphi^{IV}(x_r)\, x_r^s\right] + \text{ и т. д.}$$

По извѣстной формулѣ Эйлера—Маклорена [1]), если $f(x)$ есть любая непрерывная функція отъ x, то

$$\int f(x)dx = \Sigma[kf(x)] + \left[\frac{k}{2}f(x) - \frac{k^2}{12}f'(x) + \frac{k^4}{720}f'''(x) + \ . \ . \ . \right],$$

гдѣ, какъ сумма, такъ и члены въ квадратныхъ скобкахъ берутся по предѣламъ интегрированія (т. е. вмѣсто x подставляется верхній и нижній предѣлы интеграла и затѣмъ берется разность обоихъ выраженій).

Если кривая имѣетъ на предѣлахъ касаніе безконечно большого порядка съ осью x-овъ, то всѣ члены суммы въ квадратныхъ скобкахъ, т. е. $f(x)$, $f'(x)$, $f'''(x)$ и т. д., на предѣлахъ обращаются въ ноль, и мы имѣемъ:

---

[1]) А. Марковъ. Исчисленіе конечныхъ разностей. Изд. 2, 1911, стр. 142.

$$(30) \quad \dots\dots\dots \quad \int f(x)dx = \Sigma\,[kf(x)]\,.$$

Предположимъ теперь, что $\varphi^{(p)}(x)x^s$ есть функція, которая вмѣстѣ со всѣми своими производными исчезаетъ на обоихъ предѣлахъ распредѣленія.

Тогда, примѣняя ур-іе (30) къ (29)-му получимъ:

$$(31) \quad \dots \quad \Sigma(n_r x_r^s) = \int\varphi(x)x^s dx + \frac{k^2}{24}\int\varphi''(x)x^s dx + \\ + \frac{k^4}{1920}\int\varphi^{IV}(x)x^s dx + \dots$$

Интегрируя по частямъ, въ случаѣ $p > s$, получимъ:

$$\int\varphi^{(p)}(x)x^s dx = [\varphi^{(p-1)}(x).x^s - s\varphi^{(p-2)}(x).x^{s-1} + s(s-1)\varphi^{(p-3)}(x).x^{s-2} - \\ - \dots + (-1)^s s(s-1)(s-2)\dots 3.2.1.\varphi^{(p-s-1)}(x)] = 0,$$

т. к. на предѣлахъ интегрированія $\varphi(x) = 0$, $\varphi'(x) = 0$ и т. д.

Если $p \leqslant s$, то интегрируя по частямъ, кромѣ членовъ, обращающихся на предѣлахъ въ нуль, получимъ еще интегралъ, т. что

$$\int\varphi^{(p)}(x)x^s dx = (-1)^p.s(s-1)(s-2)\dots(s-p+1)\int\varphi(x)x^{s-p}dx,$$

т. е. сведемъ нашъ интегралъ къ $(s-p)$-му моменту.

Подставляя въ (31) вмѣсто s послѣдовательно 0, 1, 2 . . . и ограничиваясь, по доказанному, членами въ которыхъ $p \leqslant s$, получимъ:

$$(32) \quad \dots \left\{ \begin{aligned} \Sigma(n_r) &= \int\varphi(x)dx = N \\ \Sigma(n_r x_r) &= \int\varphi(x)x\,dx = N\mu'_1 \\ \Sigma(n_r x_r^2) &= \int\varphi(x)x^2dx + 2\frac{Nk^2}{24} = N\left(\mu'_2 + \frac{1}{12}k^2\right) \\ \Sigma(n_r x_r^3) &= \int\varphi(x)x^3dx + 6\frac{k^2}{24}\int\varphi(x)x\,dx = N\left(\mu'_3 + \frac{k^2\mu'_1}{4}\right) \\ \Sigma(n_r x_r^4) &= \int\varphi(x)x^4dx + 12\frac{k^2}{24}\int\varphi(x)x^2dx + \frac{24k^4}{1920}N = \\ &= N\left(\mu'_4 + \frac{1}{2}k^2\mu'_2 + \frac{1}{80}k^4\right). \end{aligned} \right.$$

Здѣсь μ'_1, μ'_2, μ'_3, μ'_4 суть истинные моменты около оси Y. Если мы ее проведемъ черезъ центръ распредѣленія, то $\mu'_1 = 0$, и, обозначая истинные центральные моменты черезъ μ_1, μ_2, μ_3, μ_4, а черезъ ν_1, ν_2, ν_3, ν_4 центральные грубые моменты, мы найдемъ въ силу равенства

$$\Sigma(n_r x_r^s) = N\nu_s, \quad \text{что}$$

$$(33) \quad \ldots\ldots \quad \begin{cases} \mu_0 = \nu_0 = 1 \\ \mu_1 = \nu_1 = 0 \\ \mu_2 = \nu_2 - \dfrac{k^2}{12} \\ \mu_3 = \nu_3 \\ \mu_4 = \nu_4 - \dfrac{k^2}{2}\nu_2 + \dfrac{7k^4}{240} \end{cases}$$

или полагая $k = 1$ (для удобства вычисленія интервалъ всегда слѣдуетъ принимать равнымъ единицѣ):

$$(34) \quad \ldots\ldots \quad \begin{cases} \mu_0 = \nu_0 = 1 \\ \mu_1 = \nu_1 = 0 \\ \mu_2 = \nu_2 - \dfrac{1}{12} \\ \mu_3 = \nu_3 \\ \mu_4 = \nu_4 - \dfrac{1}{2}\nu_2 + \dfrac{7}{240}. \end{cases}$$

Это и будутъ поправки Шеппарда къ грубымъ моментамъ.

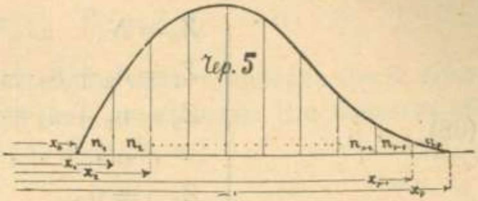
Ходъ вычисленія, слѣдовательно, таковъ:

Сначала находимъ грубые моменты, затѣмъ по формуламъ (7) находимъ грубые центральные, а затѣмъ по (34), вводя поправки Шеппарда, получаемъ истинные центральные моменты. Все вычисленіе ведемъ, принимая величину интервала (k) за единицу, а по окончаніи вычисленій умножаемъ p-ый моментъ на k^p.

(B) Если по характеру матеріала видно, что кривая распредѣленія не имѣетъ касанія безконечно высокаго порядка съ осью X, въ особенности, если кривая пересѣкаетъ ось x подъ конеч-

нымъ угломъ, то поправокъ Шеппарда прилагать нельзя. Тогда для вычисленія моментовъ рекомендуется другой способъ [1]).

Пусть $y = \varphi(x)$ будетъ, какъ раньше, кривая, изображающая распредѣленіе. На оси x отложимъ интервалы, первый отъ x_0 до x_1, второй отъ x_1 до x_2 и т. д. Величины x со значками обозначаютъ у насъ теперь, слѣдовательно, разстоянія не до середины интервала, какъ раньше, а до границъ его. Обозначимъ



черезъ n_r численность группы въ интервалѣ x_{r-1} до x_r. Эти величины $(n_1, n_2 \ldots . n_r \ldots . n_p)$ даны изъ опыта и равны, очевидно,

$$(35) \ldots n_1 = \int_{x_0}^{x_1} y dx, \quad n_2 = \int_{x_1}^{x_2} y dx, \ \ldots \ n_p = \int_{x_{p-1}}^{x_p} y dx.$$

Пусть все число наблюденій будетъ

$$N = n_1 + n_2 + \ldots + n_p .$$

Для n-аго момента около линіи перпендикулярной къ оси X *въ началѣ* координатъ имѣемъ:

$$(36) \ldots\ldots\ldots\ldots N\mu'_n = \int_{x_0}^{x_p} x^n y dx .$$

Введемъ новую перемѣнную

$$(37) \ldots\ldots\ldots\ldots Z = \int_{x}^{x_p} y dx.$$

Z представляетъ изъ себя, очевидно, часть площади нашей кривой, т. е. число индивидуумовъ съ величиной признака, заключенной между нѣкоторымъ x и верхнимъ предѣломъ x_p.

---

[1]) K. Pearson. On the systematic Fitting of Curves, Biometrica Vol. I p. 282 и сл.

Для не математика этотъ способъ окажется можетъ быть труднымъ. Вообще его слѣдуетъ примѣнять лишь при изслѣдованіяхъ требующихъ большой точности. Въ остальныхъ случаяхъ при непримѣнимости поправокъ Шеппарда можно употреблять поправки, изложенные ниже подъ (C).

Тогда

$$Z_0=\int_{x_0}^{x_p} ydx,\ Z_1=\int_{x_1}^{x_p} ydx,\ \ldots Z_p=\int_{x_p}^{x_p} ydx.$$

Очевидно, что

$$(38) \ldots\ldots \left\{ \begin{array}{l} Z_0=N \\ Z_1=n_2+n_3+\ldots+n_p \\ Z_2=n_3+n_4+\ldots+n_p \\ \ldots\ldots\ldots\ldots \\ Z_{p-1}=n_p \\ Z_p=0. \end{array} \right.$$

Дифференцируя $Z=\int_{x}^{x_p} ydx$, получаемъ

$$\frac{dZ}{dx}=-y,$$

а слѣдовательно:

$$(39) \ldots N\mu'_n=\int_{x_0}^{x_p} x^n ydx=-\int_{x_0}^{x_p} x^n\frac{dZ}{dx}dx=-\int_{x_0}^{x_p} x^n dZ.$$

Интегрируя по частямъ, будемъ имѣть:

$$N\mu'_n=-\Big[Zx^n\Big]_{x_0}^{x_p}+n\int_{x_0}^{x_p} Zx^{n-1}dx=Z_0x_0{}^n+n\int_{x_0}^{x_p} Zx^{n-1}dx.$$

Такимъ образомъ,

$$(40) \ldots\ldots\ldots \mu'_n=x_0{}^n+\frac{n}{N}\int_{x_0}^{x_p} Zx^{n-1}dx:$$

Величину признака можно измѣрять разностью между данной величиной и нѣкоторой постоянной, принятой за начало счета. Начало счета можно взять въ началѣ распредѣленія. Тогда $x_0=0$, а слѣд.:

$$(41) \ldots\ldots\ldots \mu'_n=\frac{n}{N}\int_{x_0}^{x_p} Zx^{n-1}dx.$$

Это основная формула для нахожденія истинныхъ моментовъ въ нашемъ случаѣ. Правило очевидно.

Для $x=x_0=0$, $x=x_1$, $x=x_2$, ... $x=x_p$ вычисляемъ значенія $Z_0=N$, $Z_1=N-n_1$, $Z_2=N-(n_1+n_2)$ . . . $Z_p=0$. Затѣмъ для нахожденія n-аго момента вычисляемъ рядъ величинъ:

$$Y_0=Z_0 x_0^{n-1}=0,\ \ Y_1=Z_1 x_1^{n-1},\ \ Y_2=Z_2 x_2^{n-1}\ \ldots\ldots$$

$$Y_{p-1}=Z_{p-1}x_{p-1}^{n-1},\ \ Y_p=Z_p x_p^{n-1}=0\,.$$

Величины Y_i будемъ разсматривать, какъ ординаты новой, вспомогательной кривой и опредѣлимъ ея площадь. Эта площадь (S) и будетъ интеграломъ формулы (41), т. к.

$\int_{x_0}^{x_p} Zx^{n-1}dx=\int_{x_0}^{x_p} Ydx=S$. Затѣмъ по формулѣ (41) будемъ имѣть окончательно искомый моментъ:

$$(42)\ \ldots\ldots\ldots\ldots\ \mu'_n=\frac{n}{N}S\,.$$

Площадь S можно опредѣлить по любой хорошей формулѣ квадратуръ. Для этого, соединяя вершины ординатъ хордами, найдемъ „хордальную" площадь, какъ сумму площадей трапецій:

$$S_c=k\left(\frac{1}{2}Y_0+Y_1+Y_2+\ \ldots\ +Y_{p-1}+\frac{1}{2}Y_p\right),$$

гдѣ k, какъ и раньше—интервалъ между ординатами. Т. к. у насъ $Y_0=0$ и $Y_p=0$, то

$$(43)\ \ldots\ldots\ S_c=k(Y_1+Y_2+\ \ldots\ +Y_{p-1})\,.$$

Для нахожденія по „хордальной" площади истинной площади кривой существуетъ множество формулъ. Пирсонъ предлагаетъ м. проч. слѣдующую, очень точную [1]):

$$(44)\ldots S=S_c+\frac{1}{120}\,\frac{p(15p-26)}{(p-1)(p-2)}\,[(Y_1-Y_0)-(Y_p-Y_{p-1})]\,k-$$

$$-\frac{1}{120}\,\frac{p(5p-6)}{(p-2)(p-3)}\,[(Y_2-Y_1)-(Y_{p-1}-Y_{p-2})]\,k,$$

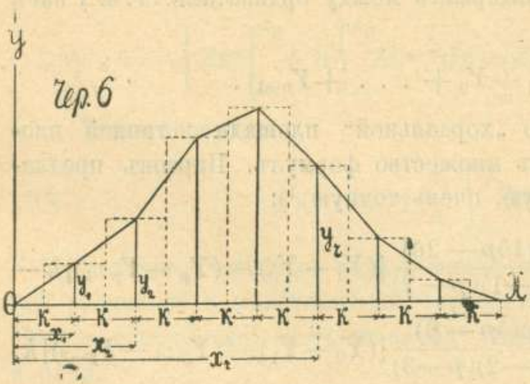
гдѣ согласно нашимъ условіямъ Y_0 и Y_p также нужно положить равными нулю.

---

[1]) Одна изъ формулъ Sheppard'a. Доказательство см. въ Lond. Math. Soc. Proc. Vol. 32 p. 270. Цитирую по Pearson'у Biometrica Vol. I p. 275.

(C) Третій случай представляется тогда, когда числа, данныя изъ опыта, располагаются настолько неправильно, что примѣненіе формулы для криволинейныхъ квадратуръ можетъ почитаться излишнимъ. Всѣ тѣ изгибы параболической кривой, которые даетъ формула вродѣ Симпсоновой, не имѣютъ все равно никакого реальнаго значенія, т. к. обязаны своимъ происхожденіемъ не существу явленія, а случайнымъ неправильностямъ въ опытномъ матеріалѣ. Тогда оказывается болѣе правильнымъ разсматривать эмпирическую кривую такъ, какъ она дана, т. е. какъ ломанную линію, и задача сводится къ нахожденію моментовъ площади, составленной изъ трапецій. Способъ, изложенный подъ *(B)*, примѣнимъ и здѣсь, но Пирсонъ далъ для этого случая готовыя формулы, позволяющія сразу перейти отъ центральныхъ грубыхъ къ истиннымъ центральнымъ моментамъ. Кромѣ того, т. к. этотъ способъ въ своемъ примѣненіи гораздо проще второго способа *(B)*, то мы можемъ употреблять его во всѣхъ случаяхъ, когда поправки Шеппарда не приложимы, и когда представляется излишней та нѣсколько бо́льшая точность, которую даетъ второй методъ.

Въ параграфѣ 1-мъ мы изображали эмпирическое распредѣленіе ступенчатой фигурой, составленной изъ прямоугольниковъ. Нѣсколько бо́льшее приближеніе къ плавной кривой частоты представляетъ изъ себя изображеніе того-же матеріала т. наз. эмпирическимъ полигономъ частоты (распредѣленія). Вообразимъ себѣ по краямъ нашей ступенчатой фигуры еще по одному прямоугольнику нулевой высоты и соединимъ затѣмъ (см. черт. 6) середины верхнихъ основаній *всѣхъ* прямоугольниковъ прямыми. Полученная фигура и будетъ полигономъ частоты.



Если величину интервала (k) примемъ за единицу, то ординаты $y_1, y_2 \dots y_r \dots$ будутъ численно равны площадямъ соотвѣтствующихъ прямоугольниковъ, а сумма:

$$1^n y_1 + 2^n y_2 + 3^n y_3 + \ldots + r^n y_r + \ldots = N\nu'_n, \text{ т. е.}$$

будетъ равняться грубому n-ому моменту около оси, проходящей черезъ начало полигона (O), умноженному на численность всей совокупности (N). Истинный-же n-ый моментъ около той-же оси будетъ моментомъ площади, составленной изъ трапецій, на которыя разбивается нашъ полигонъ. Мы теперь и займемся его вычисленіемъ. [1] [2]

Сначала найдемъ n-ый моментъ площади трапеціи, ограниченной ординатами y_1 и y_2. Обозначимъ его черезъ ${}_2\mu'_n$. Если любую промежуточную ординату обозначимъ черезъ y, то

$$N \cdot {}_2\mu'_n = \int_{x_1}^{x_2} y x^n dx \text{ или, т. к. изъ пропорціи}$$

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1} \text{ слѣдуетъ, что}$$

$$y = \frac{y_2 - y_1}{x_2 - x_1} x + \frac{y_1 x_2 - y_2 x_1}{x_2 - x_1}, \text{ то нашъ интегралъ будетъ:}$$

$$N \cdot {}_2\mu'_n = \int_{x_1}^{x_2} \frac{y_2 - y_1}{x_2 - x_1} x^{n+1} dx + \int_{x_1}^{x_2} \frac{y_1 x_2 - y_2 x_1}{x_2 - x_1} x^n dx =$$

$$= \frac{y_2 - y_1}{x_2 - x_1} \cdot \frac{x_2^{n+2} - x_1^{n+2}}{n+2} + \frac{y_1 x_2 - y_2 x_1}{x_2 - x_1} \cdot \frac{x_2^{n+1} - x_1^{n+1}}{n+1} =$$

$$= y_2 \left(\frac{x_2^{n+2} - x_1^{n+2}}{(n+2)k} - \frac{x_2^{n+1} - x_1^{n+1}}{(n+1)k} x_1 \right) -$$

$$- y_1 \left(\frac{x_2^{n+2} - x_1^{n+2}}{(n+2)k} - \frac{x_2^{n+1} - x_1^{n+1}}{(n+1)k} x_2 \right).$$

Въ множителѣ при y_2 замѣнимъ x_1 черезъ $(x_2 - k)$, а въ множителѣ при y_1 замѣнимъ x_2 черезъ $(x_1 + k)$ и разложимъ $(x_2 - k)^{n+2}$, $(x_2 - k)^{n+1}$, $(x_1 + k)^{n+2}$ и $(x_1 + k)^{n+1}$ по биному Ньютона. Послѣ простыхъ алгебраическихъ передѣлокъ получимъ:

[1] Читатель не интересующійся выводомъ можетъ обратиться прямо къ формуламъ (47).

[2] K. Pearson, Skew Variation in Homogeneous Material, Phil. Trans. Vol. 186, A. Part I, p. 348—350.

$$N._2\mu'_n = y_2\left(\frac{x_2^n k}{2!} - \frac{n}{3!}x_2^{n-1}k^2 + \frac{n(n-1)}{4!}x_2^{n-2}k^3 - \right.$$
$$\left. - \frac{n(n-1)(n-2)}{5!}x_2^{n-3}k^4 + \ldots\right) -$$
$$- y_1\left(\frac{x_1^n k}{2!} + \frac{n}{3!}x_1^{n-1}k^2 + \frac{n(n-1)}{4!}x_1^{n-2}k^3 + \right.$$
$$\left. + \frac{n(n-1)(n-2)}{5!}x_1^{n-3}k^4 + \ldots\right).$$

Если мы теперь обратимся къ чертежу (6), то замѣтимъ, что, суммируя выраженія подобныя предыдущему, мы встрѣтимъ y_r одинъ разъ въ качествѣ y_2, другой разъ въ качествѣ y_1, а т. к. крайнія ординаты равны нулю, то слѣдовательно n-ый моментъ *всей площади* относительно прежней оси будетъ:

$$N\mu'_n = \Sigma\left[2y_r\left(\frac{x_r^n k}{2!} + \frac{n(n-1)}{4!}x_r^{n-2}k^3 + \right.\right.$$
$$\left.\left. + \frac{n(n-1)(n-2)(n-3)}{6!}x_r^{n-4}k^5 + \ldots\right)\right].$$

Полагая $k = 1$ и замѣчая, что, напр.,

$$\Sigma y_r x_r^n = 1^n y_1 + 2^n y_2 + \ldots + r^n y_r + \ldots = N\nu'_n,$$

получимъ:

$$(45)\ldots \mu'_n = \nu'_n + \frac{n(n-1)}{12}\nu'_{n-2} + \frac{n(n-1)(n-2)(n-3)}{360}\nu'_{n-4} +$$
$$+ \frac{n(n-1)(n-2)(n-3)(n-4)(n-5)}{20160}\nu'_{n-6} + \ldots$$

Откуда, полагая послѣдовательно: $n = 1, 2, 3, 4$, и вспоминая (4), что $\nu'_0 = 1$, находимъ:

$$(46)\ldots\ldots \begin{cases} \mu'_1 = \nu'_1 \\ \mu'_2 = \nu'_2 + \dfrac{1}{6} \\ \mu'_3 = \nu'_3 + \dfrac{1}{2}\nu'_1 \\ \mu'_4 = \nu'_4 + \nu'_2 + \dfrac{1}{15}. \end{cases}$$

Если мы теперь воспользуемся формулами (7), то легко найдемъ:

$$(47) \ldots\ldots \begin{cases} \mu_1 = 0 \\ \mu_2 = \nu_2 + \frac{1}{6} \\ \mu_3 = \nu_3 \\ \mu_4 = \nu_4 + \nu_2 + \frac{1}{15}. \end{cases}$$

Въ этой послѣдней формѣ методъ трапецій и слѣдуетъ примѣнять. Практически дѣло обстоитъ очень просто. Полагая интервалъ равнымъ единицѣ, вычисляемъ грубые моменты ν'_1, $\nu'_2 \ldots$ Затѣмъ по формуламъ (7) находимъ центральные грубые моменты, и, наконецъ, прибавляемъ: (*a*) или поправки Шеппарда, если имѣютъ мѣсто условія, указанныя выше на стр. 25, или-же (*b*) поправки по методу трапецій, т. е. ко второму моменту прибавляемъ $+\frac{1}{6}$, къ четвертому $\nu_2 + \frac{1}{15}$. Затѣмъ, если хотимъ вернуться къ первоначальнымъ единицамъ, n-ый моментъ умножаемъ на k^n.

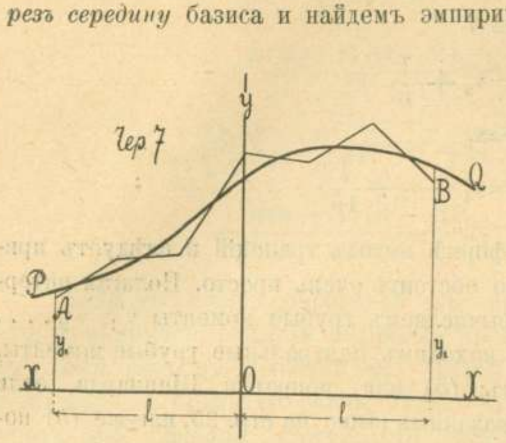
§ 8. *Нахожденіе параболическихъ кривыхъ, соотвѣтствующихъ опытнымъ даннымъ.*

Разъ найдены моменты эмпирической кривой, то задача сводится къ нахожденію коэффиціентовъ теоретической кривой, имѣющей равные моменты. Какую кривую мы выберемъ, зависитъ, какъ уже упоминалось выше, отъ общихъ соображеній, для которыхъ правила дать нельзя. Въ качествѣ кривыхъ частоты наилучшій результатъ даютъ кривыя Пирсона, но для многихъ иныхъ цѣлей часто употребляются съ успѣхомъ параболическія кривыя, типа: $y = a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots\ldots$ Мы хотимъ показать въ настоящемъ §-ѣ, какъ при помощи метода моментовъ найти коэффиціенты этой кривой. Такимъ образомъ, мы будемъ имѣть полный примѣръ приложенія метода. [1])

Пусть AB (чер. 7) будетъ эмпирической линіей, которую мы хотимъ замѣнить наиболѣе подходящей параболой n—1-го порядка. Величина площади, ограниченной этой линіей, ординатами y_1 и y_2 и осью XX, пусть будетъ N. Разстояніе между основаніями орди-

[1]) K. Pearson. On the Systematic Fitting of Curves, Biometrica, Vol. II, p. 12–16

натъ—базисъ кривой—пусть равняется $2l$. Проведемъ ось Y *черезъ середину* базиса и найдемъ эмпирическіе моменты около этой оси. Пусть это будутъ: $\mu'_1, \mu'_2, \mu'_3 \ldots\ldots$ Теперь намъ нужно найти моменты отрѣзка площади параболической кривой, заключенной въ тѣхъ же предѣлахъ, приравнять ихъ эмпирическимъ моментамъ и изъ полученныхъ ур-ій опредѣлить коэффиціенты кривой.



Ур-іе параболы напишемъ въ слѣдующей формѣ:

$$(48) \ldots y = y_0\left[e_0 + e_1\frac{x}{l} + e_2\left(\frac{x}{l}\right)^2 + \ldots + e_{n-1}\left(\frac{x}{l}\right)^{n-1}\right],$$

гдѣ $y_0 = \frac{N}{2l}$ есть средняя ордината, которую мы легко находимъ, т. к. N и l намъ извѣстны.

Умножая обѣ части ур-ія на $\left(\frac{x}{l}\right)^{2r}$ и интегрируя между $x = l$ и $x = -l$, будемъ имѣть:

$$(49) \ldots\ldots\ldots \frac{1}{l^{2r}}\int_{-l}^{+l} yx^{2r}dx = N\frac{\mu'_{2r}}{l^{2r}} =$$

$$= 2y_0 l\left(\frac{e_0}{2r+1} + \frac{e_2}{2r+3} + \ldots + \frac{1-(-1)^n}{2}\,\frac{e_{n-1}}{2r+n}\right).^{1)}$$

---

[1]) Останутся только члены, содержавшіе x въ четныхъ степеняхъ.

Напр., $$\int_{-l}^{+l} e_2\left(\frac{x}{l}\right)^2\left(\frac{x}{l}\right)^{2r}dx = \frac{e_2}{l^{2r+2}}\cdot\left(\frac{x^{2r+3}}{2r+3}\right)_{-l}^{+l} =$$

$$= \frac{e_2}{(2r+3)l^{2r+2}}\left[(+l)^{2r+3} - (-l)^{2r+3}\right] =$$

$$= \frac{e_2}{(2r+3)l^{2r+2}}\cdot 2l^{2r+3} = \frac{2e_2 l}{2r+3}.$$

Члены, содержавшіе x въ нечетныхъ степеняхъ, исчезнутъ, т. к. послѣ

Умножая-же обѣ части на $\left(\frac{x}{l}\right)^{2r+1}$ и интегрируя между тѣми-же предѣлами, найдемъ:

$$(50) \ldots \ldots \frac{1}{l^{2r+1}}\int_{-l}^{+l} yx^{2r+1}dx = N\frac{\mu'_{2r+1}}{l^{2r+1}} =$$

$$= 2y_0 l\left(\frac{e_1}{2r+3} + \frac{e_3}{2r+5} + \ldots + \frac{1+(-1)^n}{2}\,\frac{e_{n-1}}{2r+n+1}\right).$$

Положивъ для сокращенія $\lambda_s = \frac{\mu'_s}{l^s}$ и замѣтивъ, во (1), что $\lambda_0 = \frac{\mu'_0}{l^0} = 1$, во (2), что $2y_0 l$, согласно опредѣленію, $= N$, находимъ:

$$(51) \ldots \left\{\begin{array}{l} e_0 + \frac{1}{3}e_2 + \frac{1}{5}e_4 + \ldots = \lambda_0 = 1, \\ \frac{1}{3}e_0 + \frac{1}{5}e_2 + \frac{1}{7}e_4 + \ldots = \lambda_2, \\ \frac{1}{5}e_0 + \frac{1}{7}e_2 + \frac{1}{9}e_4 + \ldots = \lambda_4, \\ \ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots \\ \frac{1}{3}e_1 + \frac{1}{5}e_3 + \frac{1}{7}e_5 + \ldots = \lambda_1, \\ \frac{1}{5}e_1 + \frac{1}{7}e_3 + \frac{1}{9}e_5 + \ldots = \lambda_3, \\ \frac{1}{7}e_1 + \frac{1}{9}e_3 + \frac{1}{11}e_5 + \ldots = \lambda_5, \\ \ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots \end{array}\right.$$

Изъ этихъ ур-ій легко получить рѣшенія для частныхъ случаевъ.

интегрированія мы будемъ имѣть разность четныхъ степеней. Множитель передъ послѣднимъ членомъ $\frac{1-(-1)^n}{2}$ обращается въ единицу, когда $n-1$ четное число и, въ нуль, когда $n-1$ нечетное.

(1) Прямая линія.

Напишемъ ея ур-іе въ формѣ, соотвѣтствующей общему ур-ію параболы, т. е.

$$y = y_0\left(e_0 + e_1 \frac{x}{l}\right).$$

Тогда

$$(52) \quad \ldots\ldots\ldots \begin{cases} e_0 = \lambda_0 = 1, \\ \frac{1}{3} e_1 = \lambda_1, \end{cases}$$

и ур-іе прямой, выраженное черезъ моменты:

$$y = y_0\left(1 + 3\lambda_1 \frac{x}{l}\right).$$

(2) Парабола 2-го порядка.

$$y = y_0\left[e_0 + e_1 \frac{x}{l} + e_2 \left(\frac{x}{l}\right)^2\right].$$

Наши ур-ія теперь будутъ:

$$e_0 + \frac{1}{3} e_2 = 1,$$

$$\frac{1}{3} e_0 + \frac{1}{5} e_2 = \lambda_2,$$

$$\frac{1}{3} e_1 = \lambda_1.$$

Откуда:

$$(53) \quad \ldots\ldots\ldots \begin{cases} e_0 = \frac{3}{4}(3 - 5\lambda_2), \\ e_1 = 3\lambda_1, \\ e_2 = \frac{15}{4}(3\lambda_2 - 1). \end{cases}$$

(3) Парабола 3-го порядка (кубическая).

$$y = y_0\left[e_0 + e_1 \frac{x}{l} + e_2 \left(\frac{x}{l}\right)^2 + e_3 \left(\frac{x}{l}\right)^3\right].$$

Коэффиціенты найдутся изъ ур-ій:

$$e_0+\frac{1}{3}e_2=1,\qquad \frac{1}{3}e_1+\frac{1}{5}e_3=\lambda_1,$$

$$\frac{1}{3}e_0+\frac{1}{5}e_2=\lambda_2,\quad \frac{1}{5}e_1+\frac{1}{7}e_3=\lambda_3.$$

Они будутъ:

$$(54)\ \ldots\ \left\{\begin{array}{ll} e_0=\frac{3}{4}(3-5\lambda_2), & e_1=\frac{15}{4}(5\lambda_1-7\lambda_3), \\ e_2=\frac{15}{4}(3\lambda_2-1), & e_3=\frac{35}{4}(-3\lambda_1+5\lambda_3). \end{array}\right.$$

(4) Парабола 4-го порядка.

$$y=y_0\left[e_0+e_1\frac{x}{l}+e_2\left(\frac{x}{l}\right)^2+e_3\left(\frac{x}{l}\right)^3+e_4\left(\frac{x}{l}\right)^4\right].$$

Изъ ур-ій:

$$e_0+\frac{1}{3}e_2+\frac{1}{5}e_4=1,\qquad \frac{1}{3}e_1+\frac{1}{5}e_3=\lambda_1,$$

$$\frac{1}{3}e_0+\frac{1}{5}e_2+\frac{1}{7}e_4=\lambda_2,\qquad \frac{1}{5}e_1+\frac{1}{7}e_3=\lambda_3,$$

$$\frac{1}{5}e_0+\frac{1}{7}e_2+\frac{1}{9}e_4=\lambda_4,$$

получимъ:

$$(55)\ \ldots\ \left\{\begin{array}{ll} e_0=\frac{15}{64}(15-70\lambda_2+63\lambda_4), & e_1=\frac{15}{4}(5\lambda_1-7\lambda_3) \\ e_2=\frac{105}{32}(-5+42\lambda_2-45\lambda_4), & e_3=\frac{35}{4}(-3\lambda_1+5\lambda_3). \\ e_4=\frac{315}{64}(3-30\lambda_2+35\lambda_4), & \end{array}\right.$$

(5) Парабола 5-го порядка.

$$y=y_0\left[e_0+e_1\frac{x}{l}+e_2\left(\frac{x}{l}\right)^2+e_3\left(\frac{x}{l}\right)^3+e_4\left(\frac{x}{l}\right)^4+e_5\left(\frac{x}{l}\right)^5\right],$$

гдѣ:

$$
(56)\ \ldots \left\{
\begin{aligned}
e_0 &= \frac{15}{64}(15 - 70\lambda_2 + 63\lambda_4)\,,\\
e_2 &= \frac{105}{32}(-5 + 42\lambda_2 - 45\lambda_4)\,,\\
e_4 &= \frac{315}{64}(3 - 30\lambda_2 + 35\lambda_4)\,,\\
e_1 &= \frac{105}{64}(35\lambda_1 - 126\lambda_3 + 99\lambda_5)\,,\\
e_3 &= \frac{315}{32}(-21\lambda_1 + 90\lambda_3 - 77\lambda_5)\,,\\
e_5 &= \frac{693}{64}(15\lambda_1 - 70\lambda_3 + 63\lambda_5)\,.
\end{aligned}
\right.
$$

(6) Парабола 6-го порядка.

$$
y = y_0\left[e_0 + e_1\frac{x}{l} + e_2\left(\frac{x}{l}\right)^2 + e_3\left(\frac{x}{l}\right)^3 + e_4\left(\frac{x}{l}\right)^4 + e_5\left(\frac{x}{l}\right)^5 + e_6\left(\frac{x}{l}\right)^6\right],
$$

гдѣ:

$$
(57)\ \ldots \left\{
\begin{aligned}
e_0 &= \frac{35}{256}(35 - 315\lambda_2 + 693\lambda_4 - 429\lambda_6)\,,\\
e_2 &= \frac{315}{256}(-35 + 567\lambda_2 - 1485\lambda_4 + 1001\lambda_6)\,,\\
e_4 &= \frac{3465}{256}(7 - 135\lambda_2 + 385\lambda_4 - 273\lambda_6)\,,\\
e_6 &= \frac{3003}{256}(-5 + 105\lambda_2 - 315\lambda_4 + 231\lambda_6)\,,\\
e_1 &= \frac{105}{64}(35\lambda_1 - 126\lambda_3 + 99\lambda_5)\,,\\
e_3 &= \frac{315}{32}(-21\lambda_1 + 90\lambda_3 - 77\lambda_5)\,,\\
e_5 &= \frac{693}{64}(15\lambda_1 - 70\lambda_3 + 63\lambda_5)\,.
\end{aligned}
\right.
$$

Задача разрѣшена Пирсонамъ т. обр. разъ навсегда, и во всѣхъ практическихъ приложеніяхъ можно пользоваться уже готовыми формулами.

§ 9. *Нормальная кривая распредѣленія (кривая Гаусса). Отклоненія отъ нормальнаго типа.*

Познакомимся теперь нѣсколько ближе со свойствами нормальной кривой. При выводѣ ея ур-ія въ § 4 мы принимали за начало координатъ центръ распредѣленія, измѣряя признакъ разностью между его величиной и величиной его средней ариѳметической. Принявъ за начало координатъ любую точку, мы ур-іе (20) должны будемъ написать:

$$(58) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad y = y_0 e^{-\frac{1}{2}\frac{(x-h)^2}{a^2}} .$$

Найдемъ зависимость между коэффиціентами этого ур-ія и характерными величинами распредѣленія: численностью всей совокупности (N), средней ариѳметической $(\bar{x})$ и среднимъ отклоненіемъ (σ).

Прежде всего, разсматривая ур-іе (58), мы замѣчаемъ, что чѣмъ больше $x - h$, тѣмъ y меньше. Наибольшее значеніе y соотвѣтствуетъ $x = h$. Кромѣ того величина $(x - h)^2$ всегда положительна и y имѣетъ одинаковое значеніе, какъ въ томъ случаѣ, когда x больше, такъ и въ томъ случаѣ, когда x меньше h на одну и ту-же величину. Это значитъ что кривая симметрична относительно своей наибольшей ординаты, равныя отклоненія въ обѣ стороны отъ величины признака, равной h, встрѣчаются одинаково часто, а слѣд. эта величина является средней ариѳметической.

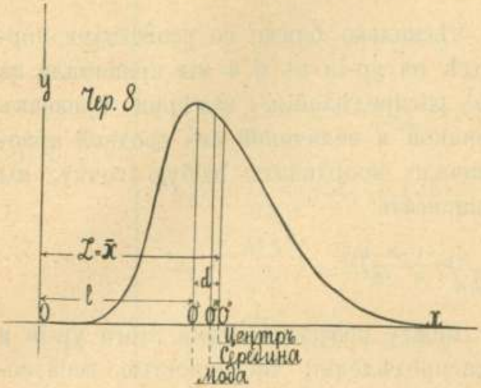
Итакъ, $$\bar{x} = h .$$

Если мы возьмемъ кривую распредѣленія болѣе общаго вида, то *центръ* распредѣленія вообще говоря не совпадетъ съ той точкой основанія кривой, которой соотвѣтствуетъ максимальная ордината. Эту послѣднюю точку будемъ по предложенію Пирсона [1]) называть *модой*, а соотвѣтствующую величину признака (абсциссу этой точки) *модальной величиной*. Разстояніе

---

[1]) Phil. Trans. Vol. 186 A, p. 345, примѣч. (Пирсонъ называетъ модой абсциссу этой точки и самую точку безразлично. Мнѣ кажется болѣе въ духѣ русскаго языка усвоить для соотвѣтствующей абсциссы названіе модальной величины. Т. обр. можно говорить о модальномъ ростѣ, модальной заработной платѣ и т. п.

центра отъ моды будемъ называть *радіусомъ ассиметріи* (d). Онъ будетъ, слѣд., положительнымъ, если центръ расположенъ на право отъ моды, и отрицательнымъ въ обратномъ случаѣ.



Отношеніе радіуса ассиметріи къ среднему отклоненію носитъ названіе *коэффиціента ассиметріи* (Skewness). Я буду обозначать его буквой α.

$$(59). \quad . \quad . \ \alpha = \frac{d}{\sigma}.$$

Кромѣ того *медіаной* называется то значеніе признака, которое дѣлитъ всю совокупность на двѣ равныхъ части. Соотвѣтствующую точку основанія кривой я буду называть *серединой распредѣленія.* [1])

Въ нормальной кривой всѣ эти три величины совпадаютъ: модальный размѣръ равенъ среднему (ариѳметическому) и равенъ серединному. Вотъ одно изъ основаній почему статистикъ не можетъ удовлетвориться одной нормальной кривой, а долженъ усвоить и ассиметричныя кривыя Пирсона.

Пойдемъ дальше.

Принимая опять за начало координатъ центръ распредѣленія, мы будемъ имѣть для нормальной кривой ур-іе

$$y = y_0 e^{-\frac{x^2}{2\sigma^2}}.$$

Найдемъ среднее отклоненіе. По опредѣленію второго момента

$$N\mu_2 = \int_{-\infty}^{+\infty} y_0 e^{-\frac{x^2}{2\sigma^2}} x^2 dx.$$

---

[1]) Середина распредѣленія находится между центромъ и модой; какъ показалъ Пирсонъ, въ большинствѣ случаевъ существуетъ приближенное равенство: разстояніе середины отъ центра=половинѣ разстоянія середины отъ моды. Phil. Trans Vol. 186 A, 375—376. См. также Phil. Trans. Vol. 190 A, p. 441—2.

Это свойство моды позволяетъ находить ее съ точностью для практическихъ приложеній обыкновенно достаточной.

Интегрируя по частямъ, по формулѣ $\int u dv = uv - \int v du$, находимъ:

$$y_0 \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2a^2}} x^2 dx = -a^2 y_0 \int_{-\infty}^{+\infty} x \left[e^{-\frac{x^2}{2a^2}} \left(-\frac{x}{a^2}\, dx \right) \right] =$$

$$= \left[-a^2 y_0 e^{-\frac{x^2}{2a^2}} . x \right]_{-\infty}^{+\infty} + a^2 y_0 \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2a^2}} dx .$$

Выраженіе въ квадратныхъ скобкахъ на обоихъ предѣлахъ обращается въ нуль, т. к. пред. $\left.\frac{x}{e^{x^2}}\right|_{x=\infty} = 0.$

Далѣе, какъ доказывается въ курсахъ интегральнаго исчисленія: [1])

$$\int_{-\infty}^{+\infty} e^{-u^2} du = \sqrt{\pi} ,$$

и нашъ интегралъ будетъ равняться:

$$(60) \ldots N\mu_2 = a^2 y_0 \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2a^2}} dx = a^3 \sqrt{2}\, y_0 \int_{-\infty}^{+\infty} e^{-\left(\frac{x}{a\sqrt{2}}\right)^2} d\left(\frac{x}{a\sqrt{2}}\right) =$$

$$= a^3 \sqrt{2\pi} y_0 .$$

Вся площадь кривой найдется уже безъ труда, т. к.

$$N = y_0 \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2a^2}} dx .$$

Этотъ интегралъ мы имѣли выше. Итакъ:

$$N = \sqrt{2}\, a y_0 \int_{-\infty}^{+\infty} e^{-\left(\frac{x}{a\sqrt{2}}\right)^2} d\left(\frac{x}{a\sqrt{2}}\right) = \sqrt{2}\, a y_0 . \sqrt{\pi} \text{, а слѣд.,}$$

$$(61) \ldots\ldots\ldots\ldots N = \sqrt{2\pi} a y_0 .$$

Дѣля теперь (60) на (61) получаемъ:

$$(62) \ldots\ldots\ldots \begin{cases} \mu_2 = \sigma^2 = a^2 \\ \sigma = a . \end{cases}$$

Подставляя вмѣсто a въ ур-іе (61) его значеніе, находимъ:

$$(63) \ldots\ldots\ldots\ldots y_0 = \frac{N}{\sqrt{2\pi}\sigma} ,$$

---

[1]) См., напр., Г. Ковалевскій, Основы дифференціальнаго и интегральнаго исчисленій. Одесса, 1911 г., стр. 356.

и получаемъ ур-іе нормальной кривой въ окончательномъ видѣ:

$$(64) \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad y = \frac{N}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}},$$

при началѣ координатъ въ центрѣ распредѣленія или

$$(65) \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad y = \frac{N}{\sqrt{2\pi}\sigma} e^{-\frac{(x-h)^2}{2\sigma^2}} \text{ при любомъ началѣ.}$$

Т. обр., зная N, среднее ариѳметическое h и среднее отклоненіе σ, мы можемъ найти ур-іе нормальной кривой, отвѣчающей даннымъ опыта.

Если мы приложимъ тотъ-же методъ интегрированія къ ур-ію (64) для нахожденія высшихъ моментовъ, то найдемъ:

$$\mu_3 = \mu_5 = \mu_7 \; . \; . \; . \; . = 0 \, .$$

Всѣ нечетные моменты для Гауссовой кривой (какъ и слѣдуетъ изъ ея симметричной формы) равны нулю. Между четными моментами существуютъ зависимости. Ограничиваясь четвертымъ моментомъ будемъ имѣть:

$$\mu_4 = 3\mu_2^2 \, .$$

Пирсонъ вводитъ обозначенія, играющія роль въ его теоріи ассиметричныхъ кривыхъ:

$$(66) \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad \begin{cases} \beta_1 = \mu_3^2/\mu_2^3 \\ \beta_2 = \mu_4/\mu_2^2 . \end{cases}$$

Величины β_1 и β_2 находятся по моментамъ, вычисленнымъ на основаніи данныхъ опыта. Очевидно, ни одно эмпирическое распредѣленіе не можетъ считаться нормальнымъ, если для него, въ предѣлахъ вѣроятныхъ ошибокъ, не удовлетворяются условія:

$$\mu_3 = 0, \; \mu_4 = 3\mu_2^2 \, .$$

Этимъ условіямъ можетъ быть придана другая форма.

Именно, изучая обобщенное ур-іе кривой распредѣленія

$$\frac{1}{y}\frac{dy}{dx} = \frac{x-a}{b_0 + b_1 x + b_2 x^2} \, ,$$

Пирсонъ нашелъ [1]) для радіуса ассиметріи и для коэффиціента ассиметріи слѣдующія выраженія:

[1]) On the General Theory of Skew Correlation and Non-Linear Regression, p. 7.

$$(67) \ldots\ldots\ldots d = \frac{\frac{1}{2}\sqrt{\beta_1}(\beta_2+3)}{5\beta_2-6\beta_1-9}\sigma.$$

$$(68) \ldots\ldots\ldots \alpha = \frac{\frac{1}{2}\sqrt{\beta_1}(\beta_2+3)}{5\beta_2-6\beta_1-9}.$$

Въ тѣхъ случахъ, когда кривая не очень значительно отличается отъ Гауссовой, эти выраженія могутъ быть, какъ показалъ Пирсонъ, упрощены, и мы будемъ имѣть:

$$(69) \ldots\ldots\ldots d = \frac{1}{2}\sigma\sqrt{\beta_1}.$$

$$(70) \ldots\ldots\ldots \alpha = \frac{1}{2}\sqrt{\beta_1}.$$

Къ этимъ выраженіямъ присоединимъ *коэффиціентъ разсьянія:*

$$(71) \ldots\ldots\ldots \eta = \beta_2 - 3$$

и мы получимъ формулы для опредѣленія того, насколько наша эмпирическая кривая отличается отъ кривой Гаусса, формулы имѣющія реальный смыслъ.

Если кривая ассиметрична, то по (69) и (70) [или въ случаѣ болѣе сильной ассиметріи по (67) и (68)] мы найдемъ характерныя для нашего распредѣленія величины радіуса ассиметріи и коэффиціента ассиметріи. Этимъ величинамъ мы будемъ придавать значеніе, конечно, лишь въ томъ случаѣ, если они болѣе или менѣе значительно превышаютъ свои вѣроятныя ошибки. Обозначая вѣроятную ошибку буквой E со значкомъ внизу, мы имѣемъ [1]) (для кривыхъ незначительно отличающихся отъ нормальнаго типа):

$$(72) \ldots\ldots \left\{ \begin{aligned} E_d &= 0{,}67449\sigma\sqrt{\frac{3}{2N}} \\ E_\alpha &= 0{,}67449\sqrt{\frac{3}{2N}} \\ E_\eta &= 0{,}67449\sqrt{\frac{24}{N}}.\ ^{2)} \end{aligned} \right.$$

[1]) Pearson and Filon, Phil. Trans. Vol. 191 A, p. 276—277. См. также Phil. Trans. Vol. 198 A, p. 278—9.

[2]) Вычисленіе удобно расположить такъ (R. Pearl, Biom. Vol. V, p. 190). По таблицамъ Gibson (Леонтовичъ, op. cit. ч. III, табл. VI) нахо-

При этомъ можетъ оказаться, что кривая достаточно симметрична (α и d менѣе своихъ удвоенныхъ вѣроятныхъ ошибокъ), но тѣмъ не менѣе кривую нельзя признать нормальной, т. к. не осуществляется равенство $\beta_2 = \frac{\mu_4}{\mu_2^2} = 3$. Реальное значеніе коэффиціента η таково. Если крайнія группы представлены сильнѣе, чѣмъ въ нормальной кривой, то это скажется соотвѣтственнымъ увеличеніемъ четвертаго момента по сравненію со вторымъ, и мы будемъ имѣть, какъ признакъ *сверхнормальнаго* разсѣянія, $\eta > 0$.

Въ противномъ случаѣ, мы будемъ имѣть *поднормальное разсѣяніе:* $\eta < 0$.

Максимальная ширина англійскихъ мужскихъ череповъ, носовая ширина тѣхъ-же череповъ, головной указатель старобаварскихъ череповъ, ушная высота женскихъ „накадскихъ" череповъ дали сверхнормальное разсѣяніе. Изъ 12 опредѣленій роста мужчинъ и женщинъ, изслѣдованныхъ Powys, въ 11 случаяхъ было найдено наоборотъ поднормальное разсѣяніе. [1])

§ 10. Вычисленіе коэффиціентовъ кривыхъ Пирсона.

Убѣдившись, напр., способами предыдущаго параграфа, что нормальная кривая не подходитъ къ данному случаю, мы—при желаніи получать теоретическую модель явленія—должны будемъ заняться нахожденіемъ ур-ія ассиметричной кривой. Въ крайнемъ случаѣ можно ограничиться вычисленіемъ радіуса ассиметріи, коэффиціента ассиметріи и коэф. разсѣянія (формулы 69, 70 и 71).

Ходъ вычисленія коэффиціентовъ кривой былъ разъясненъ выше на примѣрѣ параболы: изъ ур-ія кривой интегрированіемъ его опредѣляемъ теоретическіе моменты, а затѣмъ приравниваемъ ихъ эмпирическимъ. Получаемъ достаточное число ур-ій для опредѣленія коэффиціентовъ. Эта работа, впрочемъ, разъ навсегда была выполнена Пирсономъ, т. что статистику можно пользоваться

димъ $\chi_1 = \frac{0{,}67449}{\sqrt{N}}$, а затѣмъ:

$$E_\alpha = \chi_1 \cdot \sqrt{3/2} = \chi_1 \cdot 1{,}2247449$$

$$E_d = \sigma E_\alpha$$

$$E_\eta = 4E_\alpha.$$

[1]) См. Biometrica, IV p. 175.

готовыми рецептами. Я не буду останавливаться на выводѣ формулъ кривыхъ и ур-ій для нахожденія ихъ коэффиціентовъ. Каждый математикъ въ состояніи это сдѣлать, исходя изъ ур-ія

$$\frac{1}{y}\,\frac{dy}{dx}=\frac{x-a}{b_0+b_1x+b_2x^2}$$

и слѣдуя указанному пути. Кромѣ того въ русской литературѣ уже имѣется изложеніе почти всѣхъ относящихся сюда выводовъ въ книгѣ Р. Орженцкаго (Сводные признаки, 1910, стр. 226—281). Тамъ-же читатель найдетъ и рядъ численныхъ примѣровъ приложенія этого метода.

Мнѣ кажется только небезполезнымъ дать сопоставленіе всѣхъ относящихся сюда формулъ.

Итакъ, прежде всего мы должны, какъ можно тщательнѣе, опредѣлить моменты эмпирическаго распредѣленія. По нимъ вычисляются постоянныя:

(73) . . . . $\beta_1=\mu_3^2/\mu_2^3 \qquad \beta_2=\mu_4/\mu_2^2$.

(74) . . . . $s=\dfrac{6(\beta_2-\beta_1-1)}{3\beta_1-2\beta_2+6}$ и

(75) . . . . . . $k=\dfrac{\beta_1(\beta_2+3)^2}{4(4\beta_2-3\beta_1)(2\beta_2-3\beta_1-6)}$ [1]).

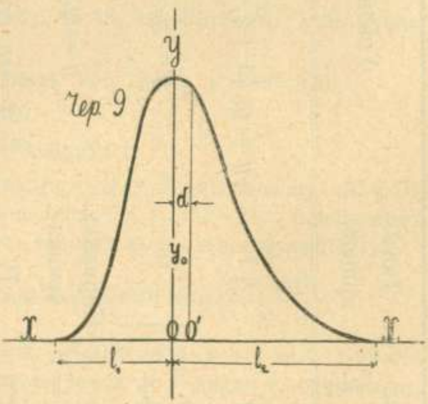
Величина k служитъ критеріемъ типа кривой.
См. таблицу на стр. 48. [2])

Типъ I.

(76) . . $y=y_0\left(1+\dfrac{x}{l_1}\right)^{m_1}\left(1-\dfrac{x}{l_2}\right)^{m_2}$.

Начало—въ модальной точкѣ (O), y_0 есть модальная (максимальная, иногда минимальная) ордината. O' центръ распредѣленія. Радіусъ ассиметріи $=\bar{x}-x_{\text{mod.}}=d$, $l=l_1+l_2$ есть базисъ распредѣленія. a — коэф. ассиметріи.

Чер. 9



[1]) См. Phil. Trans. Vol. 186 A, p. 368 и Vol. 197 A, p. 414.
[2]) Phil. Trans. Vol. 197 p. 445.

Величина критерія k.	Типъ кривой.	Уравненіе ея.	Philos. Trans. A, Vol	Протяженіе.	Ассиметрія.
$k<0$	I	$y=y_0\left(1+\frac{x}{l_1}\right)^{m_1}\cdot\left(1-\frac{x}{l_2}\right)^{m_2}$	186 p. 367	Ограничена съ обѣихъ сторонъ	Ассиметрична
$k=0$ $\beta_1=0,\ \beta_2\neq 3$	II	$y=y_0\left(1-\frac{x^2}{l^2}\right)^m$	186 p. 372	Ограничена съ обѣихъ сторонъ	Симметрична
$k=0$ $\beta_1=0,\ \beta_2=3$	Нормальн. кривая	$y=y_0 e^{-\frac{x^2}{2\sigma^2}}$		Неограничена	"
$0<k<1$	IV	$y=y_0\,(\cos\theta)^{2m}.\,e^{-\nu\theta}$, гдѣ $\theta=arctg\,\frac{x}{a}$	186 p. 376	Неограничена	Ассиметрична
$k=1$	V	$y=y_0x^{-p}.\,e^{-\frac{\gamma}{x}}$	197 p. 446	Ограничена съ одной стороны	"
$1<k<\infty$	VI	$y=y_0\,\frac{(x-l)^{q_2}}{x^{q_1}}$	197 p. 448	"	"
$k=\infty$ (практически, когда k даже умѣренно велико).	III	$y=y_0\left(1+\frac{x}{l}\right)^p e^{-\gamma x}$	186 p. 373	"	"

Величины, характеризующія распредѣленіе этого типа найдутся по формуламъ:

$$(77) \ldots d = \frac{\mu_3}{2\mu_2} \cdot \frac{s+2}{s-2}; \ \alpha = d/\sigma, \ ^{1})$$

$$(78) \ldots l = \frac{\sigma}{2} \sqrt{\beta_1 (s+2)^2 + 16 (s+1)}, \ ^{2})$$

$$(79) \ldots l_1 = \frac{1}{2}(l - ds), \quad l_2 = \frac{1}{2}(l + ds), \ ^{3})$$

$$(80) \ldots m_1 = \frac{l_1}{l}(s-2), \quad m_2 = \frac{l_2}{l}(s-2), \ ^{4})$$

$$(81) \ldots y_0 = \frac{N}{l} \cdot \frac{m_1^{m_1} \cdot m_2^{m_2}}{(m_1+m_2)^{m_1+m_2}} \cdot \frac{\Gamma(m_1+m_2+2)}{\Gamma(m_1+1) \cdot \Gamma(m_2+1)},$$

гдѣ N число индивидуумовъ = площади кривой. [5]

Приближенное выраженіе для y_0:

$$(81') \ldots y_0 = \frac{N}{l} \cdot \frac{(m_1+m_2+1)\sqrt{m_1+m_2}}{\sqrt{2\pi m_1 m_2}} \cdot e^{\frac{1}{12}\left(\frac{1}{m_1+m_2} - \frac{1}{m_1} - \frac{1}{m_2}\right)}. \ ^{6)}\ ^{7)}$$

---

1) Можно вывести изъ формулы Пирсона, Phil. Trans. Vol. 186 A. p. 370.

2) Phil. Trans. Vol. 186 A, p. 369.

3) Davenport, Statistical Methods, p. 32; ср. Пирсонъ, Phil. Trans. Vol. 186 A, p. 369—370.

4) Davenport, p. 32. Ср. Phil. Trans. Vol. 186 A, p. 369—370.

5) Phil. Trans. Vol. 186 A, p. 369.

6) Phil. Trans. Vol. 186 A, p 369, примѣч.

7) Таблицы для вычисленія функціи Γ см. у Леонтовича op. cit. ч. III, табл. VIII. Кромѣ того Пирсонъ (Biometrica, VI p. 118—119) предложилъ формулу для приближеннаго, но очень точнаго вычисленія функціи Γ:

$$log \frac{\Gamma(x+1)}{x^x e^{-x}} = 0{,}3990899 + \frac{1}{2} log x + 0{,}080929 \,.\, Sin \left(\frac{25^0{,}623}{x}\right).$$

Ошибка равна 1/25000 при $x+1=2$; 1/50000 при $x+1=3$; 1/900000 при $x+1=7$; при $x+1=11$ формула точна до 7 знака включительно. Такимъ обр., формула годится для всѣхъ значеній x, которыхъ нѣтъ въ готовыхъ таблицахъ. Также очень точна формула Forsyth'а $\Gamma(n+1) = \sqrt{2\pi}\left(\frac{\sqrt{n^2+n+1/6}}{e}\right)^{n+1/2}$. (ibid., p. 118). Ошибка $< \frac{1}{240 n^3}$.

Типъ II.

Это частный случай перваго, когда $l_1 = l_2$ и $m_1 = m_2$.

Ур-іе кривой:

$$(82) \ldots\ldots \quad y = y_0\left(1 - \frac{x^2}{l^2}\right)^m,$$

гдѣ l есть половина базы. Кривая симметрична. Слѣд. $d = 0$, $\alpha = 0$.

Въ этомъ случаѣ:

$$(83) \ldots\ldots \quad s = \frac{3(\beta_2 - 1)}{3 - \beta_2}, \text{ ибо } \beta_1 = 0.$$

$$(84) \ldots\ldots \quad l = \sigma\sqrt{s+1},$$

$$(85) \ldots\ldots \quad m = \frac{1}{2}(s-2),$$

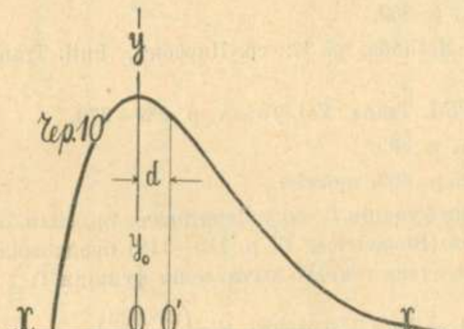
$$(86) \ldots\ldots \quad y_0 = \frac{N}{l}\,\frac{\Gamma(m+1{,}5)}{\sqrt{\pi}\,\Gamma(m+1)}.\ ^{1)}$$

или приближенно:

$$(86') \ldots\ldots \quad y_0 = \frac{N}{\sigma\sqrt{2\pi}} \cdot \frac{s-1}{\sqrt{(s+1)(s-2)}}\, e^{-\frac{1}{4(s-2)}}\ ^{2)}$$

Типъ III.

$$(87) \ldots\ldots \quad y = y_0\left(1 + \frac{x}{l}\right)^p . e^{-\gamma x}, \text{ гдѣ } p = \gamma l.$$



Хотя теоретически этотъ типъ требуетъ, чтобы критерій $k = \infty$, но даже при умѣренныхъ положительныхъ значеніяхъ его эта кривая даетъ хорошіе результаты.

Базисъ кривой ограниченъ только съ одной стороны. Начало координатъ въ модѣ. l есть разстояніе отъ начала распредѣленія до моды. Для вычисленія достаточно знанія только трехъ первыхъ моментовъ.

---

1) Phil. Trans. Vol. 186 A, p. 372.

2) Davenport, op. cit., p. 33.

(88) . Радіусъ ассиметріи $d=\frac{\mu_3}{2\mu_2}$, коэф. ассиметріи $\alpha=\frac{d}{\sigma}$.

(89) . . . . $l=\frac{\mu_2}{d}\quad d,$

(90) . . . . $\gamma=\frac{1}{d}$,

(91) . . . . $p=\frac{l}{d}$,

(92) . . . . $y_0=\frac{N}{l}\ \frac{p^{p+1}}{e^p\Gamma(p+1)}$. [1])

Если дано начало распредѣленія, то мы знаемъ разстояніе между этимъ началомъ и центромъ, т. е. намъ извѣстно

$$L=l+d.$$

Тогда изъ (89) находимъ:

(93) . . . . $d=\frac{\mu_2}{l+d}=\frac{\mu_2}{L}$.

(94) . . . . $l=L-d$.

γ, p и y_0 находимъ изъ (90), (91) и (92).

Этотъ способъ можетъ служить иной разъ для контроля; хотя онъ и менѣе точенъ, но зато упрощаетъ вычисленія, ибо не нужно находить μ_3. Впрочемъ, не зная моментовъ, нельзя быть увѣреннымъ, что кривая подойдетъ именно подъ этотъ типъ.

Типъ IV.

(95) . . . . $y=y_0(cos\theta)^{2m}.\,e^{-\nu\theta}$,

гдѣ $\theta=arctg\left(\frac{x}{a}\right)$ $\left(\text{въ дуговыхъ единицахъ, т. е. } \theta=\frac{\theta^0}{180^0}\pi\right)$.

Кривая ассиметрична, но простирается безгранично въ обѣ стороны. „a“ есть постоянная, которая принимается за величину положительную.

Начало координатъ находится въ точкѣ $x=0$, $\theta=0$, $y=y_0$ и не совпадаетъ ни съ модой, ни съ центромъ. Разстояніе моды

[1]) Cp. Phil. Trans. Vol. 186 A, p. 373—374.

отъ начала обозначимъ черезъ l, (см. чер. 11) а центра отъ начала $L = \mu'_1 = l + d$. Вмѣсто s принимаемъ въ основу вычисленій функцію обратную по знаку.

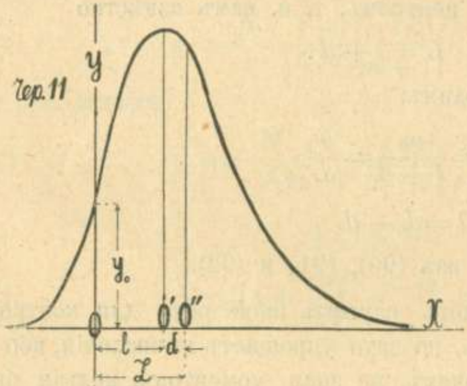
$$(96) \ldots\ldots\ldots\ r = -s = \frac{6(\beta_2 - \beta_1 - 1)}{2\beta_2 - 3\beta_1 - 6},\ ^{1)}$$

$$(97) \ldots\ldots\ldots\ m = \frac{1}{2}(r+2),$$

$$(98) \ldots\ldots\ldots\ d = \frac{\mu_3}{2\mu_2}\,\frac{r-2}{r+2},$$

$$(99) \ldots\ldots\ldots\ a = \frac{\sigma}{4}\sqrt{16(r-1) - \beta_1(r-2)^2},$$

Чер. 11



$$(100)\ \nu = -\frac{\mu_3}{4\mu_2}\cdot\frac{r(r-2)}{a},$$

$$(101)\ \ldots\ L = md,$$

$$(102)\ \ldots\ l = L - d,$$

$$(103)\ y_0 = \frac{Ne^{1/2\,\nu\pi}}{a\int_0^\pi Sin^r\theta . e^{\nu\theta} d\theta}$$

или приближенно:

$$(104) \ldots\ldots\ y_0 = \frac{N}{a}\sqrt{\frac{r}{2\pi}}\ \frac{e^{\frac{Cos^2\varphi}{3r} - \frac{1}{12r} - \varphi r tg\varphi}}{(Cos\varphi)^{r+1}},\ \text{гдѣ}\ tg\varphi = \frac{\nu}{r}.\ ^{2)}$$

---

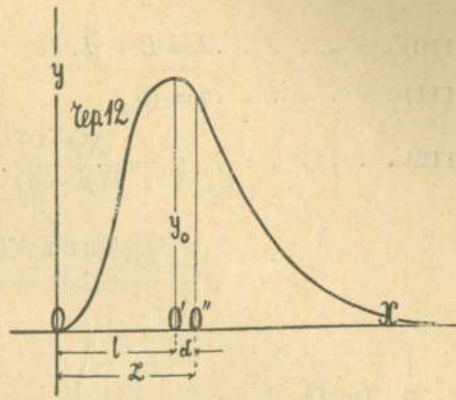
[1]) r всегда положит. и >3 для кривыхъ этого типа. См. Phil. Trans. Vol. 186 A, p. 379.

[2]) Эти формулы получаются изъ формулъ Пирсона (Phil. Trans. Vol. 186 A, 377–380) рядомъ простыхъ преобразованій:

(97)-я дана у Пирсона. l. c. p. 378. (98)-я получается изъ формулы для коэффиціента ассиметріи (ibid.): Skewness $= \frac{1}{2}\sqrt{\beta_1}\,\frac{r-2}{r+2}$, умноженіемъ на σ и подстановкой вмѣсто β_1 его значенія μ_3^2/μ_2^3. Знакъ d опредѣляется знакомъ μ_3, что прямо выражено въ формулѣ (98). Формулу (99) получимъ, если въ формулу Пирсона $a = r\sqrt{\frac{\mu_2(r-1)}{z}}$ подставимъ

Типъ V.

(105) . . $y = y_0 x^{-p} e^{-\gamma/x}$.



Базисъ ограниченъ съ одной только стороны. Начало координатъ въ началѣ базиса. y_0 максимальная ордината. У кривыхъ этого типа вспомогательная величина s всегда отрицательна. Вводя, какъ и въ кривой IV типа

$$(106) \; . \; . \; . \; . \; . \; . \; . \; r = -s = \frac{6(\beta_2 - \beta_1 - 1)}{2\beta_2 - 3\beta_1 - 6},$$

мы изъ общей формулы (67) легко найдемъ:

$$(107) \; . \; . \; . \; . \; . \; . \; . \; d = \frac{\mu_3}{2\mu_2}\,\frac{r-2}{r+2}, \;\; \alpha = \frac{d}{\sigma}, \text{ затѣмъ:}$$

$$(108) \; . \; . \; . \; . \; . \; . \; . \; p = r + 2,$$

$z = \dfrac{r^2}{1 - \dfrac{\beta_1}{16}\,\dfrac{(r-2)^2}{r-1}}$ (ibid. p. 378). Изъ формулы Пирсона $\nu = \sqrt{z - r^2}$ (ibid.) подстановкой z получаемъ выраженіе, приводимое у Орженцкаго l. cit. p. 270, а затѣмъ воспользовавшись формулой (99), упрощаемъ его въ формулу (100). Знакъ минусъ ставимъ передъ формулой въ силу замѣчанія Пирсона, что ν и μ_3 имѣютъ обратные знаки, что, впрочемъ, прямо вытекаетъ изъ выраженія $\mu_3 = -\dfrac{4a^3\nu(r^2+\nu^2)}{r^3(r-1)(r-2)}$ (Pearson l. cit. p. 377). Формулу (101) получаю изъ формулы Пирсона $\mu'_1 = -\dfrac{a\nu}{r}$ (l. cit p. 377), находя $\left(-\dfrac{a\nu}{r}\right)$ изъ (100) и пользуясь затѣмъ (97) и (98). (102)-я вытекаетъ изъ опредѣленія радіуса ассиметріи. (103)-я и (104)-я даны у Pearson'a въ этомъ-же видѣ l. cit. p. 378, 380. Цѣль этихъ преобразованій: по возможности упростить формулы, поставить ихъ въ послѣдовательности наиболѣе удобной для вычисленій и придать имъ такую форму, чтобы правильные знаки опредѣлялись сами собой путемъ простой подстановки надлежащихъ величинъ въ формулы.

$$(109) \ldots\ldots \quad L = \frac{1}{2}dp,$$

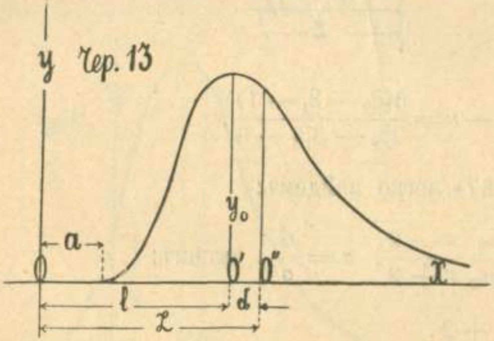
$$(110) \ldots\ldots \quad l = L - d,$$

$$(111) \ldots\ldots \quad \gamma = lp,$$

$$(112) \ldots\ldots \quad y_0 = \frac{N\gamma^{p-1}}{\Gamma(p-1)}.$$ [1])

Типъ VI.

$$y = y_0 \frac{(x-a)^{q_2}}{x^{q_1}}.$$



Базисъ ограниченъ съ одной только стороны. Кривая начинается на разстояніи a отъ начала координатъ. Обозначивъ какъ и въ двухъ предыдущихъ типахъ черезъ r величину обратную s по знаку (здѣсь $r > 0$), будемъ имѣть:

$$(113) \ldots\ldots \quad d = \frac{\mu_3}{2\mu_2}\,\frac{r-2}{r+2}, \quad \alpha = \frac{d}{\sigma},$$

$$(114) \ldots\ldots \quad a = \sigma\sqrt{\frac{1}{4}\beta_1(r-2)^2 - 4(r-1)},$$

---

[1]) Пирсонъ для нахожденія p даетъ квадратное ур-іе $(p-4)^2 - \frac{16}{\beta_1}(p-4) - \frac{16}{\beta_1} = 0$, положительный корень котораго доставляетъ искомое рѣшеніе. (См. Phil. Trans. Vol. 197 A, p. 447). Общая формула для d (наша 67-я) была найдена имъ позже. При помощи ея можно найти p, не прибѣгая къ рѣшенію квадратнаго ур-ія. Изъ (форм. VIII) (l. cit. p. 447) находимъ $\mu_3/\mu_2 = 4\gamma/(p-2)(p-4)$, подставляемъ въ выраженіе для d (107 выше) и приравниваемъ другому выраженію того-же d (форм. XVI-я l. cit. p. 448). Такимъ путемъ опредѣлится p. Остальныя формулы легко получить изъ формулъ Пирсона данныхъ тамъ-же.

$$(115) \ldots\ldots q_1 = \frac{1}{2}(r+2)\left(\frac{d}{a}r+1\right),$$

$$(116) \ldots\ldots q_2 = \frac{1}{2}(r+2)\left(\frac{d}{a}r-1\right),$$

$$(117) \ldots\ldots L = \frac{a}{r}(q_1-1),$$

$$(118) \ldots\ldots l = L - d.$$

$$(119) \ldots\ldots y_0 = \frac{Na^{r+1}.\,\Gamma(q_1)}{\Gamma(r+1)\,\Gamma(q_2+1)}.\ ^{1)}$$

---

1) Пирсонъ (l. c) обозначаетъ буквой r величину, обозначенную у насъ черезъ s. Для q_1 и q_2 онъ даетъ (p. 450) квадратное ур-іе (въ моемъ обознач.) $Z^2+rZ+\varepsilon=0$, однимъ корнемъ котораго будетъ $1-q_1$, другимъ $1+q_2$, и гдѣ $\varepsilon=\frac{r^2}{4+{}^1/_4\,\beta_1(r-2)^2/(1-r)}$. Зная d, мы можемъ рѣшить это ур-іе. Именно, на основаніи свойства корней квадратнаго ур-ія $(1-q_1)(1+q_2)=\varepsilon$, а $q_1-q_2-2=r$. Замѣняя въ форм. XXIV (l. cit p. 450) $(1-q_1)(1+q_2)$ черезъ ε и подставляя его значеніе, находимъ данную выше формулу (114) для a. Затѣмъ Пирсонъ даетъ (p. 450) $d=\frac{a(q_1+q_2)}{(q_1-q_2)(q_1-q_2-2)}$. Подставляя сюда r вмѣсто (q_1-q_2-2), и замѣчая, что $q_1-q_2=r+2$, находимъ: $q_1+q_2=\frac{d}{a}r(r+2)$. Складывая и вычитая послѣднія два выраженія, находимъ наши формулы (115) и (116). Формула (117) соотвѣтствуетъ первой изъ формулъ (XXII) у Пирсона (l. c. p. 449), а (119) тождественна съ XXV-й, только q_1-q_2-2 и q_1-q_2-1 замѣнены соотвѣтственно черезъ r и черезъ $r+1$. Произведенныя измѣненія въ формулахъ, дѣлая излишнимъ рѣшеніе квадратнаго ур-ія съ многозначными коэффиціентами, должны значительно упростить примѣненіе кривой этого типа.

Часть II.

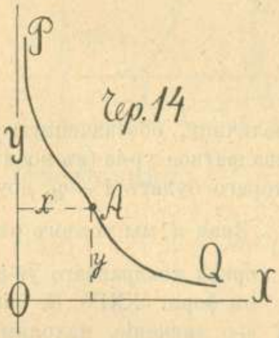
Теорія корреляціи.

ГЛАВА I.

Корреляція между двумя величинами.

§ 1. Понятіе корреляціонной зависимости.

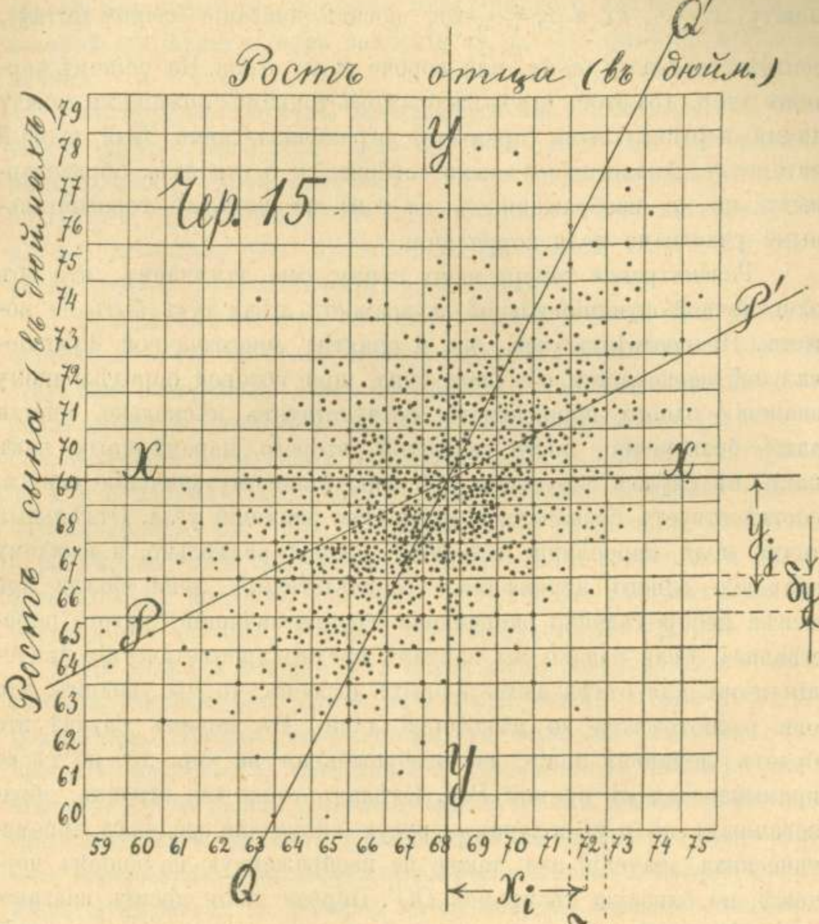
Основнымъ типомъ зависимостей, съ которыми имѣютъ дѣло т. наз. точныя науки, служитъ *однозначная функціональная зависимость.* Каждому значенію одной величины (x) соотвѣтствуетъ одно опредѣленное значеніе второй величины (y). Если мы пару значеній обѣихъ величинъ изобразимъ точкой плоскости такъ, чтобы абсцисса точки равнялась значенію одной, а ордината ея значенію другой величины, то (чер. 14) совокупность этихъ точекъ расположится на нѣкоторой линіи, (PQ), которая и будетъ служить изображеніемъ данной зависимости.



Нашему изученію подлежатъ соотношенія совершенно иного типа. Допустимъ, что мы хотимъ найти зависимость между ростомъ отца (x) и ростомъ сына (y). Каждую пару значеній этихъ величинъ, соотвѣтствующую каждой такой парѣ индивидуумовъ, мы можемъ по прежнему изобразить точкой, но всѣ полученныя нами точки не расположатся на одной линіи, а дадутъ картину, схематически изображенную на чер. 15 [1]). Чтобы въ ней легче

[1]) Въ основу нашей схемы положены дѣйствительныя измѣренія, собранныя Пирсономъ (K. Pearson and A. Lee, On the Laws of Inheritance in Man, Biometrika, Vol. II Table I p. 370 и Table XXII p. 415). Такъ какъ отдѣльныя измѣренія въ цитированной работѣ не даны, а даны лишь числа

разобраться, раздѣлимъ все т. наз. *поле корреляціи* на вертикальныя и горизонтальныя полосы, соотвѣтствующія равнымъ интер-



валамъ величинъ x и y. Значенія x-а и y-а, отвѣчающія серединамъ интерваловъ, переномеруемъ по порядку и будемъ называть

паръ индивидуумовъ въ группахъ, соотвѣтствующихъ нашимъ клѣточкамъ, то отдѣльныя точки поставлены на угадъ; число ихъ въ каждой клѣточкѣ соотвѣтствуетъ однако дѣйствительности (хотя тоже не абсолютно, ибо оказалось затруднительнымъ установить положенія точекъ на границахъ интерваловъ). Во всякомъ случаѣ позволительно утверждать, что *общее впечатлѣніе* отъ схематическаго чер. 15 вполнѣ соотвѣтствуетъ дѣйствительному положенію вещей.

варіантами. Совокупность случаевъ, соотвѣтствующихъ i-ой варіантѣ x-а, т. е. совокупность случаевъ, когда x заключается между $x_i - \frac{1}{2}\delta x$ и $x_i + \frac{1}{2}\delta x$, носитъ названіе *строя* (array), соотвѣтствующаго x_i-му, или короче x_i-го строя. На нашемъ чертежѣ (чер. 15) этотъ строй изображенъ точками, лежащими между двумя вертикальными прямыми, ограничивающими i-ый x-овый интервалъ. Аналогично можно говорить и о строяхъ, образованныхъ по y, изображенныхъ на томъ-же чертежѣ горизонтальными участками поля корреляціи.

Разсматривая теперь нашу схему, мы замѣчаемъ, что объ однозначной функціональной зависимости рѣчи тутъ быть не можетъ. Не соотвѣтствуетъ она и понятію многозначной функціональной зависимости обычнаго типа, при которой опредѣленному значенію одного перемѣннаго соотвѣтствуетъ нѣсколько, иногда даже безконечно много, значеній второго перемѣннаго, какъ напр. въ случаѣ синусоиды, гдѣ опредѣленному значенію синуса, соотвѣтствуетъ безконечное множество значеній угла. Отдѣльныя точки поля корреляціи всѣ расположены раздѣльно, и каждому значенію одного перемѣннаго соотвѣтствуетъ лишь болѣе или менѣе неопредѣленно очерченная группа значеній другого перемѣннаго. Если однако мы найдемъ среднія ариѳметическія значенія y-овъ для отдѣльныхъ x-овыхъ строевъ, то мы увидимъ, что они расположатся по нѣкоторой линіи. Въ нашемъ случаѣ это будетъ ломанная линія, не изображенная на чер. 15, но тѣсно примыкающая къ прямой PP'. Сдѣлавъ то-же для строевъ, образованныхъ по y-у, получимъ другую линію для среднихъ ариѳметическихъ значеній x-а, также не изображенную на нашемъ чертежѣ, но близкую къ прямой QQ'. Первая линія носитъ названіе ***линіи регрессіи y-а относительно x-а***, вторая—***линіи регрессіи x-а относительно y-а***. Так. обр., хотя мы и не можемъ по одной величинѣ опредѣлить другую въ каждомъ отдѣльномъ случаѣ, но за то мы въ состояніи указать среднее значеніе каждой величины, соотвѣтствующее опредѣленному значенію другой. Кромѣ того, разсматривая нашу схему и остановивъ свое вниманіе на какомъ либо отдѣльномъ строѣ, мы замѣтимъ, что гуще всего точки расположены около линіи регрессіи, т. е. около своего средняго значенія. Чѣмъ дальше отъ него, тѣмъ онѣ становятся болѣе рѣдкими, и на нѣкоторомъ разстояніи мы ихъ уже совсѣмъ или почти совсѣмъ не встрѣчаемъ. Так. обр. и частота, съ которой встрѣчается каждое значеніе одной величины, оказывается функціей другой величины, или выражая то-же немного иначе: частота, съ которой встрѣчается пара значеній x_i, y_j, есть функція этихъ величинъ.

Эта функція можетъ быть изображена способомъ, аналогичнымъ употреблявшемуся нами въ I-ой части для изображенія распредѣленія одной величины. Пусть число индивидуумовъ въ подгруппѣ, соотвѣтствующей одновременно i-ому интервалу x-а и j-му интервалу y-а будетъ n_{ij}; на нашей схемѣ n_{ij} равно числу точекъ въ соотвѣтствующей клѣткѣ.

Вообразимъ теперь на каждой такой клѣткѣ по параллелепипеду, объемъ котораго пропорціоналенъ численности соотвѣтствующей подгруппы. Легко сообразить, что высота каждаго параллелепипеда изображаетъ число индивидуумовъ (случаевъ), приходящихся на единицу площади, т. е. на единицу интервала по x-у и единицу интервала по y-у. Верхнія площадки нашихъ параллелепипедовъ, при безграничномъ увеличеніи ихъ числа и размѣровъ всей совокупности, въ предѣлѣ сольются и дадутъ поверхность, носящую названіе поверхности распредѣленія или поверхности частоты. Общее ур-іе ея будетъ $Z = f(x,y)$.

Мы можемъ теперь дать опредѣленіе корреляціонной зависимости, обобщая его сразу на случай какого угодно числа перемѣнныхъ. Именно, мы говоримъ, что *нѣсколько величинъ находятся въ корреляціи, если каждой совокупности значеній всѣхъ величинъ кромѣ одной соотвѣтствуетъ цѣлый комплексъ значеній этой послѣдней, причемъ средняя ариѳметическая величина каждаго перемѣннаго измѣняется въ зависимости отъ значеній остальныхъ, и частота, съ которой встрѣчается каждая комбинація значеній перемѣнныхъ, есть функція этихъ значеній.*

Если средняя величина какой либо перемѣнной остается постоянной для всѣхъ строевъ, образованныхъ по другимъ перемѣннымъ, то говорятъ, что корреляціи нѣтъ, или, что корреляція равна нулю. Если съ увеличеніемъ одной величины среднее ариѳметическое значеніе другой также увеличивается, то корреляція будетъ положительной, въ противномъ-же случаѣ она будетъ отрицательной. Чѣмъ тѣснѣе отдѣльныя значенія величины примыкаютъ къ линіи регрессіи, чѣмъ меньше, слѣдовательно, разницы между среднимъ значеніемъ величины въ каждомъ строѣ и отдѣльными значеніями той-же величины въ предѣлахъ этого строя, тѣмъ полнѣе будетъ корреляція. Вообразимъ себѣ, что точки нашей схемы (чер. 15) все тѣснѣе и тѣснѣе группируются около одного направленія; линіи регрессіи при этомъ должны будутъ все болѣе и болѣе сближаться и, когда всѣ точки совокупности расположатся по одной линіи, тогда и линіи регрессіи сольются съ нею. Мы будемъ имѣть случай совершенной корреляціи, т. е. переходъ корреляціонной связи въ обычную функціональную.

Полное изслѣдованіе корреляціонной зависимости должно заключать въ себѣ: во (1), нахожденіе линій регрессіи, во (2), отысканіе мѣры для степени корреляціонной связи и въ (3) опредѣленіе ур-ія поверхности распредѣленія, дающаго возможность по одной величинѣ найти вѣроятность каждаго значенія другой.

Послѣдняя задача разрѣшена до сихъ поръ лишь для частнаго случаѣ т. наз. нормальной поверхности распредѣленія, соотвѣтствующей нормальной кривой Гаусса, съ которой мы имѣли дѣло въ I-ой части.

§ 2. Корреляціонная таблица.

Изображеніе каждаго отдѣльнаго случая особой точкой, вполнѣ пригодное для выясненія характера изучаемыхъ нами отношеній, не удобно однако для цѣлей статистической практики. Не нужно вѣдь забывать, что явленія, изучаемыя статистическими методами, формируются подъ воздѣйствіемъ безчисленнаго множества причинъ, и что одна изъ важнѣйшихъ задачъ изслѣдованія заключается въ томъ, чтобы уловить основныя тенденціи въ изучаемыхъ соотношеніяхъ, освободивъ ихъ по сколько возможно отъ примѣси случайнаго элемента. Поэтому ни одна характеристика массового явленія не можетъ почитаться достаточной, если не сопровождается указаніемъ на ея вѣроятную ошибку, и всякія два метода, приводящіе къ результатамъ, разнящимся другъ отъ друга меньше чѣмъ на величину своей вѣроятной ошибки, должны быть признаны равноцѣнными. Изъ нихъ заслуживаетъ предпочтенія, очевидно, тотъ, который требуетъ меньшей вычислительной работы. Изъ этихъ соображеній вытекаютъ всѣ тѣ упрощенные пріемы, которые примѣняются современной статистикой.

Основной изъ нихъ знакомъ намъ уже по первой части. Онъ заключается въ томъ, что мы соединяемъ отдѣльные случаи въ группы и каждую группу разсматриваемъ, какъ состоящую изъ индивидуумовъ тождественныхъ другъ съ другомъ, приписывая всѣмъ имъ величину соотвѣтствующую серединѣ подлежащаго интервала. На такихъ-же основаніяхъ и поле корреляціи, съ точнымъ указаніемъ значеній величины въ каждомъ отдѣльномъ случаѣ, замѣняется корреляціонной таблицей, въ каждомъ подраздѣленіи которой обозначается лишь общее число случаевъ, приходящихся на данный интервалъ. Такую таблицу мы могли бы получить изъ схематическаго чер. 15, если бы въ каждую клѣточку мы вписали число, равное числу находящихся въ ней точекъ. При этомъ, если какая нибудь точка находится на границѣ двухъ подраздѣленій, то ее приходится разбить пополамъ и засчитать въ каждую изъ группъ по 1/2 случая, а точка, находящаяся на границѣ четырехъ подраздѣленій, даетъ подобнымъ-же образомъ начало четвертямъ случаевъ [1]). Нѣсколько такихъ таблицъ, послужившихъ основаніемъ для приводимыхъ ниже иллюстрирующихъ методъ вычисленій, читатель найдетъ въ приложеніи и, конечно, сразу замѣтитъ, что корреляціонная таблица есть не что иное, какъ обыкновенная, хорошо каждому статистику знакомая комбинаціонная таблица. Извлечь изъ сгруппированнаго въ ней сырого матеріала все, что онъ можетъ дать, и есть задача метода корреляціи.

§ 3. Линіи регрессіи.

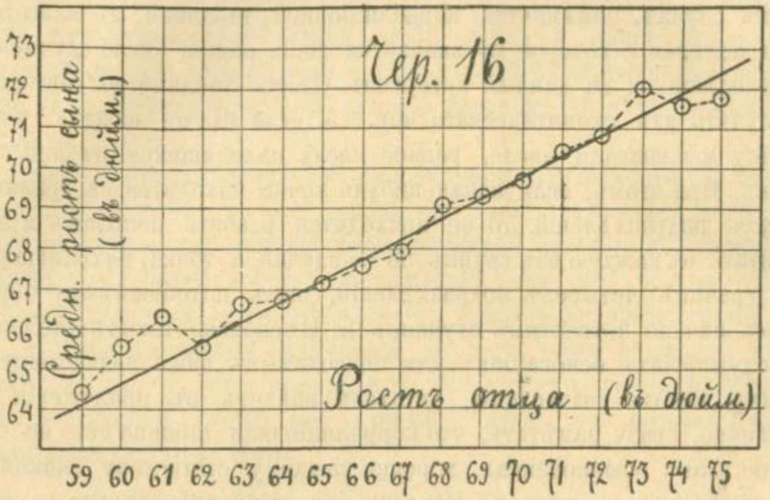
Первый шагъ къ изученію корреляціонной зависимости состоитъ въ томъ, чтобы найти среднія ариѳметическія отдѣльныхъ строевъ. Изобразивъ эти величины точками и соединивъ послѣднія прямыми, мы получимъ эмпирическую линію регрессіи, разсмотрѣніе которой сразу даетъ цѣнныя указанія относительно зависимости, подлежащей нашему изученію. Какъ однако ни цѣнна такая діаграмма сама по себѣ, однимъ разсматриваніемъ ея статистикъ удовлетвориться не можетъ, т. к. она только наталкиваетъ на

---

[1]) Если установленіе группировки предшествуетъ измѣренію или вычисленію величинъ, относящихся къ отдѣльнымъ случаямъ, то почти всегда возможно это вычисленіе или измѣреніе въ сомнительномъ случаѣ произвести съ такою точностью, чтобы избѣжать усложняющаго окончательный подсчетъ дробленія единицъ. (Способъ этотъ рекомендованъ U. Yule'ем. См. Journ. of the Roy. Stat. Soc. 1899, p. 257).

цѣлый рядъ вопросовъ, на которые нельзя отвѣтить безъ дальнѣйшей математической обработки матеріала.

Разсмотримъ для примѣра чер. 16. Онъ построенъ на основаніи измѣреній, собранныхъ Пирсономъ, охватившихъ 1078 паръ



отцовъ-сыновей [1]). Пунктирная линія, соединяющая отмѣченныя кружками точки, есть линія регрессіи. Она показываетъ, какъ измѣняется средній ростъ сыновей съ измѣненіемъ роста отцовъ. Такъ мы видимъ, что у отцовъ, ростъ которыхъ не больше чѣмъ на полдюйма отличается отъ 59 дюймовъ, оказались сыновья со среднимъ ростомъ въ 64,1″. Росту 60″ для отца соотвѣтствуетъ средній ростъ сыновей въ 65,6″ и т. д. Вообще говоря, увеличеніе роста отцовъ связано со значительнымъ увеличеніемъ средняго роста сыновей, въ чемъ и выражается извѣстный фактъ наслѣдственной передачи признаковъ. Законъ этой передачи на первый взглядъ не отличается простотой. Слѣдуя за всѣми зигзагами линіи регрессіи, мы должны констатировать, что въ группѣ наиболѣе низкорослыхъ отцовъ (59″—61″) увеличенію роста отца на 1″ соотвѣтствуетъ въ среднемъ увеличеніе роста сыновей также на 1″. Затѣмъ мы имѣемъ, т. сказ., аномальный интервалъ (61″—62″), въ

[1]) Чер. 16 заимствованъ у Пирсона—On the Laws of Inheritance in Man, Biom Vol. II p. 362—и основанъ на данныхъ табл. XXII (p. 415), послужившей намъ уже для составленія схематическаго чер. 15.

которомъ увеличенію роста отца соотвѣтствуетъ *уменьшеніе* средняго роста сыновей. Въ слѣдующемъ интервалѣ повторяется картина, наблюдавшаяся нами въ двухъ первыхъ интервалахъ, а дальше слѣдуетъ обширная группа отцовъ (63″—67″) съ ростомъ, приближающимся къ среднему, и съ ослабленной способностью къ наслѣдственной передачѣ. Увеличенію роста отца на 1″ здѣсь соотвѣтствуетъ увеличеніе средняго роста сыновей всего на 0,4″.

Изъ анализа линіи регрессіи мы могли бы получить еще много такихъ „законовъ", которыхъ я не привожу здѣсь вовсе не потому, что имъ мѣсто не въ нашей работѣ, а въ трактатѣ по теоріи наслѣдственности. Послѣдней съ ними также нечего дѣлать, ибо они вовсе не существуютъ. Если бы мы измѣрили другую тысячу паръ отцовъ-сыновей, то—можно утверждать съ полной увѣренностью — въ этомъ новомъ матеріалѣ не осталось бы ни слѣда отъ нашихъ многихъ законовъ. Общее направленіе линіи регрессіи оказалось бы прежнимъ, но отдѣльные зигзаги расположились бы, можетъ быть, совершенно иначе. Въ самомъ дѣлѣ, коррелятивная связь выражается только въ среднихъ величинахъ и, чтобы уловить ее, необходимо большое число наблюденій, т. к. въ маленькихъ группахъ общая тенденція можетъ быть совершенно затушевана причинами случайнаго характера. Обратимся къ чер. 15; мы увидимъ, что крайніе зигзаги линіи регрессіи легко находятъ свое объясненіе въ малочисленности соотвѣтствующихъ случаевъ. Въ частности три крайнихъ лѣвыхъ среднихъ ариѳметическихъ выведены на основаніи 3, $3^1/_2$ и 8 случаевъ. Не нужно быть большимъ теоретикомъ, чтобы сразу почувствовать весьма малую достовѣрность соотвѣтствующихъ деталей линіи регрессіи.

Соображенія, высказанныя только что по поводу частнаго случая, имѣютъ, очевидно, общее значеніе. При изслѣдованіи всякой зависимости всегда необходимо умѣть отвлечься отъ чертъ, присущихъ индивидуальному матеріалу, эллиминируя случайныя отклоненія, затемняющія дѣйствіе общихъ тенденцій. Единственнымъ для этого средствомъ въ статистикѣ является нахожденіе числового выраженія для всякой характеристики и сравненіе его съ его вѣроятной ошибкой. Только этотъ путь гарантируетъ намъ полученіе результатовъ, на которые можно положиться.

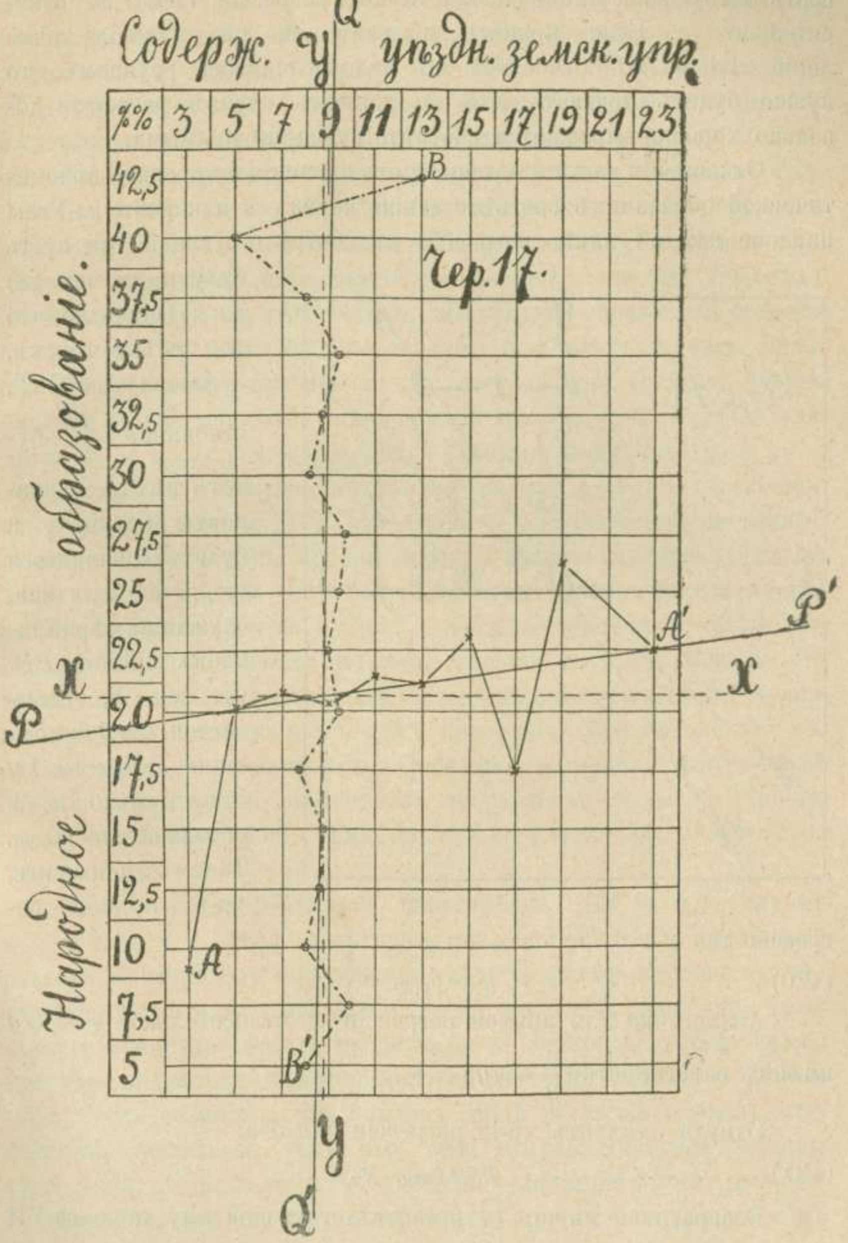
Въ частности, задача использованія линій регрессіи стоитъ, очевидно, такъ: эти линіи нужно преобразовать, сгладивъ ихъ зигзаги и выявивъ основную заключающуюся въ каждой изъ нихъ тенденцію. Эта цѣль достигается нахожденіемъ нѣкоторой плавной линіи, которая должна какъ можно тѣснѣе примыкать къ точкамъ эмпирической линіи регрессіи, и которую можно назвать теоретической линіей регрессіи. Если отклоненія эмпирической линіи отъ теоретической не выходятъ въ своей совокупности за предѣлы вѣроятныхъ случайныхъ отклоненій, то теоретическая линія регрессіи можетъ разсматриваться, какъ вполнѣ адэкватное изображеніе дѣйствительныхъ отношеній. Въ противномъ случаѣ теоретической линіей регрессіи хотя и можно пользоваться, но не надо забывать при этомъ, что она является только болѣе или менѣе грубымъ приближеніемъ.

Линія регрессіи, адэкватная въ предѣлахъ вѣроятныхъ ошибокъ эмпирическому матеріалу, можетъ служить критеріемъ для установленія типовъ регрессіи. Въ этомъ смыслѣ различаютъ линейную регрессію, геометрическимъ образомъ которой является прямая, и криволинейную, выражаемую обыкновенно какой нибудь параболической кривой. Теорія криволинейной регрессіи до сихъ поръ разработана недостаточно, и мы въ дальнѣйшемъ будемъ имѣть дѣло главнымъ образомъ съ первымъ типомъ, къ счастью, по крайней мѣрѣ въ извѣстномъ приближеніи, преобладающимъ въ ряду случаевъ, съ которыми имѣетъ дѣло статистическая практика.

§ 4. Иллюстраціи.

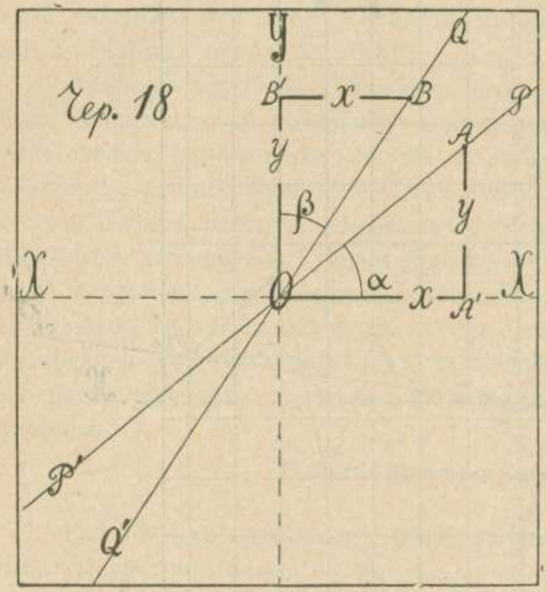
Разсмотримъ графическое изображеніе корреляціонной зависимости, представленное на черт. 17, и ту таблицу (см. приложеніе, табл. VII), на основаніи которой онъ построенъ. Мы имѣемъ здѣсь дѣло съ изслѣдованіемъ вопроса о томъ, существуетъ-ли зависимость между двумя статьями расходнаго бюджета уѣздныхъ земствъ, именно, процентомъ расхода на народное образованіе и процентомъ расхода на содержаніе самой земской управы. Матеріаломъ послужили данныя за 1901 годъ, относящіяся ко всѣмъ 359 уѣзднымъ земствамъ.

Оси координатъ XX и YY проведены черезъ центръ распредѣленія, т. е. черезъ точку, соотвѣтствующую среднимъ ариѳметическимъ той и другой величины. Если мы, согласно съ принципами, развитыми въ предыдущемъ параграфѣ, найдемъ теоретическія прямыя регрессіи, пользуясь способомъ, изложеннымъ ниже въ § 6, то окажется, что эти прямыя также пройдутъ черезъ



центръ распредѣленія. Это будутъ—PP', соотвѣтствующая эмпирической линіи регрессіи (AA') y относительно x, и QQ', соотвѣтствующая эмпирической линіи регрессіи (BB') x относительно y. Если принять во вниманіе, что крайнія точки линій AA' и BB' относятся къ малочисленнымъ группамъ, то нужно будетъ признать, что въ данномъ случаѣ регрессія довольно хорошо выражается соотвѣтствующими прямыми.

Отклоненія каждой величины отъ значенія ея средней ариѳметической обозначимъ соотвѣтственно черезъ x и черезъ y. Углы наклона каждой линіи регрессіи къ соотвѣтствующей оси пусть будутъ (см. чер. 18) α и β. Очевидно, что для всѣхъ точекъ, лежащихъ на PP', отношеніе $\frac{y}{x}$ будетъ имѣть одинаковую величину и будетъ равняться $tg\alpha$. Эта величина, служащая мѣрой наклона прямой PP' къ оси X, называется *коэффиціентомъ регрессіи* y относительно x и обозначается $\rho_{y(x)}$. Такимъ образомъ, ур-іе прямой регрессіи для y-а будетъ



$$(120) \quad y = \rho_{y(x)} \cdot x\,.$$

Аналогично для прямой регрессіи x относительно y (QQ') имѣемъ зависимость $\frac{x}{y} = tg\beta = \rho_{x(y)}$.

Откуда слѣдуетъ ур-іе регрессіи для x-а:

$$(121) \quad x = \rho_{x(y)} \cdot y\,.$$

Возвратимся къ чер. 17 и соотвѣтствующей ему таблицѣ VII (см. прилож.). Первое, что здѣсь бросается въ глаза, такъ это почти полное отсутствіе зависимости между подлежащими величинами. Процентъ расхода на народное образованіе можетъ быть, какъ очень низкимъ, такъ и очень высокимъ, не смотря на величину расходовъ по управленію. Очевидно, что заботливость земства о народномъ образованіи совсѣмъ или почти совсѣмъ не связана съ тѣмъ обстоятельствомъ, какую часть всѣхъ расходовъ составляетъ содержаніе самой уѣздной управы. Если мы способомъ, изложеннымъ ниже, найдемъ коэффиціентъ регрессіи для народнаго образованія относительно расходовъ на управленіе, то онъ окажется равнымъ $\rho_{y(x)} = 0{,}14$. Это значитъ, что земства, у которыхъ расходъ на управленіе превышаетъ на величину x средній для всѣхъ земствъ расходъ по этой статьѣ, на народное образованіе тратятъ въ среднемъ на $0{,}14x$ больше средняго для всѣхъ земствъ. Зависимости этой можно придать еще и другую форму. Пусть x_1 будетъ расходъ на управленіе въ одной группѣ земствъ, а x_2 — въ другой. Тогда $y_1 = \rho_{y(x)} x_1$, а $y_2 = \rho_{y(x)} x_2$. Вычитая, получимъ: $y_2 - y_1 = \rho_{y(x)} (x_2 - x_1)$, откуда, подставляя вмѣсто $\rho_{y(x)}$ его численное значеніе, найдемъ: $y_2 - y_1 = 0{,}14\,(x_2 - x_1)$. Возьмемъ численный примѣръ. Пусть одна група земствъ тратитъ на управленіе на 10% больше другой. Изъ нашей формулы слѣдуетъ, что на народное образованіе она будетъ тратить въ среднемъ также нѣсколько большую часть бюджета, именно, на $1{,}4\%$ больше. По сравненію съ тѣми громадными различіями, которыя существуютъ вообще между земствами въ дѣлѣ народнаго образованія, эта величина совершенно ничтожна. Очевидно, величина расходовъ на народное образованіе зависитъ въ несравненно большей степени отъ другихъ факторовъ, чѣмъ отъ того, съ которымъ мы ее здѣсь сопоставили.

Еще слабѣе обратная зависимость. Въ этомъ случаѣ $\rho_{x(y)} = 0{,}02$, т. е., если взять сразу числовой примѣръ, увеличеніе расхода на народное образованіе на 10% смѣты связано съ увеличенія расхода на управленіе всего на $0{,}2\%$. Конечно, даже такая слабая зависимость представляетъ интересъ, однако нужно еще удостовѣриться, дѣйствительно-ли она существуетъ на самомъ дѣлѣ. Вѣдь возможно, что наклонъ линій регрессіи вызванъ случайными причинами, такъ что, если бы мы составили таблицы корреляціи для ряда лѣтъ, то на одной—линіи регрессіи оказались бы наклоненными въ одну сторону, на другой — въ другую и въ среднемъ для большого промежутка времени совпали бы съ направленіемъ осей координатъ, указывая тѣмъ на отсутствіе всякой зависимости. Мы опять приходимъ къ признанію необходимости имѣть вѣроятныя ошибки числовыхъ характеристикъ явленія. Ихъ не можетъ замѣнить даже распространеніе изслѣдованія на рядъ лѣтъ, о которомъ мы только что говорили. Если бы для одного года мы получили положительный, для другого отрицательный коэффиціентъ регрессіи, но если бы каждый превышалъ свою вѣроятную ошибку, напр., въ 10 разъ, то мы должны были бы съ громадной вѣроятностью—практически совпадающей съ полной увѣренностью—вывести заключеніе, что зависимость между явленіями существовала какъ въ одномъ, такъ и въ другомъ случаѣ, но что только типъ ея—вслѣдствіе какихъ то причинъ—претерпѣлъ измѣненіе.

Разсмотримъ теперь другой примѣръ. Чер. 19 и табл. II (см. прилож.) изображаютъ зависимость между средними мѣсячными цѣнами въ Ельцѣ и въ Самарѣ за 11-ти лѣтній періодъ

1893—1903 г.г. Картина, представляющаяся здѣсь читателю, рѣзко отличается отъ предыдущей. Числа табл. II не разбросаны по всей таблицѣ, а тянутся въ опредѣленномъ направленіи сравнительно неширокой полосой. Прямыя регрессіи, какъ показываетъ чер. 19, сближены и наклонены къ соотвѣтствующимъ осямъ подъ значительными углами. Коэффиціенты регрессіи въ этомъ случаѣ должны, слѣдовательно, имѣть значительно бóльшую величину. Въ самомъ дѣлѣ, $\rho_{y(x)} = 1{,}13$, а $\rho_{x(y)} = 0{,}68$. Такимъ обр., увеличеніе средней мѣсячной цѣны ржи въ Ельцѣ, напр., на 10 коп. связано съ увеличеніемъ такой-же цѣны въ Самарѣ въ среднемъ на 11,3 коп; повышенію-же цѣны ржи въ Самарѣ на 10 коп. соотвѣтствуетъ среднее повышеніе цѣны въ Ельцѣ на 6,8 коп. Связь между разсматриваемыми цѣнами оказывается значительно болѣе тѣсной, чѣмъ связь между явленіями предыдущаго примѣра.

§ 5. *Коэффиціентъ корреляціи.*

Корреляціонная зависимость между явленіями можетъ быть, какъ мы видѣли, и болѣе и менѣе тѣсной. Начиная отъ полной независимости и переходя черезъ рядъ градацій, она въ предѣлѣ превращается въ строго функціональную связь двухъ величинъ. Какъ уже упоминалось выше, степень корреляціонной зависимости отражается на положеніи линій регрессіи: при отсутствіи корреляціи онѣ должны совпадать съ осями координатъ, а въ случаѣ перехода корреляціонной зависимости въ функціональную онѣ должны слиться въ одну линію, которая будетъ прямой, если зависимость носитъ линейный характеръ.

Приведенные примѣры имѣли своей задачей иллюстрировать эти положенія и придать имъ нѣкоторую наглядность. Читатель вѣроятно согласится теперь съ тѣмъ, что величина отдѣльнаго коэффиціента регрессіи не можетъ еще служить мѣрой тѣсноты корреляціонной зависимости, короче, мѣрой корреляціи. Во первыхъ, коэффиціентовъ регрессіи два; въ то время, какъ одинъ можетъ имѣть довольно значительную величину, другой можетъ быть близокъ къ нулю, и корреляція будетъ все еще далека отъ строгой функціональной зависимости. Во вторыхъ, коэффиціентъ регрессіи есть величина именованная и зависитъ поэтому отъ выбора единицъ и принятаго масштаба. Напр., въ случаѣ корреляціи между цѣною хлѣба и смертностью коэффиціенты регрессіи будутъ имѣть различную величину въ зависимости отъ того, будемъ-ли мы опредѣлять смертность въ процентахъ или промилляхъ, а цѣну въ коп. за фунтъ, или въ копейкахъ за пудъ, или въ рубляхъ за четверть и т. д. Мѣра-же корреляціи должна быть, очевидно, числомъ отвлеченнымъ.

Разсмотримъ корень квадратный изъ произведенія коэффиціентовъ регрессіи. Это будетъ величина отвлеченная, независящая отъ выбора единицъ измѣренія изучаемыхъ величинъ. Въ самомъ дѣлѣ, взявъ на одной прямой регрессіи (см. чер. 18) любую точку A, а на другой любую точку B, будемъ имѣть:

$$\rho_{y(x)} = \frac{AA'}{OA'} \text{ и } \rho_{x(y)} = \frac{BB'}{OB'},$$

слѣдовательно: $\sqrt{\rho_{y(x)} \cdot \rho_{x(y)}} = \sqrt{\frac{AA'}{OA'} \cdot \frac{BB'}{OB'}} = \sqrt{\frac{AA'}{OB'} \cdot \frac{BB'}{OA'}}$. Это будетъ число отвлеченное, ибо таковыми будутъ отношенія

$$\frac{AA'}{OB'} \text{ и } \frac{BB'}{OA'}.$$

Такъ какъ $\rho_{y(x)} = tg\alpha$, а $\rho_{x(y)} = tg\beta$, то, слѣдовательно, $\sqrt{\rho_{y(x)} \cdot \rho_{x(y)}} = \sqrt{tg\alpha \, . \, tg\beta}$.

Если корреляціи нѣтъ, и прямыя регрессіи совпадаютъ съ осями координатъ, то $tg\alpha = 0$, $tg\beta = 0$, а слѣдовательно, и $\sqrt{\rho_{y(x)} \cdot \rho_{x(y)}} = 0$.

Если корреляція переходитъ въ линейную функціональную зависимость, то прямыя регрессіи сливаются, и углы α и β въ суммѣ даютъ 90^0.

Въ этомъ случаѣ $\sqrt{\rho_{y(x)} \cdot \rho_{x(y)}} = \sqrt{tg\alpha \, . \, tg(90-\alpha)} =$
$= \sqrt{tg\alpha \, . \, ctg\,\alpha} = 1$.

Указанныя свойства дѣлаютъ среднее геометрическое изъ коэффиціентовъ регрессіи удобной мѣрой степени корреляціонной зависимости. Величина эта играетъ большую роль въ теоріи корреляціи, обозначается буквой r съ подписными значками, указывающими, къ какимъ величинамъ она относится, и называется *коэффиціентомъ корреляціи*.

Итакъ,

(122) . . . . . $r_{xy} = \pm\sqrt{\rho_{y(x)} \cdot \rho_{x(y)}}$.

Коэффиціенты регрессіи имѣютъ всегда одинаковые знаки. Можно условиться ставить передъ корнемъ въ выраженіи (122) знакъ (+), если они оба положительны и (—), если они оба отрицательны. Въ дальнѣйшемъ изложеніи мы придемъ къ другой формулѣ для того-же коэффиціента корреляціи, которая можетъ быть принята и обыкновенно принимается за основную, и изъ которой знакъ коэффиціента корреляціи опредѣляется самъ собой. Вообще, всѣ разсужденія, съ которыми мы имѣли дѣло до сихъ поръ, мы можемъ считать предварительными, направленными на уясненіе основныхъ понятій теоріи корреляціи. Къ строгому доказательству ихъ и послѣдовательному выводу ряда положеній и формулъ теоріи мы переходимъ въ слѣдующемъ параграфѣ.

§ 6. Выводъ формулъ для коэффиціентовъ регрессіи и коэффиціента корреляціи.

Прямая регрессіи должна указывать намъ нѣкоторое среднее направленіе эмпирической линіи и можетъ быть получена вообще различными способами. Напримѣръ, ее можно провести такъ, чтобы сумма разстояній отъ нея всѣхъ точекъ эмпирической линіи регрессіи сдѣлалась наименьшей, считая всѣ разстоянія по ихъ абсолютной величинѣ, независимо отъ знака. Можно также найти прямую, для которой сумма квадратовъ разстояній была бы наименьшей, или сумма (независимо отъ знака) третьихъ, или сумма четвертыхъ степеней и т. д. Въ выборѣ одного изъ этихъ способовъ мы до извѣстной степени свободны, и каждый изъ нихъ будетъ хорошъ, если дастъ сравнительно простой результатъ, и если всѣ *согласятся* примѣнять его. Эти замѣчанія необходимы для того, чтобы подчеркнуть то условное, что присуще принятому въ наукѣ *методу наименьшихъ квадратовъ*. Во всякомъ случаѣ, важно отмѣтить, что, употребляя этотъ методъ, мы не дѣлаемъ никакихъ допущеній относительно характера распредѣленія отдѣльныхъ значеній нашихъ величинъ, такъ что всѣ формулы теоріи корреляціи сохраняютъ свое значеніе при всякомъ „законѣ“ распредѣленія.

Часть обозначеній, относящихся къ корреляціонной таблицѣ намъ уже извѣстна. Сопоставимъ ихъ и введемъ недостающія. Общее число случаевъ обозначается буквой N. Численность i-ого x-оваго строя (вертикальнаго столбца таблицы) будетъ n_{x_i}, j-ого y-оваго строя (горизонтальной строки таблицы)—n_{y_j}. Численность подгруппы, принадлежащей одновременно x_i-му и y_j-му строямъ,—

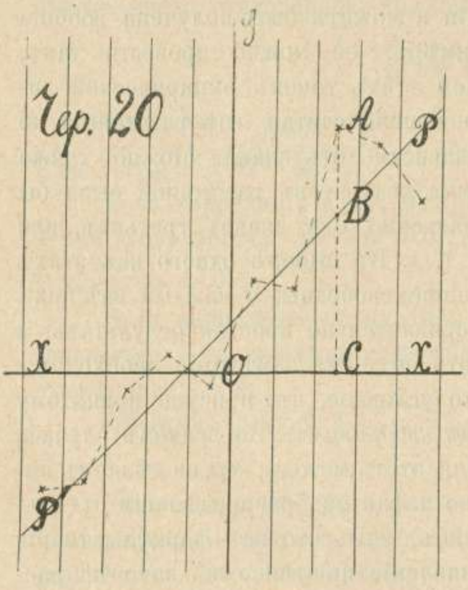
$n_{x_i y_j}$, или короче n_{xy}, или n_{ij}. Среднія ариѳметическія для всей совокупности пусть будутъ $\bar{x}$ и $\bar{y}$, или h_x, h_y, а среднія отклоненія — σ_x, σ_y. Для совокупности величинъ, составляющихъ отдѣльный строй, также можно найти постоянныя распредѣленія; для строевъ, образованныхъ по x, это будутъ: средняя ариѳметическая—y_{x_i}, или короче y_x, и среднее отклоненіе $\sigma_{n_{x_i}}$, или σ_{n_x}; для строевъ по y аналогично: x_{y_j}, или x_y, и $\sigma_{n_{y_j}}$, или σ_{n_y}.

Перейдемъ теперь къ выводу ур-ія прямой регрессіи для y относительно x. [1])

Пусть это ур-іе будетъ:

(123) . . . . $Y = a + bX$,

гдѣ Y и X—координаты прямой регрессіи, измѣряемыя отъ осей, проходящихъ черезъ центръ распредѣленія; $b = \rho_{y(x)}$ есть тангенсъ угла, образуемаго этою прямою (PP') съ осью X-овъ, а a — отрѣзокъ, отсѣкаемый ею на оси Y. Звѣздочками (см. черт. 20) отмѣчены центры распредѣленія отдѣльныхъ x-овыхъ строевъ.



Обозначимъ буквою d разстояніе (измѣренное по вертикальному направленію) между точками линіи регрессіи и прямою PP'. Напр., для точки A (чер. 20) $d = AB = AC - BC$. Т. к. AC есть средняя ариѳметическая величина y для даннаго строя, а BC ордината прямой PP', то мы будемъ имѣть:

$$d = y_x - Y.$$

Прямую PP' нужно опредѣлить изъ условія, чтобы сумма квадратовъ подобныхъ величинъ была наименьшая. При этомъ нужно

[1]) См. U. Yule, On the Significance of Bravais' Formulae for Regression, etc. in the case of Skew Correlation, Proc. of the Roy. Soc. Vol. 60, p. 477—489.

принять во вниманіе, что численность различныхъ строевъ различна, слѣд., вычисляя сумму квадратовъ, нужно различнымъ d придать различный вѣсъ. Т. обр., въ окончательномъ видѣ условіе, изъ котораго опредѣляется прямая регрессіи, будетъ таково:

$$(124) \quad \sum n_x d^2 = \text{minimum}.$$

Это условіе можно свести къ другому, представляющему свой особый интересъ съ точки зрѣнія теоріи корреляціи. Обозначимъ черезъ Σ_i операцію суммированія величинъ, принадлежащихъ къ i-му x-овому строю, и найдемъ сумму квадратовъ разностей между всѣми значеніями y въ предѣлахъ этого строя и величиной соотвѣтствующей ординаты (Y) прямой регрессіи. Мы получ.:

$$(125) \quad \sum_i (y-Y)^2 = \sum_i [(y-y_x)+(y_x-Y)]^2 = \sum_i (y-y_x)^2 + \\ + 2\sum_i (y-y_x)(y_x-Y) + \sum_i (y_x-Y)^2.$$

Первая сумма представляетъ изъ себя сумму квадратовъ отклоненій отдѣльныхъ значеній величины отъ ея средняго ариѳметическаго и равняется, слѣдовательно, $n_x \sigma_{n_x}^2$. Во второй суммѣ величина $(y_x - Y)$, какъ общій множитель, можетъ быть вынесена за знакъ суммы; остающееся-же выраженіе — $\Sigma(y - y_x)$ — по свойству средней ариѳметической равняется нулю (см. § 2 стр. 8 форм. (3)). Третья сумма состоитъ изъ величины $(y_x - Y)^2$, постоянной въ предѣлахъ строя и повторенной число разъ, равное его численности; она равна, слѣдовательно, $n_x(y_x - Y)^2 = n_x d^2$. Кромѣ того, т. к. въ предѣлахъ одного и того-же строя x, соотвѣтствующее каждому y, можетъ быть принято за величину постоянную и равную X, т. е. абсциссѣ прямой регрессіи, то, очевидно, что въ выраженіи $\sum (y-Y)^2$ мы можемъ замѣнить Y его значеніемъ изъ ур-ія (123)-го, т. е. выраженіемъ $a + bX$, а въ послѣднемъ — величину X абсциссой, общей всѣмъ индивидуумамъ строя, т. е. величиной x. Так. обр., равенство (125) преобразуется къ слѣдующему виду:

$$(126) \quad \sum_i [y-(a+bx)]^2 = n_x \sigma_{n_x}^2 + n_x d^2.$$

Составляя такія-же выраженія для всѣхъ строевъ и суммируя ихъ, получимъ:

$$\sum [y-(a+bx)]^2 = \sum n_x \sigma_{n_x}^2 + \sum n_x d^2, \text{ откуда:}$$

$$(127) \quad \sum n_x d^2 = \sum [y-(a+bx)]^2 - \sum n_x \sigma_{n_x}^2.$$

Сумма $\sum n_x \sigma_{n_x}^2$ зависитъ только отъ характера распредѣленія и, конечно, не измѣняется, какъ бы мы ни провели прямую регрессіи (PP'), т. е. какое бы значеніе мы ни дали величинамъ a и b. Отсюда слѣдуетъ, что сумма $\sum n_x d^2$ будетъ наименьшей въ томъ случаѣ, когда осуществится условіе:

$$(128) \quad \ldots\ldots\ldots \sum [y-(a+bx)]^2 = \text{minimum}.$$

Этотъ результатъ бросаетъ новый свѣтъ на то условіе, которому мы подчинили прямую регрессіи. Ея ур-іе (123) есть $Y=a+bX$. Если бы всѣ точки совокупности были расположены на этой прямой, то, зная величину одного перемѣннаго (x), мы для каждаго отдѣльнаго случая могли бы найти величину и другого перемѣннаго (y), пользуясь ур-іемъ $y=a+bx$. Но, т. к. одному значенію x соотвѣтствуетъ рядъ случаевъ съ различными значеніями второго перемѣннаго, то, опредѣляя величину послѣдняго изъ ур-ія регрессіи, мы искомой величины не получимъ. Иными словами, примѣняя ур-іе регрессіи къ отдѣльнымъ случаямъ, мы каждый разъ будемъ получать болѣе или менѣе ошибочный результатъ, и ошибка его будетъ равна $y-(a+bx)$. Условіе (124), которому мы первоначально, слѣдуя методу наименьшихъ квадратовъ, подчинили прямую PP', равносильно, какъ мы видимъ теперь другому условію (128): *найти такую линейную зависимость между x и y, чтобы пользуясь ею для опредѣленія y по x сдѣлать ошибки, сумма квадратовъ которыхъ была бы наименьшей.*

Подобнымъ же образомъ, если мы найдемъ среднія ариѳметическія y-овыхъ строевъ, то получимъ вторую эмпирическую линію регрессіи, не совпадающую съ первой; прямая регрессіи, опредѣленная т. обр., чтобы сумма квадратовъ *горизонтальныхъ* разстояній ея отъ эмпирической линіи регрессіи была наименьшей, будетъ имѣть по доказанному ур-іе, пользуясь которымъ для опредѣленія индивидуальныхъ x-овъ по соотвѣтствующимъ y-амъ, мы сдѣлаемъ ошибки, сумма квадратовъ которыхъ также будетъ наименьшей.

Найдемъ теперь окончательный видъ ур-ія прямой регрессіи для y, т. е. опредѣлимъ его коэффиціенты a и b. Для этого, какъ было только что доказано, нужно найти минимумъ выраженія $\sum[y-(a+bx)]^2$. По правиламъ дифференціальнаго исчисленія возьмемъ производныя отъ него по a и по b и приравняемъ ихъ нулю. Т. обр. получатся у насъ два ур-ія для опредѣленія неизвѣстныхъ коэффиціентовъ ур-ія регрессіи. Дифференцируя [1]), будемъ имѣть:

(129) . . . . $\sum[y-(a+bx)]=0$ и

(130) . . . . $\sum x[y-(a+bx)]=0$

Изъ (129), суммируя, находимъ:

$$\sum y-\sum a-b\sum x=0\,,\ \text{или, т. к.}\ \sum a=Na\,,$$

(131) . . . . . $a=\frac{1}{N}\left(\sum y-b\sum x\right)=0\,,$

ибо $\sum x=0$ и $\sum y=0$, т. к. x и y представляютъ изъ себя отклоненія отъ своихъ среднихъ ариѳметическихъ (§ 2 стр. 8).

Результатъ (131) очень важенъ. Онъ показываетъ, что когда $x=0$, то и y въ среднемъ также равняется нулю, т. е. когда первая величина имѣетъ свое среднее значеніе, то и вторая (въ среднемъ для ряда случаевъ) также совпадаетъ со своей средней. Такъ какъ то-же самое оказывается справедливымъ и для случая многихъ перемѣнныхъ (см. § 27), то мы заключаемъ, что въ случаѣ линейной регрессіи представленіе *типеческаго*, какъ сочетанія среднихъ ариѳметическихъ, вполнѣ допустимо. „Средній человѣкъ" Кетле, имѣющій средній для своего возраста ростъ, средніе размѣры различныхъ органовъ, среднія способности и т. д. не представляетъ изъ себя ничего нереальнаго, ибо какъ показалъ рядъ статистическихъ изслѣдованій (особенно школы Пирсона), въ антропологіи можно къ самымъ разнообразнымъ признакамъ примѣнять линейныя формулы съ очень небольшой ошибкой. Однако, уже зависимость роста отъ возраста совершенно нелинейна [2]). А въ случаѣ нелинейной регрессіи вполнѣ возможно, чтобы средняя величина одного признака была ассоціирована (въ среднемъ) не со средней величиной другого и наоборотъ. Въ этомъ случаѣ индивидуумъ, имѣющій всѣ признаки

---

[1]) Читатель, не знакомый съ дифференціальнымъ исчисленіемъ, долженъ просто повѣрить, что условіе (128) сводится къ (129) и (130), и можетъ слѣдить за дальнѣйшими выводами.

[2]) См., напр., A. O. Powys, Data for the Problem of Evolution in Man, Biometrica Vol. I p. 47.

среднихъ размѣровъ, можетъ быть чрезвычайно мало вѣроятнымъ, а значитъ и не типичнымъ индивидуумомъ [1]).

Обращаясь къ уравненію (130) и полагая въ немъ $a=0$, находимъ:

$$\sum xy - b\sum x^2 = 0 \text{ , откуда}$$

$$b = \frac{\sum xy}{\sum x^2}.$$

Если мы въ формулѣ (8) § 3 замѣнимъ $x-\bar{x}$ черезъ x, т. е. примемъ, что x есть отклоненіе отъ средняго, затѣмъ, если мы вмѣсто того, чтобы умножать x^2 на n_x, просто повторимъ x^2 слагаемымъ нужное число разъ, то мы и получимъ сумму $\left(\sum x^2\right)$, стоящую въ знаменателѣ полученнаго выраженія для b. Такимъ образомъ $\sum x^2 = N\sigma_x^2$. Вспоминая кромѣ того, что b есть тангенсъ угла, образуемаго прямой регрессіи съ осью X, и что, слѣдовательно, $b=\rho_{y(x)}$, мы для коэффиціента регрессіи y относительно x получаемъ слѣдующее выраженіе:

$$(132) \; . \; . \; . \; . \; . \; . \; . \; . \quad \rho_{y(x)} = \frac{\sum xy}{N\sigma_x{}^2}.$$

Аналогичнымъ путемъ для второй прямой регрессіи найдемъ:

$$(133) \; . \; . \; . \; . \; . \; . \; . \quad \rho_{x(y)} = \frac{\sum xy}{N\sigma_y{}^2},$$

и для коэффиціента корреляціи, равнаго $\sqrt{\rho_y \rho_x}$ (см. выше форм. 122) будемъ имѣть выраженіе:

$$(134) \; . \; . \; . \; . \; . \; . \; . \; . \quad r_{xy} = \frac{\sum xy}{N\sigma_x\sigma_y},$$

представляющее изъ себя *основную формулу метода корреляціи*. Замѣняя въ (132) и (133) $\sum xy$ черезъ ея значеніе, найденное изъ (134), мы получимъ для коэффиціентовъ регрессіи общеупотребительныя простыя выраженія:

$$(135) \; . \; . \; . \quad \left\{ \begin{aligned} \rho_{y(x)} &= \frac{\sigma_y}{\sigma_x} r_{xy}, \\ \rho_{x(y)} &= \frac{\sigma_x}{\sigma_y} r_{xy}, \end{aligned} \right.$$

---

[1]) K. Pearson, On the General Theory of Skew Correlation and Non-Linear Regression, Drapers' Company Research Memoirs, Biom. Ser. II 1905 p. 29.

а, слѣдовательно, ур-ія прямыхъ регрессіи будутъ:

$$(136) \ldots\ldots \begin{cases} Y = \frac{\sigma_y}{\sigma_x} r_{xy} x \quad \text{и} \\ X = \frac{\sigma_x}{\sigma_y} r_{xy} y . \end{cases}$$

§ 7. Другія формулы для коэффиціента корреляціи.

Прежде чѣмъ мы перейдемъ къ анализу полученныхъ выраженій, слѣдуетъ остановиться еще немного на математической сторонѣ дѣла и вывести рядъ формулъ, съ которыми мы будемъ встрѣчаться въ дальнѣйшемъ.

(*A*) Сумма $\sum xy$ представляетъ изъ себя выраженіе въ нѣсколькихъ отношеніяхъ неудобное для вычисленія. Прежде всего, при большомъ числѣ случаевъ чрезвычайно обременительно находить всѣ произведенія xy для каждой пары значеній въ отдѣльности. Такой порядокъ вычисленія долженъ быть однако рекомендованъ, если все число случаевъ, входящихъ въ таблицу, очень не велико, напр. 20, 30, 50. Когда оно болѣе или менѣе значительно, слѣдуетъ для упрощенія вычисленій разбить всю таблицу на клѣтки, какъ было показано выше, и вести вычисленіе въ предположеніи, что во всѣхъ случаяхъ, относящихся къ каждой такой подгруппѣ (клѣткѣ), всѣ значенія x равны соотвѣтствующей x-овой варіантѣ, а всѣ значенія y, соотвѣтствующей y-овой. Тогда сумма произведеній xy для случаевъ одной такой подгруппы равна будетъ произведенію варіантъ, соотвѣтствующихъ подгруппѣ, на численность ея, и для коэффиціента корреляціи получится выраженіе:

$$(137) \ldots\ldots\ldots r = \frac{\sum n_{xy} xy}{N \sigma_x \sigma_y} .$$

(*B*) Это суммированіе можно производить въ различномъ порядкѣ. Напримѣръ, составляя соотвѣтствующія произведенія сначала для первой вертикальной колоны, замѣтимъ, что x_1 войдетъ во всѣ произведенія общимъ множителемъ и что, слѣдовательно,

$$\sum\nolimits_1 n_{xy} xy = x_1 \sum\nolimits_1 n_{xy} y .$$

Но среднее значеніе y для первой колоны равняется

$$y_{x_1} = \frac{\sum_1 n_{xy} y}{n_{x_1}} . \text{ Слѣдовательно,}$$

$$\sum_1 n_{xy}xy = n_{x_1}y_{x_1}x_1 .$$

Продолжая суммированіе въ томъ же порядкѣ для второго, третьяго и такъ дальше столбцовъ и складывая полученныя выраженія, найдемъ, что для всей таблицы:

$$(138) \ldots\ldots\ldots \sum n_{xy}xy = \sum n_x y_x x .$$

Аналогичнымъ путемъ, измѣняя порядокъ суммированія и идя по отдѣльнымъ горизонтальнымъ строкамъ, получимъ:

$$(139) \ldots\ldots\ldots \sum n_{xy}xy = \sum n_y x_y y .$$

Эти формулы даютъ указаніе относительно наиболѣе удобнаго практически порядка суммированія. Ихъ теоретическое приложеніе встѣтится намъ ниже.

(*C*) Вспомнимъ, что при выводѣ нашихъ формулъ для коэффиціента корреляціи мы измѣряли наши величины ихъ отклоненіями отъ ихъ среднихъ значеній. Чтобы вернуться къ обычному способу измѣренія нужно вмѣсто x подставить $x-\bar{x}$, а вмѣсто y взять $y-\bar{y}$. Тогда формула для коэффиціента корреляціи будетъ имѣть слѣдующій видъ:

$$(140) \ldots\ldots r_{xy} = \frac{\sum n_{xy}(x-\bar{x})(y-\bar{y})}{N\sigma_x\sigma_y}, \quad ^{1)}$$

а ур-ія регрессіи (136) примутъ свою наиболѣе употребительную форму:

$$(141) \ldots\ldots \left\{ \begin{aligned} Y-\bar{y} &= \frac{\sigma_y}{\sigma_x} r_{xy}(x-\bar{x}) \\ X-\bar{x} &= \frac{\sigma_x}{\sigma_y} r_{xy}(y-\bar{y}) . \end{aligned} \right.$$

(*D*) Выраженіе (140) все еще не удобно для вычисленія, такъ какъ въ него входятъ произведенія многозначныхъ чиселъ $x-\bar{x}$ и $y-\bar{y}$, которыя оказываются таковыми вслѣдствіе того, что среднія ариѳметическія $\bar{x}$ и $\bar{y}$ очень рѣдко и только случайно бываютъ цѣлыми числами. Чтобы найти формулу удобную для вычисленій, перемножимъ $(x-\bar{x})$ и $(y-\bar{y})$, а затѣмъ найдемъ соотвѣтствующія суммы. Мы будемъ имѣть:

---

[1]) Если отклоненія двухъ величинъ отъ средняго значенія обозначить соотвѣтственно δx и δy, то получится также часто встрѣчающаяся формула: $r_{xy} = \frac{\Sigma \delta x \delta y}{N\sigma_x\sigma_y}$, которую пишутъ еще и такъ: $N\sigma_x\sigma_y r_{xy} = \Sigma \delta x \delta y$.

$$(142) \quad \ldots \quad \sum n_{xy}(x-\bar{x})(y-\bar{y}) =$$
$$= \sum n_{xy}xy - \sum n_{xy}x\bar{y} - \sum n_{xy}y\bar{x} + \sum n_{xy}\bar{x}\,\bar{y} =$$
$$= \sum n_{xy}xy - \bar{y}\sum n_{xy}x - \bar{x}\sum n_{xy}y + \bar{x}\bar{y}\sum n_{xy} =$$
$$= \sum n_{xy}xy - \bar{y}\,.\,N\bar{x} - \bar{x}\,.\,N\bar{y} + \bar{x}\bar{y}\,.\,N =$$
$$= \sum n_{xy}xy - N\bar{x}\bar{y}.$$

Подставляя найденное выраженіе въ (140), получимъ:

$$(143) \quad \ldots\ldots \quad r = \frac{\frac{1}{N}\sum n_{xy}xy - \bar{x}\,\bar{y}}{\sigma_x\sigma_y}.$$

Это и есть формула наиболѣе удобная для вычисленія коэффиціента корреляціи, въ особенности если $\Sigma n_{xy}xy$ замѣнить однимъ изъ выраженій: (138) или (139).

§ 8. Средняя ошибка уравненія регрессіи.

Если двѣ величины связаны строгой функціональной зависимостью линейнаго типа, то ур-іе

$$y = a + bx$$

даетъ возможность по одной величинѣ опредѣлить значеніе другой. Въ случаѣ зависимости корреляціонной такое опредѣленіе невозможно, и чтобы убѣдиться въ этомъ, стоитъ только бросить бѣглый взглядъ на любую таблицу корреляціи. Мы увидимъ, что ростъ сыновей далеко не опредѣляется еще ростомъ отцовъ, что при одинаковыхъ цѣнахъ въ одномъ пунктѣ цѣны въ другомъ могутъ колебаться въ довольно широкихъ предѣлахъ и т. д. Однако, нетрудно сдѣлать и дальнѣйшее наблюденіе. Именно, тѣ предѣлы, въ которыхъ колеблется одна изъ величинъ *при опредѣленномъ значеніи другой,* сужены по сравненію съ ея колебаніями въ общемъ случаѣ, т. е. по сравненію съ тѣми колебаніями ея значеній, которымъ она можетъ подвергаться *при всевозможныхъ значеніяхъ другой величины.*

Это уменьшеніе можетъ быть настолько велико, что разностями значеній, которыя величина принимаетъ при опредѣленномъ значеніи другой, для нѣкоторыхъ цѣлей можно даже вовсе пренебречь. Формула регрессіи будетъ тогда служить для опредѣленія значенія одной величины по данному значенію другой, съ тѣмъ отличіемъ отъ формулы, выражающей строгую функціональную зависимость, что результатъ вычисленія, независимо отъ точности самого измѣренія, будетъ только въ большей или меньшей степени приближеннымъ. Эти сображенія можно распространить не только на случаи корреляціи, *близкой* къ функціональной зависимости, но и на *всѣ* случаи вообще. Формула регрессіи даетъ намъ величину средняго значенія одной величины при данномъ значеніи другой. Индивидуальное значеніе величины въ отдѣльномъ случаѣ будетъ отклоняться отъ этого средняго, и, если бы мы могли найти среднюю величину и законъ распредѣленія этихъ отклоненій, мы могли бы примѣнять формулу регрессіи къ отдѣльнымъ случаямъ. Наше сужденіе носило бы тогда такой характеръ: если ростъ отца равенъ x сантиметровъ, то ростъ сына будетъ y сантиметровъ $\pm$ такая то средняя ошибка. Если распредѣленіе слѣдуетъ закону Гаусса или вообще какому нибудь извѣстному намъ закону, то наше предсказаніе можетъ быть приведено къ другой болѣе опредѣленной формѣ, т. к. мы можемъ указать тогда, что, напр., въ $^1/_2$, $^3/_4$, $90^0/_0$ всѣхъ случаевъ ростъ сына будетъ отличаться отъ y не больше, чѣмъ на такую-то величину.

Ошибка, которую мы сдѣлаемъ, прилагая формулу регрессіи къ отдѣльному случаю, очевидно, будетъ равняться

$$y-(a+bx),$$

а *средняя квадратичная ошибка* всѣхъ такихъ опредѣленій равна будетъ:

$$(144) \quad \Sigma_y=\sqrt{\frac{\sum[y-(a+bx)]^2}{N}}.$$

Ур-іе регрессіи выведено было въ § 6 подъ тѣмъ условіемъ, чтобы оно обращало въ минимумъ выраженіе

$$\sum[y-(a+bx)]^2.$$

Слѣдовательно, пользуясь этимъ ур-іемъ для опредѣленія y въ отдѣльныхъ частныхъ случаяхъ, мы сдѣлаемъ при этомъ рядъ ошибокъ, сумма квадратовъ которыхъ будетъ наименьшая. Такимъ образомъ ур-іе регрессіи есть наилучшая изъ всѣхъ возможныхъ формулъ зависимости линейнаго типа.

Самую величину средней ошибки найти не трудно. Пусть x и y будутъ отклоненіями отъ среднихъ значеній. Тогда ур-іе регрессіи приметъ простую форму $y = \frac{\sigma_y}{\sigma_x} rx$, и для средней квадратичной ошибки мы получимъ:

$$N\Sigma_y^2 = \sum\left(y - \frac{\sigma_y}{\sigma_x} rx\right)^2 = \sum y^2 - 2\frac{\sigma_y}{\sigma_x} r \sum xy + \frac{\sigma_y^2}{\sigma_x^2} r^2 \sum x^2 =$$

$$= N\sigma_y^2 - 2\frac{\sigma_y}{\sigma_x} r \,.\, N\sigma_x\sigma_y r + \frac{\sigma_y^2}{\sigma_x^2} r^2 \,.\, N\sigma_x^2 = N\sigma_y^2(1 - r^2),$$

откуда:

(145) . . . . $\Sigma_y^2 = \sigma_y^2(1 - r^2)$ и

(146) . . . . $\Sigma_y = \sigma_y \sqrt{1 - r^2}$.

Выраженіе (145) позволяетъ сдѣлать рядъ важныхъ заключеній относительно величины коэффиціента корреляціи. Одно изъ нихъ мы имѣли выше (стр. 70), но приведенное тамъ доказательство нельзя счесть строгимъ.

Σ_y^2 представляетъ изъ себя сумму квадратовъ, слѣдовательно, это есть величина всегда положительная, равно какъ и σ_y^2. Мы заключаемъ отсюда, что

$$1 - r^2 \geqslant 0,$$

$$r^2 \leqslant 1 \text{ и}$$

(147) . . . . $-1 \leqslant r \leqslant +1$.

Итакъ, *коэффиціентъ корреляціи по своему численному значенію не можетъ быть больше единицы.*

Далѣе. Если $r = 1$, то $1 - r^2 = 0$, и $\Sigma_x^2 = 0$. Но Σ_x^2 есть сумма положительныхъ величинъ и можетъ равняться нулю только въ томъ случаѣ, когда каждая изъ нихъ въ отдѣльности равна нулю. Мы имѣемъ, такимъ образомъ, для каждаго значенія y равенство:

$$\left(y - \frac{\sigma_y}{\sigma_x} rx\right)^2 = 0 \text{ или}$$

$$y = \frac{\sigma_y}{\sigma_x} rx.$$

Такимъ образомъ, *коэффиціентъ корреляціи равняется положительной или отрицательной единицѣ только въ томъ случаѣ, когда ур-іе регрессіи удовлетворяется каждой парой значеній находящихся въ корреляціи величинъ, т. е., когда корреляція переходитъ въ строгую функціональную и притомъ линейную зависимость.*

Геометрическое изображеніе этого случая было также указано выше на стр. 60, гдѣ мы говорили, что по мѣрѣ приближенія корреляціи къ типу строго-функціональной зависимости, точки поля корреляціи будутъ группироваться все тѣснѣе и тѣснѣе около одного направленія и наконецъ расположатся (при линейной регрессіи) на одной прямой линіи. Средняя ошибка опредѣленія y по x въ этомъ случаѣ обращается въ ноль.

Формула (146) показываетъ, что въ томъ случаѣ, когда величина коэффиціента корреляціи незначительна, средняя ошибка (Σ) почти не будетъ отличаться отъ средняго отклоненія (σ). Это значитъ, что распредѣленіе значеній y въ каждомъ отдѣльномъ строѣ мало разнится отъ распредѣленія ихъ въ предѣлахъ всей совокупности, и что, слѣдовательно, предсказаніе, которое мы можемъ сдѣлать по данной величинѣ x относительно y должно быть весьма несовершеннымъ. Во всякомъ случаѣ оно почти не будетъ отличаться по своей точности отъ сужденія, которое мы можемъ составить себѣ по величинѣ средняго ариѳметическаго и средняго отклоненія, позволяющихъ, напр., при нормальномъ характерѣ распредѣленія указывать предѣлы, въ которыхъ съ извѣстной вѣроятностью можно ожидать встрѣтить значенія величины. Примѣромъ можетъ служить табл. VII-я и чер. 17.

Другая картина получается, если коэффиціентъ корреляціи близокъ къ единицѣ. Пусть, напр., $x = 200$ см., $y = 200$ см., $\sigma_x = 50$ см., $\sigma_y = 50$ см., $r_{xy} = 0{,}999$, и пусть распредѣленіе слѣдуетъ закону Гаусса. Тогда каждая изъ величинъ будетъ вообще измѣняться въ довольно широкихъ предѣлахъ, т. к. между $200 + 50$ и $200 - 50$ будетъ лежать около $^2/_3$ всѣхъ случаевъ, но при данной величинѣ одного перемѣннаго распредѣленіе величинъ другого будетъ весьма сжатымъ. Въ самомъ дѣлѣ, $\Sigma_y = \sigma_y\sqrt{1 - r^2} = 50 \,.\, \sqrt{1 - 0{,}999^2} = 50 \,.\, 0{,}045 = 2{,}25$ см. Такимъ образомъ, измѣнчивость величины редуцирована до 4,5% своей общей величины, и если мы изъ ур-ія регрессіи, напр., для $x = 210$ см. найдемъ среднее значеніе y-овъ, то оно будетъ равно $y_x = 200 + 0{,}999 \,.\, (210 - 200) = 209{,}99$ см., и всѣ отдѣльныя значенія y расположатся около этой величины такъ, что двѣ трети всѣхъ случаевъ будутъ лежать въ сравнительно узкихъ границахъ между $209{,}99 - 2{,}25$ см. и $209{,}99 + 2{,}25$ см., т. е. между 207,74 см. и 212,24 см.

Формула (146) даетъ намъ среднюю ошибку въ опредѣленіи y для *всѣхъ* случаевъ, входящихъ въ составъ совокупности, но среднія ошибки въ различныхъ строяхъ вообще не равны другъ другу. Это обстоятельство находится въ связи съ тѣмъ, одинаковы-ли среднія отклоненія во всѣхъ строяхъ или нѣтъ. Въ случаѣ равенства ихъ мы будемъ имѣть *равноизмѣнчивую* совокупность (homoscedastic), въ противномъ случаѣ, *разноизмѣнчивую* (heteroscedastic)[1]. Распредѣленіе можетъ принадлежать, какъ къ одному, такъ и къ другому типу, несмотря на характеръ регрессіи, хотя обыкновенно линейная регрессія сопровождается равноизмѣнчивостью, а криволинейная разноизмѣнчивостью.

Если регрессія линейна, и среднія отклоненія (σ_{n_x}) во всѣхъ строяхъ одинаковы, то и среднія ошибки въ нихъ также будутъ равны. Въ самомъ дѣлѣ, обозначая символомъ $\sum_i$ суммированіе величинъ, принадлежащихъ i-ому строю, и буквой ${}_i\Sigma_y$ среднюю ошибку въ томъ-же строѣ, будемъ имѣть

$$
{}_i\Sigma_y^2 = \frac{1}{n_{x_i}} \sum\nolimits_i (y - \rho_{y(x)} x)^2 = \frac{1}{n_{x_i}} \sum\nolimits_i [(y - y_x) + (y_x - \rho_{y(x)} x)]^2 =
$$

$$
= \frac{1}{n_{x_i}} \sum\nolimits_i (y - y_x)^2 = \sigma_{n_{x_i}}^2,
$$

ибо $y_x - \rho_{y(x)} x = 0$ въ силу линейности регрессіи.

Средняя ошибка для всей совокупности равна будетъ той-же величинѣ. Въ самомъ дѣлѣ, суммированіе отдѣльныхъ ошибокъ можно произвести по отдѣльнымъ строямъ и мы будемъ имѣть:

$$
N\Sigma_y^2 = \sum (y - \rho_{y(x)} x)^2 = \sum\nolimits_1 (y - \rho_{y(x)} x)^2 + \sum\nolimits_2 (y - \rho_{y(x)} x)^2 + \ldots
$$

$$
+ \sum\nolimits_p (y - \rho_{y(x)} x)^2 =
$$

$$
= n_{x_1} \cdot {}_1\Sigma_y^2 + n_{x_2} \cdot {}_2\Sigma_y^2 + \ldots + n_{x_p} \cdot {}_p\Sigma_y^2.
$$

Но по доказанному всѣ среднія ошибки отдѣльныхъ строевъ равны другъ другу, равны σ_{n_x}. Поэтому будемъ имѣть:

$$
\Sigma_y = \sigma_{n_x},
$$

а, слѣдовательно, въ силу (146) въ случаѣ равноизмѣнчиваго распредѣленія и линейной регрессіи *среднее отклоненіе* каждаго строя

$$
(148) \quad \ldots\ldots\ldots \quad \sigma_{n_x} = \sigma_y \sqrt{1 - r^2}.
$$

---

[1] Терминологія Пирсона (On the General Theory of Skew Correlation etc., p. 22).

Если распредѣленіе, кромѣ того, не далеко отклоняется отъ нормальнаго, то мы можемъ найти и вѣроятную ошибку, и, такимъ образомъ, для опредѣленія *индивидуальнаго* значенія одной величины по другой, коррелятивно съ ней связанной, получимъ формулу:

$$(149) \quad . \; . \; . \; . \; . \quad y = \bar{y} + \rho_{y(x)}(x - \bar{x}) \pm 0{,}67449\, \sigma_y \sqrt{1 - r^2}.$$

Примѣръ приложенія этихъ формулъ встрѣтится намъ ниже.

§ 9. *Прямыя регрессіи.*

Ур-іе регрессіи можно представить въ слѣдующемъ видѣ:

$$\frac{Y}{\sigma_y} = r\frac{x}{\sigma_x},$$

гдѣ Y и x — отклоненія отъ среднихъ ариѳметическихъ. Если каждую изъ этихъ величинъ мы измѣримъ при помощи ея средняго отклоненія, т. е. если положимъ:

$$\frac{Y}{\sigma_y} = \eta, \quad \frac{x}{\sigma_x} = \xi,$$

то ур-іе регрессіи для y относительно x приметъ чрезвычайно простую форму

$$(150) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad \eta = r\xi,$$

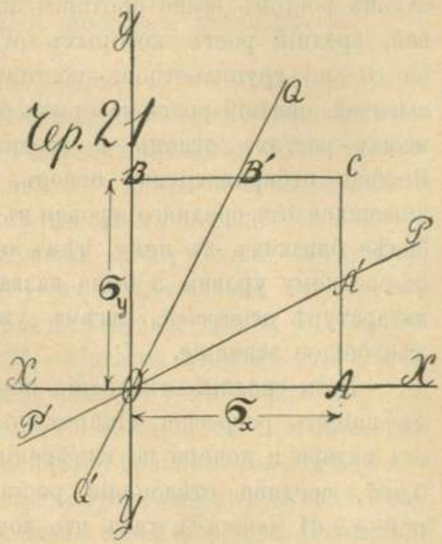
и аналогично ур-іе регрессіи для x относительно y будетъ:

$$(151) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad \xi = r\eta.$$

Геометрически r представляетъ изъ себя здѣсь тангенсъ угла, образуемаго линіей регрессіи съ соотвѣтствующей осью, и мы видимъ, что если для каждой величины принять за единицу ея среднее отклоненіе, то эти углы будутъ равны, какъ показано на чертежѣ 21.

Отложимъ на оси X отрѣзокъ OA, равный $\sigma_x = 1$, и на оси Y — отрѣзокъ $OB = \sigma_y = 1$ и построимъ на этихъ отрѣзкахъ квадратъ $ACBO$. Прямая регрессіи для y отсѣчетъ на AC отрѣзокъ AA', а прямая регрессіи для x отсѣчетъ на BC отрѣзокъ BB'. Такъ какъ $\frac{AA'}{OA} = \frac{BB'}{OB} = r$, то, очевидно, и отношеніе $\frac{AA'}{AC} = \frac{BB'}{BC} = r$. Каждая линія регрессіи дѣлитъ соотвѣтствующую сторону квадрата въ отношеніи равномъ коэффиціенту корреляціи. Мы имѣемъ здѣсь геометрическое изображеніе замѣчательнаго соотношенія, выражаемаго формулами (150) и (151). Если для удобства взять численный примѣръ, то соотношеніе это можно выразить слѣдующимъ образомъ. Пусть коэффиціентъ корреляціи равняется 0,5. Тогда отклоненію величины x отъ своего средняго на величину средняго отклоненія будетъ соотвѣтствовать *въ среднемъ* отклоненіе величины y на половину своего средняго отклоненія. Наоборотъ, отклоненію величины y на величину своего средняго отклоненія будетъ соотвѣтствовать отклоненіе x въ среднемъ также на половину величины своего средняго отклоненія. Если x отклонится, напр., на $^1/_4$ своего средняго отклоненія, то среднее ариѳметическое соотвѣтствующихъ отклоненій y-а равно будетъ $^1/_8$ средняго отклоненія y-а, и т. д.

Чер. 21



Если коэффиціентъ корреляціи равенъ нулю, то линіи регрессіи совпадаютъ съ осями XX и YY. Среднее значеніе отклоненій одной величины равно нулю при всякомъ отклоненіи другой. Если коэффиціентъ корреляціи равенъ единицѣ, то (форм. 150 и 151) $\eta = \xi$: отклоненія обѣихъ величинъ, выраженныя въ единицахъ ихъ среднихъ отклоненій, будутъ равны, и линіи регрессіи сольются, совпадая съ прямой, наклоненной къ осямъ подъ угломъ 45°.

Если мы теперь перейдемъ къ обычнымъ единицамъ измѣренія каждой величины, то подобное симметричное соотношеніе уклоненій и симметричное расположеніе прямыхъ регрессіи найдемъ только въ томъ случаѣ, когда среднія отклоненія обѣихъ величинъ равны будутъ другъ другу. Такъ приблизительно обстоитъ дѣло въ области наслѣдственности, т. к. здѣсь среднія отклоненія величины признака у родителей и дѣтей отличаются лишь незначительно. Если мы пренебрежемъ этой все-таки наблюдаемой разницей, то получимъ очень простое соотношеніе: например, при коэффиціентѣ корреляціи равномъ 0,5 (типичная средняя величина въ этой области) среднiй размѣръ отклоненія величины признака у сыновей окажется равнымъ половинѣ отклоненія величины этого признака у отца. Такимъ образомъ, группа отцовъ ростомъ выше средняго на 20 см. будетъ имѣть сыновей, среднiй ростъ которыхъ отличается отъ средняго только на 10 см.; группа отцовъ ростомъ ниже средняго будетъ имѣть сыновей, среднiй ростъ которыхъ окажется занимающимъ середину между ростомъ отцовъ и среднимъ ростомъ всего населенія. Вообще, отбирая группу отцовъ, мы получимъ сыновей, отклоняющихся отъ средняго уровня въ томъ-же направленіи, но только болѣе близкихъ къ нему, чѣмъ отцы. Это какъ бы возвращеніе къ среднему уровню и было названо первоначально въ англійской литературѣ *регрессіей,* затѣмъ уже терминъ этотъ получилъ болѣе общее значеніе.

Если среднія отклоненія неравны, то неравны будутъ и коэффиціенты регрессіи. Напримѣръ, коэффиціентъ корреляціи роста матери и дочери по измѣреніямъ Пирсона оказался равнымъ 0,507, среднее отклоненіе роста матерей — 2,39 дюйма, дочерей — 2,61 дюйма [1]), такъ что дочери оказались болѣе измѣнчивыми, чѣмъ матери [2]). Въ результатѣ мы имѣемъ такое на первый взглядъ странное соотношеніе, что дочери болѣе похожи на матерей, чѣмъ матери на дочерей [3]). Въ самомъ дѣлѣ, коэффиціентъ регрессіи для дочерей равенъ $\frac{2,61}{2,39} . 0,507 = 0,55$, а для матерей $= \frac{2,39}{2,61} . 0,507 = 0,46$. Поэтому группа матерей ростомъ на 10 см. выше средняго (всѣхъ матерей) имѣетъ дочерей, среднiй ростъ которыхъ на 5,5 см. больше средняго роста всѣхъ дочерей. Наоборотъ, группа дочерей, ростъ которыхъ на 10 см. выше

---

1) K. Pearson and A. Lee, On the Laws of Inheritance in Man. Biometrica, Vol. II, p. 370, 378.

2) Причины этого могли быть различны, напримѣръ, тутъ могло оказать вліяніе то обстоятельство, что матери представляютъ изъ себя болѣе узкую группу, ибо не всѣ дочери дѣлаются въ свою очередь матерями.

3) K. Pearson, Regression, Heredity and Panmixia, Phil. Trans. Vol. 187 A, p. 276.

средняго, имѣетъ своими матерями женщинъ, средній ростъ которыхъ только на 4,6 см. выше средняго. Короче, дочери въ среднемъ больше приближаются къ матерямъ и дальше удалены отъ средняго общаго уровня, чѣмъ то можно сказать о матеряхъ опредѣленной группы дочерей.

§ 10. Примѣръ вычисленія таблицы корреляціи.

Чтобы облегчить примѣненіе формулъ, мы иллюстрируемъ способъ вычисленія коэффиціента корреляціи на слѣдующемъ примѣрѣ, всѣ числа котораго вполнѣ вымышлены и подобраны такимъ образомъ, чтобы, поскольку возможно, упростить ариѳметическія операціи; для той-же цѣли число группъ взято меньше обычнаго. [1])

Чтобы получить для воображенія точку опоры, допустимъ, что въ этой таблицѣ (см. стр. 88) мы имѣемъ дѣло съ цѣнами. Въ верхней горизонтальной строкѣ указаны цѣны на одномъ рынкѣ (x), въ лѣвой вертикальной — цѣны на второмъ (y). Цѣны, соотвѣтствующія серединамъ интерваловъ, (онѣ проставлены въ скобкахъ у краевъ таблицы) примемъ за варіанты соотвѣтствующихъ строевъ и допустимъ, какъ первое приближеніе, что всѣ цѣны, относящіяся къ случаямъ, зарегистрированнымъ въ отдѣльной колонѣ или строкѣ, одинаковы и равны своей варіантѣ. Вычисленные на основаніи такого допущенія моменты будутъ „грубыми".

Если общее число случаевъ велико, сдѣланныя ошибки отчасти взаимно уничтожатся: къ первымъ моментамъ и къ моменту произведенію (Σxy) поправка не нужна, вторые-же моменты должны быть исправлены или по системѣ Шеппарда [Ч. I форм. (34)],

---

[1]) При всѣхъ вычисленіяхъ, связанныхъ съ примѣненіемъ метода корреляціи, незамѣнимыя услуги оказываетъ ариѳмометръ. Хотя въ окончательномъ результатѣ рѣдко бываетъ возможно, да и нужно, удерживать больше двухъ, трехъ десятичныхъ знаковъ, однако тѣ величины, которыя играютъ промежуточную роль и нужны въ дальнѣйшихъ вычисленіяхъ, должны быть найдены съ большей точностью для того, чтобы ошибки вычисленія не накопились до величины сравнимой, или даже большей, чѣмъ вѣроятная ошибка результата. При употребленіи ариѳмометра не составляетъ труда вести вычисленіе съ 5-ю или даже съ 6-ю десятичными знаками, гарантируя, такимъ образомъ, точное совпаденіе результатовъ вычисленій двухъ изслѣдователей одного и того-же матеріала. Если можно удовлетвориться меньшей точностью, то при обширныхъ вычисленіяхъ допустимо пользованіе хорошей логариѳмической линейкой для сокращенія труда, который иначе могъ бы сдѣлаться непосильнымъ.

если числа ряда сходятъ постепенно на нѣтъ, или по методу трапецій (Ч. I, форм. 47).

Одну изъ среднихъ варіантъ примемъ за 0, а остальныя обозначимъ: —1, —2; +1, +2. Эти *условныя варіанты* вписаны во второй сверху горизонтальной и во второй слѣва вертикальной строкахъ и кромѣ того для удобства повторены съ другой стороны таблицы; (ихъ слѣдуетъ вписывать краснымъ черниломъ, а здѣсь они набраны жирнымъ шрифтомъ). Такимъ образомъ, мы условно

			(45)	(55)	(65)	(75)	(85)		a_2	b_2	c_2	d_2
			40—50	50—60	60—70	70—80	80—90	y	n_y	$n_y x_y$	$n_y x_y y$	x_y
			—2	**—1**	**0**	**+1**	**+2**					
(75)	80—70	**+2**			1	2	1	**+2**	*4*	*+ 4*	*+ 8*	*+ 1,0000*
(65)	70—60	**+1**		2	5	4	5	**+1**	*16*	*+12*	*+12*	*+0,7500*
(55)	60—50	**0**	1	5	6	4	1	**0**	*17*	*— 1*	*0*	*—0,0588*
(45)	50—40	**—1**	5	3	1	1		**—1**	*10*	*—12*	*+12*	*— 1,2000*
(35)	40—30	**—2**	2	1				**—2**	*3*	*— 5*	*+10*	*— 1,6667*
	x		**—2**	**—1**	**0**	**+ 1**	**+ 2**					
a_1	n_x		*8*	*11*	*13*	*11*	*7*		*50*			
b_1	$n_x y_x$		*— 9*	*- 3*	*+ 6*	*+ 7*	*+ 7*					
c_1	$n_x y_x x$		*+ 18*	*+ 3*	*0*	*+ 7*	*+14*				*+42*	
d_1	y_x		*—1,1250*	*—0,2727*	*+0,4615*	*+0,6364*	*+1,0000*					

обозначили нулемъ x равное 65 коп. и y равное 55 коп. Затѣмъ x равное 55 коп. мы приняли за —1, 45 коп. за —2, 75 коп. за +1, 85 коп. за +2; подобнымъ-же образомъ поступили съ y-омъ. Наши *условныя единицы* равны, слѣдовательно, $k_x = 10$ коп. и $k_y = 10$ к.

Справа и внизу отведемъ по 4 свободныхъ строки (обозначенныхъ a_1, b_1, c_1, d_1 и a_2, b_2, c_2, d_2).

Найдемъ *итоги вертикалей и горизонталей*, то есть, численности всѣхъ x-овыхъ и y-овыхъ строевъ и впишемъ ихъ въ строки (a_1) и (a_2). Строка (a_1) дастъ намъ распредѣленіе первой цѣны (x), а строка (a_2) — распредѣленіе второй (y). Найдемъ *итоги* строкъ (a_1) *и* (a_2); они должны совпадать (провѣрка!) и равняться всему числу взятыхъ случаевъ. Этотъ итогъ впишемъ въ клѣткѣ пересѣченія (a_1) и (a_2). Мы найдемъ: $N = 50$.

Вычислимъ теперь *среднія ариѳметическія*. Для этого каждое n_x и n_y умножимъ на условную варіанту и сумму раздѣлимъ на N. Мы найдемъ:

(I)

$13 . 0 = 0$	
$11 . (-1) = -11$	$11 . (+1) = +11$
$8 . (-2) = -16$	$7 . (+2) = +14$
-27	$+25$

$$\bar{x} = -2 : 50 = -0{,}04 .$$

(II)

$17 . 0 = 0$	
$10 . (-1) = -10$	$16 . (+1) = 16$
$3 . (-2) = -6$	$4 . (+2) = 8$
-16	$+24$

$$\bar{y} = +8 : 50 = +0{,}16 .$$

Это будутъ (въ условныхъ единицахъ) среднія ариѳметическія, равныя первымъ грубымъ моментамъ $\nu'_{1(x)}$ и $\nu'_{1(y)}$.

Найдемъ теперь *вторые грубые моменты*. Умножая соотвѣтствующія n_x и n_y на квадраты условныхъ варіантъ, будемъ имѣть:

(I)	(II)
$8 . 4 = 32$	$4 . 4 = 16$
$11 . 1 = 11$	$16 . 1 = 16$
$13 . 0 = 0$	$17 . 0 = 0$
$11 . 1 = 11$	$10 . 1 = 10$
$7 . 4 = 28$	$3 . 4 = 12$
$\nu'_{2(x)} = 82 : 50 = 1{,}64 .$	$\nu'_{2(y)} = 54 : 50 = 1{,}08 .$

Вычитая изъ полученныхъ грубыхъ вторыхъ моментовъ квадраты первыхъ грубыхъ моментовъ [Ч. I форм. (7)], получаемъ *вторые центральные грубые моменты.* Затѣмъ придаемъ поправки; въ настоящемъ случаѣ по системѣ трапецій, то есть, *прибавляя* по $^1/_6 = 0{,}166667$ (при возможности примѣненія поправокъ Шеппарда нужно *вычесть* $^1/_{12} = 0{,}083333$). Получимъ центральные истинные моменты ($\mu_{2(x)}$ и $\mu_{2(y)}$). Извлекая квадратный корень, будемъ имѣть главныя отклоненія σ_x и σ_y. Вычисленіе, составляя продолженіе предыдущаго, располагается такъ:

(I)

$$\left.\begin{array}{r} \nu'_{2(x)} = 1{,}640000 \\ \bar{x}^2 = 0{,}001600 \end{array}\right\} -$$

$$\left.\begin{array}{r} \nu_{2(x)} = 1{,}638400 \\ +\,^1/_6 = 0{,}166667 \end{array}\right\} +$$

$$\mu_{2(x)} = 1{,}805067 = \sigma_x^2$$

$$\sigma_x = \sqrt{\mu_{2(x)}} = 1{,}344.$$

(II)

$$\left.\begin{array}{r} \nu'_{2(x)} = 1{,}080000 \\ \bar{y}^2 = 0{,}025600 \end{array}\right\} -$$

$$\left.\begin{array}{r} \nu_{2(y)} = 1{,}054400 \\ +\,^1/_6 = 0{,}166667 \end{array}\right\} +$$

$$\mu_{2(y)} = 1{,}221067 = \nu_y^2$$

$$\sigma_y = \sqrt{\mu_{2(y)}} = 1{,}105$$

Находимъ также нужныя для дальнѣйшаго

$$\bar{x}\,\bar{y} = -\,0{,}0064 \quad \text{и} \quad \sigma_x\sigma_y = 1{,}485\,.$$

Приступаемъ теперь къ вычисленію *коэффиціента корреляціи.* Для этого найдемъ сумму $\sum n_{xy}xy$ и для контроля—двумя способами: сначала по формулѣ $\sum n_{xy}xy = \sum n_x y_x x$, затѣмъ по формулѣ $\sum n_{xy}xy = \sum n_y x_y y$. Замѣтимъ, что y_{x_i} есть средняя величина y-а въ i-ой вертикальной колонѣ, слѣд., равна $\sum_i n_{x_iy} y / n_{x_i}$, откуда $n_{x_i} y_{x_i} = \sum_i n_{x_iy} y$.

Начинаемъ съ лѣвой колоны. Число въ каждой клѣткѣ умножаемъ на соотвѣтствующую условную варіанту y-а и сумму вписываемъ въ этой же колонѣ въ строку (b_1). Такъ-же поступаемъ и съ прочими колонами.

Вычисленіе $n_x y_x$:

-2	-1	0	$+1$	$+2$
1.0=0 5.(−1)=−5 2.(−2)=−4	2.(+1)=+2 5.0=0 3.(−1)=−3 1.(−2)=−2	1.(+2)=+2 5.(+1)=+5 6.0=0 1.(−1)=−1	2.(+2)=+4 4.(+1)=+4 4.0=0 1.(−1)=−1	1.(+2)=+2 5.(+1)=+5 1.0=0
$n_{x_{-2}} y_{x_{-2}}=-9$	$n_{x_{-1}} \cdot y_{x_{-1}}=-3$	$n_{x_0} y_{x_0}=+6$	$n_{x_1} y_{x_1}=+7$	$n_{x_2} y_{x_2}=+7$

Аналогичное вычисленіе проводимъ и для строкъ, заполняя при этомъ колону (b_2). Именно:

$+2$	$+1$	0	-1	-2
1.0=0 2.(+1)=+2 1.(+2)=+2	2.(−1)=−2 5.0=0 4.(+1)=+4 5.(+2)=+10	1.(−2)=−2 5.(−1)=−5 6.0=0 4.(+1)=+4 1.(+2)=+2	5.(−2)=−10 3.(−1)=−3 1.0=0 1.(+1)=+1	2.(−2)=−4 1.(−1)=−1
$n_{y_2} \cdot x_{y_2}=+4$	$n_{y_1} \cdot x_{y_1}=+12$	$n_{y_0} x_{y_0}=-1$	$n_{y_{-1}} x_{y_{-1}}=-12$	$n_{y_{-2}} x_{y_{-2}}=-5$

Чтобы получить сумму $\sum n_x y_x x$, умножаемъ каждое число строки (b_1) на соотвѣтствующую варіанту и полученныя числа вписываемъ въ строку (c_1); получимъ:

(I)

$$
\begin{aligned}
(-9)\cdot(-2) &= +18 \\
(-3)\cdot(-1) &= +3 \\
(+6)\cdot 0 &= 0 \\
(+7)\cdot(+1) &= +7 \\
(+7)\cdot(+2) &= +14 \\
\hline
\sum n_x y_x x &= +42
\end{aligned}
$$

Находимъ такимъ-же способомъ сумму $\sum(n_y x_y . y)$, заполняя колону (c_2)

$$\text{(II)}$$

$$(+4).(+2) = +8$$
$$(+12).(+1) = +12$$
$$(-1).0 = 0$$
$$(-12).(-1) = +12$$
$$(-5).(-2) = +10$$
$$\sum n_y x_y y = +42 \quad ^{1)}$$

Суммы сошлись, и мы можемъ сразу найти коэффиціентъ корреляціи [форм. (143)]:

$$r = \frac{\dfrac{\sum n_{xy} xy}{N} - \bar{x}\bar{y}}{\sigma_x \sigma_y} = \frac{+\dfrac{42}{50} - (-0{,}0064)}{1{,}485} = +0{,}5700$$

и коэффиціенты регрессіи [форм. (135)]:

$$\rho_{y(x)} = \frac{\sigma_y}{\sigma_x} r = +0{,}469 \qquad \rho_{x(y)} = \frac{\sigma_x}{\sigma_y} r = +0{,}693\,.$$

Вѣроятныя ошибки полученныхъ величинъ найдутся по формуламъ, которыя мы даемъ въ слѣдующемъ параграфѣ. Если мы вообще хотимъ получить окончательные результаты не въ условныхъ единицахъ, то *переходъ къ обычнымъ единицамъ* удобнѣе сдѣлать до вычисленія вѣроятныхъ ошибокъ. Вспомнимъ, что мы приняли за условныя единицы $k_x = 10$ коп. и $k_y = 10$ коп. Чтобы вернуться къ обычному способу счета, нужно $\bar{x}$ и σ_x помножить на k_x, а $\bar{y}$ и σ_y на k_y; $\rho_{y(x)}$ придется помножить на $\frac{k_y}{k_x}$, а $\rho_{x(y)}$ на $\frac{k_x}{k_y}$. Можно, конечно, коэффиціенты регрессіи не вычислять въ условныхъ единицахъ вовсе, а найти ихъ по величинамъ среднихъ отклоненій, выраженныхъ уже въ обычныхъ мѣрахъ. Это однако возможно не всегда, такъ такъ, если мы хотимъ по-

[1]) Все вычисленіе, хотя и отнимаетъ—въ особенности въ случаѣ большой таблицы—довольно много времени, все-же менѣе громоздко, чѣмъ это можетъ показаться. Въ частности при нахожденіи $\sum n_{xy} xy$ всѣ вычисленія легко производятся въ умѣ, причемъ отдѣльныя произведенія можно откидывать на счетахъ; писать приходится только результаты, заносимыя въ соотвѣтствующія графы таблицы, какъ указано выше.

строить графику регрессіи въ условномъ масштабѣ, намъ нужно и коэффиціенты регрессіи имѣть въ томъ-же масштабѣ.

Въ нашемъ случаѣ будемъ имѣть:

$$\overline{x} = -0{,}04 \times 10 = -0{,}4 \text{ коп.}; \qquad \overline{y} = +0{,}16 \times 10 = +1{,}6 \text{ коп.}$$
$$\sigma_x = 1{,}344 \times 10 = 13{,}44 \text{ коп.}; \qquad \sigma_y = 1{,}105 \times 10 = 11{,}05 \text{ коп.}$$

За условный нуль у насъ были приняты цѣны $x = 65$ и $y = 55$ коп., такъ что $\overline{x}$ и $\overline{y}$ суть отклоненія отъ этихъ величинъ. Обозначая среднія ариѳметическія, измѣренныя отъ „настоящаго" нуля черезъ h_x и h_y, получимъ окончательный результатъ:

$$h_x = 65 - 0{,}4 = 64{,}6 \text{ коп.}$$
$$h_y = 55 + 1{,}6 = 56{,}6 \text{ коп.}$$

Коэффиціенты регрессіи останутся безъ измѣненія, такъ какъ въ нашемъ случаѣ $k_y/k_x = 1$, а, слѣдовательно, ур-ія теоретическихъ прямыхъ регрессіи будутъ (см. форм. (141)):

$$Y = 56{,}6 + 0{,}469(x - 64{,}6)$$
$$X = 64{,}6 + 0{,}693(y - 56{,}6)\,.$$

Обратимся теперь къ *нахожденію линій регрессіи.* Раздѣливъ числа строки (b_1) на соотвѣтствующія числа строки (a_1), получимъ среднія ариѳметическія отдѣльныхъ x-овыхъ строевъ, то есть, величины y_x. Мы ихъ впишемъ въ строку (∂_1). Аналогичнымъ образомъ найдемъ x_y, подѣливъ числа колоны (b_2) на числа колоны (a_2). Ихъ впишемъ въ колону (∂_2). Отложивъ эти числа отъ условнаго нуля на средней линіи соотвѣтствующихъ строкъ или колонъ и соединивъ полученныя точки прямыми, получимъ эмпирическія линіи регрессіи. Чтобы найти теоретическія прямыя регрессіи, поступимъ такъ: отъ условнаго y-оваго нуля, то есть отъ середины нулевой строки, отложимъ по вертикальному направленію $\overline{y}$ (въ нашемъ случаѣ внизъ 0,04); черезъ найденную точку проведемъ горизонтальную прямую, которая будетъ центральной осью XX. Аналогично, отложивъ отъ середины нулевой колоны $\overline{x}$, найдемъ центральную ось YY. Точка ихъ пересѣченія (O) будетъ центромъ распредѣленія нашей совокупности. Затѣмъ (см. чер. 18), отъ точки O отложимъ произвольный отрѣзокъ OA' и на перпендикулярѣ къ нему въ точкѣ A' отложимъ отрѣзокъ $A'A = \rho_{y(x)} \cdot OA'$. Прямая OA и будетъ искомой прямой регрессіи для y по x. Аналогично найдемъ и другую прямую OB.

§ 11. Генеральная совокупность и пробная группа.

Предположимъ, что мы имѣемъ очень обширную совокупность случаевъ, изъ которыхъ каждый характеризуется парой величинъ x и y. Вся эта совокупность въ цѣломъ пусть будетъ недоступна перечисленію и измѣренію, мы назовемъ ее *генеральной совокупностью* (general population) и поставимъ своей задачай узнать характеризующія ее величины по *пробной группѣ* (random sample), взятой въ качествѣ образца изъ генеральной совокупности и составленной для этой цѣли изъ индивидуумовъ, взятыхъ на удачу (at random)

Генеральная совокупность характеризуется: средними ариѳметическими $\bar{h}_1$ и $\bar{h}_2$, средними ариѳметическими $\bar{\sigma}_1$ и $\bar{\sigma}_2$, коэффиціентами регрессіи $\bar{\rho}_{1(2)}$ и $\bar{\rho}_{2(1)}$, коэффиціентомъ корреляціи $\bar{r}_{12}$ и т. д. Вслѣдствіе случайности состава пробной группы мы не можемъ, конечно, ожидать, чтобы ея постоянныя распредѣленія ($h_1, h_2, \sigma_1, \sigma_2, \rho_{1(2)}, \rho_{2(1)}, r_{12}$ и т. д.) совпадали съ соотвѣтствующими величинами для генеральной совокупности въ каждомъ отдѣльномъ случаѣ. Только по мѣрѣ того, какъ мы будемъ брать все большее и большее число пробныхъ группъ, для каждой постоянной будетъ получаться все большее и большее число значеній, и среднее ариѳметическое ихъ будетъ приближаться къ той величинѣ, которую эта постоянная имѣетъ въ генеральной совокупности. Для безконечно большого числа пробныхъ группъ всѣ значенія каждой постоянной, напр. r, составятъ совокупность со среднимъ значеніемъ $\bar{r}$ и отклоненіями, въ каждомъ отдѣльномъ случаѣ равными $\delta r = r - \bar{r}$. Также и для другихъ постоянныхъ. Въ качествѣ перваго приближенія можно допустить, что распредѣленіе величинъ δr, $\delta\rho$, $\delta\sigma$, δh будетъ удовлетворять закону Гаусса [1]). Именно, меньшія отклоненія будутъ встрѣчаться чаще,

---

[1]) См. K. Pearson and L. Filon, On the Probable Errors of Frequency Constants, Phil. Trans., Vol. 191 A, 1898. Выводы, къ которымъ пришли авторы этого замѣчательнаго мемуара, носятъ, какъ они и сами это отмѣчаютъ (p. 234), лишь приближенный характеръ. Это обстоятельство должно стать яснымъ, по крайней мѣрѣ относительно величины r, даже и для неспеціалиста въ этихъ вопросахъ, если сообразить, что въ Гауссовомъ (нормальномъ) распредѣленіи возможны всѣ величины отклоненій отъ $+\infty$ до $-\infty$; величина-же r можетъ измѣняться лишь отъ $+1$ до -1. Законъ распредѣленія ея будетъ другой. Для пробныхъ группъ изъ большого

большія рѣже, и вѣроятность всякаго отклоненія можетъ быть найдена изъ обычныхъ таблицъ интеграла вѣроятностей, если мы знаемъ среднее отклоненіе для данной величины (напр. для δr). Если бы мы произвели на самомъ дѣлѣ опытъ составленія множества пробныхъ группъ изъ одной и той-же генеральной совокупности, то мы могли бы получить среднее отклоненіе для каждой изъ ошибокъ δh, $\delta \sigma$, $\delta \rho$, δr эмпирически, примѣнивъ указанные въ I ч. способы вычисленія. Но этотъ путь слишкомъ труденъ, и поэтому теорія ошибокъ стремится вывести эти величины à priori на основаніи различныхъ теоретическихъ соображеній. Часть ея выводовъ носитъ общій характеръ, часть справедлива лишь въ приближеніи, такъ какъ основывается на допущеніи, что сама генеральная совокупность слѣдуетъ въ своемъ распредѣленіи закону Гаусса. Эта предпосылка, однако, умаляетъ цѣнность выводовъ лишь въ весьма слабой степени, такъ какъ среднія отклоненія и вѣроятныя ошибки подлежащихъ величинъ обыкновенно очень малы, и особенная точность въ опредѣленіи ихъ не играетъ большой роли при оцѣнкѣ результатовъ.

Разъ намъ извѣстно полученное теоретическимъ путемъ среднее отклоненіе для какой либо ошибки, напримѣръ, для δr, то мы можемъ найти и вѣроятную ошибку, умноживъ его на 0,67449. Так. обр., мы имѣемъ $E_r = 0{,}67449\Sigma_r$, $E_\rho = 0{,}67449\Sigma_\rho$ и т. д.

Еще на одно обстоятельство слѣдуетъ обратить вниманіе. Теорія показываетъ, что ошибки отдѣльныхъ постоянныхъ въ большинствѣ случаевъ не независимы, а находятся другъ съ другомъ въ корреляціи. То есть, если мы отберемъ изъ нашихъ пробныхъ группъ тѣ, въ которыхъ, напримѣръ, h_1 будетъ выше средняго, то среднее значеніе всѣхъ h_2 для тѣхъ-же группъ окажется не равно среднему своему значенію ($\bar{h}_2$) для всей совокупности пробныхъ группъ, а будетъ больше или меньше его въ зависимости отъ знака коэффиціента корреляціи $R_{h_1 h_2}$. Если бы мы на самомъ дѣлѣ составили большое число пробныхъ группъ,

---

числа индивидуумовъ это значенія не имѣетъ, такъ какъ значительныя отклоненія чрезвычайно маловѣроятны. Для очень малыхъ пробныхъ группъ и въ вопросахъ, требующихъ принятія въ разсчетъ большихъ отклоненій, съ указаннымъ обстоятельствомъ необходимо, однако, считаться. См. „Student", Probable Error of a Correlation Coefficient, Biometrica Vol. VI, p. 302—310.

то мы могли бы найти этотъ коэффиціентъ корреляціи обычнымъ путемъ по формулѣ:

$$(152) \ldots\ldots \quad M\Sigma_{h_1}\Sigma_{h_2}R_{h_1h_2} = \sum \delta h_1 \delta h_2 ,$$

гдѣ M есть число пробныхъ группъ, а Σ_{h_1} и Σ_{h_2} — среднія отклоненія для h_1 и h_2, варіирующихся вслѣдствіе случайныхъ причинъ отъ одной пробной группы къ другой. Той-же формулой пользуются и при теоретическомъ выводѣ коэффиціента корреляціи ошибокъ. Примѣръ мы будемъ имѣть ниже.

§ 12. *Вѣроятныя ошибки и коэффиціенты корреляціи между постоянными въ случаѣ нормальнаго распредѣленія.*

Выводъ вѣроятныхъ ошибокъ слишкомъ сложенъ и по соображеніямъ мѣста и времени намъ приходится отказаться отъ этой задачи. Я приведу лишь важнѣйшіе результаты относящихся сюда изслѣдованій [1]). Для полноты повторяю нѣкоторыя формулы, приведенныя въ I ч.

(a) Вѣроятныя ошибки:

$$(153) \ldots\ldots \quad E_h = 0{,}67449 \frac{\sigma}{\sqrt{N}}$$

$$(154) \ldots\ldots \quad E_\sigma = 0{,}67449 \frac{\sigma}{\sqrt{2N}}$$

$$(155) \ldots\ldots \quad E_r = 0{,}67449 \frac{1 - r^2}{\sqrt{N}}$$

$$(156) \ldots \begin{cases} E_{\rho_{12}} = 0{,}67449 \dfrac{\sigma_1}{\sigma_2} \sqrt{\dfrac{1 - r^2}{N}} , \\[2ex] E_{\rho_{21}} = 0{,}67449 \dfrac{\sigma_2}{\sigma_1} \sqrt{\dfrac{1 - r^2}{N}} \end{cases}$$

---

[1]) Кромѣ цитированнаго въ примѣч. на стр. 91 мемуара K. Pearson'a и L. Filon'a нужно назвать еще слѣдующія работы: W. F. Sheppard, On the Application of the Theory of Error to Cases of Normal Distribution and Normal Correlation, Phil. Trans. Vol. 192 A, 1899. K. Pearson, On the Mathematical Theory of Errors of Judgment, Phil. Trans. Vol. 198 A, 1902 p. 276—279. K. Pearson, On the General Theory of Skew Correlation and Non-Linear Regression, Drapers' Comp. Research Memoirs, Biom. Ser. II 1905 и небольшая, но вслѣдствіе популярности изложенія незамѣнимая редакціонная статья въ томѣ II Біометрики: On the Probable Errors of Frequency Constants.

(b) Коэффиціенты корреляціи.

(157) . . . . $R_{h_1h_2}=r_{12}$

(158) . . . . $R_{\sigma_1\sigma_2}=r^2_{12}$

(159) . . . . $R_{\sigma_1 r_{12}}=R_{\sigma_2 r_{12}}=\frac{r_{12}}{\sqrt{2}}$

(160) . . . $R_{h_1\sigma_1}=R_{h_2\sigma_2}=R_{h_1\sigma_2}=R_{h_2\sigma_1}=R_{h_1r_{12}}=R_{h_2r_{12}}=0.$

По поводу этихъ формулъ необходимо сдѣлать нѣкоторыя замѣчанія. Прежде всего очевидно, что вѣроятная ошибка каждой величины уменьшается съ увеличеніемъ численности совокупности. Кромѣ того, вѣроятная ошибка коэффиціентовъ регрессіи и корреляціи уменьшается съ увеличеніемъ r. Поэтому, чѣмъ больше корреляція, тѣмъ меньше можетъ быть число случаевъ, достаточное для того, чтобы съ несомнѣнностью опредѣлить наличность корреляціонной связи и ея величину. При $r=0{,}9$ и $N=25$ $E_r=0{,}026$,—вѣроятная ошибка составляетъ т. обр. неполныхъ 3% самой величины. Чтобы получить такое-же отношеніе вѣроятной ошибки къ коэффиціенту корреляціи, когда послѣдній равенъ 0,1, нужно располагать матеріаломъ численностью около 100000. Если допустить, что для достовѣрности опредѣленія величины необходимо, чтобы она по крайней мѣрѣ въ 5 разъ превышала свою вѣроятную ошибку, то, какъ не трудно найти, при $r=0{,}1$ N должно быть все-же не меньше 1000.

Какъ было уже замѣчено выше, приведенныя формулы носятъ приближенный характеръ. По изслѣдованіямъ „Student'a" (цитир. выше) формулу (155) для вѣроятной ошибки коэффиціента корреляціи можно примѣнять уже при $N=30$. Для группъ меньшей численности на нее полагаться не слѣдуетъ, и приходится употреблять другой способъ разсчета, котораго по его сложности я не привожу, отсылая читателя къ цитированной работѣ. Какъ практическое правило, можно принять, что при N между 20 и 30 коэффиціентъ корреляціи долженъ быть не менѣе 0,5, чтобы о самомъ существованіи корреляціонной связи можно было говорить съ нѣкоторой увѣренностью. По разсчету „Student'a" при отсутствіи корреляціи въ генеральной совокупности и при $N=21$, коэффиціентъ корреляціи можетъ случайно оказаться больше $+0{,}5$

и меньше — 0,5 лишь 2 раза на 100 пробныхъ группъ [1]). Въ концѣ концовъ нужно къ сожалѣнію признать, что до дальнѣйшаго усовершенствованія теоріи (или по крайней мѣрѣ до составленія таблицъ на основѣ формулъ „Student'a") статистику не слѣдуетъ примѣнять методъ корреляціи къ группамъ меньше, чѣмъ изъ 20 случаевъ.

Переходимъ къ формуламъ для коэффиціентовъ корреляціи между постоянными распредѣленія. Чтобы дать представленіе о ихъ значеніи, мы коснемся ихъ отношенія къ теоріямъ наслѣдственности и подбора. Если допустить, что распредѣленіе признаковъ индивидуумовъ опредѣленнаго біологическаго вида слѣдуетъ нормальному закону распредѣленія (а это если не для всѣхъ, то для многихъ видовъ и признаковъ не далеко отъ истины), то формулы (157)—(160) сразу-же дадутъ намъ рядъ важныхъ указаній. Пусть значекъ (1) относится къ размѣрамъ одного органа, (2) — къ размѣрамъ другого; r_{12} будетъ коэффиціентомъ органической корреляціи между ними, корреляціи, которую мы можемъ легко опредѣлить, измѣривъ подлежащіе размѣры для нѣсколькихъ сотенъ индивидуумовъ. Какъ будетъ дѣйствовать подборъ, направленный на измѣненіе среднихъ размѣровъ одного изъ органовъ? напр., естественный подборъ, при которомъ больше шансовъ выжить имѣютъ индивидуумы съ размѣромъ даннаго органа, соотвѣтствующимъ новымъ условіямъ, отличнымъ отъ прежнихъ? Формулы (160) показываютъ, что при этомъ останется безъ измѣненія абсолютная измѣнчивость вида по данному признаку (такъ какъ $R_{h_1\sigma_1} = 0$), измѣнчивость другихъ признаковъ ($R_{h_1\sigma_i} = 0$) и коэффиціенты корреляціи между подбираемымъ и остальными признаками. Но средніе размѣры другихъ органовъ должны будутъ измѣниться, причемъ величина измѣненія можетъ быть нами предсказана заранѣе, если мы знаемъ r_{12}, $r_{13} \ldots$, которые, какъ сказано, всегда можно опредѣлить.

Другая картина получится, если подборъ направленъ на величину средняго отклоненія, напримѣръ, если при измѣнившихся

---

[1]) Op. cit. p. 308. См. также Journ. of the Roy. Stat. Soc. 1907: Hooker, Correlation of the Weather and Crops, p. 6, и замѣчанія Edgeworth'a и Yule'я въ протоколѣ засѣданія, посвященнаго обсужденію доклада Hooker'a. Вычисленія „Student'a" въ нѣсколько разъ менѣе благопріятны для оцѣнки результатовъ работы Hooker'a, чѣмъ его собственныя соображенія.

условіяхъ наиболѣе благопріятной величиной какого либо органа останется прежній средній его размѣръ, и лишь уклоненія отъ него станутъ болѣе вредными. Средняя величина органа останется прежней ($R_{\sigma_1 h_1}=0$), не измѣнятся и среднія величины другихъ органовъ ($R_{\sigma_1 h_2}=0$), но должны будутъ измѣниться среднія ихъ отклоненія ($R_{\sigma_1\sigma_2}=r_{12}^2$) и коэффиціентъ корреляціи между ними $\left(R_{\sigma_1 r_{12}}=\frac{r_{12}}{\sqrt{2}}\right)$. При этомъ, такъ какъ r^2 есть величина сравнительно небольшая и притомъ быстро убывающая съ убываніемъ r, то, очевидно, во 1-хъ, что вліяніе подбора средняго отклоненія одного органа на величину средняго отклоненія другого органа вообще меньше, чѣмъ вліяніе подбора одного средняго размѣра на другой средній размѣръ, во 2-хъ, что вліяніе это сколько нибудь замѣтно можетъ сказываться только на органахъ со сравнительно высокой степенью корреляціонной связи. При этомъ слѣдуетъ, однако, замѣтить, что какъ бы незначительно ни было это вліяніе въ нѣкоторыхъ случаяхъ, коэффиціентъ корреляціи $R_{\sigma_i\sigma_j}=r^2$ и, слѣдовательно, представляетъ собою величину всегда положительную. Поэтому увеличеніе измѣнчивости одного органа всегда связано съ увеличеніемъ, а уменьшеніе измѣнчивости его — съ уменьшеніемъ измѣнчивости всѣхъ другихъ органовъ. Напр., если случайныя обстоятельства (или искусственный подборъ) выдѣлятъ группу, члены которой будутъ болѣе похожи другъ на друга въ одномъ какомъ либо отношеніи, то они будутъ болѣе похожи другъ на друга и во всѣхъ прочихъ отношеніяхъ [1]).

Мы далеко не исчерпали всѣхъ даже самыхъ непосредственныхъ выводовъ, которые можно сдѣлать изъ приведенныхъ формулъ, не пускаясь ни въ какія болѣе сложныя теоретическія соображенія. Но и сказаннаго, я думаю, достаточно, чтобы читатель могъ почувствовать, къ какимъ важнымъ проблемамъ въ этой области подводитъ теорія корреляціи.

§ 13. Вѣроятная ошибка разности.

Знаніе коэффиціентовъ корреляціи между постоянными распредѣленія даетъ возможность выводить дальнѣйшія формулы вѣ-

---

[1]) См. K. Pearson and L. Filon, op. cit., Phil. Trans., Vol. 191 A., p. 241 примѣч.

роятныхъ ошибокъ. Общій принципъ мы иллюстрируемъ прежде всего на важномъ случаѣ *вѣроятной ошибки разности.*

Пусть $z_0 = x_0 - y_0$ будетъ разностью двухъ постоянныхъ генеральной совокупности. Въ какой нибудь пробной группѣ онѣ получаютъ значенія z, x, y, отличающіяся отъ истинныхъ на δz, δx, δy.

Очевидно, что

$$\delta z = \delta x - \delta y, \text{ откуда}$$

$$(\delta z)^2 = (\delta x)^2 - 2\delta x \,.\, \delta y + (\delta y)^2 \,.$$

Если мы возьмемъ сумму всѣхъ подобныхъ выраженій для M пробныхъ группъ, то получимъ:

$$\sum(\delta z)^2 = \sum(\delta x)^2 - 2\sum\delta x \,.\, \delta y + \sum(\delta y)^2 \,.$$

Эти суммы намъ извѣстны и могутъ быть выражены черезъ среднія отклоненія и коэффиціенты корреляціи. Мы получимъ:

$$M\sigma_z^2 = M\sigma_x^2 - 2M\sigma_x\sigma_y r_{xy} + M\sigma_y^2 \,,$$

или, сокращая на M и извлекая квадратный корень:

(161) . . . . . . $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2 - 2\sigma_x\sigma_y r_{xy}}$.

Если величины x и y независимы, т. е., если $r_{xy} = 0$, то мы будемъ имѣть:

(162) . . . . $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$.

Если обѣ части равенствъ (161) и (162) умножить на 0,67449, то каждое среднее отклоненіе замѣнится соотвѣтствующей вѣроятной ошибкой. Слѣдовательно, формулы справедливы и для послѣднихъ. Замѣтимъ, что если $r_{xy} \gtrless 0$, то вѣроятная ошибка, которую даетъ формула (161), будетъ $\frac{\text{меньше}}{\text{больше}}$, чѣмъ вѣроятная ошибка по формулѣ (162). Слѣдовательно, статистикъ, пользуясь формулой (162) въ случаѣ зависимыхъ величинъ, рискуетъ: при наличности положительной корреляціи непризнать существенной разницу, которая является таковой, или, что еще хуже, при наличности отрицательной корреляціи признать сущесущественнымъ несущественное различіе.

Въ частности, вѣроятная ошибка разности среднихъ ариѳметическихъ выразится [см. форм. (157)] слѣд. образомъ:

(163) . . . . . . $E_{h_1 - h_2} = \sqrt{E_{h_1}^2 + E_{h_2}^2 - 2E_{h_1}E_{h_2}r_{12}}$,

а для вѣроятной ошибки разности среднихъ отклоненій получится выраженіе (форм. 158):

$$(164) \quad . \;.\;.\;. \quad E_{\sigma_1-\sigma_2}=\sqrt{E_{\sigma_1}^2+E_{\sigma_2}^2-2E_{\sigma_1}E_{\sigma_2}r_{12}^2}\,.$$

Чтобы узнать, существенна-ли *разница между двумя коэффиціентами корреляціи*, нужно вычислить $E_{r_a-r_b}$ по формулѣ:

$$(165) \quad . \;.\;.\;. \quad E_{r_a-r_b}=\sqrt{E_{r_a}^2+E_{r_b}^2-2E_{r_a}E_{r_b}R_{r_ar_b}}\,,$$

которая представляетъ собою частный видъ выраженія (161). Здѣсь E_{r_a} и E_{r_b} вычисляются изъ (155), а $R_{r_ar_b}$ — коэффиціентъ корреляціи между коэффиціентами корреляціи—по формуламъ [1]:

$$(166) \quad . \;. \quad R_{r_{12}r_{13}}=r_{23}-\frac{1}{2}r_{12}r_{13}\frac{1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}r_{13}r_{23}}{(1-r_{12}^2)(1-r_{13}^2)}\,,$$

$$(167) \quad . \;.\;.\;. \quad R_{r_{12}r_{34}}=$$

$$=\frac{\left\{\begin{array}{l}(r_{13}-r_{12}r_{23})(r_{24}-r_{23}r_{34})+(r_{14}-r_{13}r_{34})(r_{23}-r_{12}r_{13})+\\+(r_{13}-r_{14}r_{34})(r_{24}-r_{12}r_{14})+(r_{14}-r_{12}r_{24})(r_{23}-r_{24}r_{34})\end{array}\right\}}{2(1-r_{12}^2)(1-r_{34}^2)}\,.$$

Первая примѣняется, какъ то видно по значкамъ при r, въ случаѣ изслѣдованія разности въ корреляціи одной величины съ двумя другими. Вторая—въ случаѣ разности между коэффиціентами корреляціи двухъ различныхъ паръ величинъ. Если нѣкоторые коэффиціенты корреляціи окажутся равными нулю, выраженіе (167) можетъ значительно упроститься. Что касается (166), то пользоваться имъ можно безъ большого труда, такъ какъ входящія въ составъ его выраженія должны быть—при изслѣдованіи корреляціи между тремя величинами—вычислены для другихъ цѣлей, какъ мы это увидимъ ниже.

Примѣръ. На стр. 11 были приведены нѣкоторыя данныя относительно цѣнъ ржи въ трехъ центрахъ: (1) Москвѣ, (2) Ельцѣ и (3) Самарѣ. Теперь мы можемъ оцѣнить разницы среднихъ ариѳметическихъ и среднихъ отклоненій. Вѣроятныхъ ошибокъ первыхъ мы вычислять не будемъ, такъ какъ, очевидно, что здѣсь мы имѣемъ дѣло съ существенными разностями, а найдемъ только вѣроятныя ошибки разностей среднихъ отклоненій. Коэф-

---

[1]) K. Pearson and L. Filon, Probable Errors of Frequency Constants, loc. cit., p. 259, 262.

фиціенты корреляціи, найденные изъ табл. I, II и III (см. прилож.) будутъ:

$$r_{12} = 0{,}792 \qquad (r^2_{12} = 0{,}6275)$$
$$r_{13} = 0{,}768 \qquad (r^2_{13} = 0{,}5901)$$
$$r_{23} = 0{,}878 \qquad (r^2_{23} = 0{,}7708)\,.$$

Примѣняя формулу (164) получимъ:

$$\sigma_3 - \sigma_1 = 1{,}20 \pm 0{,}51$$
$$\sigma_1 - \sigma_2 = 1{,}90 \pm 0{,}44$$
$$\sigma_3 - \sigma_2 = 3{,}10 \pm 0{,}38\,.$$

Первая разность больше своей вѣроятной ошибки въ 2,4 раза, вторая въ 4,3 раза и третья въ 8,2 раза. Это значитъ, что за существованіе различія (не случайнаго) въ первомъ случаѣ можно поставить 18 противъ 1, во второмъ — 534 : 1, въ третьемъ — 53.000.000 : 1 [1]). Иначе говоря, утверждая существованіе различія въ случаяхъ, подобныхъ первому, мы ошиблись бы 1 разъ изъ 19, въ случаяхъ подобныхъ второму 1 разъ изъ 535, въ случаяхъ третьяго типа 1 разъ изъ 53.000.000. Существованіе различія между средними отклоненіями въ Москвѣ и въ Самарѣ можно считать поэтому вѣроятнымъ, вторую разницу (Москва—Елецъ) можно считать почти, а третью (Елецъ—Самара) безусловно достовѣрной.

---

[1]) Изъ таблицы Енке (Леонтовичъ ч. I табл. VI) находимъ, что вѣроятность отклоненія, *не* превышающаго, какъ въ первомъ случаѣ, больше чѣмъ въ 2,4 раза своей вѣроятной ошибки равна 0,89450; отсюда вѣроятность большаго положительнаго *или* отрицательнаго отклоненія равна $1 - 0{,}89450 = 0{,}10550$, а вѣроятность одного большаго положительнаго $= 0{,}10550 : 2 = 0{,}05275$. Противоположная вѣроятность$=0{,}94725$, и искомое число шансовъ $= \frac{0{,}94725}{0{,}05275} = 18 : 1$.

Третій случай выходитъ уже за предѣлы таблицы Енке. Мы поступимъ такъ: если величина превышаетъ свою вѣроятную ошибку въ 8,2 раза, то среднее отклоненіе она превыситъ въ $0{,}67449 \,.\, 8{,}2$, т. е. въ 5,5 раза. Изъ таблицъ Шеппарда (Леонтовичъ ч. III табл. XX) находимъ, что вѣроятность того, что величина не дастъ *положительнаго* отклоненія, превышающаго свое среднее отклоненіе больше, чѣмъ въ 5,5 раза, равна 0,999.999.9810. Обратная вѣроятность равна 0,000.000.0190. Искомое число шансовъ будетъ $\frac{0{,}999.999.9810}{0{,}000.000.0190} = 53 \,.\, 10^6 : 1$.

§ 14. *Вѣроятныя ошибки въ случаѣ ненормальнаго распредѣленія.*

Если распредѣленіе не удовлетворяетъ закону Гаусса, то формулы вѣроятныхъ ошибокъ, данныя въ § 12, могутъ разсматриваться лишь какъ приближенныя. Въ зависимости отъ степени приближенія самого распредѣленія къ нормальному измѣняется и степень приближенія, которую они даютъ.

Въ частности выраженіе для вѣроятной ошибки средняго ариѳметическаго остается справедливымъ для всякаго распредѣленія, такъ что мы всегда имѣемъ:

$$E_h = 0{,}67449 \frac{\sigma}{\sqrt{N}} \cdot$$

Общее выраженіе для вѣроятной ошибки средняго отклоненія будетъ:

$$(168) \quad . \; . \; . \; . \; E_\sigma = 0{,}67449 \sqrt{\frac{\mu_4 - \mu_2^2}{4\mu_2}/N}, \quad \text{или}$$

$$(169) \quad . \; . \; . \; . \; E_\sigma = 0{,}67449 \frac{\sigma}{\sqrt{2N}} \sqrt{1 + {}^1\!/_2\eta},$$

гдѣ $\eta = \beta_2 - 3$ есть коэффиціентъ разсѣянія (см. ч. I, стр. 45). Если η незначительна, то $\sqrt{1 + {}^1\!/_2\eta} \doteq 1 + {}^1\!/_4\eta$, и легко разсчитать, что для того, чтобы ошибка въ E_σ была не больше 5%, обычную формулу (154) нужно замѣнить (169) въ томъ случаѣ, когда $\eta > +0{,}2$ или $\eta < -0{,}2$, то есть, когда $\beta_2 > 3{,}2$ или $< 2{,}8$. Однако, *если* вѣроятная ошибка мала по сравненію со среднимъ отклоненіемъ, то въ большинствѣ случаевъ можно удовлствориться и меньшей точностью. Можно замѣтить поэтому, что даже въ случаяхъ сравнительно рѣдкихъ, когда $\beta_2 = 4$ или $\beta_2 = 2$, ошибка, проистекающая изъ пользованія формулой (154), составляетъ лишь соотвѣтственно 19% и 29% истинной вѣроятной ошибки [1]).

Корреляція между средней ариѳметической и среднимъ отклоненіемъ также не будетъ равна нулю, т. к. въ общемъ случаѣ

$$(170) \quad . \; . \; . \; . \; . \; . \; . \; . \; \Sigma_h \Sigma_\sigma R_{h\sigma} = \frac{\mu_3}{2\sigma}/N.$$

---

[1]) Cp. Biometrica Vol. VI p. 117.

Знакъ этого выраженія зависитъ отъ знака третьяго момента. При $\mu_3 > 0$ увеличеніе σ связано съ увеличеніемъ h и увеличеніе h съ увеличеніемъ σ. При $\mu_3 < 0$ зависимость между ними обратная.

Въ качествѣ примѣра приложенія этой формулы найдемъ вѣроятную ошибку коэффиціента измѣнчивости

$$V = 100\frac{\sigma}{h}.$$

Логариѳмируя это выраженіе и беря дифференціалы отъ обѣихъ частей, находимъ: [1])

$$(171) \; . \; . \; . \; . \; . \; . \; . \; . \quad \frac{\delta V}{V} = \frac{\delta\sigma}{\sigma} - \frac{\delta h}{h}.$$

Возвышая въ квадратъ, суммируя для всѣхъ пробныхъ группъ и дѣля на число ихъ, получимъ:

$$\frac{(\delta V)^2}{V^2} = \frac{(\delta\sigma)^2}{\sigma^2} + \frac{(\delta h)^2}{h^2} - 2\frac{\delta\sigma\delta h}{\sigma h},$$

$$\frac{1}{V^2}\sum(\delta V)^2 = \frac{1}{\sigma^2}\sum(\delta\sigma)^2 + \frac{1}{h^2}\sum(\delta h)^2 - \frac{2}{\sigma h}\sum\delta\sigma\delta h,$$

откуда

$$(172) \; . \; . \; . \; . \quad \frac{1}{V^2}\Sigma_V^2 = \frac{1}{\sigma^2}\Sigma_\sigma^2 + \frac{1}{h^2}\Sigma_h^2 - \frac{2}{\sigma h}\Sigma_\sigma\Sigma_h R_{\sigma h}.$$

Если распредѣленіе нормальное, то $R_{\sigma h} = 0$, $\Sigma_h = \frac{\sigma}{\sqrt{N}}$, $\Sigma_\sigma = \frac{\sigma}{\sqrt{2N}}$ (см. форм. 153, 154, 160) и мы получимъ:

---

[1]) То-же можно получить и элементарнымъ путемъ. Пусть въ какой либо пробной группѣ наши величины превратятся въ $V + \delta V = 100\frac{\sigma + \delta\sigma}{h + \delta h}$. Это можно представить такъ: $V\left(1 + \frac{\delta V}{V}\right) = 100\frac{\sigma}{h}\,\frac{1 + \delta\sigma/\sigma}{1 + \delta h/h}$ Дѣля на $1 + \delta h/h$ и отбрасывая члены, въ которые входятъ очень малыя величины: $\left(\frac{\delta h}{h}\right)^2$, $\left(\frac{\delta h}{h}\right)^3$ и т. д., получимъ $V\left(1 + \frac{\delta V}{V}\right) = V\left(1 + \frac{\delta\sigma}{\sigma} - \frac{\delta h}{h}\right)$, откуда уже прямо слѣдуетъ выраженіе (171).

$$(173) \quad . \quad . \quad . \quad . \quad E_V = 0{,}67449\,\Sigma_V = 0{,}67449\,\frac{V}{\sqrt{2N}}\sqrt{1+2\left(\frac{V}{100}\right)^2},$$

выраженіе, которое дано было въ I ч. (форм. 12).

Если распредѣленіе отличается отъ нормальнаго, то $R_{\sigma h} \neq 0$, и мы должны вычислить все выраженіе (172) полностью. Пользуясь (153), (169) и (170), легко находимъ послѣ простыхъ преобразованій, что

$$(174) \; . \; . \; E_V = 0{,}67449\,\frac{V}{\sqrt{2N}}\sqrt{1+2\left(\frac{V}{100}\right)^2+\left[\frac{1}{2}\,\eta-2\,\frac{\mu_3}{h\sigma^2}\right]}.$$

Выраженіе въ квадратныхъ скобкахъ обыкновенно бываетъ мало во всѣхъ случаяхъ, когда распредѣленіе не очень далеко отъ нормальнаго. Поэтому обычной формулой (173) можно пользоваться въ широкихъ предѣлахъ.

Намъ остается еще привести выраженіе вѣроятной ошибки коэффиціента корреляціи для случаевъ ненормальнаго распредѣленія. Оно довольно сложно.

Введемъ обозначеніе:

$$(175) \quad . \quad . \quad . \quad . \quad p_{qs} = \sum\left[n_{xy}(x-\bar{x})^q(y-\bar{y})^s\right]/N,$$

съ частными случаями котораго:

$$p_{20} = \sum\left[n_{xy}(x-\bar{x})^2\right]/N = \sigma_x^2$$

$$p_{02} = \sum\left[n_{xy}(y-\bar{y})^2\right]/N = \sigma_y^2$$

$$p_{11} = \sum\left[n_{xy}(x-\bar{x})\,(y-\bar{y})\right]/N = \sigma_x\sigma_y r_{xy}\,,$$

мы встрѣчались выше.

Наиболѣе общее выраженіе для вѣроятной ошибки коэффиціента корреляціи, правильное при *всякомъ* распредѣленіи, найдено было Шеппардомъ (въ цитир. мемуарѣ, Ph. Trans. Vol. 192) и нѣсколько упрощенно Пирсономъ. Оно будетъ [1]):

---

[1]) K. Pearson, On further Methods of Determining Correlation. Drap. Comp. Research Memoirs, Biom. Ser. IV, 1907, p. 25. (Въ послѣднемъ членѣ у Пирсона опечатка, бросающаяся сразу въ глаза вслѣдствіе нарушенія симметріи (вмѣсто P_{02} стоитъ P_{20}). Ср. также On the General Theory of Skew Correlation etc. p. 20).

$$(176) \ldots E_r = 0{,}67449 \frac{r}{\sqrt{N}} \sqrt{\frac{p_{22}}{p_{11}^2} + \frac{1}{2}\frac{p_{22}}{p_{20}p_{02}} + \frac{1}{4}\frac{p_{40}}{p_{20}^2} + \frac{1}{4}\frac{p_{04}}{p_{02}^2} - \frac{p_{31}}{p_{11}p_{20}} - \frac{p_{13}}{p_{11}p_{02}}}.$$

Вычисленіе этого выраженія представляетъ собою нелегкую работу; формула (176) не принадлежитъ поэтому къ употребительнымъ.

Къ счастью для статистики опытъ приложенія ея выяснилъ, что даже въ случаяхъ, сильно отклоняющихся отъ нормальнаго типа, она даетъ результаты достаточно близкіе къ получаемымъ изъ обычной формулы — $0{,}67449 \frac{1-r^2}{\sqrt{N}}$. Поэтому въ статистической практикѣ послѣдняя можетъ считаться достаточно надежной для употребленія во всѣхъ обычныхъ случаяхъ [1]).

Примѣръ. Возьмемъ такое совершенно произвольное и такое абсолютно несхожее съ нормальнымъ распредѣленіе, какъ на прилагаемой таблицѣ. Среднія ариѳметическія значенія x и y совпадаютъ съ нулевыми

	−2	−1	0	+1	Всего
+1				5	*5*
0					*0*
−1	2	1	2		*5*
Всего	*2*	*1*	*2*	*5*	*10*

варіантами, что упрощаетъ вычисленіе. Послѣднее расположимъ въ слѣдующей таблицѣ:

---

[1]) Только при наличности очень сильной ассиметріи, сомнительности результатовъ приложенія обычной формулы и настоятельной важности строгой оцѣнки значенія полученнаго коэффиціента корреляціи — стоитъ примѣнять общую формулу (176). При вычисленіи моментовъ-произведеній p_{qs} нужно поступать также, какъ мы поступали и съ обычными моментами. Именно, разлагая $(x-\bar{x})^q$ и $(y-\bar{y})^s$ по биному Ньютона, перемножая результаты и суммируя, мы сведемъ центральные моменты p_{qs} къ нецентральнымъ, взятымъ около любыхъ осей, типа $\pi_{qs} = \Sigma\, n_{xy} x^q y^s / N$. Сначала вычисляются нецентральные моменты (π_{qs}), а затѣмъ уже находятся по выведеннымъ формуламъ центральные.

n_{xy}	x	y	$n_{xy}x$	$n_{xy}y$	$n_{xy}x^2$	$n_{xy}y^2$	$n_{xy}xy$	$n_{xy}x^2y^2$	$n_{xy}x^4$	$n_{xy}y^4$	$n_{xy}x^3y$	$n_{xy}xy^3$
2	−2	−1	−4	−2	8	2	+4	8	32	2	+16	+4
1	−1	−1	−1	−1	1	1	+1	1	1	1	+1	+1
2	0	−1	0	−2	0	2	0	0	0	2	0	0
5	+1	+1	+5	+5	5	5	+5	5	5	5	+5	+5
			0	0	14	10	10	14	38	10	22	10

Раздѣляя суммы послѣдней строки на $N=10$, получимъ нужныя постоянныя:

p_{10}	p_{01}	p_{20}	p_{02}	p_{11}	p_{22}	p_{40}	p_{04}	p_{31}	p_{13}
0	0	1,4	1	1	1,4	3,8	1	2,2	1

Откуда вычисляемъ:

$$\sigma_x=\sqrt{p_{20}}=1{,}183212,\quad \sigma_y=\sqrt{p_{02}}=1,\quad r_{xy}=\frac{p_{11}}{\sigma_x\sigma_y}=\frac{1}{1{,}183212}=0{,}84515,$$

$$r^2=\frac{1}{1{,}4}=0{,}714286,\quad 1-r_{xy}^2=0{,}285714,\quad E'_r=0{,}67449\,\frac{1-r_{xy}^2}{\sqrt{N}}=0{,}06094.$$

Теперь примѣнимъ общую формулу для вѣроятной ошибки коэффиціента корреляціи. Подставляя найденныя уже величины въ (176), будемъ имѣть:

$$E_r=0{,}67449\cdot\frac{0{,}84515}{\sqrt{10}}\cdot\sqrt{\frac{1{,}4}{1^2}+\frac{1}{2}\,\frac{1{,}4}{1{,}4\,.\,1}+\frac{1}{4}\,\frac{3{,}8}{1{,}4^2}+\frac{1}{4}\,\frac{1}{1^2}-\frac{2{,}2}{1\,.\,1{,}4}-\frac{1}{1\,.\,1}}=$$

$$=0{,}67449\cdot\frac{0{,}84515}{\sqrt{10}}\sqrt{0{,}063265}=0{,}04534.$$

Такимъ образомъ, приближенное значеніе (E'_r) только на 1/3 отличается отъ истиннаго ($E'_r\,/\,E_r=1{,}34$); это не много, если принять во вниманіе характеръ распредѣленія, разнящійся отъ нормальнаго гораздо больше, чѣмъ это имѣетъ мѣсто въ большинствѣ случаевъ, съ которыми приходится встрѣчаться на практикѣ.

§ 15 Разностный способъ нахожденія коэффиціента корреляціи.

Въ § 11 мы нашли, что $\sigma^2_{x-y} = \sigma^2_x + \sigma^2_y - 2\sigma_x\sigma_y r_{xy}$. Если этому выраженію придать нѣсколько иной видъ, то мы получимъ удобную формулу, которая во многихъ случаяхъ можетъ значительно сократить работу нахожденія коэффиціента корреляціи.

Пусть x и y будутъ значенія величинъ, входящихъ въ таблицу корреляціи и измѣренныхъ каждая отъ нѣкотораго условнаго нуля, вообще не совпадающаго съ соотвѣтствующимъ центромъ распредѣленія. Тогда

$$(177) \; . \; . \; . \; . \; . \; \sum(x-y)^2 = \sum x^2 + \sum y^2 - 2\sum xy.$$

Здѣсь $\sum x^2 = N\nu'_{2(x)}$ и $\sum y^2 = N\nu'_{2(y)}$, гдѣ $\nu'_{2(x)}$ и $\nu'_{2(y)}$ обозначаютъ, какъ всегда, грубые нецентральные моменты. Кромѣ того, какъ мы знаемъ, $\sum(x-\overline{x})(y-\overline{y}) = \sum xy - N\overline{x}\,\overline{y}$, а, слѣдоват., $\sum xy = \sum(x-\overline{x})(y-\overline{y}) + N\overline{x}\,\overline{y} = N\sigma_x\sigma_y r_{xy} + N\overline{x}\,\overline{y}$. Подставляя указанныя величины въ формулу (177), найдемъ:

$$\sum(x-y)^2 = N\nu'_{2(x)} + N\nu'_{2(y)} - 2N\overline{x}\,\overline{y} - 2N\sigma_x\sigma_y r_{xy}$$

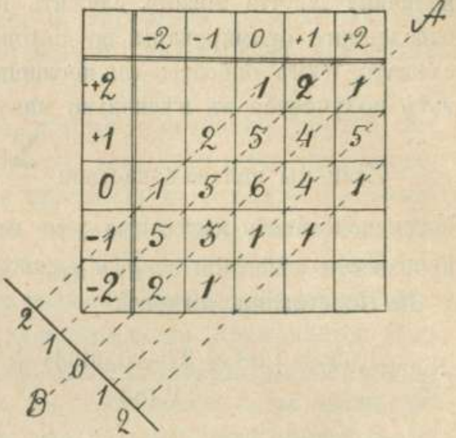
и, опредѣляя отсюда r_{xy}, получимъ:

$$(178) \; . \; . \; r_{xy} = \frac{\nu'_{2(x)} + \nu'_{2(y)} - 2\overline{x}\,\overline{y} - \frac{1}{N}\sum(x-y)^2}{2\sigma_x\sigma_y}.$$

Изъ величинъ, входящихъ въ правую часть этого равенства, $\overline{x}$, $\overline{y}$, σ_x и σ_y должны быть найдены, даже если мы не имѣемъ въ виду нахожденія коэффиціента корреляціи, какъ основныя постоянныя статистической группы; $\nu'_{2(x)}$ и $\nu'_{2(y)}$ получаются, какъ вспомогательныя величины при нахожденіи σ_x и σ_y. Слѣдовательно, нахожденіе коэффиціента корреляціи требуетъ только вычисленія $\sum(x-y)^2$. Въ большинствѣ случаевъ, оно будетъ легче, чѣмъ нахожденіе суммы произведеній $\sum xy$.

Ходъ работы легче всего показать на примѣрѣ. Возьмемъ таблицу корреляціи, съ которой мы уже имѣли дѣло выше (см. стр. 88).

Чер. 22



Проведемъ діагональ AB черезъ клѣтки, имѣющія одинаковыя варіанты: [(+2, +2); (+1, +1); (0, 0); (—1, —1); (—2, —2)]. Разность варіантъ для этихъ клѣтокъ будетъ, очевидно, равняться нулю, что мы и отмѣтимъ, поставивъ ноль на продолженіи этой діагонали. Проведемъ остальныя діагонали и отмѣтимъ ихъ цифрами 1,2; 1,2, какъ показано на чер. 22. Легко сообразить, что эти наши отмѣтки равны разностямъ между варіантами клѣтокъ, лежащихъ на соотвѣтственной діагонали; знакъ разности при этомъ безразличенъ, такъ какъ мы эти числа будемъ возвышать въ квадратъ.

Составимъ теперь слѣдующую таблицу:

$(x-y)^2$	m_i	$m_i(x-y)^2$
4	6	24
1	28	28
0	16	0
Всего	50	52

Здѣсь въ первомъ столбцѣ стоятъ квадраты чиселъ, которыми мы отмѣтили діагонали, т. е. квадраты разностей $(x-y)$. Во второмъ столбцѣ помѣщены суммы чиселъ клѣтокъ первоначальной таблицы, расположенныхъ по діагоналямъ, отмѣченнымъ одинаковыми числами. Именно, 6 равно суммѣ чиселъ, стоящихъ на діагоналяхъ (2) и (2); 28 — суммѣ чиселъ на діагоналяхъ (1) и (1); 16 — суммѣ чиселъ на нулевой діагонали. Перемножая числа перваго и второго столбца, получимъ суммы для клѣтокъ съ одинаковой разностью, а складывая найденныя такимъ образомъ числа третьяго столбца, будемъ имѣть искомую величину $\sum(x-y)^2 = 52$. (См. нижнюю клѣтку крайней правой колоны).

Замѣтимъ, что сумма чиселъ второго столбца должна дать величину N, что можетъ служить провѣркой правильности произведеннаго суммированія по діагоналямъ. Все вычисленіе легче сдѣлать, чѣмъ описать: для неслишкомъ большой таблицы результатъ получается въ нѣсколько минутъ.

Если мы найдемъ частное $\frac{\sum(x-y)^2}{N} = \frac{52}{50} = 1{,}04$, то останется подставить извѣстныя уже величины въ формулу (178) и произвести дѣйствія. $\bar{x}$, $\bar{y}$, $\nu'_{2(x)}$, $\nu'_{2(y)}$, σ_x и σ_y найдены были въ § 10. Подставляя, имѣемъ:

$$r_{xy} = \frac{1{,}64 + 1{,}08 - 2(-0{,}0064) - 1{,}04}{2\,.\,1{,}485} = \frac{0{,}8464}{1{,}485} = 0{,}5700\,,$$

т. е., въ точности то-же значеніе, которое было найдено раньше другимъ способомъ.

Такъ какъ выраженіе (178) получилось, какъ выводъ изъ ряда тождествъ, то величина, найденная при его помощи, также должна быть всегда тождественна съ величиной, полученной по методу произведеній. Формула вѣроятной ошибки остается поэтому также прежней [1]).

§ 16. *Криволинейная регрессія.*

Прямая регрессіи можетъ служить вполнѣ пригодной теоретической моделью явленія до тѣхъ лишь поръ, пока уклоненія эмпирической линіи регрессіи настолько незначительны, что ихъ можно посчитать случайными. Но хотя громадное множество яв-

---

[1]) Изложенный методъ представляетъ изъ себя нѣкоторое видоизмѣненіе метода разностей, употреблявшагося Пирсономъ для замѣны метода произведеній. Въ прежней формѣ этотъ методъ не былъ свободенъ отъ недостатковъ, такъ какъ давалъ вообще нѣсколько иныя значенія для r_{xy}, чѣмъ основной и наиболѣе надежный изъ всѣхъ возможныхъ методъ произведеній. (См. A. Wright, A. Lee, and K. Pearson, Biometrica Vol. V, p. 410 и A. Harris, Biometrica Vol. VII p. 214—218). Въ изложенной модификаціи этотъ недостатокъ устраненъ.

леній вполнѣ удовлетворительно могутъ изображаться линейными формулами, тѣмъ не менѣе не рѣдки случаи и криволинейной регрессіи. Тогда, если изслѣдователь не захочетъ ограничиться эмпирической линіей регрессіи, т. е., простымъ констатированіемъ фактическаго состоянія его матеріала, а пожелаетъ выдѣлить основныя, *не случайныя* черты характера изслѣдуемой зависимости, ему придется приступить къ дальнѣйшей обработкѣ своихъ данныхъ.

Первый методъ, наиболѣе старый и наиболѣе грубый, заключается въ томъ, чтобы провести отъ руки плавную кривую, которая бы какъ можно тѣснѣе прилегала къ отдѣльнымъ эмпирическимъ точкамъ. Кривая, мы сказали, должна быть плавной, это значитъ, что кривизна ея должна мѣняться съ возможной постепенностью и число перегибовъ должно быть какъ можно меньше. Методъ этотъ при всей своей элементарности примѣняется и до сихъ поръ [1]) и можетъ въ цѣломъ рядѣ случаевъ дать достаточно удовлетворительные результаты, въ особенности, когда число случаевъ въ совокупности незначительно, и опредѣленіе кривой болѣе совершенными методами все равно точныхъ результатовъ дать не могло бы. Кромѣ того часто бо́льшая точность и не требуется, такъ что примѣненіе болѣе совершенныхъ методовъ, сопряженныхъ съ затратой большой вычислительной работы было бы попросту потерей времени.

Второй методъ состоитъ въ томъ, что теоретическія кривыя подбираются къ отдѣльнымъ отрѣзкамъ эмпирической линіи регрессіи. Выгода такого разбиванія задачи на части заключается въ бо́льшей простотѣ кривыхъ, которыми можно при этомъ пользоваться, и особенно ощутительна въ случаѣ регрессіи сложнаго характера, къ которой плохо подходятъ параболическія кривыя 2-ой и 3-ей степени.

Идти дальше параболы третьей, въ крайнемъ случаѣ, четвертой степени врядъ-ли можно совѣтовать, ибо вычисленіе коэффиціентовъ ихъ сопряжено съ нахожденіемъ высшихъ моментовъ, имѣющихъ большія вѣроятныя ошибки, да и требуетъ кромѣ того слишкомъ большой затраты времени на вычисленія,—затраты, которая можетъ все-равно не привести къ удачному результату.

---

[1]) См., напр., K. Pearson, On the Change in Expectation of Life in Man during a period of circa 2000 years, Biom. Vol. I, p. 261—264.

Лучше, по здравомъ обсужденіи дѣла, разбить таблицу корреляціи на части и воспользоваться болѣе простыми кривыми для каждой изъ нихъ въ отдѣльности [1]).

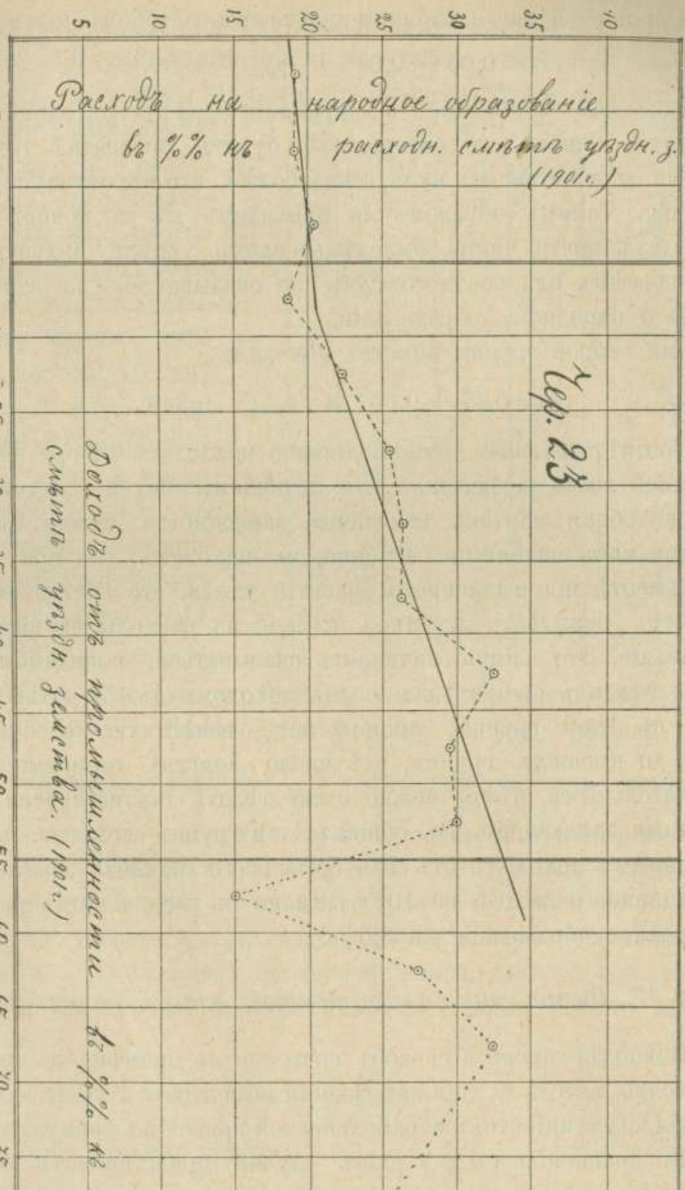
Примѣръ. Разсмотримъ зависимость между относительной высотой расхода уѣзднаго земства на народное образованіе (въ $^0/_0{^0}/_0$ къ расходной смѣтѣ) и относительной высотой дохода, получаемаго земствомъ отъ промышленности и торговли. Наши данныя (см. табл. VIII прилож.) относятся къ 1901 г. и охватываютъ всѣ 359 уѣздныхъ земствъ. Подъ понятіе торгово-промышленнаго дохода подведены рубрики: (*a*) доходъ съ документовъ на право торговли и промысловъ и (*b*) доходъ, получаемый отъ обложенія заводскихъ, фабричныхъ и торгово-промышленныхъ помѣщеній.

Бѣглое даже разсмотреніе эмпирической линіи регрессіи (чер. 23) показываетъ, что прямая линія не является въ этомъ случаѣ наиболѣе подходящимъ образомъ. Въ самомъ дѣлѣ, въ то время какъ общая тенденція явно клонится къ росту расходовъ на народное образованіе по мѣрѣ роста торгово-промышленнаго дохода, въ низшихъ группахъ ничего подобнаго не наблюдается. Такъ какъ низшія группы имѣютъ значительную численность, то случайнымъ это обстоятельство признать врядъ-ли возможно. Скорѣе мы имѣемъ здѣсь дѣло какъ разъ съ одной изъ характерныхъ чертъ явленія.

Чтобы избѣжать сложныхъ вычисленій, связанныхъ съ нахожденіемъ кривой регрессіи, мы раздѣлимъ нашу таблицу на три части. Къ первой отнесемъ земства съ торгово-промышленнымъ доходомъ отъ 0 до 15%, ко второй земства, для которыхъ эта часть дохода составляетъ отъ 15 до 55%, и въ третью группу выдѣлимъ остальныя земства. Въ этой послѣдней группѣ всего 4 земства; каждое изъ нихъ стоитъ изолировано отъ другихъ, такъ что о среднемъ ариѳметическомъ и о „строяхъ“ можно говорить по отношенію къ нимъ лишь cum grano salis. Мы предпочли поэтому совсѣмъ откинуть ихъ и ограничиться двумя первыми группами.

---

[1]) См., напр., A. O. Powys, Data for the Problem of Evolution in Man, Biom. Vol. I, p. 49. (Это, собственно говоря, коллективная работа; указанное мѣсто принадлежитъ Пирсону). См. также статью того-же автора подъ тѣмъ-же заглавіемъ: Biom. Vol. IV, p. 233—285.


Расходъ на народное образованіе въ %% къ расходн. смѣтъ уѣздн. з. (1901 г.)
Черт. 23
Доходъ отъ промышленности въ %% къ доходн. смѣтъ уѣздн. земства. (1901 г.)

Для первой части таблицы обычнымъ способомъ получилось:

$$r = 0{,}05 \pm 0{,}04 \qquad \rho_{y(x)} = 0{,}09.$$

Такъ какъ коэффиціентъ корреляціи едва превосходитъ свою вѣроятную ошибку, то осторожнѣе будетъ признать, что онъ врядъ-ли отличается отъ нуля, тѣмъ болѣе, что вѣроятная ошибка не велика. Такимъ образомъ мы приходимъ къ заключенію, что въ уѣздахъ почти чисто земледѣльческихъ успѣхи индустріализаціи совсѣмъ или почти совсѣмъ не сказываются на заботахъ земства о народномъ образованіи.

Для второй группы земствъ находимъ:

$$r = 0{,}47 \pm 0{,}06 \quad \text{и} \quad \rho_{y(x)} = 0{,}349\,.$$

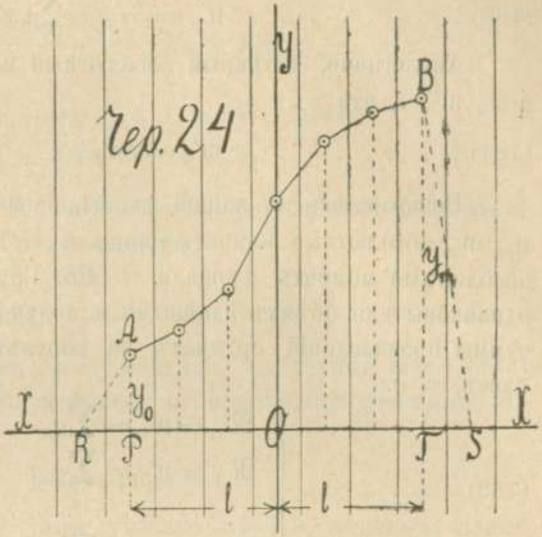
Соотвѣтствующая прямая хорошо подходитъ къ средней части нашей линіи регрессіи, т. что въ общемъ получается довольно правдоподобная картина измѣненія зависимости между интересующими насъ явленіями. Именно, мы находимъ, что при извѣстной высотѣ промышленнаго развитія уѣзда, это обстоятельство начинаетъ оказывать замѣтное вліяніе на расходы на народное образованіе. Это вліяніе начинаетъ сказываться, повидимому, съ группы уѣздовъ съ промышленнымъ доходомъ въ 15—20% и, начиная съ этой группы, пріобрѣтаетъ извѣстную устойчивость вплоть до высшихъ группъ, гдѣ число земствъ становится уже незначительнымъ, чтобы можно было дѣлать статистически обоснованныя заключенія. Въ общемъ для группъ земствъ съ промышленнымъ доходомъ отъ 15—55% всего бюджета возрастаніе промышленнаго дохода на 10% связано съ увеличеніемъ расхода на народное образованіе на 3,5%.

§ 17. Вычисленіе коэффиціентовъ кривой регрессіи.

Наиболѣе простой способъ состоитъ въ примѣненіи правилъ § 7, именно метода *C*, для нахожденія моментовъ и затѣмъ въ вычисленіи коэффиціентовъ параболической кривой по формуламъ § 8.

По сравненію съ § 7 нашъ случай представляетъ однако нѣкоторыя особенности, на которыхъ мы и должны остановиться.

Пусть AB (чер. 24) будетъ эмпирическая линія регрессіи, y_0 и y_m — первая и послѣдняя ординаты. Задача заключается въ томъ, чтобы найти моменты площади $APTB$ около нѣкоторой оси. Чтобы примѣнить формулы § 8, за начало координатъ примемъ середину разстоянія между крайними ординатами (O) и положимъ $OT = -OP = l$. Разстояніе между двумя сосѣдними ординатами примемъ за единицу. Особенность нашего случая по сравненію со случаемъ, разсмотрѣннымъ въ § 7, состоитъ въ томъ, что крайнія ординаты здѣсь не равны нулю. Чтобы свести нашу задачу къ предыдущему случаю, возьмемъ на оси XX двѣ точки R и S такъ, чтобы $RP = TS = 1$ и найдемъ моменты площади $RABS$; затѣмъ вычтемъ моменты площадей двухъ треугольниковъ RAP и SBT [1]).



Наши формулы упростятся, если мы введемъ нѣкоторыя новыя обозначенія. Именно, мы называли раньше грубыми моментами выраженія типа $\nu'_p = \frac{1}{N}\sum n_x x^p$, гдѣ n_x есть численность подгруппы, а x — разстояніе *середины* соотвѣтствующаго интервала отъ начала координатъ. Теперь (какъ и въ § 8) мы имѣемъ дѣло не съ численностями, а съ площадями и отыскиваемъ моменты площадей. Вмѣсто n_x въ нашу формулу войдетъ, слѣдовательно, ордината y, умноженная на величину интервала, то есть, какъ мы приняли, на единицу, и выраженіе для грубаго момента площади будетъ:

$$(179) \quad . \; . \; . \; . \; . \; . \; . \; . \quad \nu'_p = \frac{1}{S}\sum y x^p,$$

гдѣ S есть площадь фигуры.

Эту величину можно называть *относительнымъ моментомъ* по сравненію съ *абсолютнымъ моментомъ*, который равняется

[1]) См. K. Pearson, On the systematic Fitting of Curves, Biom., Vol. II, p. 7—9.

той-же суммѣ $\left(\sum yx^p\right)$, не раздѣленной на S. Обозначая абсолютные грубые моменты буквой $\mathfrak{n}'$ [1]) будемъ имѣть:

$$(180) \quad \mathfrak{n}'_p = S\nu'_p = \sum yx^p.$$

Аналогично, истинный абсолютный моментъ обозначимъ черезъ $\mathfrak{m}'$, т. что

$$(181) \quad \mathfrak{m}'_p = S\mu'_p,$$

Возвращаясь къ нашей задачѣ, обозначимъ малыми буквами $\mathfrak{n}_p$, $\mathfrak{m}_p$ абсолютные моменты площади $APTB$, большими—$\mathfrak{N}_p$, $\mathfrak{M}_p$ абсолютные моменты площади $RABS$. Грубые моменты, очевидно, одинаковы для обѣихъ площадей и получатся, если мы составимъ суммы произведеній ординатъ на соотвѣтствующія степени абсциссъ, т. е.:

$$(182) \quad \begin{cases} \mathfrak{N}'_0 = \mathfrak{n}'_0 = \sum y_x \\ \mathfrak{N}'_1 = \mathfrak{n}'_1 = \sum y_x x \\ \dots\dots\dots \\ \mathfrak{N}'_p = \mathfrak{n}'_p = \sum y_x x^p. \end{cases}$$

Затѣмъ найдемъ истинные моменты площади $RABS$ по формулѣ (45) § 7, которая, какъ легко сообразить, годится и для вычисленія абсолютныхъ моментовъ. Мы получимъ:

$$(183) \quad \begin{cases} \mathfrak{M}'_1 = \mathfrak{n}'_1 & \mathfrak{M}'_4 = \mathfrak{n}'_4 + \mathfrak{n}'_2 + \frac{1}{15} \\ \mathfrak{M}'_2 = \mathfrak{n}'_2 + \frac{1}{6} & \mathfrak{M}'_5 = \mathfrak{n}'_5 + \frac{5}{3}\mathfrak{n}'_3 + \frac{1}{3}\mathfrak{n}'_1 \\ \mathfrak{M}'_3 = \mathfrak{n}'_3 + \frac{1}{2}\mathfrak{n}'_1 & \mathfrak{M}'_6 = \mathfrak{n}'_6 + \frac{5}{2}\mathfrak{n}'_4 + \mathfrak{n}'_2 + \frac{1}{28} \end{cases}$$

и т. д.

Чтобы получить теперь истинные абсолютные моменты для истинной площади нашей линіи регрессіи, т. е. для площади $APTB$, изъ найденныхъ выраженій нужно еще вычесть соотвѣтствующіе моменты площадей двухъ лишнихъ треугольниковъ. По вычисленію Пирсона [2]) будемъ имѣть:

---

[1]) Черточка показываетъ, что это *нецентральные* моменты.

[2]) l. cit., p. 8.

Для нечетныхъ моментовъ (1-го, 3-го, 5-го):

(184 a) . . . . . $\mathfrak{m}'_n = \mathfrak{M}'_n - L_n(y_m - y_0)$

а для четныхъ (2-го, 4-го, 6-го и т. д.):

(184 b) . . . . . $\mathfrak{m}'_n = \mathfrak{M}'_n - L_n(y_m + y_0)$,

гдѣ

(185) . . . . $L_n = \dfrac{(l+1)^{n+2} - (n+2)l^{n+1} - l^{n+2}}{(n+1)(n+2)}$.

Площадь $APTB$ получаемъ, суммируя площади составляющихъ ее трапецій. Она будетъ равна:

(186) . . . . . $S = \sum y - \frac{1}{2}(y_m + y_0)$.

Если мы раздѣлимъ на эту величину найденные выше (184) абсолютные моменты, то получимъ относительные истинные моменты:

(187) . . . . $\mu'_n = \mathfrak{m}'_n / S$.

Затѣмъ находимъ, какъ указано въ § 8, вспомогательные величины:

$$y_0 = \frac{S}{2l}$$

$$\lambda_n = \frac{\mu'_n}{l^n}$$

и по формуламъ § 8 легко вычисляемъ коэффиціенты параболической кривой регрессіи.

Для упрощенія вычисленій привожу таблицу (см. слѣд. стр.) значеній L_n для величинъ n отъ 1 до 5 и l отъ 3 до 20 [1]).

§ 18. Вычисленіе коэффиціентовъ кривой регрессіи (продолженіе).

Способъ, изложенный въ предыдущемъ §-ѣ, по своей сравнительной простотѣ можетъ найти себѣ примѣненіе во многихъ случаяхъ [2]). Однако онъ обладаетъ въ приложеніи къ нашей задачѣ

---

[1]) l. cit. p. 9.

[2]) Онъ употребленъ, между прочимъ, въ неоднократно цитированной работѣ Powys'a. См. Biom. Vol. IV, p. 236.

Таблица значеній дополнительнаго члена для поправки моментовъ трапецоидальной площади.

$$L_n = \frac{(l+1)^{n+2} - (n+2)l^{n+1} - l^{n+2}}{(n+1)\ (n+2)}.$$

l	L_1	L_2	L_3	L_4	L_5
3	1,667	5,583	18,800	63,633	216,524
4	2,167	9,417	41,050	179,500	787,357
5	2,667	14,250	76,300	409,367	2200,857
6	3,167	20,083	127,550	811,233	5167,024
7	3,667	26,917	197,800	1455,100	10715,857
8	4,167	34,750	290,050	2422,967	20257,357
9	4,667	43,583	407,300	3808,833	35641,524
10	5,167	53,417	552,550	5718,700	59218,357
11	5,667	64,250	728,800	8270,567	93897,857
12	6,167	76,083	939,050	11594,433	143210,024
13	6,667	88,917	1186,300	15832,300	211364,857
14	7,167	102,750	1473,550	21138,167	303312,357
15	7,667	117,583	1803,800	27678,033	424802,524
16	8,167	133,417	2180,050	35629,900	582445,357
17	8,667	150,250	2605,300	45183,767	783770,857
18	9,167	168,083	3082,550	56541,633	1037289,024
19	9,667	186,917	3614,800	69917,500	1352549,857
20	10,167	206,750	4205,050	85537,367	1740203,357

важнымъ недостаткомъ, заключающимся въ томъ, что всѣ ординаты оказываютъ одинаковое вліяніе на результатъ, хотя величины однѣхъ болѣе, другихъ менѣе достовѣрны. Вотъ почему при выводѣ ур-ія прямой регрессіи этимъ методомъ не пользуются, а примѣняютъ другія формулы, основанныя на принципахъ метода наименьшихъ квадратовъ [1]).

Примѣненіе этого метода (точнѣе метода моментовъ) и къ криволинейной регрессіи разработано Пирсономъ въ его мемуарѣ On the general Theory of Skew Correlation and Non-Linear Regression, но вслѣдствіе своей сложности методъ этотъ не можетъ быть здѣсь изложенъ.

Но мы можемъ подойти къ дѣлу проще. Именно, если мы оставимъ методъ моментовъ, и примѣнимъ традиціонный методъ наименьшихъ квадратовъ, то мы получимъ вполнѣ раціональное и притомъ сравнительно простое рѣшеніе. Вопросъ только заключается въ томъ, какіе вѣса придать отдѣльнымъ ординатамъ эмпирической линіи регрессіи. Если, какъ раньше, численность x_i-ого строя будетъ n_{x_i}, средняя ариѳметическая величина y въ этомъ строѣ (ордината линіи регрессіи) — y_{x_i}, среднее отклоненіе для того-же строя — $\sigma_{n_{x_i}}$, то вѣроятная ошибка y_{x_i} будетъ:

$$0{,}67449\ \sigma_{n_{x_i}} \Big/ \sqrt{n_{x_i}}\,.$$

Точность, съ которой намъ извѣстна величина отдѣльныхъ ординатъ, тѣмъ больше, слѣдовательно, чѣмъ меньше $\sigma_{n_{x_i}}$, и чѣмъ больше $\sqrt{n_{x_i}}$. Мы, поэтому, и будемъ считать вѣса отдѣльныхъ ординатъ пропорціональными $\sqrt{n_{x_i}}$ и обратно пропорціональными среднимъ отклоненіямъ строевъ. Вѣса *квадратовъ* разностей будемъ считать пропорціональными, слѣдовательно, квадратамъ предыдущихъ величинъ.

Обозначая вѣсъ буквой p_i будемъ имѣть:

$$p_i = \frac{n_{x_i}}{\sigma^2_{n_{x_i}}}\,. \qquad (188a)$$

Если среднія отклоненія въ различныхъ строяхъ равны, или почти равны другъ другу, то вѣса будутъ пропорціональны простымъ численностямъ [2]), т. е., можно будетъ положить:

---

[1]) K. Pearson, On the General Theory of Skew Correlation etc., p. 52—53.

[2]) Какъ и принимаетъ Пирсонъ въ цитир. работѣ.

(188 b) . . . . $p_i = n_{x_i}$.

Пусть ур-іе кривой регрессіи будетъ

(189) . . . . $Y = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m$.

Тогда по методу наименьшихъ квадратовъ мы должны найти коэффиціенты изъ условія, чтобы взвѣшенная сумма квадратовъ разностей была наименьшей, т. е., чтобы

(190) . . $\sum p_i (y_{x_i} - a_0 - a_1 x_i - a_2 x_i^2 - \ldots - a_m x_i^m)^2 = \text{minimum}$.

Дифференцируя это выраженіе по a_r, получимъ:

$$\sum p_i x_i^r (y_{x_i} - a_0 - a_1 x_i - a_2 x_i^2 - \ldots - a_m x_i^m) = 0$$

или

(191) . . . $a_0 . \sum p_i x_i^r + a_1 \sum p_i x_i^{r+1} + \ldots +$

$$+ a_m \sum p_i x_i^{r+m} = \sum p_i y_{x_i} x_i^r.$$

Мы положимъ для сокращенія:

(192) . . . . . . $\begin{cases} \sum p_i x_i^r = \mathfrak{m}_r \\ \sum p_i y_x x_i^r = \mathfrak{n}_r. \end{cases}$

Давая r въ (191) всѣ значенія отъ 0 до m, получимъ $m+1$ ур-іе:

(193) . . $\begin{cases} \mathfrak{m}_0 a_0 + \mathfrak{m}_1 a_1 + \mathfrak{m}_2 a_2 + \ldots + \mathfrak{m}_m a_m = \mathfrak{n}_0 \\ \mathfrak{m}_1 a_0 + \mathfrak{m}_2 a_1 + \mathfrak{m}_3 a_2 + \ldots + \mathfrak{m}_{m+1} a_m = \mathfrak{n}_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \mathfrak{m}_m a_0 + \mathfrak{m}_{m+1} a_1 + \mathfrak{m}_{m+2} a_2 + \ldots + \mathfrak{m}_{2m} a_m = \mathfrak{n}_m. \end{cases}$

Рѣшая эти ур-ія, находимъ $a_0, a_1, a_2 \ldots a_m$ — коэффиціенты искомой параболы [1]).

[1]) Величины $\mathfrak{m}_0, \mathfrak{m}_1, \mathfrak{m}_2 \ldots\ldots$ — очень большія числа. Рѣшенія ур-ій (193) поэтому всего, кажется, удобнѣе расположить слѣдующимъ образомъ: сначала ограничиваемся не слишкомъ большимъ числомъ знаковъ и находимъ приближенныя величины $a'_0, a'_1 \ldots a'_m$. Полагая $a_0 = a'_0 + \alpha'_0$, $a_1 = a'_0 + \alpha'_1$ и т. д., подставимъ эти величины въ ур-ія (193). Тогда будемъ имѣть:

$$\mathfrak{m}_0 \alpha'_0 + \mathfrak{m}_1 \alpha'_1 + \ldots + \mathfrak{m}_m \alpha'_m = \mathfrak{n}_0 - (\mathfrak{m}_0 a'_0 + \mathfrak{m}_1 a'_1 + \ldots + \mathfrak{m}_m a'_m) = \mathfrak{n}'_0$$

и аналогично для остальныхъ ур-ій. По величинамъ $\mathfrak{n}'_0, \mathfrak{n}'_1 \ldots \mathfrak{n}'_m$ судимъ, насколько приблизительно велики ошибки, и, удерживая опять въ коэффиціентахъ $\mathfrak{m}_0, \mathfrak{m}_1, \mathfrak{m}_2 \ldots \mathfrak{m}_m$ лишь не много знаковъ, рѣшаемъ

Этотъ методъ долженъ дать результаты во всякомъ случаѣ не худшія, чѣмъ методъ Пирсона, въ случаѣ-же неравноизмѣнчиваго распредѣленія даже лучшія.

§ 19. *Корреляціонное отношеніе.*

Случай криволинейной регрессіи требуетъ и своей особой мѣры для степени корреляціонной зависимости. Какъ мы показали въ § 8 II ч., коэффиціентъ корреляціи можетъ равняться единицѣ только въ томъ случаѣ, когда корреляція совершенная и регрессія строго линейная.

Если первое условіе выполнено, но регрессія нелинейная, коэффиціентъ корреляціи будетъ все-же меньше единицы.

Затѣмъ, равенство коэффиціента корреляціи нулю можетъ свидѣтельствовать объ отсутствіи корреляціи тоже только лишь при условіи линейности. Въ самомъ дѣлѣ, т. к. $r = \sqrt{\rho_x \rho_y}$, то $r = 0$, если одинъ изъ коэффиціентовъ регрессіи равенъ нулю, т. е. когда одна изъ прямыхъ регрессіи горизонтальна. Но это возможно и при наличности корреляціи и даже при наличности совершенной корреляціи, тождественной со строгой функціональной зависимостью. Пусть, напр., зависимость между двумя величинами выражается параболой или другой какой-либо симметричной кривой съ вертикальной осью и двумя вѣтвями, одной восходящей, другой нисходящей. Прямая, всего тѣснѣе примыкающая къ точкамъ такой кривой, будетъ горизонтальна, а слѣд., коэффиціентъ корреляціи будетъ равняться нулю.

Слѣдующія соображенія даютъ возможность установить мѣру корреляціи и для случая криволинейной регрессіи [1]).

систему ур-ій для поправокъ $\alpha'_0, \alpha'_1 \ldots \alpha'_m$. Если нужно, находимъ такимъ-же способомъ третьи поправки и т. д. Обыкновенно, второе приближеніе удовлетворитъ требованіямъ нужной точности. Самое рѣшеніе, кажется, проще всего вести слѣдующимъ образомъ (предполагая многозначность коэффиціентовъ и возможность пользоваться ариѳмометромъ): Коэффиціенты перваго ур-ія дѣлимъ на $\mathfrak{m}_0$, второго на $\mathfrak{m}_1$, и т. д. Затѣмъ, вычитая попарно, исключаемъ a_0 и получаемъ систему ур-ій съ числомъ неизвѣстныхъ на единицу меньшимъ. Съ этой системой поступаемъ по прежнему и, такимъ образомъ, доходимъ до 1 ур-ія съ однимъ неизвѣстнымъ. Остальныя неизвѣстныя найдутся подстановкой.

[1]) См. K. Pearson, On the general Theory of Skew Correlation etc., p. 9—11.

Если между двумя признаками не существуетъ корреляціонной зависимости, то группы, образованные по одному изъ нихъ, должны показывать одинаковое распредѣленіе второго признака, такое-же какъ и въ общей совокупности. Пренебрегая случайными отклоненіями, т. е. допуская, что совокупность наша очень велика, а въ соотвѣтствіи съ этимъ всѣ вѣроятныя ошибки настолько малы, что ими можно пренебречь, мы будемъ имѣть наше первое основное положеніе:

Положеніе I. *Въ случаѣ отсутствія корреляціи средняя ариѳметическая величина признака въ каждомъ строѣ равна средней ариѳметической для всей совокупности, т. е.*

$$(194) \quad \ldots\ldots\ldots\ldots \quad y_x = \overline{y},$$

и среднее отклоненіе, вычисленное для каждаго строя въ отдѣльности, равно среднему отклоненію для всей совокупности, т. е.,

$$(195) \quad \ldots\ldots\ldots\ldots \quad \sigma_{n_x} = \sigma_y .$$

Разсмотримъ тотъ случай, когда корреляція становится совершенной, т. е. превращается въ строгую функціональную зависимость.

Тогда опредѣленному значенію перемѣнной x соотвѣтствуетъ одно опредѣленное значеніе второй перемѣнной (y). Отклоненія величины y отъ этого единственнаго значенія равны нулю (мы интервалы предполагаемъ безконечно малыми), а слѣд., и среднее отклоненіе y въ каждомъ отдѣльномъ x-овомъ строѣ также равно нулю. Итакъ, мы формулируемъ второе основное положеніе:

Положеніе II. *Въ случаѣ совершенной корреляціи, т. е., въ случаѣ перехода корреляціонной зависимости въ строгую функціональную всѣ среднія отклоненія отдѣльныхъ строевъ равны нулю, т. е.*

$$(196) \quad \ldots\ldots\ldots\ldots \quad \sigma_{n_x} = 0 .$$

Если мы назовемъ σ_{n_x} частнымъ среднимъ отклоненіемъ, а σ_y общимъ среднимъ отклоненіемъ, то величина

$$(197) \quad \ldots\ldots\ldots \quad \sigma_a = \sqrt{\frac{\sum n_x \sigma_{n_x}^2}{N}}$$

можетъ быть названа *среднимъ квадратичнымъ среднихъ частныхъ отклоненій.* Она получится, если мы квадратъ каждаго средняго частнаго отклоненія помножимъ на численность соотвѣтствующаго строя, полученную сумму раздѣлимъ на численность всей совокупности и изъ частнаго извлечемъ квадратный корень.

Разсмотримъ отношеніе

$$\frac{\sigma_a^2}{\sigma_y^2}.$$

При отсутствіи корреляціи эта величина обращается въ единицу, т. к. въ этомъ случаѣ, каждое $\sigma_{n_x} = \sigma_y$, а слѣдовательно,

$$\sigma_a^2 = \frac{1}{N}\sum n_x \sigma_{n_x}^2 = \sigma_{n_x}^2 \frac{\sum n_x}{N} = \sigma_y^2.$$

Когда корреляція совершенная, то каждое $\sigma_{y_x} = 0$, а слѣдовательно и $\sigma_a = 0$, а слѣдовательно и $\sigma_a^2 / \sigma_y^2 = 0$.

Но за мѣру корреляціи удобнѣе принять такую величину, которая съ увеличеніемъ степени корреляціи также увеличивается, а съ уменьшеніемъ ея уменьшается. На этомъ основаніи Пирсонъ предложилъ принять за мѣру корреляціи не $\frac{\sigma_a^2}{\sigma_y^2}$, а величину η, опредѣляемую ур-іемъ:

$$(198) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad \eta^2 = 1 - \frac{\sigma_a^2}{\sigma_y^2},$$

и назвалъ ее корреляціоннымъ отношеніемъ.

На основаніи сказаннаго выше относительно величины $\frac{\sigma_a^2}{\sigma_y^2}$ мы можемъ считать доказаннымъ:

Положеніе III. *Въ случаѣ отсутствія корреляціи корреляціонное отношеніе $\eta = 0$, въ случаѣ совершенной корреляціи оно равно $\eta = 1$.*

Корреляціонное отношеніе можетъ быть представлено еще и въ другой притомъ очень удобной формѣ. Но для совершенія этого преобразованія мы должны доказать одно вспомогательное предложеніе.

Если мы назовемъ *разность между среднимъ ариѳметическимъ для отдѣльнаго строя и среднимъ ариѳметическимъ для всей совокупности* $(y_x - \bar{y})$ *отклоненіемъ линіи регрессіи отъ центральной оси распредѣленія*, квадратъ каждаго такого отклоненія помножимъ на численность соотвѣтствующаго строя и затѣмъ сумму такихъ произведеній раздѣлимъ на численность всей совокупности, то величина

$$(199) \quad \sigma_м = \sqrt{\frac{n_x(y_x - \bar{y})^2}{N}}$$

можетъ быть названа *среднимъ квадратичнымъ отклоненіемъ линіи регрессіи отъ центральной оси*. Докажемъ теперь:

Положеніе IV:

$$(200) \quad \sigma_y^2 = \sigma_a^2 + \sigma_м^2 .$$

Мы знаемъ, что $\sigma_y^2 = \frac{1}{N} \sum n_x(y - \bar{y})^2$. Т. к. $n_x(y - \bar{y})^2 = = (y - \bar{y})^2 + (y - \bar{y})^2 + \ldots . (y - \bar{y})^2$ всего n_x разъ, то то-же можно, очевидно, написать и иначе, именно

$$\sigma_y^2 = \frac{1}{N} \sum (y - \bar{y})^2.$$

Будемъ обозначать знакомъ $\sum_i$ сумму, относящуюся къ величинамъ, принадлежащимъ къ i-ому x-овому строю, тогда, очевидно,

$$(201) \quad \sigma_y^2 = \frac{1}{N}\left[\sum_1 (y - \bar{y})^2 + \sum_2 (y - \bar{y})^2 + \ldots + \sum_p (y - \bar{y})^2\right].$$

Каждая изъ этихъ суммъ, напр. i-ая, можетъ быть представлена такъ:

$$(202) \quad \sum_i (y - \bar{y})^2 = \sum_i [(y - y_x) + (y_x - \bar{y})]^2 =$$
$$= \sum_i (y - y_x)^2 + \sum_i (y_x - \bar{y})^2 + 2 \sum_i (y - y_x)(y_x - \bar{y}) .$$

y_x есть средняя ариѳметическая разсматриваемаго строя, а $(y - y_x)^2$ есть квадратъ уклоненія отдѣльной величины отъ средней для этого строя, слѣдовательно:

$$\sum_i (y - y_x)^2 = n_x \sigma_{n_x}^2 .$$

$y_x - \bar{y}$ для даннаго строя есть величина постоянная, повторяется въ суммѣ $\sum_i (y_x - \bar{y})^2$ столько разъ, сколько индивидуумовъ въ этомъ строѣ, слѣдовательно,

$$\sum_i (y_x - \bar{y})^2 = n_x (y_x - \bar{y})^2 .$$

Послѣдняя сумма равна нулю, ибо

$$\sum_i (y - y_x)(y_x - \bar{y}) = (y_x - \bar{y}) \sum_i (y - y_x) = 0 .$$

Так. обр., выраженіе (202) принимаетъ слѣдующую форму:

(203) . . . . $\sum_i (y-\bar{y})^2 = n_x\sigma_{n_x}^2 + n_x(y_x-\bar{y})^2$.

Подставляя это выраженіе въ (201), т. е., находя сумму подобныхъ выраженій для всѣхъ строевъ, мы въ силу опредѣленій (197) и (199) получимъ:

$$\sigma_y^2 = \frac{1}{N}\left[\sum n_x\sigma_{n_x}^2 + \sum n_x(y_x-\bar{y})^2\right] = \frac{1}{N}\left[N\sigma_a^2 + N\sigma_м^2\right],$$

откуда имѣемъ:

(204) . . . . $\sigma_y^2 = \sigma_a^2 + \sigma_м^2$.

Раздѣливъ обѣ части на σ_y^2 найдемъ:

$$\frac{\sigma_м^2}{\sigma_y^2} = 1 - \frac{\sigma_a^2}{\sigma_y^2} = \eta^2,$$

откуда

(205) . . . . $\eta = \frac{\sigma_м}{\sigma_y}$.

Это выраженіе удобнѣе для вычисленія корреляціоннаго отношенія, оно-же даетъ намъ возможность доказать теорему, обратную положенію III-му.

Положеніе V. *Если $\eta = 0$, корреляціи нѣтъ. Если $\eta = 1$, корреляція—совершенная.*

Въ самомъ дѣлѣ, т. к. $\eta^2 = \frac{\frac{1}{N}\sum n_x(y_x-\bar{y})^2}{\sigma_y^2}$, и т. к. числитель этого выраженія представляетъ собою сумму однихъ только положительныхъ чиселъ, то η можетъ равняться нулю единственно въ томъ случаѣ, когда каждое $y_x = \bar{y}$, т. е. когда корреляціи не существуетъ.

Затѣмъ, т. к. $\eta^2 = 1 - \frac{\sigma_a^2}{\sigma_y^2}$, то, когда $\eta = 1$, σ_a^2 должно равняться нулю. Но $\sigma_a^2 = \frac{1}{N}\sum n_x\sigma_{n_x}^2$ и можетъ равняться нулю, только когда всѣ $\sigma_{n_x} = 0$, т. е. когда корреляція сдѣлается совершенной.

Изъ равенства (204) слѣдуетъ, что σ_y^2 равняется суммѣ двухъ существенно положительныхъ величинъ, слѣдовательно,

$\sigma_м = \sigma_y$ только тогда, когда $\sigma_a = 0$, т. е. въ случаѣ совершенной корреляціи, въ остальныхъ случаяхъ $\sigma_м^2 < \sigma_y^2$. Отсюда слѣдуетъ, что въ случаѣ несовершенной корреляціи η всегда меньше единицы и что она никогда не можетъ быть больше единицы.

Что касается вѣроятной ошибки корреляціоннаго отношенія, то она также была найдена Пирсономъ въ цитированномъ мемуарѣ [1]). Полное выраженіе ея слишкомъ сложно, но съ достаточной точностью можно принять для нея приближенную формулу:

$$(206) \; . \; . \; . \; . \; . \; . \; . \; . \; . \; E_\eta = 0{,}67449 \, \frac{1 - \eta^2}{\sqrt{N}} \, .$$

Примѣръ. Разсмотримъ чер. 25 [2]), изображающій зависимость между средними мѣсячными цѣнами ржи и чугуна въ Германіи въ періодъ 1879—1900 г.г. Т. к. регрессія явно криволинейная, то можно думать, что коэффиціентъ корреляціи мало пригоденъ для характеристики степени зависимости. Въ самомъ дѣлѣ, онъ равняется $r = 0{,}19 \pm 0{,}04$, т. е. въ $4^3/_4$ раза превышаетъ свою вѣроятную ошибку, но все-же представляетъ изъ себя величину незначительную.

Если мы вычислимъ корреляціонное отношеніе, то будемъ имѣть:

$$\eta_y = \frac{\sigma_{м_y}}{\sigma_y} = 0{,}265 \pm 0{,}039 \, ,$$

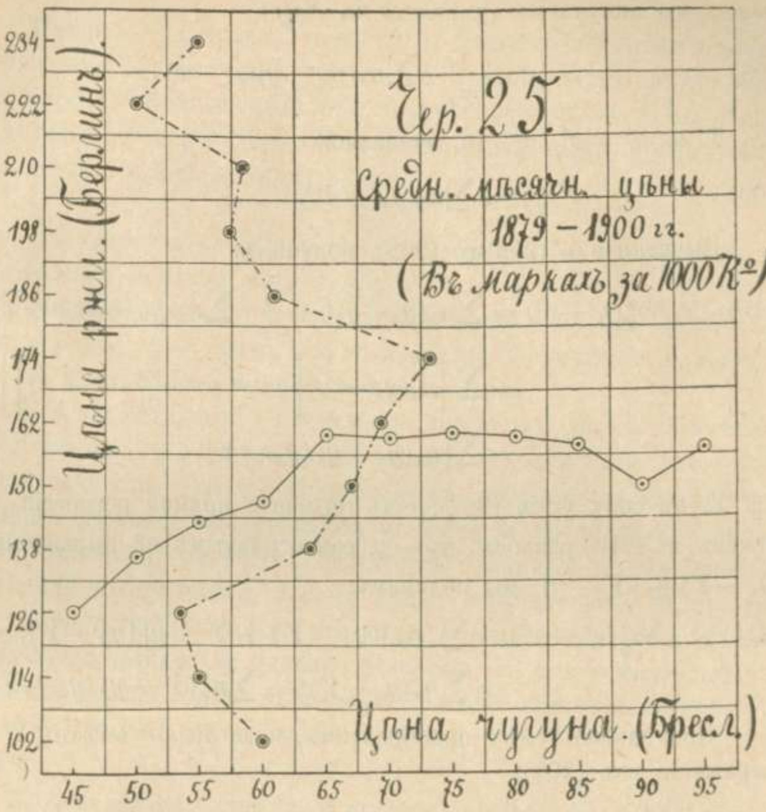
$$\eta_x = \frac{\sigma_{м_x}}{\sigma_x} = 0{,}576 \pm 0{,}028 \, .$$

Мы видимъ, что корреляціонная зависимость между цѣнами ржи и чугуна довольно высока, во много разъ превышаетъ свою вѣроятную ошибку и, слѣдовательно, не случайна. Кромѣ того, мы замѣчаемъ, что зависимость средней цѣны чугуна отъ цѣны ржи значительно (больше чѣмъ вдвое) выше зависимости средней цѣны ржи отъ цѣны чугуна,—обстоятельство, значеніе котораго съ точки зрѣнія экономической науки можетъ быть выяснено лишь дальнѣйшимъ изслѣдованіемъ.

---

1) On the general Theory of Skew Correlation etc. p. 11—19.

2) См. также табл. IX прилож.



§ 20. *Зависимость между корреляціоннымъ отношеніемъ* (η) *и коэффиціентомъ корреляціи* (r).

Мы видѣли выше, что прямая регрессіи есть прямая, наиболѣе тѣсно примыкающая къ эмпирической линіи регрессіи. Ея ур-іе

(207) . . . . . . $Y = \bar{y} + \frac{\sigma_y}{\sigma_x} r (x - \bar{x})$, гдѣ

$$r = \frac{\sum n_{xy}(x - \bar{x})(y - \bar{y})}{N\sigma_x\sigma_y}$$

или, какъ было доказано:

$$= \frac{\sum n_x(y_x - \bar{y})(x - \bar{x})}{N\sigma_x\sigma_y}.$$

Отсюда мы получаемъ (умножая на $N\sigma_y^2 r$)

$$(208) \quad N\sigma_y^2 r^2 = \frac{\sigma_y}{\sigma_x} r \sum n_x(y_x - \bar{y})(x - \bar{x}).$$

Т. к. $\eta^2 = \sigma_м^2 / \sigma_y^2$, то, очевидно:

$$(209) \quad N\sigma_y^2 \eta^2 = \sum n_x (y_x - \bar{y})^2.$$

Вычитая изъ (209)-го (208), получимъ:

$$(210) \quad N\sigma_y^2(\eta^2 - r^2) = \sum n_x(y_x - \bar{y})^2 - \frac{\sigma_y}{\sigma_x} r \sum n_x (y_x - \bar{y})(x - \bar{x}) =$$
$$= \sum \left\{ n_x(y_x - \bar{y}) \left[y_x - \bar{y} - \frac{\sigma_y}{\sigma_x} r(x - \bar{x}) \right] \right\} =$$
$$= \sum \{ n_x(y_x - \bar{y}) \, [y_x - Y] \},$$

гдѣ Y, въ силу ур-ія (207), есть ордината прямой регрессіи. Замѣнивъ въ (210) разность $y_x - \bar{y}$ равносильнымъ ей выраженіемъ $(y_x - Y) + (Y - \bar{y})$, мы получимъ:

$$(211) \quad N\sigma_y^2(\eta^2 - r^2) = \sum \{ n_x \, [(y_x - Y) + Y - \bar{y}] \, [y_x - Y] \} =$$
$$= \sum n_x (y_x - Y)^2 + \sum n_x (Y - \bar{y})(y_x - Y).$$

Послѣднюю сумму преобразуемъ, подставляя вмѣсто Y его выраженіе изъ (207):

$$\sum \left\{ n_x \left[\bar{y} + \frac{\sigma_y}{\sigma_x} r(x - \bar{x}) - \bar{y} \right] \left[y_x - \bar{y} - \frac{\sigma_y}{\sigma_x} r(x - \bar{x}) \right] \right\} =$$
$$= \frac{\sigma_y}{\sigma_x} r \sum n_x (x - \bar{x})(y_x - \bar{y}) - \frac{\sigma_y^2}{\sigma_x^2} r^2 \sum n_x (x - \bar{x})^2 =$$
$$= \frac{\sigma_y}{\sigma_x} r \,.\, N\sigma_x \sigma_y r - \frac{\sigma_y^2}{\sigma_x^2} r^2 \,.\, N\sigma_x^2 = N\sigma_y^2 r^2 - N\sigma_y^2 r^2 = 0.$$

Так. обр., выраженіе (211) сводится къ одной первой суммѣ, и мы окончательно имѣемъ:

$$(212) \quad N\sigma_y^2(\eta^2 - r^2) = \sum n_x (y_x - Y)^2.$$

Это выраженіе чрезвычайно важно. Оно показываетъ намъ, что $\sigma_y^2(\eta^2 - r^2)$ представляетъ собою средній квадратъ отклоненій линіи регрессіи отъ прямой, которая наиболѣе тѣсно къ ней примыкаетъ, и оно-же позволяетъ намъ сдѣлать нѣкоторыя заключенія относительно величины корреляціоннаго отношенія.

Именно, т. к. правая часть (212) состоитъ изъ однѣхъ только положительныхъ величинъ, то всегда должно имѣть мѣсто неравенство $\eta^2 > r^2$. Такъ какъ при этомъ корреляціонное отношеніе мы должны считать величиной положительной, то изъ предыдущаго неравенства будетъ слѣдовать, что

$$\eta > |r|,$$

т. е. что *корреляціонное отношеніе всегда больше абсолютной величины коэффиціента корреляціи.*

Равенство между этими величинами возможно только въ томъ случаѣ, когда (какъ это показываетъ формула 212) *каждое* $y_x = Y$, *т. е. когда регрессія строго линейна.*

Всѣ предыдущія разсужденія имѣли въ своемъ основаніи то допущеніе, что численность нашей совокупности настолько велика, что вѣроятными ошибками можно пренебречь. Въ дѣйствительной статистической практикѣ мы никогда не встрѣтимъ абсолютно линейной регрессіи, вслѣдствіе наличности случайныхъ отклоненій въ предѣлахъ нашего всегда ограниченнаго матеріала. При наличности нѣкоторой практики можно, разумѣется, судить на глазъ, достаточно-ли линейна регрессія, но для сомнительныхъ случаевъ важно знать вѣроятную ошибку критерія линейности $(\eta^2 - r^2)$, который мы обозначимъ буквой

(213) . . . . $\zeta = \eta^2 - r^2$.

Точное значеніе вѣроятной ошибки найдено J. Blackman'омъ [1]), но вслѣдствіе его большой сложности допустимо пользованіе приближенной формулой, данной имъ-же:

(214) . . . . . . $E_\zeta = 0{,}67449 \, . \, 2\sqrt{\frac{\zeta}{N}}$.

Если необходима бо́льшая точность, можно поступить слѣдующимъ образомъ. Очевидно, что на ряду съ разностью $\zeta = \eta^2 - r^2$, критеріемъ линейности можетъ служить также и разность абсолютныхъ значеній обѣихъ величинъ

(215) . . . . $\vartheta = |\eta| - |r|$.

Blackman для вѣроятныхъ ошибокъ ζ и ϑ даетъ въ качествѣ дальнѣйшаго приближенія слѣдующія формулы:

(216) . . $\frac{\zeta}{E_\zeta} = \frac{\sqrt{N}}{0{,}67449} \cdot \frac{1}{2}\sqrt{\zeta} \cdot \frac{1}{\sqrt{1 + (1-\eta^2)^2 - (1-r^2)^2}}$ и

[1]) Biometrika Vol. IV, On Tests for Linearity of Regression in Frequency Distributions, p. 332—350.

$$(217) \ldots \frac{\vartheta}{E_\vartheta} = \frac{\sqrt{N}}{0{,}67449} \cdot \frac{1}{2}\sqrt{\zeta} \cdot \frac{\sqrt{4\eta_1|r|}}{\eta_1 + |r|} \cdot \frac{1}{\sqrt{1 + \frac{|r|(1-\eta_1^2)^2 - \eta_1(1-r^2)^2}{\eta_1 + |r|}}}.$$

Обыкновенно можно пользоваться формулой (214) или вытекающей изъ нея

$$(218) \ldots\ldots \frac{\zeta}{E_\zeta} = \frac{\sqrt{N}}{0{,}67449} \cdot \frac{1}{2}\sqrt{\zeta}.$$

Если эта величина не больше 2 или $2^1/_2$, регрессію можно считать линейной. Въ сомнительныхъ случаяхъ, слѣдуетъ обратиться къ формуламъ (216) и (217). Если они дадутъ близкіе результаты, то на этомъ можно успокоиться. При значительномъ разногласіи между ними придется обратиться къ точнымъ выраженіямъ для вѣроятныхъ ошибокъ ζ и ϑ (Biometrica Vol. IV, p. 337 и 339), которыхъ я не привожу вслѣдствіе ихъ сложности.

§ 21. Корреляція и причинная зависимость.

Вопросъ, которымъ мы намѣреваемся закончить эту главу, заслуживаетъ, конечно, гораздо болѣе вниманія, чѣмъ мы можемъ здѣсь удѣлить ему. Мы принуждены однако ограничиться лишь немногими замѣчаніями.

Какъ корреляціонная, такъ и строго функціональная зависимость, разумѣется, не тождественны съ причинной связью между явленіями. Это явствуетъ хотя бы уже изъ того обстоятельства, что функціональная зависимость и корреляція могутъ имѣть мѣсто въ тѣхъ сферахъ, гдѣ было бы безсмысленно говорить о причинѣ и слѣдствіи, именно, въ мірѣ идеальныхъ геометрическихъ и иныхъ подобныхъ построеній.

Логическая природа этихъ понятій глубоко различна, и мы можемъ замѣтить это, даже не касаясь теоретикопознавательной стороны вопроса. Прежде всего, всякая функціональная зависимость, какъ строгая, такъ и нестрогая [1]) (корреляціонная) двух-

[1]) Терминъ „не строго опредѣленной“, или „не совершенной“ зависимости (связи, взаимоотношенія) встрѣчается у П. А. Некрасова; см. его „Теорію вѣроятностей“ 2 изд. 1912 г, напр., стр. 427, 439. Этимъ въ высокой степени интереснымъ трудомъ авторъ не могъ воспользоваться, т. к. получилъ его уже во время печатанія настоящей работы.

стороння въ томъ смыслѣ, что, какое перемѣнное разсматривать какъ независимое, а какое какъ зависимое, зависитъ отъ нашего произвола. Относительно-же причины и слѣдствія мы связаны фактическимъ состояніемъ вещей и должны подчиниться результатамъ изслѣдованія.

Затѣмъ, сужденіе о наличности функціональной связи и сужденіе о причинѣ даже въ приложеніи къ одному и тому-же матеріалу имѣютъ различное содержаніе. Разсмотримъ, напр., зависимость между массой, объемомъ, температурой, и давленіемъ газа. Она выражается, какъ извѣстно, формулой

$$\frac{pv}{t+273}=CM,$$

гдѣ M — масса, v — объемъ, p — давленіе, t — температура, а C — нѣкоторая постоянная. Ур-іе это связываетъ 4 величины. Каждую изъ нихъ мы можемъ найти, если извѣстны 3 другихъ. Вотъ основной смыслъ этого ур-ія. Если двѣ величины будутъ оставаться постоянными, а третья будетъ мѣняться, то будетъ извѣстнымъ образомъ измѣняться и четвертая. Всѣ четыре величины функціонально связаны, и нелѣпо спрашивать, какая изъ нихъ причина, какая слѣдствіе.

Этотъ вопросъ однако вполнѣ законенъ, когда мы отнесемъ его не къ зависимости величинъ *вообще*, а къ конкретнымъ физическимъ процессамъ измѣненія. Такъ, пусть v и M остаются постоянными; тогда, зная p, мы можемъ найти t, зная t, можемъ найти p. Каждая изъ этихъ величинъ съ чисто математической точки зрѣнія можетъ быть разсматриваема, какъ независимая перемѣнная, но физически обѣ онѣ играютъ существенно различную роль. Именно, если объемъ и масса остаются постоянными, то мы не можемъ измѣнить давленія иначе, какъ посредствомъ измѣненія температурнаго состоянія газа. Поэтому измѣненіе температуры будетъ причиной, измѣненіе давленія всегда только слѣдствіемъ.

Если мы будемъ сжимать газъ, явленіе будетъ сложнѣе, такъ какъ при этомъ должны измѣняться въ общемъ случаѣ и упругость и температура. Взявъ для простоты измѣненіе упругости при постоянной температурѣ, мы найдемъ, что оно является очень простой функціей объема, если судить только съ точки зрѣнія чисто функціональнаго взаимоотношенія величинъ. То-же измѣненіе упругости съ причинной точки зрѣнія будетъ результатомъ сложной комбинаціи явленій: затраты работы на уменьшеніе объема и оттока энергіи, поддерживающаго температуру на прежнемъ уровнѣ.

Что касается корреляціи, то здѣсь дѣло еще сложнѣе. Изъ двухъ коррелятивно связанныхъ между собою явленій ни одно, вообще говоря, не можетъ разсматриваться какъ полная причина другого. Если бы это было такъ, то мы имѣли бы не корреляціонную, а строгую функціональную зависимость. Наличность корреляціи свидѣтельствуетъ поэтому лишь о томъ, что одно изъ двухъ явленій есть частичная причина другого или частичное слѣдствіе, или что оба они суть слѣдствія частично-общихъ причинъ. Возможны и комбинаціи этихъ случаевъ. Такъ, напр., корреляція цѣнъ ржи въ Москвѣ и въ Самарѣ устанавливается на основаніи многократныхъ наблюденій фактовъ, въ основѣ которыхъ лежатъ сложные процессы. Во первыхъ, цѣны въ обоихъ городахъ складываются подъ общими вліяніями, вдущими изъ множества другихъ пунктовъ рынка: въ этомъ смыслѣ несомнѣнно мы имѣемъ дѣло съ общими слѣдствіями. Затѣмъ, по всей вѣроятности, имѣетъ мѣсто и прямое вліяніе цѣнъ, причемъ одинъ разъ повышеніе цѣны въ Самарѣ можетъ служить стимуломъ повышенія цѣнъ и въ Москвѣ, въ другой разъ, повышеніе цѣны въ Москвѣ можетъ повести къ повышенію цѣнъ и въ Самарѣ. Коэффиціентъ корреляціи погашаетъ всю эту живую игру экономическихъ силъ, переводя лишь результатъ всѣхъ этихъ скрещивающихся взаимно вліяній на языкъ точныхъ цифръ. Вотъ почему статистикъ, какъ таковой, будучи вполнѣ компетентнымъ въ установленіи корреляціи между любыми величинами, къ какой бы области онѣ ни принадлежали, не компетентенъ въ высказываніи причинныхъ сужденій. Для этого мало быть статистикомъ, а нужно быть біологомъ, медикомъ, метеорологомъ, экономистомъ и т. д. смотря по области изслѣдованія.

Если два явленія внѣшняго міра связаны функціонально, это еще не значитъ, что между ними существуетъ прямая или косвенная причинная связь. Наличность функціональной связи можетъ быть и чисто случайной или, лучше сказать, чисто идеографической. Напр., между движеніемъ земли по ея орбитѣ и движеніемъ Сиріуса по своей—нѣтъ никакой причинной связи: ни прямой, ни косвенной. Однако, выведя изъ наблюденій закономѣрность въ движеніи Сиріуса, астрономъ приводитъ его въ функціональную зависимость съ движеніемъ земли, ибо по положенію ея (кажущемуся положенію солнца) опредѣляется время, а въ функціи этого солнечнаго времени опредѣляется положеніе Сиріуса.

Точно также и наличность корреляціи вовсе не служитъ еще сама по себѣ признакомъ существованія причинной связи, какъ прямой, такъ и косвенной. То обстоятельство, что коэффиціентъ корреляціи во много разъ превышаетъ свою вѣроятную ошибку, служитъ лишь критеріемъ, при помощи котораго мы выдѣляемъ основныя черты явленія, эллиминируя вліяніе неопредѣленнаго множества причинъ, имѣющихъ тенденцію при большомъ числѣ испытаній взаимно уничтожиться. Противъ систематической ошибки малость вѣроятной ошибки насъ не застраховываетъ. Точно также не застраховываетъ она насъ и отъ возможнаго случайнаго совпаденія во времени двухъ самостоятельныхъ рядовъ причинъ. Напр., если бы мы захотѣли изслѣдовать зависимость между движеніемъ крыльевъ двухъ любыхъ птицъ и для этого запротоколировали какъ нибудь эти движенія, напр. при помощи синематографа, то мы могли бы получить результаты двоякаго рода въ зависимости отъ величины періода наблюденія. Если бы мы стали отмѣчать наклонъ крыльевъ по прошествіи каждой секунды и вели бы наблюденіе въ теченіе болѣе или менѣе продолжительнаго времени, то коэффиціентъ корреляціи оказался бы равнымъ нулю. Но если бы мы отмѣчали наклонъ крыльевъ въ продолженіе секунды черезъ промежутки равные тысячнымъ долямъ этого періода, то мы нашли бы, что коэффиціентъ корреляціи не равняется нулю, а представляетъ изъ себя величину довольно значительную. При этомъ онъ былъ бы положительнымъ, если бы начало нашихъ наблюденій *случайно* совпало съ моментомъ, когда обѣ птицы подымали или опускали свои крылья и отрицательнымъ, если бъ въ этотъ короткій промежутокъ времени одна птица опускала, другая подымала крылья.

Нѣчто подобное и встрѣчается намъ при изслѣдованіи корреляціи различныхъ соціальныхъ явленій, движущихся во времени. Здѣсь мы часто наблюдаемъ, во первыхъ, медленныя многолѣтнія измѣненія величины (цѣны, рождаемости, смертности и т. д.) и кромѣ того годичныя, мѣсячныя и даже суточныя колебанія около этого медленно измѣняющагося уровня.

Можетъ случиться, что паденіе или поднятіе уровня каждой изъ этихъ величинъ вызвано двумя рядами независимыхъ причинъ, т. что коэффиціентъ корреляціи, который мы вычислимъ, введетъ насъ только въ заблужденіе относительно существованія причинной связи, если мы не учтемъ этого обстоятельства.

Мы нашли, напр., довольно высокую величину ($0{,}19 \pm 0{,}04$) для корреляціоннаго отношенія цѣны чугуна и цѣны ржи. Разсмотрѣніе чер. 26 [1]), представляющаго намъ картину движенія мѣсячныхъ цѣнъ обоихъ товаровъ въ теченіе 21 года, приводитъ къ тому заключенію, что обѣ явленія — въ предѣлахъ изучаемаго промежутка времени по крайней мѣрѣ — подвержены были періодическимъ колебаніямъ съ очень большой амплитудой. Приблизительнаго совпаденія періодичности однако уже достаточно, чтобы корреляція оказалась на лицо. Это совпаденіе могло быть случайнымъ, колебанія могли раньше не совпадать, они могутъ по прошествіи извѣстнаго времени разойтись, мы объ этомъ по тѣмъ двумъ періодамъ, которые здѣсь имѣемъ, еще судить не можемъ [2]).

Итакъ, нашей задачей является раздѣлить и раздѣльно изслѣдовать два момента: (1) взаимную зависимость медленныхъ измѣненій уровня величинъ и (2) взаимную зависимость колебаній ихъ около этого уровня. Первая задача до сихъ поръ не получила еще надлежащаго разрѣшенія, и мы ее оставимъ въ сторонѣ [3]), для второй-же нѣкоторые методы уже выработаны.

§ 22. *Методъ моментальнаго средняго и методъ послѣдовательныхъ отклоненій.*

Первый методъ былъ предложенъ Гукеромъ [4]). Онъ замѣтилъ, что если начертить рядомъ кривую брачности и кривую внѣшней торговли (импорта или экспорта), то положительная зависимость между этими двумя явленіями, конечно, бросается сразу въ глаза, но что зависимость эта существуетъ повидимому только между отклоненіями этихъ величинъ отъ нѣкотораго медленно измѣняющагося уровня. Въ то время какъ отдѣльные зигзаги

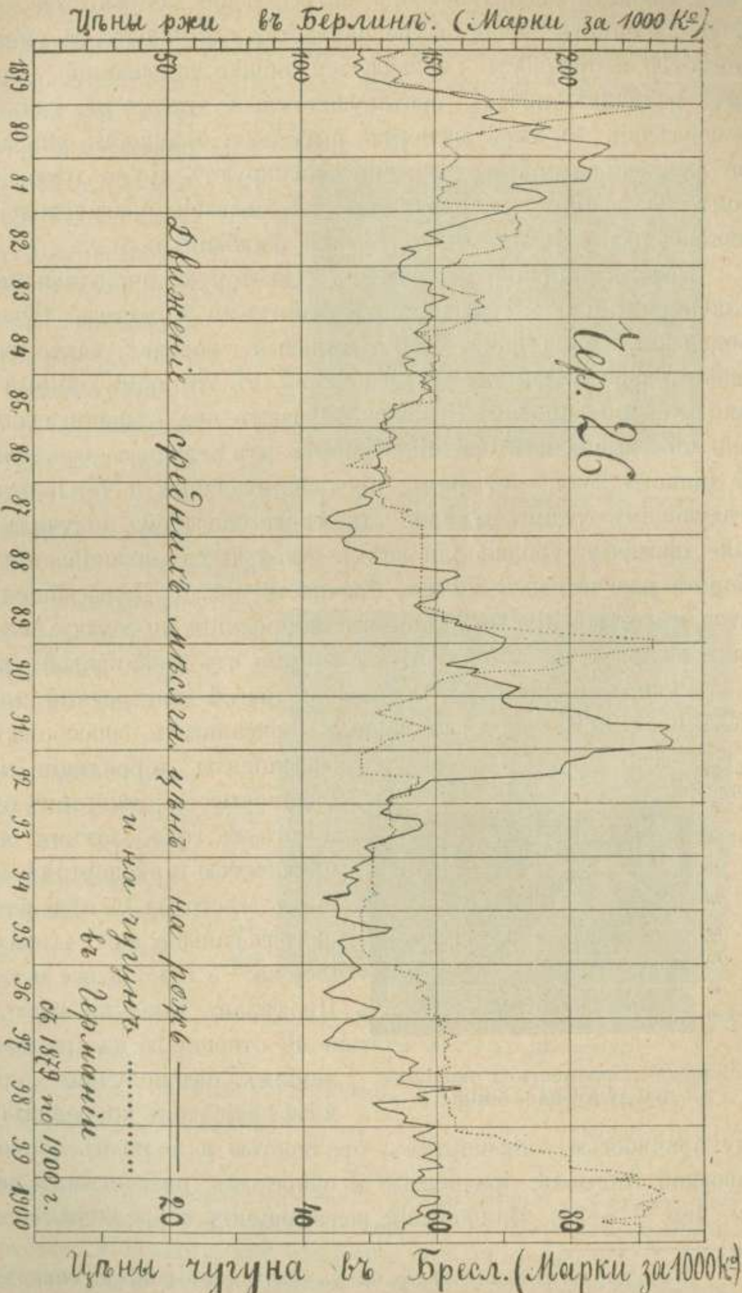
---

[1]) Составленъ, также какъ и таблица IX приложенія, по даннымъ, сгруппированнымъ у Otto Schmitz'a, Die Bewegung der Waarenpreise in Deutschland, Berlin, S. 65—70 и 217—222.

[2]) Другой прим. см. R. H. Hooker, the Suspension of the Berlin Produce Exchange and its Effect upon Corn Prices. Journ. of the Roy. Stat. Society, 1901, p. 603

[3]) Позволю себѣ замѣтить, что эта задача ожидаетъ еще изслѣдователя, который бы попытался приложить къ ея рѣшенію методы *гармоническаго анализа*, въ другихъ областяхъ оказывающаго великія услуги.

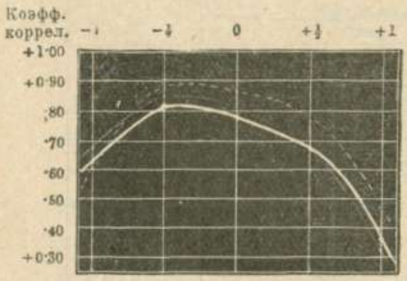
[4]) R. H. Hooker, Correlation of the Marriage-Rate with Trade, Journ. of the Roy. Stat. Soc. 1901, p. 485—492.


Цѣны ржи въ Берлинѣ. (Марки за 1000 Kº)
Черт. 26
Движенiе среднихъ мѣсячн. цѣнъ на рожь ——
и на чугунъ ……
въ Германiи
съ 1879 по 1900 г.
Цѣны чугуна въ Бресл. (Марки за 1000 kº)

этихъ кривыхъ обнаруживаютъ значительное соотвѣтствіе, т. что съ увеличеніемъ экспорта увеличивается и брачность, съ уменьшеніемъ его и брачность уменьшается, общее направленіе движенія этихъ величинъ носитъ противуположный характеръ, такъ какъ за послѣднія 40 лѣтъ экспортъ выросъ, а брачность упала. Эти два разнохарактерныхъ движенія маскируютъ другъ друга, и поэтому коэффиціентъ корреляціи оказывается незначительнымъ, равняясь всего $+0{,}18$ съ вѣроятной ошибкой въ 0,9.

Оба явленія обнаруживаютъ нѣкоторую періодичность, съ величиной періода, равной приблизительно 9 годамъ. Гукеръ и предложилъ опредѣлять „моментальный уровень", какъ среднее ариѳметическое для такого ряда лѣтъ, въ которомъ данный годъ находится по серединѣ. Такимъ образомъ при рѣшеніи своей задачи онъ вычислялъ средній экспортъ изъ величинъ, относящихся къ данному году, четыремъ предшествующимъ и четыремъ послѣдующимъ годамъ, и точно такимъ-же способомъ получилъ кривыя движенія уровня для всѣхъ остальныхъ явленій: импорта, оборота разсчетныхъ палатъ, брачности и т. д. Дальнѣйшее сводится къ отысканію коэффиціента корреляціи не между абсолютными величинами, а между отклоненіями отъ этой кривой уровня.

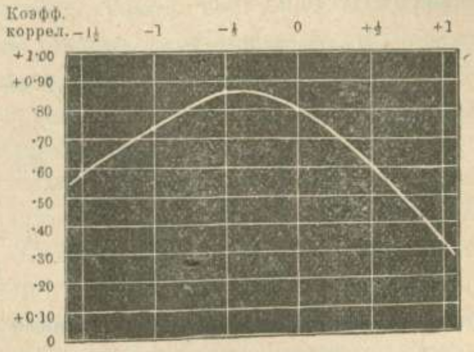
Гукеръ примѣнилъ тамъ-же и другой интересный методъ изслѣдованія. Именно, онъ нашелъ (описаннымъ способомъ) коэффиціенты корреляціи между брачностью и экспортомъ одного и того-же года, затѣмъ между брачностью и экспортомъ, имѣвшимъ мѣсто на $^1/_2$ года раньше, на годъ раньше, на $1^1/_2$ года раньше, на $^1/_2$ года позже и т. д. [1]). Продѣлавъ такія-же вычисленія и по отношенію къ другимъ явленіямъ, онъ получилъ „кривыя коэффиціентовъ корреляціи" между брачностью и экспортомъ, брачностью и полнымъ оборотомъ вывозной торговли, брачностью и оборотомъ разсчетныхъ палатъ (см. чер. 27—29). Наибольшій коэффиціентъ корреляціи (вершина



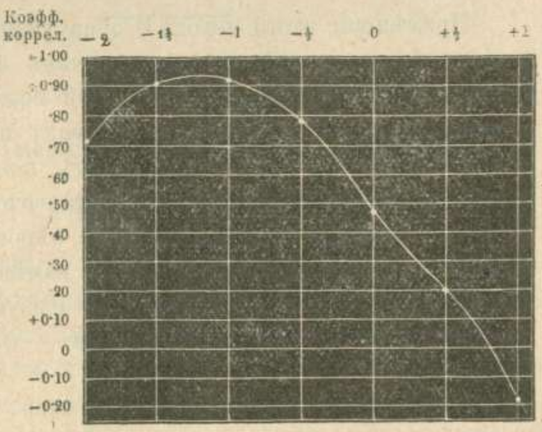
Черт. 27. Брачность и экспортъ на душу населенія.

[1]) Экспортъ, имѣвшій мѣсто, напр., полугодомъ раньше, опредѣлялся какъ среднее ариѳметическое экспорта даннаго и экспорта предыдущаго года.

соотвѣтствующей кривой) отмѣчаетъ величину періода, по истеченіи котораго данное явленіе проявляетъ максимальное свое вліяніе на размѣры брачности. Изъ приложенныхъ чертежей видно, что вліяніе измѣненій въ экспортѣ и импортѣ на брачность съ наибольшей силой сказывается приблизительно черезъ 5 мѣсяцевъ, вліяніе-же измѣненій въ оборотахъ разсчетныхъ палатъ черезъ 1 годъ $2^1/_2$ мѣсяца. Эти результаты представляютъ изъ себя во всякомъ случаѣ нѣчто новое, чего простымъ сопоставленіемъ и разсматриваніемъ кривыхъ получить никакъ не возможно [1]).



Чер 28. Брачность и импортъ на душу населенія.



Чер. 29. Брачность и оборотъ разсчетной палаты на душу населенія.

Тому-же автору принадлежитъ и методъ, который мы назвали въ заголовкѣ параграфа методомъ послѣдовательныхъ отклоненій. Онъ состоитъ въ томъ, что коэффиціентъ корреляціи находится не между самими величинами, а между разностями ихъ послѣдовательныхъ значеній.

---

[1]) Методъ „сглаживанія" колебаній, употребленный Гукеромъ, конечно, не совершененъ. Соціальныя явленія не отличаются правильной періодичностью, и поэтому при среднемъ періодѣ, напр., въ 9 лѣтъ на нѣкоторыя девятилѣтія можетъ упасть по два максимума или по два минимума, повышая или понижая безъ достаточнаго основанія уровень даннаго года. Поэтому кривая, полученная такимъ

Такъ, пусть $x_0, x_1, x_2, \ldots x_n$ и $y_0, y_1, y_2, \ldots y_n$ будутъ двѣ серіи наблюденій, сдѣланныхъ черезъ равные промежутки времени (напр. ежедневныя цѣны на двухъ какихъ нибудь рынкахъ). Пусть $d_{x_1} = x_1 - x_0$, $d_{x_2} = x_2 - x_1, \ldots d_{x_n} = x_n - x_{n-1}$; $d_{y_1} = y_1 - y_0$, $d_{y_2} = y_2 - y_1, \ldots d_{y_n} = y_n - y_{n-1}$.

Для среднихъ ариѳметическихъ мы будемъ имѣть очень простыя выраженія:

$$(219) \ldots\ldots \begin{cases} \bar{d}_x = \frac{1}{n} \sum d_{x_i} = \frac{x_n - x_0}{n}, \\ \bar{d}_y = \frac{1}{n} \sum d_{y_i} = \frac{y_n - y_0}{n}, \end{cases}$$

что дастъ величины обыкновенно мало разнящіяся отъ нуля.

Дальнѣйшія вычисленія, т. е. нахожденіе среднихъ отклоненій и коэффиціента корреляціи $r_{d_x d_y}$, производятся обычными способами.

Примѣненіе этого метода [1]) обнаруживаетъ, что между коэффиціентомъ корреляціи самихъ величинъ и коэффиціентомъ корреляціи послѣдовательныхъ разностей можетъ быть значительная разница. Напр., цѣна маиса на фермахъ штата Айова и весь урожай маиса въ Соед. Штатахъ даютъ коэффиціентъ корреляціи равный всего $-0{,}28 \pm 0{,}14$, что нужно счесть величиной близкой къ нулю, если принять во вниманіе вѣроятную ошибку. Зависимость-же между послѣдовательными измѣненіями одной и другой величины отъ года къ году характеризуется весьма высокимъ коэффиціентомъ корреляціи, равнымъ $-0{,}84$, указывающимъ по

способомъ, можетъ быть еще сглажена, напр., отъ руки. Этимъ однако вносится въ изслѣдованіе субъективный моментъ, котораго лучше избѣгать. Поэтому наилучшимъ методомъ „сглаживанія" будетъ употребленіе, подходящей кривой, напр., параболы, ур-іе которой не трудно опредѣлить пользуясь методами, изложенными въ §§ 7—8 I ч. и въ § 17 II ч. При нахожденіи коэффиціентовъ корреляціи по методу Гукера и при пользованіи кривой, сглаженной отъ руки, во всякомъ случаѣ *безусловно* необходимо при печатаніи результатовъ изслѣдованія прилагать и чертежъ „сглаженной" кривой въ достаточномъ масштабѣ, т. к. безъ него невозможно составить представленіе о правильности выполненія работы.

[1]) Онъ былъ предложенъ Гукеромъ въ 1901 г. (см. Journ. of the Roy. Stat. Soc., 1901 p. 604) и подробнѣе разсмотренъ и примѣненъ въ его-же статьѣ: On the Correlation of Successive Observations; Illustrated by Corn Prices, Journ. of the Roy. Stat. Soc., 1905, p. 696—703.

мнѣнію Гукера на то, что измѣненія въ количествѣ маиса, производимаго въ С.-Штатахъ, являются однимъ изъ наиболѣе важныхъ факторовъ, опредѣляющихъ цѣну, выплачиваемую въ Айовѣ [1]).

Примѣненіе метода корреляціи къ явленіямъ, измѣняющимся во времени, находится во всякомъ случаѣ еще въ стадіи первыхъ попытокъ, дѣлаемыхъ чисто эмпирически, ощупью. Этотъ вопросъ ждетъ еще систематической разработки, и для изслѣдователей открывается здѣсь обширное поле. Прежде всего важно выяснить, при какихъ условіяхъ возможенъ здѣсь переходъ отъ простого констатированія корреляціонной связи къ тѣмъ или инымъ типамъ каузальныхъ сужденій. Затѣмъ, и въ связи съ первымъ вопросомъ, необходимо подвергнуть всю проблему разсмотрѣнію съ точки зрѣнія теоріи вѣроятности, которая одна только можетъ дать ключь къ разрѣшенію цѣлаго ряда частныхъ вопросовъ, относящихся къ указанной проблемѣ [2]).

ГЛАВА II.

Корреляція между тремя и болѣе величинами.

§ 23. Основная теорема теоріи линейной регрессіи.

Группу значеній величинъ $x_1, x_2, \ldots x_p$ назовемъ для краткости *индивидуумомъ совокупности*. $x_1, x_2, x_3 \ldots$ могутъ быть размѣрами различныхъ органовъ одного и того-же существа, или x_1 можетъ быть размѣромъ органа, принадлежащаго потомку, $x_2, x_3, \ldots x_p$ размѣрами того-же или другихъ опредѣленныхъ органовъ его послѣдовательныхъ предковъ. Это могутъ быть цѣны одного и того же товара въ различныхъ пунктахъ рынка въ одинъ и тотъ-же моментъ или цѣны различныхъ товаровъ въ

---

[1]) l. cit. p. 703.

[2]) Литературныя указанія относительно (немногочисленныхъ пока еще) попытокъ приложенія метода корреляціи къ экономическимъ и соціальнымъ вопросамъ см. G. U. Yule, The Applications of the Method of Correlation to Social and Economic Statistics, Journ. of the Roy. Stat. Soc., 1909, p. 721—730. Болѣе полная библіографія, главнымъ образомъ, теоретическихъ и біологическихъ работъ, принадлежащихъ новому направленію, имѣется въ цитир. на стр. 1 работѣ А. В. Леонтовича, ч. II, стр. 191—214.

одномъ и томъ-же пунктѣ и т. д. и т. д. Способы, при помощи которыхъ объединяются въ одну группу опредѣленныя значенія величинъ, могутъ быть разнообразны до безконечности, но, конечно, въ одномъ изслѣдованіи основаніе ассоціаціи величинъ должно быть выдержано отъ начала до конца.

Каждая изъ величинъ $x_1, x_2, \ldots x_p$ можетъ принимать опредѣленное значеніе въ различныхъ группахъ, или, какъ мы говоримъ, у различныхъ индивидуумовъ совокупности. Если бы связь между этими величинами была строгой функціональной, то, зная значенія всѣхъ величинъ кромѣ одной, мы могли бы опредѣлить и значеніе этой послѣдней. Въ случаѣ корреляціи связь носитъ болѣе свободный характеръ. Совокупности опредѣленныхъ значеній $x_2, x_3 \ldots x_p$ соотвѣтствуетъ не одно, а множество значеній x_1, причемъ среднее ариѳметическое этихъ значеній является функціей значеній остальныхъ величинъ. Это среднее ариѳметическое x_1-го для *строя*, соотвѣтствующаго опредѣленному комплексу значеній $x_2, x_3, \ldots x_p$, мы обозначимъ черезъ x_{1m}, численность строя — черезъ $n_{x_1 \ldots x_p}$ или короче — черезъ $n_{(x_i)}$, среднее отклоненіе x_1 въ этомъ строѣ — черезъ σ_{1m}.

Ур-іе, связывающее среднее значеніе одного перемѣннаго со значеніями остальныхъ перемѣнныхъ, будемъ, какъ и раньше, называть ур-іемъ регрессіи. Если всѣ перемѣнныя входятъ въ это ур-іе въ первой степени, то мы будемъ имѣть дѣло съ линейной, въ противномъ случаѣ съ криволинейной регрессіей. Вслѣдствіе наличности случайныхъ причинъ мы никогда въ дѣйствительности не получимъ строго линейной регрессіи и намъ придется довольствоваться только приближеніемъ. Поэтому, какъ и раньше, будемъ различать эмпирическую регрессію и теоретическую. Въ случаѣ трехъ величинъ геометрическимъ образомъ, соотвѣтствующимъ линейной регрессіи, служитъ плоскость, криволинейная-же можетъ быть изображена нѣкоторой иной поверхностью. Въ случаѣ бо́льшого числа величинъ наглядное геометрическое представленіе невозможно, но математикъ имѣетъ суррогатъ наглядности въ представленіи многомѣрнаго пространства и различныхъ мыслимыхъ въ немъ образовъ, аналогичныхъ образамъ, представимымъ въ пространствѣ 3-хъ измѣреній. Намъ однако нѣтъ надобности обращаться къ этой концепціи, такъ какъ мы эмпирическія взаимоотношенія нашихъ величинъ можемъ мыслить въ формѣ ряда комбинаціонныхъ таблицъ, а теоретическую регрессію будемъ представлять аналитически ур-іемъ

$$(220) \ldots X_1 = a_{11} + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1p}x_p$$

и аналогично для другихъ перемѣнныхъ.

Перейдемъ теперь къ основной теоремѣ теоріи линейной регрессіи, доказанной выше въ § 6 для частнаго случая двухъ перемѣнныхъ [1]).

Обозначивъ разность между эмпирическимъ значеніемъ и теоретическимъ черезъ d_1, т. е. положивъ

$$(221) \ldots\ldots d_1 = x_{1m} - X_1,$$

мы опредѣлимъ коэффиціенты ур-ія (220) такимъ образомъ, чтобы

$$(222) \ldots\ldots \sum n_{(x_1)} d_1^2 = \text{minimum}.$$

Возьмемъ нижеслѣдующую сумму для i-ого строя и распространимъ ее на всѣ значенія x_1-го; мы найдемъ:

$$\sum_i [x_1 - (a_{11} + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1p}x_p)]^2 =$$
$$= \sum_i (x_1 - X_1)^2 = \sum_i [(x_1 - x_{1m}) + (x_{1m} - X_1)]^2 =$$
$$= \sum_i (x_1 - x_{1m})^2 + 2\sum (x_1 - x_{1m})(x_{1m} - X_1) + \sum (x_{1m} - X_1)^2 =$$
$$= n_{(x_1)}\sigma_{1m}^2 + n_{(x_1)}(x_{1m} - X_1)^2 = n_{(x_1)}\sigma_{1m}^2 + n_{x_1}d_1^2, \text{ откуда:}$$
$$n_{(x_1)}d_1^2 = \sum_i [x_1 - (a_{11} + a_{12}x_2 + \ldots + a_{1p}x_p)]^2 - n_{(x_1)}\sigma_{1m}^2,$$

и искомая сумма для всей совокупности равняется

$$(223) \ldots \sum n_{(x_1)}d_1^2 = \sum [x_1 - (a_{11} + a_{12}x_2 + \ldots + a_{1p}x_p)]^2 - \\ - \sum n_{(x_1)}\sigma_{1m}^2.$$

---

[1]) Случай n перемѣнныхъ былъ изслѣдованъ впервые Edgeworth'омъ, Phil. Magasine, Vol. 34, 1892, Correlated Averages, p. 190—204, а затѣмъ разработанъ дальше K. Pearson'омъ въ мемуарѣ: Regression, Heredity, and Panmixia, Phil. Trans., Vol. 187 A, p. 253—318. Наше изложеніе основано на оригинальной работѣ U. Yule'я, освободившаго теорію линейной регрессіи отъ излишнихъ допущеній. См. Proc. of the Roy. Soc. Vol. 60, On the Significance of Bravais' Formulae for Rergession, etc., in the Case of Skew Correlation, p. 477—489. Yule, впрочемъ, ограничивается только случаями двухъ, трехъ и четырехъ перемѣнныхъ.

Такъ какъ $\sum n_{(x_1)}\sigma_{1m}^2$ не зависитъ отъ $a_{11}, a_{12}, \ldots a_{1p}$, а только отъ характера самой совокупности, то, очевидно, $\sum n_{(x_1)}d_1^2$ будетъ наименьшей, когда будетъ наименьшей сумма

$$\sum[x_1-(a_{11}+a_{12}x_2+\ldots\ldots+a_{1p}x_p)]^2.$$

Такимъ образомъ, *основное условіе метода наименьшихъ квадратовъ*, какъ и для случая двухъ перемѣнныхъ, *сводится къ опредѣленію такой линейной функціи* (220), *чтобы, примѣнивъ ее къ опредѣленію величины перваго перемѣннаго по остальнымъ въ каждомъ индивидуальномъ случаѣ, мы получили ошибки, сумма квадратовъ которыхъ была бы наименьшей.*

§ 24. Случай трехъ величинъ.

Ур-ія регрессіи въ этомъ случаѣ будутъ:

$$(224)\ \ldots\ldots \begin{cases} X_1=a_{11}+a_{12}x_2+a_{13}x_3, \\ X_2=a_{22}+a_{21}x_1+a_{23}x_3, \\ X_3=a_{33}+a_{31}x_1+a_{32}x_2. \end{cases}$$

Найдемъ коэффиціенты перваго ур-ія; коэффиціенты остальныхъ можно будетъ написать по соображеніямъ симметріи.

Какъ показано было въ предыдущемъ параграфѣ, ур-іе регрессіи должно удовлетворять условію:

$$(225)\ \ldots\ \sum[x_1-(a_{11}+a_{12}x_2+a_{13}x_3)]^2=\text{minimum}.$$

Дифференцируя послѣдовательно по a_{11}, a_{12}, a_{13}, получимъ ур-ія:

$$(226)\ \ldots \begin{cases} \sum x_1-\sum a_{11}-\sum a_{12}x_2-\sum a_{13}x_3=0, \\ \sum x_1x_2-\sum a_{11}x_2-\sum a_{12}x_2^2-\sum a_{13}x_2x_3=0, \\ \sum x_1x_3-\sum a_{11}x_3-\sum a_{12}x_2x_3-\sum a_{13}x_3^2=0. \end{cases}$$

Первое ур-іе даетъ сразу $a_{11}=0$, ибо $\sum x_1=0$, $\sum x_2=0$, $\sum x_3=0$, потому что мы опять за x_1, x_2 и x_3 принимаемъ отклоненія соотвѣтствующихъ величинъ отъ ихъ среднихъ ариѳметическихъ.

Остальныя суммы состоятъ изъ членовъ извѣстнаго намъ типа. Именно, $\sum a_{12}x_2^2=a_{12}\sum x_2^2=a_{12}\,.\,N\sigma_2^2$; $\sum a_{13}x_3^2=Na_{13}\sigma_3^2$;

$\sum x_1 x_2 = N\sigma_1\sigma_2 r_{12}$; $\sum x_1 x_3 = N\sigma_1\sigma_3 r_{13}$, а $\sum x_2 x_3 = N\sigma_2\sigma_3 r_{23}$, гдѣ r_{12}, r_{23} и r_{13} — коэффиціенты корреляціи для перемѣнныхъ, взятыхъ попарно, вычисленные извѣстнымъ намъ способомъ. Такимъ образомъ, ур-ія (226) сводятся къ слѣдующимъ (мѣняемъ знаки, переносимъ часть членовъ въ правую часть и сокращаемъ на N):

$$(227) \quad . \; . \; . \quad \begin{cases} a_{12}\sigma_2^2 + a_{13}\sigma_2\sigma_3 r_{23} = \sigma_1\sigma_2 r_{12}, \\ a_{12}\sigma_2\sigma_3 r_{23} + a_{13}\sigma_3^2 = \sigma_1\sigma_3 r_{13}. \end{cases}$$

Рѣшая эти ур-ія, находимъ:

$$(228) \quad . \; . \; . \quad a_{12} = \frac{\sigma_1}{\sigma_2}\,\frac{r_{12} - r_{23} r_{13}}{1 - r_{23}^2};\; a_{13} = \frac{\sigma_1}{\sigma_3}\,\frac{r_{13} - r_{12} r_{23}}{1 - r_{23}^2},$$

гдѣ a_{12} и a_{13} будутъ коэффиціентами регрессіи. Ур-іе теоретической регрессіи будетъ:

$$(229) \quad . \; . \; . \quad X_1 = \frac{r_{12} - r_{23} r_{13}}{1 - r_{23}^2}\,\frac{\sigma_1}{\sigma_2}\,x_2 + \frac{r_{13} - r_{23} r_{12}}{1 - r_{23}^2}\,\frac{\sigma_1}{\sigma_3}\,x_3 \quad \text{или,}$$

если мы вмѣсто отклоненій отъ среднихъ ариѳметическихъ подставимъ величины x_1, x_2, x_3, измѣренныя отъ обычнаго нуля:

$$(230) \quad . \; . \; . \quad X_1 = \bar{x}_1 + \frac{r_{12} - r_{23} r_{13}}{1 - r_{23}^2}\,\frac{\sigma_1}{\sigma_2}\,(x_2 - \bar{x}_2) + $$

$$+ \frac{r_{13} - r_{23} r_{12}}{1 - r_{23}^2}\,\frac{\sigma_1}{\sigma_3}\,(x_3 - \bar{x}_3).$$

Если регрессія линейна, то это ур-іе дастъ величины среднихъ ариѳметическихъ (x_{1m}) для строевъ x_1-ого, образованныхъ по x_2-му и x_3-му. Если регрессія уклоняется отъ линейнаго типа, то это ур-іе дастъ все таки нѣкоторыя указанія относительно зависимости между величинами, такъ какъ оно представляетъ собою нѣкоторую среднюю линейную зависимость, примѣняя которую къ отдѣльнымъ частнымъ случаямъ, мы сдѣлаемъ ошибки, сумма квадратовъ которыхъ будетъ наименьшей.

Средняя (квадратичная) величина этой ошибки представляетъ извѣстный интересъ. Мы обозначимъ ее черезъ Σ_1 и, очевидно, будемъ имѣть (см. форм. 225):

$$(231) \quad . \; . \; . \; . \quad \Sigma_1^2 = \frac{1}{N}\sum (x_1 - a_{12}x_2 - a_{13}x_3)^2.$$

Если мы подставимъ сюда значенія a_{12} и a_{13} изъ (228), то произведя рядъ простыхъ алгебраическихъ передѣлокъ получимъ:

$$(232') \; . \; . \; . \; . \; \Sigma_1^2 = \sigma_1^2 \frac{1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2 r_{12} r_{23} r_{13}}{1 - r_{23}^2} \,.\ ^{1)}$$

Остальныя формулы можно написать по симметріи, перемѣнивъ соотвѣтствующимъ образомъ значки; мы будемъ имѣть:

$$(232'') \; . \; . \; . \; \Sigma_2^2 = \sigma_2^2 \frac{1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2 r_{12} r_{23} r_{13}}{1 - r_{13}^2} \,,$$

$$(232''') \; . \; . \; . \; \Sigma_3^2 = \sigma_3^2 \frac{1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2 r_{12} r_{23} r_{13}}{1 - r_{12}^2} \,.$$

Эти выраженія имѣютъ большое значеніе, такъ какъ при помощи ихъ можно оцѣнить среднюю величину тѣхъ ошибокъ, кото-

---

[1]) Въ самомъ дѣлѣ, $\Sigma (x_1 - a_{12} x_2 - a_{13} x_3)^2 = N\sigma_1^2 +$

$$+ \frac{(r_{12} - r_{23} r_{13})^2}{(1 - r_{23}^2)^2} \frac{\sigma_1^2}{\sigma_2^2} \cdot N\sigma_2^2 + \frac{(r_{13} - r_{23} r_{12})^2}{(1 - r_{23}^2)^2} \frac{\sigma_1^2}{\sigma_3^2} \cdot N\sigma_3^2 - 2 \frac{r_{12} - r_{23} r_{13}}{1 - r_{23}^2} \frac{\sigma_1}{\sigma_2} \cdot$$

$$\cdot N\sigma_1 \sigma_2 r_{12} - 2 \frac{r_{13} - r_{23} r_{12}}{1 - r_{23}^2} \frac{\sigma_1}{\sigma_3} \cdot N\sigma_1 \sigma_3 r_{13} + 2 \frac{r_{12} - r_{23} r_{13}}{1 - r_{23}^2} \frac{\sigma_1}{\sigma_2} \cdot$$

$$\cdot \frac{r_{13} - r_{23} r_{12}}{1 - r_{23}^2} \frac{\sigma_1}{\sigma_3} \cdot N\sigma_2 \sigma_3 r_{23} = N\sigma_1^2 \left\{ 1 + \frac{1}{(1 - r_{23}^2)^2} \Big[(r_{12} - r_{23} r_{13}) \left[r_{12} - r_{23} r_{13} - \right.\right.$$

$$\left. - 2 r_{12} (1 - r_{23}^2) + (r_{13} - r_{23} r_{12}) r_{23} \right] + (r_{13} - r_{23} r_{12}) \left[r_{13} - r_{23} r_{12} - 2 r_{13} (1 - r_{23}^2) + \right.$$

$$\left.\left. + (r_{12} - r_{23} r_{13}) r_{23} \right] \Big] \right\} = N\sigma_1^2 \left\{ 1 + \frac{1}{(1 - r_{23}^2)^2} \Big[(r_{12} - r_{23} r_{13}) (r_{12} r_{23}^2 - r_{12}) + \right.$$

$$\left. + (r_{13} - r_{23} r_{12}) (r_{13} r_{23}^2 - r_{13}) \Big] \right\} = N\sigma_1^2 \left\{ 1 + \frac{1}{(1 - r_{23}^2)^2} \Big[(r_{12} - r_{23} r_{13}) r_{12} (r_{23}^2 - 1) + \right.$$

$$\left. + (r_{13} - r_{23} r_{12}) \, r_{13} (r_{23}^2 - 1) \Big] \right\} = N\sigma_1^2 \left\{ 1 - \frac{1}{1 - r_{23}^2} \Big[r_{12}^2 + r_{13}^2 - 2 r_{12} r_{23} r_{13} \Big] \right\} =$$

$= N\sigma_1^2 \dfrac{1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2 r_{12} r_{23} r_{13}}{1 - r_{23}^2}$, откуда въ силу (231) получаемъ (232′).

рыя мы сдѣлаемъ, если будемъ пользоваться ур-іями регрессіи (229 и 230) для опредѣленія величины признака въ отдѣльныхъ случаяхъ. Если распредѣленіе слѣдуетъ нормальному закону (Гаусса), то среднія отклоненія во всѣхъ строяхъ будутъ равны, и наши среднія квадратичныя ошибки (232) превратятся въ среднія отклоненія отдѣльныхъ строевъ:

$$\Sigma_1 = \sigma_{1m}\,,\ \Sigma_2 = \sigma_{2m}\,,\ \Sigma_3 = \sigma_{3m}\,.$$

Пользуясь таблицами интеграла вѣроятности, можно будетъ тогда опредѣлить и вѣроятность любого отклоненія отъ величины, полученной изъ ур-ія регрессіи.

Между тремя коэффиціентами корреляціи существуетъ нѣкоторая зависимость. Именно, такъ какъ лѣвая часть каждаго изъ выраженій (232) представляетъ собою сумму квадратовъ ошибокъ и, слѣдовательно, не можетъ быть отрицательной, то общій числитель выраженій, стоящихъ въ правой части долженъ быть положительнымъ. Такимъ образомъ, три коэффиціента корреляціи связаны неравенствомъ:

$$(233)\ .\ .\ .\ .\ 1 - r_{12}^2 - r_{23}^2 - r_{13}^2 + 2r_{12}r_{23}r_{13} > 0.$$

Прибавляя и вычитая $r_{12}^2 r_{13}^2$, чтобы выдѣлить полный квадратъ разности $r_{23} - r_{12}r_{13}$, получимъ:

$$r_{23}^2 - 2r_{23}\,.\,r_{12}r_{13} + r_{12}^2 r_{13}^2 < 1 - r_{12}^2 - r_{13}^2 + r_{12}^2 r_{13}^2, \text{ или}$$

$$(r_{23} - r_{12}r_{13})^2 < 1 - r_{12}^2 - r_{13}^2 + r_{12}^2 r_{13}^2,$$

откуда находимъ предѣлы, въ которыхъ при данныхъ r_{12} и r_{13} можетъ заключаться r_{23}:

$$(234)\ .\ .\ r_{12}r_{13} - \sqrt{1 - r_{12}^2 - r_{13}^2 + r_{12}^2 r_{13}^2} < r_{23} < r_{12}r_{13} + \\ + \sqrt{1 - r_{12}^2 - r_{13}^2 + r_{12}^2 r_{13}^2}.$$

Это неравенство дастъ намъ рядъ небезинтересныхъ указаній, если мы при помощи его вычислимъ предѣлы для r_{23} при нѣкоторыхъ значеніяхъ r_{12} и r_{13}:

Значенія r_{12} и r_{13}:	Соотвѣтствующія значенія r_{23}:
$r_{12}=r_{13}=0$,	$-1 \leqslant r_{23} \leqslant +1$, [1]
$r_{12}=r_{13}=\pm 1$,	$r_{23}=+1$,
$r_{12}=+1, \quad r_{13}=-1$,	$r_{23}=-1$,
$r_{12}=0; \quad r_{13}=\pm 1$,	$r_{23}=0$,
$r_{12}=0, \quad r_{13}=\pm r$,	$-\sqrt{1-r^2} < r_{23} < +\sqrt{1-r^2}$,
$r_{12}=r_{13}=\pm r$,	$2r^2-1 < r_{23} < +1$,
$r_{12}=+r, \quad r_{13}=-r$,	$-1 < r_{23} < 1-2r^2$, [2]
$r_{12}=r_{13}=\pm\sqrt{0{,}5}=\pm 0{,}707$,	$0 < r_{23} < 1$,
$r_{12}=+\sqrt{0{,}5}, \quad r_{13}=-\sqrt{0{,}5}$.	$-1 < r_{23} < 0$.

Можно было бы думать, что, если x_1 находится въ положительной корреляціи съ x_2 и съ x_3, то и между x_2 и x_3 должна быть положительная корреляціонная связь. Однако это не необходимо. Напр., если $r_{12}=+{}^1/_4$ и $r_{13}=+{}^1/_4$, то r_{23} можетъ имѣть любыя значенія между $+1$ и $2\left(\frac{1}{4}\right)^2-1$, т. е. между $+1$ и $-0{,}875$; если $r_{12}=+\frac{7}{10}$, а $r_{13}=-\frac{7}{10}$, то r_{23} можетъ имѣть значенія отъ -1 до $+\frac{1}{50}$. Только когда r_{12}, равное численно r_{13}, больше $\sqrt{0{,}5}$, r_{23} всегда должно имѣть опредѣленный знакъ, будучи положительнымъ, когда r_{12} и r_{13} имѣютъ одинаковые знаки, и отрицательнымъ, когда — разные.

Всегда-ли увеличивается точность опредѣленія величины изъ ур-ія регрессіи, когда мы переходимъ отъ корреляціи двухъ величинъ къ корреляціи трехъ?

Средняя ошибка опредѣленія x_1 въ первомъ случаѣ равна, согласно формулѣ (146), величинѣ

$$\sigma_1\sqrt{1-r_{12}^2},$$

во второмъ-же (фор. 232) она равняется:

---

[1]) У G. U. Jule'я (op. cit. p. 486) ошибочно (вѣроятно опечатка!) указанъ 0.

[2]) Тамъ-же вмѣсто правильной величины $1-2r^2$ данъ невѣрный предѣлъ $2r^2-1$.

$$\sigma_1 \sqrt{\frac{1-r_{12}^2-r_{23}^2-r_{13}^2-2r_{12}r_{23}r_{13}}{1-r_{23}^2}}.$$

Условіемъ увеличенія точности опредѣленія служитъ, очевидно, неравенство:

$$\frac{1-r_{12}^2-r_{23}^2-r_{13}^2+2r_{12}r_{23}r_{13}}{1-r_{23}^2}<1-r_{12}^2,$$

которое можно представить въ видѣ:

$$\frac{r_{12}^2+r_{13}^2-2r_{12}r_{23}r_{13}}{1-r_{23}^2}>r_{12}^2.$$

Знаменатель лѣвой части всегда >0, слѣдовательно, при умноженіи обѣихъ частей на $1-r_{23}^2$ знакъ неравенства сохраняется, и мы получаемъ:

$$r_{12}^2+r_{13}^2-2r_{12}r_{23}r_{13}>r_{12}^2-r_{12}^2r_{23}^2,$$

откуда:

$$r_{13}^2+r_{12}^2r_{23}^2-2r_{12}r_{23}r_{13}>0, \quad \text{т. е.}$$

(235) . . . . $(r_{13}-r_{12}r_{23})^2>0$.

Выраженіе въ лѣвой части неравенства не можетъ быть отрицательнымъ, но можетъ быть нулемъ. Такимъ образомъ, точность опредѣленія отъ прибавленія новаго перемѣннаго уменьшиться не можетъ, но можетъ остаться безъ перемѣны. Это будетъ въ томъ случаѣ, когда соотвѣтствующій коэффиціентъ регрессіи, числителемъ котораго является $r_{13}-r_{12}r_{23}$ (см. форм. 229 и 230), окажется равнымъ нулю.

Напр., пусть $r_{12}=+0{,}8$, $r_{23}=+0{,}5$, $r_{13}=+0{,}4$.

Опредѣляя x_1 по одному x_2-му, будемъ имѣть ур-іе:

$$x_1=\bar{x}_1+\frac{\sigma_1}{\sigma_2}\cdot 0{,}8\,(x_2-\bar{x}_2)$$

со средней ошибкой

$$\Sigma_1=\sigma_1\sqrt{1-0{,}8^2}=0{,}6\,\sigma_1.$$

Можно-бы думать, что прибавленіе третьяго перемѣннаго, связаннаго съ первыми двумя довольно значительной корреляціонной связью, увеличитъ точность опредѣленія. Однако, вычисляя коэффиціенты регрессіи, находимъ:

$$\rho_{1(2)} = \frac{r_{12} - r_{23} r_{13}}{1 - r_{23}^2} \frac{\sigma_1}{\sigma_2} = \frac{+0,8 - 0,4\,.\,0,5}{1 - 0,5^2} \frac{\sigma_1}{\sigma_2} = +0,8 \frac{\sigma_1}{\sigma_2}$$

по прежнему, а

$$\rho_{1(3)} = \frac{r_{13} - r_{23} r_{12}}{1 - r_{23}^2} \frac{\sigma_1}{\sigma_3} = \frac{+0,4 - 0,8\,.\,0,5}{1 - 0,5^2} \frac{\sigma_1}{\sigma_3} = 0\,.$$

Средняя ошибка должна по доказанному остаться безъ перемѣны. Въ самомъ дѣлѣ:

$$\Sigma_1 = \sigma_1 \sqrt{\frac{1 - 0,8^2 - 0,5^2 - 0,4^2 + 2\,.\,0,8\,.\,0,5\,.\,0,4}{1 - 0,5^2}} = 0,6\,\sigma_1\,.$$

Такимъ образомъ, корреляитивная связь первыхъ двухъ величинъ съ третьей (x_3) не оказываетъ никакого вліянія на наши формулы и не увеличиваетъ достовѣрности и точности нашихъ заключеній.

Другой крайній случай представляетъ переходъ корреляитивной связи въ строгую функціональную зависимость. Такой переходъ будетъ, очевидно, имѣть мѣсто въ томъ только случаѣ, когда каждой парѣ значеній двухъ перемѣнныхъ x_2 и x_3 будетъ соотвѣтствовать *только одно* значеніе x_1, т. е., когда *средняя квадратичная ошибка* опредѣленія x_1-ого по x_2-му и x_3-му обратится въ ноль.

Чтобы яснѣе представить себѣ этотъ случай, мы упростимъ его, предположивъ, что корреляція первой величины со второй равна корреляціи ея съ третьей, т. е., что

$$r_{12} = r_{13} = r, \text{ и положимъ } \rho = r_{23}.$$

Тогда средняя квадратичная ошибка (Σ_1) будетъ въ точности равна нулю, если

$$\frac{1 - r_{23}^2 - r_{12}^2 - r_{13}^2 + 2 r_{12} r_{23} r_{13}}{1 - r_{23}^2} = 0, \text{ т. е. если}$$

$$1 - \frac{2r^2 - 2r^2\rho}{1 - \rho^2} = 0\,;$$

откуда находимъ:

$$1 - \frac{2r^2}{1 + \rho} = 0\,, \text{ и, наконецъ:}$$

$$r = \sqrt{\frac{1}{2}(1 + \rho)}\,.$$

Напр., если $\rho = 0{,}7572$, то при $r = \sqrt{0{,}8786} = 0{,}9373$ корреляція превратится въ строгую функціональную линейную зависимость.

§ 25. Иллюстраціи.

Величина 0,7572 представляетъ собою коэффиціентъ корреляціи между барометрическими высотами въ Southampton'ѣ (южный берегъ Англіи) и въ Laudale (западный берегъ Шотландіи). Разстояніе между этими станціями равно 444 милямъ, и почти по серединѣ его расположена метеорологическая станція Stonyhurst. Коэффиціенты корреляціи между Stonyhurst'омъ и Laudale'емъ и Stonyhurst'омъ и Southampton'омъ по приблизительному разсчету Пирсона должны быть почти равны другъ другу и оба близки къ величинѣ 0,95 — 0,94. Удаляясь отъ линіи, соединяющей Laudale и Southampton т. обр., чтобы коэффиціенты корреляціи между точкой нашего пути и обѣими станціями оставались одинаковыми, мы дойдемъ до такого пункта, гдѣ эти коэффиціенты уменьшатся до найденной выше величины (0,9373). Метеорологическая станція, построенная въ этомъ мѣстѣ, (Пирсонъ предполагаетъ, что это будетъ гдѣ-то недалеко отъ Whitby) будетъ обладать тѣмъ замѣчательнымъ свойствомъ, что ея барометрическія высоты будутъ точной линейной функціей барометрическихъ высотъ въ Southampton'ѣ и Laudale'ѣ [1]). Разработка метеорологическихъ данныхъ по методу корреляціи могла бы усовершенствовать дѣло предсказанія погоды. Нужно было бы только произвести большую вычислительную работу опредѣленія коэффиціентовъ корреляціи между факторами погоды въ однихъ пунктахъ въ предшествующій моментъ времени и факторами погоды въ другихъ пунктахъ въ послѣдующіе моменты. При надлежащемъ выборѣ пунктовъ и промежутковъ времени могли бы получиться, вѣроятно, интересные результаты.

Работа К. Пирсона и миссъ Ли имѣла своей цѣлью только обратить вниманіе спеціалистовъ въ области метеорологіи на новый методъ. Нахожденіе десятковъ и сотенъ коэффиціентовъ корреляціи по ежедневнымъ метеорологическимъ датамъ за рядъ

---

[1]) Prof. Karl Pearson and Miss A. Lee, On the Distribution of Frequency of the Barometric Height at Divers Stations, Phil. Trans., Vol. 190 A, 1898, p. 458—459.

лѣтъ—работа доступная только цѣлому учрежденію, а не отдѣльнымъ лицамъ. Чтобы показать однако, какъ близко при подходящемъ выборѣ станцій могутъ совпадать вычисленныя величины съ дѣйствительными, Пирсонъ употребилъ упрощенный методъ. Ур-іе регрессіи линейно. Допустивъ линейную зависимость между барометрической высотой въ Stonyhurst'ѣ съ одной стороны и Southampton'ѣ и Laudale'ѣ съ другой, Пирсонъ положилъ:

$$H_{\mathrm{St.}} = xH_{\mathrm{South.}} + yH_{\mathrm{Laud}} + z\,.$$

Чтобы найти постоянныя этого ур-ія: x, y и z, были выбраны 12 наблюденій, относящихся къ 15-му дню каждаго мѣсяца, и обработаны по методу наименьшихъ квадратовъ. Полученныя числовыя значенія, будучи подставлены въ вышеприведенное ур-іе, дали возможность вычислить по $H_{\mathrm{South.}}$ и $H_{\mathrm{Laud.}}$ барометрическія высоты въ Stonyhurst'ѣ. Вычисленіе на дѣлѣ было произведено для 50 моментовъ, взятыхъ черезъ равные 2-хъ недѣльные промежутки, со слѣдующимъ результатомъ (барометрическія высоты даны въ дюймахъ):

Высота барометра въ Stonyhurst'ѣ: [1]		
Наблюденная.	Вычисленная.	Разность.
30,58	30,61	+0,03
30,17	30,14	—0,03
30,03	29,98	—0,05
30,06	30,11	+0,05
29,92	29,94	+0,02
30,35	30,38	+0,03
29,85	29,84	—0,01
30,12	30,11	—0,01
30,13	30,13	0,00
30,19	30,20	+0,01
30,51	30,48	—0,03

[1]) op. cit. p. 461.

Высота барометра въ Stonyhurst'ѣ:		
Наблюденная.	Вычисленная.	Разность.
30,65	30,62	— 0,03
30,17	30,21	+ 0,04
29,52	29,55	+ 0,03
30,51	30,50	— 0,01
29,91	29,90	— 0,01
29,94	29,94	0,00
30,17	30,18	+ 0,01
30,13	30,10	— 0,03
30,34	30,33	— 0,01
30,74	30,73	— 0,01
30,18	30,25	+ 0,07
30,16	30,18	+ 0,02
30,41	30,38	— 0,03
30,01	30,00	— 0,01
29,15	29,18	+ 0,03
30,28	30,26	— 0,02
29,79	29,85	+ 0,06
29,93	29,90	— 0,03
29,91	29,93	+ 0,02
30,11	30,11	0,00
29,99	29,92	— 0,07
29,43	29,45	+ 0,02
30,15	30,15	0,00
30,22	30,22	0,00
30,16	30,13	— 0,03
30,27	30,29	+ 0,02

Высота барометра въ Stonyhurst'ѣ:		
Наблюденная.	Вычисленная.	Разность.
29,56	29,61	+ 0,05
29,66	29,85	+ 0,19
30.23	30,25	+ 0,02
29,02	28,99	— 0,03
30,35	30,30	— 0,05
29,92	29,92	0,00
29,99	29,95	— 0,04
29,95	29,96	+ 0,01
29,58	29,62	+ 0,04
30,11	30,11	+ 0,00
30,18	30,22	+ 0,06
29,92	29,89	— 0,03
29,49	29,46	— 0,03

Мы видимъ, такимъ образомъ, что разности очень малы, распредѣлены почти поровну по обѣ стороны отъ нуля, и средняя величина ихъ достигаетъ приблизительно лишь $^1/_{40}$ дюйма, т. е. 0,635 mm. Въ 38 случаяхъ ошибка меньше 1 mm., въ 11 больше 1 mm. и меньше 2-хъ mm. и только въ одномъ случаѣ ошибка больше 2-хъ mm. Пирсонъ имѣетъ основаніе разсматривать эти цифры, какъ подтвержденіе своего взгляда, что не особенно далеко отъ Stonyhurst'а должна найтись въ Ланкаширѣ станція, для которой корреляціонная зависимость превращается въ однозначную функціональную.

Мы можемъ попытаться оцѣнить достигнутую степень приближенія еще и другимъ путемъ. Именно, мы можемъ найти величину фиктивнаго коэффиціента корреляціи, который долженъ былъ бы имѣть мѣсто, чтобы барометрическія высоты въ Stonyhurst'ѣ можно было опредѣлять съ такой-же точностью не по двумъ, какъ это сдѣлалъ Пирсонъ, а по одной величинѣ.

По даннымъ вышеприведенной таблицы можно вычислить среднюю квадратичную ошибку. Она равна 0,041″. Среднее отклоненіе барометрическихъ высотъ въ Stonyhurst'ѣ равняется 0,3503 (l. c., p. 435). Между этими величинами должна существовать зависимость:

$$0{,}041 = 0{,}3503\sqrt{1-r^2},$$

откуда находимъ: $r = 0{,}993$, что представляетъ уже значительное приближеніе къ строгой функціональной зависимости. Не только въ области явленій наслѣдственности, гдѣ нормальной величиной является $r = 0{,}5$, но даже въ области органической корреляціи мы не встрѣчаемъ такой тѣсной связи. Такъ, изъ органовъ человѣческаго тѣла наибольшая корреляція существуетъ между размѣрами длинныхъ костей. Напр., для мужчинъ коэффиціентъ корреляціи между tibia и humerus наибольшій и равенъ 0,86, а для женщинъ замѣчена болѣе тѣсная связь между размѣрами tibia и femur, именно, $r = 0{,}89$ [1]). Для различныхъ, напр., судебно-медицинскихъ, полицейскихъ, научныхъ цѣлей можно опредѣлять ростъ и размѣры органовъ человѣческаго тѣла по величинѣ одного или двухъ органовъ съ значительнымъ успѣхомъ, но, какъ читатель видитъ, мы здѣсь далеки еще отъ точности достигнутой Пирсономъ въ предсказаніи барометрическихъ высотъ.

Послѣдній примѣръ возьмемъ изъ области русскихъ хлѣбныхъ цѣнъ. Коэффиціентъ корреляціи между средней мѣсячной цѣной ржи въ Самарѣ за текущій и тамъ-же за предыдущій мѣсяцъ $r_{C,\,C-1} = +0{,}93292$, коэффиціентъ корреляціи между средними цѣнами ржи въ Ельцѣ и въ Самарѣ за одинъ и тотъ-же мѣсяцъ $r_{E,\,C} = 0{,}87796$, и, наконецъ, коэффиціентъ корреляціи между средней мѣсячной въ Ельцѣ и средней за предыдущій мѣсяцъ въ Самарѣ $r_{E,\,C-1} = 0{,}85593$. Среднія ариѳметическія и среднія отклоненія приведены въ приложеніи. По этимъ даннымъ ур-іе регрессіи для средней мѣсячной самарской цѣны будетъ:

$$x_C = 47{,}04 \text{ коп.} + 0{,}383(x_E - 52{,}64) \text{ коп.} + 0{,}674(x_{C-1} - 47{,}16) \text{ коп.}$$

Изъ этого ур-ія были вычислены среднія мѣсячныя цѣны въ Самарѣ за 36 мѣс. 1894—96 гг., которыя и приводятся въ слѣдующей таблицѣ [2]):

[1]) K. Pearson, On the Reconstruction of the Stature of Prehistoric Races, Phil. Trans, Vol. 192, p 181.

[2]) Цыфры этой таблицы вычислялись при помощи логариѳмической линейки, т что за послѣднюю цыфру ручаться нельзя. Возможная неточность однако такова, что не можетъ вліять на дѣлаемые ниже выводы.

Дѣйств. цѣна.	Вычисл. цѣна.	Разность.	Дѣйств. цѣна.	Вычисл. цѣна.	Разность.
43,3 коп.	43,5 коп.	+ 0,2 коп.	28,7 коп.	31,5 коп.	+ 2,8 коп.
44,4 „	44,1 „	— 0,3 „	32,3 „	29,5 „	— 2,8 „
44,6 „	44,7 „	+ 0,1 „	30,0 „	30,8 „	+ 0,8 „
45,1 „	44,1 „	— 1,0 „	26,8 „	28,9 „	+ 2,1 „
36,6 „	41,0 „	+ 4,4 „	28,6 „	26,3 „	— 2,3 „
35,5 „	35,8 „	+ 0,3 „	28,8 “	27,2 „	— 1,6 „
31,3 „	33,6 „	+ 2,3 „	29,6 „	28,4 „	— 1,2 „
30,4 „	30,8 „	+ 0,4 „	31,2 „	29,4 „	— 1,8 „
34,8 „	30,8 „	— 4,0 „	31,2 „	30,1 „	— 1,1 „
26,8 „	33,6 „	+ 6,8 „	29,0 „	29,9 „	+ 0,9 „
30,0 „	28,4 „	— 1,6 „	27,3 „	27,1 „	— 0,2 „
27,2 „	30,6 „	+ 3,4 „	23,0 „	25,3 „	+ 2,3 „
28,9 „	28,5 „	— 0,4 „	24,4 „	22,9 „	— 1,5 „
29,6 „	30,1 „	+ 0,5 „	24,7 „	25,8 „	+ 1,1 „
30,1 „	30,8 „	+ 0,7 „	29,2 „	26,7 „	— 2,5 „
35,1 „	31,3 „	— 3,8 „	26,9 „	30,9 „	+ 4,0 „
35,8 „	35,1 „	— 0,7 „	28,0 „	30,0 „	+ 2,0 „
31,7 „	35,0 „	+ 3,3 „	28,0 „	30,3 „	+ 2,3 „

Какъ читатель видитъ, вычисленныя величины близки къ фактическимъ, уклоняясь отъ нихъ въ среднемъ всего на 1,9 коп. Это даже меньше, чѣмъ мы могли бы ожидать, слѣдуя указаніямъ теоріи, ибо средняя квадратичная ошибка нашихъ предсказаній теоретически равна $\sigma_C \sqrt{\frac{1 - r_C^2 - r_{C-1}^2 - r_E^2 + 2 r_C r_{C-1} r_E}{1 - r_{C-1,E}^2}} = 4{,}5$ к., а на дѣлѣ для 3 лѣтъ (1894—95—96) оказывается всего 2,4 коп. Это заставляетъ предполагать, во (1), что другіе годы должны дать бо́льшую ошибку, во (2), что распредѣленіе цѣнъ не подчиняется закону Гаусса (см. ниже § 30). Во всякомъ случаѣ было бы благодарной задачей изучить въ большомъ масштабѣ колебанія цѣнъ и корреляцію между ними въ различныхъ пунктахъ и за различные періоды. Теоретически, конечно, очень трудно учесть всѣ факторы образованія цѣнъ, принимаемые практиками въ разсчетъ при заключеніи срочныхъ сдѣлокъ, хотя чрезвычайная точность предвидѣнія, обнаруживаемая ими, и не позволяетъ считать подобную задачу невозможной. Можно думать однако, что задача предсказанія цѣнъ допускаетъ сильное упрощеніе, если воспользоваться въ качествѣ чувствительныхъ инструментовъ самими практиками и изслѣдовать корреляціонную зависимость между цѣнами, положенными въ основаніе срочныхъ сдѣлокъ, и дѣйствительно къ соотвѣтствующему сроку наступившими цѣнами. Этимъ путемъ можно было бы вѣроятно значительно подняться надъ уже достигнутымъ уровнемъ точности эмпирическихъ предсказаній.

§ 26. *Частные коэффиціенты корреляціи.*

Выше было уже указано, что корреляціонная связь не тождественна съ причинной зависимостью, хотя и можетъ давать цѣнныя указанія относительно наличности послѣдней. Ур-ія регрессіи, связывающія нѣсколько величинъ, позволяютъ произвести болѣе глубокій анализъ, давая возможность отдѣлить другъ отъ друга вліянія различныхъ факторовъ.

Пусть B и C будутъ явленія, которыя по нашему предположенію,—до основаній котораго намъ сейчасъ нѣтъ дѣла, являются частичными причинами явленія A.

Чтобы убѣдиться въ этомъ, мы должны были бы изслѣдовать, какъ измѣняется A съ измѣненіемъ одного B при неизмѣнномъ состояніи всѣхъ прочихъ возможныхъ причинъ, и какъ измѣняется A съ измѣненіемъ C при аналогичномъ условіи. Неизмѣнность *всѣхъ* возможныхъ факторовъ кромѣ одного есть однако требованіе, котораго мы осуществить не можемъ. Все, что въ нашихъ силахъ,—это изслѣдовать измѣненіе A при неизмѣнности *нѣкоторыхъ* опредѣленныхъ факторовъ, вліяніе которыхъ мы хотѣли бы исключить, и во многихъ случаяхъ этого одного уже достаточно для того, чтобы получить возможность сдѣлать иной разъ очень цѣнныя заключенія. Въ самомъ дѣлѣ, въ чемъ заключается основная невыгода положенія наблюдателя соціальныхъ явленій, какъ не въ томъ, что онъ только наблюдатель, что онъ не можетъ экспериментировать, по своему произволу дѣлая постоянной ту или иную группу условій? А теорія корреляціи и даетъ ему эту возможность.

Разсмотримъ съ этой точки зрѣнія ур-іе регрессіи:

(236) . . $X_A - h_A = \rho_{A(B)} \cdot (x_B - h_B) + \rho_{A(C)} \cdot (x_C - h_C)$.

h_A есть средняя величина для x_A при *всевозможныхъ* значеніяхъ B и C. X_A есть среднее изъ значеній, которыя принимаетъ x_A при *опредѣленныхъ* значеніяхъ B и C. Чтобы изслѣдовать измѣненіе A въ зависимости отъ одного только B при неизмѣнномъ C, нужно въ ур-іи (236) принять x_C за постоянную величину. Полагая при этомъ для сокращенія:

$$h_A + \rho_{A(C)} \cdot (x_C - h_C) = h'_A = \text{Const.},$$

мы будемъ имѣть:

$$X_A - h'_A = \rho_{A(B)} \cdot (x_B - h_B).$$

Въ этомъ выраженіи h'_A представляетъ собою среднюю ариѳметическую всѣхъ значеній, которыя принимаетъ x_A при данномъ значеніи x_C и при всѣхъ какихъ угодно значеніяхъ x_B. Мы видимъ, слѣдовательно, что при неизмѣнности фактора C отклоненія x_A отъ своего средняго пропорціональны отклоненіямъ x_B отъ своего средняго при коэффиціентѣ пропорціональности $\rho_{A(B)}$.

Совершенно аналогично, при неизмѣнности явленія B, отклоненія x_A отъ своей средней величины пропорціональны отклоненіямъ x_C при коэффиціентѣ пропорціональности $\rho_{A(C)}$, и мы имѣемъ:

$$X_A - h''_A = \rho_{A(C)} (x_C - h_C).$$

Такимъ образомъ, коэффиціенты регрессіи $\rho_{A(B)}$ и $\rho_{A(C)}$ служатъ мѣрой, указывающей зависимость между явленіемъ A и явленіями B и C *въ отдѣльности*.

Въ качествѣ иллюстраціи разсмотримъ слѣдующій примѣръ. Коэффиціентъ корреляціи между средними мѣсячными цѣнами ржи въ Москвѣ и въ Самарѣ довольно высокій, именно, $r_{M,C} = 0{,}77$. Если мы захотимъ узнать, дѣйствительно-ли московская цѣна находится въ такой тѣсной причинной связи съ одновременной самарской, мы должны будемъ изслѣдовать, какъ измѣняется цѣна въ Москвѣ при неизмѣнности *всѣхъ* прочихъ факторовъ кромѣ самарской цѣны. Исключить *всѣ* факторы невозможно, и мы удовлетворимся сейчасъ исключеніемъ лишь одного изъ нихъ, именно, средней цѣны ржи въ Самарѣ-же, но за предыдущій мѣсяцъ. Индексъ самарской цѣны за предыдущій мѣсяцъ пусть будетъ, какъ раньше, „$C-1$“. Тогда:

$$r_{M,C} = +0{,}77;\ r_{M,C-1} = +0{,}79;\ r_{C,C-1} = +0{,}93.$$

Подставляя эти величины въ формулу двойной регрессіи (230) (величины среднихъ отклоненій см. въ приложеніи), получимъ:

$$X_M - h_M = 0{,}49(x_{C-1} - h_{C-1}) + 0{,}23(x_C - h_C).$$

Разсматривая это ур-іе, мы видимъ, что на московскую цѣну оказываетъ болѣе сильное вліяніе не одновременная цѣна въ Самарѣ, а цѣна въ Самарѣ за предыдущій мѣсяцъ. Если бы мы взяли случаи, когда цѣна за предыдущій мѣсяцъ въ Самарѣ была одинаковой, то превышеніе одновременной самарской цѣны надъ ея среднимъ на 10 коп. связано было бы со среднимъ повышеніемъ московской цѣны всего на 2,3 коп., въ то время какъ при неизмѣнныхъ одновременныхъ цѣнахъ, превышеніе самарской цѣны предыдущаго мѣсяца на 10 коп. надъ ея средней связано было-бы со среднимъ повышеніемъ московской цѣны на 4,9 коп.

Т. к. мы не исключили прочихъ вліяній, то врядъ-ли безъ ближайшаго изслѣдованія возможно говорить здѣсь о причинной зависимости. Осторожнѣе будетъ сказать: средняя цѣна за предыдущій мѣсяцъ въ Самарѣ служитъ показателемъ вліяній на московскую цѣну болѣе сильныхъ, чѣмъ средняя цѣна въ томъ-же городѣ за одновременный мѣсяцъ.

Такимъ образомъ, коэффиціентъ корреляціи $r_{M,C}$ оказывается сравнительно высокимъ только потому, что и московская и самарская цѣны, сравнительно слабо вліяя другъ на друга непосредственно, находятся подъ сильнымъ одновременнымъ вліяніемъ другихъ факторовъ. Если бы мы захотѣли опредѣлить степень зависимости только между московской и самарской цѣнами, то намъ нужно было бы исключить вліяніе этихъ другихъ факторовъ, изслѣдовавъ случаи, когда они оказываются постоянными. Исключая для примѣра вліяніе цѣны въ Самарѣ за предыдущій мѣсяцъ, мы будемъ разсуждать такъ:

Ур-іе регрессіи для Москвы относительно C и $C-1$:

$$X_M - h_M = 0{,}49(x_{C-1} - h_{C-1}) + 0{,}23(x_C - h_C),$$

а ур-іе регрессіи для Самары относительно M и $C-1$ (найденное аналогичнымъ способомъ):

$$X_C - h_C = 0{,}86\,(x_{C-1} - h_{C-1}) + 0{,}10\,(x_M - h_M).$$

При *неизмѣнной* цѣнѣ въ Самарѣ за предыдущій мѣсяцъ, коэффиціентъ регрессіи Москвы относительно Самары ${}_{C-1}\rho_{M(C)} =$ $= 0{,}23$, а коэффиціентъ регрессіи Самары относительно Москвы при томъ-же условіи ${}_{c-1}\rho_{c(M)} = 0{,}10$. Коэффиціентъ корреляціи между ними при $x_{c-1} = \mathrm{const}$. называется частнымъ коэффиціентомъ корреляціи и равняется

$$ {}_{c-1}r_{MC} = \sqrt{{}_{c-1}\rho_{M(C)} \cdot {}_{c-1}\rho_{C(M)}} = \sqrt{0{,}23 \,.\, 0{,}10} = 0{,}15. $$

Общая формула для частнаго коэффиціента корреляціи будетъ, слѣдовательно:

$$ (237) \quad . \quad . \quad . \quad . \quad {}_3r_{12} = \frac{r_{12} - r_{13} r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}, $$

а остальные легко написать по соображеніямъ симметріи.

Вѣроятная ошибка частнаго коэффиціента корреляціи выражается той-же формулой, что и полнаго, т. е.,

$$ (238) \quad . \quad . \quad . \quad . \quad . \quad E_{{}_3r_{12}} = 0{,}67449 \, \frac{1 - {}_3r_{12}^2}{\sqrt{N}} \,. $$ [1])

Второй примѣръ я возьму изъ чрезвычайно интересной и богатой содержаніемъ работы: R. H. Hooker, Correlation of the Weather and Crope [2]).

Ближайшей цѣлью ея было выяснить вопросъ, метеорологическія условія какого періода оказываютъ наибольшее вліяніе на послѣдующій урожай. За періодъ былъ принятъ промежутокъ времени въ 8 недѣль, причемъ періоды взяты были отчасти налегающіе другъ на друга: 9—16 недѣля года, 13—20-я, 17—24-я и т. д. Для рѣшенія поставленной задачи нужно было найти коэффиціенты корреляціи между метеорологическими условіями каждаго періода и послѣдующимъ урожаемъ. Важнѣйшія метеорологическіе факторы урожая—дождь и теплота; поэтому Гукеръ вычислялъ коэффиціенты корреляціи, съ одной стороны, между размѣромъ урожая и количествомъ влаги, выпавшей въ теченіе каждаго 8-ми недѣльнаго періода, а съ другой стороны, между размѣромъ урожая и т. наз. „накопленной" $\frac{\text{свыше}}{\text{ниже}}$ 42° F. температурой. Оказалось (я разсматриваю здѣсь только урожай пшеницы), что изъ всѣхъ періодовъ очень замѣтно выдѣляется одинъ—37—44 недѣли, т. е. періодъ посѣва и ближайшихъ послѣ посѣва недѣль, т. к. избытокъ дождя въ это время оказываетъ очень замѣтное

---

1) U. Yule, Proc. of the Roy. Soc. Vol. 79, p. 182 и сл,; болѣе элементарный выводъ: D. Heron, Biometrica, Vol. VII, p. 411—412.

2) Journ. of the Roy. Stat. Soc., 1907, p. 1—42.

отрицательное вліяніе на послѣдующій урожай. Эти 8 недѣль дали наибольшій коэффиціентъ корреляціи между количествомъ влаги и урожаемъ пшеницы, именно, r_{wr} [1]) = — 0,66. Два смежныхъ періода 33 — 40 и 41 — 48 недѣли дали также значительныя величины—0,55 и —0,47. Изъ болѣе позднихъ восьми-недѣльныхъ періодовъ только 1 — 8 недѣли дали r_{wr} равное — 0,47, а остальные коэффиціенты корреляціи оказались значительно меньше.

Если мы остановимъ наше вниманіе на этихъ двухъ главныхъ періодахъ, то замѣтимъ [2]), что коэффиціентъ корреляціи между „накопленной" температурой и урожаемъ пшеницы r_{wa} будетъ по Гукеру: для 37 — 44-ой нед. + 0,36, для 1—8-ой + 0,52. Хотя вѣроятныя ошибки коэффиціентовъ корреляціи у Гукера и велики (онъ разсматривалъ періодъ въ 21 годъ всего), но + 0,36 представляетъ изъ себя все таки величину, которая, если не съ достовѣрностью, то все-же съ извѣстной долей вѣроятности, позволяетъ казалось бы предположить, что *кромѣ* дождя на урожай будущаго года оказываетъ вліяніе и „накопленная" къ періоду посѣва температура. Однако это предположеніе было-бы неосновательно. Положительный коэффиціентъ r_{wa} = + 0,36 говоритъ намъ только, что между температурными условіями періода посѣва и будущимъ урожаемъ есть коррелятивная связь, но этотъ коэффиціентъ не уполномочиваетъ насъ на заключеніе, что эта связь будетъ существовать при постоянствѣ другихъ факторовъ; а вѣдь эту-то мысль мы и выразили выше, подчеркнувъ слово „кромѣ". Вѣдь можетъ быть, что температурныя условія служатъ только показателемъ другихъ, напр., большей или меньшей дождливости осени, а при одинаковой дождливости этого періода не оказываютъ на величину будущаго урожая никакого вліянія. Такимъ образомъ, мы приходимъ къ необходимости раздѣлить вліянія на урожай обоихъ факторовъ и вычислить частные коэффиціенты корреляціи, съ одной стороны, для урожая и дождя при постоянной температурѣ ($_{a}r_{wr}$), съ другой стороны, для урожая и температуры при постоянномъ количествѣ осадковъ ($_{r}r_{wa}$).

---

[1]) Гукеръ обозначаетъ коэффиціенты корреляціи, пользуясь подписными значками: w (англ. wheat)—пшеница, r (англ. rain)—дождь, и a (англ. accumulated temperature)—накопленная температура.

[2]) l. cit., p. 30.

По формулѣ (237) имѣемъ (для 37—44 недѣли) [1]):

$$ {}_a r_{wr} = \frac{r_{wr} - r_{wa}\, r_{ar}}{\sqrt{(1-r_{wa}^2)(1-r_{ar}^2)}} = $$

$$ = \frac{-0{,}66 - (+0{,}36)(-0{,}54)}{\sqrt{(1-0{,}36^2)(1-0{,}54^2)}} = -0{,}59\,; $$

$$ {}_r r_{wa} = \frac{r_{wa} - r_{wr}\, r_{ar}}{\sqrt{(1-r_{wr}^2)(1-r_{ar}^2)}} = $$

$$ = \frac{+0{,}36 - (-0{,}66)(-0{,}54)}{\sqrt{(1-0{,}66^2)(1-0{,}54^2)}} = +0{,}006 \; ^{2)}. $$

Такимъ образомъ оказывается, что температурныя условія въ періодъ посѣва сами по себѣ не имѣютъ значенія, а важно лишь количество осадковъ.

Для періода 1 — 8 недѣли (янв., февр.) имѣютъ значеніе оба фактора, ибо

$$ {}_a r_{wr} = -0{,}55, \qquad {}_r r_{wa} = +0{,}49. $$

§ 27. Общій случай. Корреляція n перемѣнныхъ [3]).

Ур-іе линейной регрессіи въ общемъ случаѣ n перемѣнныхъ можно написать слѣдующимъ образомъ:

$$ (239) \; . \; . \; . \; . \; X_1 = a_{11} + a_{12}x_2 + a_{13}x_3 + \; . \; . \; . + a_{1n}x_n \, . $$

Здѣсь X_1 есть вѣроятное среднее значеніе всѣхъ x_1 при данныхъ значеніяхъ $x_2, x_3, \ldots x_n$; причемъ всѣ эти величины

---

[1]) Коэффиціенты корреляціи между дождемъ и температурой см. l. cit. табл IV, p. 42.

[2]) Коэффиціенты корреляціи вычислялись Гукеромъ первоначально съ 3-мя десятичными знаками, и эти ихъ величины служили для нахожденія частныхъ коэффиціентовъ корреляціи. Въ напечатанныхъ таблицахъ полные коэффиціенты корреляціи даны только съ двумя знаками, вслѣдствіе чего и происходитъ разница между вычисленными выше значеніями коэффиціентовъ ${}_a r_{wr}$ и ${}_r r_{wa}$ и значеніями ихъ, даныными у Гукера. Послѣднія, болѣе точныя, равны соотвѣтственно −0,62 и —0,00 (l cit., p 30).

[3]) Этотъ отдѣлъ требуетъ отъ читателя знакомства съ теоріей детерминантовъ Элементарныя свѣдѣнія, которыя можно найти, напр., въ извѣстной книгѣ Лоренца, Элементы высшей математики, т. II стр. 401—414, почти достаточны.

измѣряются величинами отклоненій отъ соотвѣтствующихъ среднихъ ариѳметическихъ для всей совокупности. Эти среднія мы будемъ обозначать, какъ и раньше, черезъ $h_1, h_2, \ldots h_n$. Для каждаго изъ остальныхъ перемѣнныхъ можно написать ур-іе регрессіи аналогичное 239-му, но ихъ нѣтъ нужды изслѣдовать, т. к. соотвѣтствующія формулы легко получаются одна изъ другой по соображеніямъ симметріи.

Въ § 23 было доказано, что коэффиціенты ур-ія (239) должны удовлетворять условію:

$$(240) \ldots \sum(x_1 - a_{11} - a_{12}x_2 - a_{13}x_3 - \ldots - a_{1n}x_n)^2 = \text{minimum}.$$

Дифференцируя это выраженіе по $a_{11}, a_{12}, a_{13}, \ldots a_{1n}$ и приравнивая частныя производныя нулю, мы найдемъ систему ур-ій, изъ которыхъ опредѣлятся всѣ эти коэффиціенты:

$$(241) \ldots \left\{ \begin{array}{l} \sum(x_1 - a_{11} - a_{12}x_2 - \ldots - a_{1n}x_n) = 0, \\ \sum x_2(x_1 - a_{11} - a_{12}x_2 - \ldots - a_{1n}x_n) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \sum x_n(x_1 - a_{11} - a_{12}x_2 - \ldots - a_{1n}x_n) = 0. \end{array} \right.$$

Первое изъ ур-ій (241) даетъ:

$$Na_{11} = \sum x_1 - a_{12}\sum x_2 - \ldots - a_{1n}\sum x_n = 0,$$

ибо $\sum x_i$, какъ сумма отклоненій отъ средняго ариѳметическаго, равна нулю.

Если мы въ остальныхъ ур-іяхъ (241) произведемъ суммированія, сократимъ на N, a_{11} положимъ равнымъ нулю и членъ, несодержащій ни одного изъ искомыхъ коэффиціентовъ, перенесемъ въ правую часть, то получимъ:

$$(242) \ldots \left\{ \begin{array}{l} \sigma_2^2 a_{12} + \sigma_2\sigma_3 r_{23} a_{13} + \sigma_2\sigma_4 r_{24} a_{14} + \ldots + \sigma_2\sigma_n r_{2n} a_{1n} = \sigma_1\sigma_2 r_{12}, \\ \sigma_3\sigma_2 r_{32} a_{12} + \sigma_3^2 a_{13} + \sigma_3\sigma_4 r_{34} a_{14} + \ldots + \sigma_3\sigma_n r_{3n} a_{1n} = \sigma_1\sigma_3 r_{13}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sigma_n\sigma_2 r_{n2} a_{12} + \sigma_n\sigma_3 r_{n3} a_{13} + \sigma_n\sigma_4 r_{n4} a_{14} + \ldots + \sigma_n^2 a_{1n} = \sigma_1\sigma_n r_{1n}. \end{array} \right.$$

Чтобы упростить эти ур-ія, положимъ:

$$(243) \ldots a_{12} = b_{12}\frac{\sigma_1}{\sigma_2}, \qquad a_{13} = b_{13}\frac{\sigma_1}{\sigma_3} \ldots\ldots a_{1n} = b_{1n}\frac{\sigma_1}{\sigma_n}.$$

Подставивъ эти выраженія въ (242) и сокративъ первое ур-іе на $\sigma_1\sigma_2$, второе на $\sigma_1\sigma_3$ и т. д., будемъ имѣть слѣдующую болѣе простую систему:

$$(244) \quad \left\{\begin{array}{l} b_{12} + r_{23}b_{13} + r_{24}b_{14} + \ldots\ldots + r_{2n}b_{1n} = r_{12}, \\ r_{32}b_{12} + b_{13} + r_{34}b_{14} + \ldots\ldots + r_{3n}b_{1n} = r_{13}, \\ r_{42}b_{12} + r_{43}b_{13} + b_{14} + \ldots\ldots + r_{4n}b_{1n} = r_{14}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ r_{n2}b_{12} + r_{n3}b_{13} + r_{n4}b_{14} + \ldots\ldots + b_{1n} = r_{1n}. \end{array}\right.$$

Рѣшая эту систему ур-ій при помощи детерминантовъ, находимъ, напр., для i-ого коэффиціента (b_{1i}) слѣдующее выраженіе:

$$(245) \quad b_{1i} = \frac{\begin{vmatrix} 1 & r_{23} & r_{24} & \ldots & r_{2,i-1} & r_{12} & r_{2,i+1} & \ldots & r_{2n} \\ r_{32} & 1 & r_{34} & \ldots & r_{3,i-1} & r_{13} & r_{3,i+1} & \ldots & r_{3n} \\ r_{42} & r_{43} & 1 & \ldots & r_{4,i-1} & r_{14} & r_{4,i+1} & \ldots & r_{4n} \\ \ldots & & & & & & & & \\ r_{n2} & r_{n3} & r_{n4} & \ldots & r_{n,i-1} & r_{1n} & r_{n,i+1} & \ldots & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_{23} & r_{24} & \ldots & r_{2n} \\ r_{32} & 1 & r_{34} & \ldots & r_{3n} \\ r_{42} & r_{43} & 1 & \ldots & r_{4n} \\ \ldots & & & & \\ r_{n2} & r_{n3} & r_{n4} & \ldots & 1 \end{vmatrix}}.$$

Это выраженіе можно представить въ гораздо болѣе сокращенномъ видѣ. Обозначимъ черезъ R симметричный детерминантъ, составленный изъ всѣхъ коэффиціентовъ корреляціи съ единицами на діагонали (замѣтимъ при этомъ, что $r_{12} = r_{21}$ и т. д.); мы будемъ, слѣд., имѣть:

$$(246) \quad R = \begin{vmatrix} 1 & r_{12} & r_{13} & \ldots & r_{1n} \\ r_{21} & 1 & r_{23} & \ldots & r_{2n} \\ r_{31} & r_{32} & 1 & \ldots & r_{3n} \\ \ldots & & & & \\ r_{n1} & r_{n2} & r_{n3} & \ldots & 1 \end{vmatrix}.$$

Если мы зачеркнемъ въ этомъ детерминантѣ i-ую строку и j-ую колону, т. е. строку и колону, на пересѣченіи которыхъ стоитъ элементъ съ подписными значками ij, а оставшійся детерминантъ помножимъ на $(-1)^{i+j}$, т. е. поставимъ передъ нимъ знакъ $(+)$, если $i+j$ число четное, и $(-)$, если нечетное, то мы получимъ детерминантъ, носящій названіе *минора* ij-ого элемента. Этотъ миноръ мы будемъ обозначать черезъ R_{ij}.

Детерминантъ, стоящій въ числителѣ (245), преобразуемъ, перенося вертикаль, содержащую элементы $r_{12}, r_{13}, \ldots r_{1n}$, или какъ мы будемъ писать ее $r_{21}, r_{31}, \ldots r_{n1}$, на первое мѣсто. При этомъ детерминантъ перемѣнитъ знакъ столько разъ, сколько разъ для достиженія этой цѣли нужно сдѣлать перестановокъ, мѣняя мѣстами за разъ только двѣ вертикали. Легко разсчитать, что перемѣнъ знака будетъ $i-2$, и что, такимъ образомъ, числитель въ выраженіи (245) будетъ равняться:

$$(247) \quad . \quad . \; (-1)^{i-2} . \begin{vmatrix} r_{21} & 1 & r_{23} & \ldots & r_{2,i-1} & r_{2,i+1} & \ldots & r_{2n} \\ r_{31} & r_{32} & 1 & \ldots & r_{3,i-1} & r_{3,i+1} & \ldots & r_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{n1} & r_{n2} & r_{n3} & \ldots & r_{n,i-1} & r_{n,i+1} & \ldots & 1 \end{vmatrix} .$$

Если въ детерминантѣ R (246) зачеркнуть 1-ую строку и i-ую колону, то мы, какъ легко замѣтить, получимъ такой-же детерминантъ, т. что соотвѣтствующій миноръ будетъ равняться:

$$(248) \quad . \quad . \; R_{1i} = (-1)^{i+1} \begin{vmatrix} r_{21} & 1 & r_{23} & \ldots & r_{2,i-1} & r_{2,i+1} & \ldots & r_{2n} \\ r_{31} & r_{32} & 1 & \ldots & r_{3,i-1} & r_{3,i+1} & \ldots & r_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{n1} & r_{n2} & r_{n3} & \ldots & r_{n,i-1} & r_{n,i+1} & \ldots & 1 \end{vmatrix} .$$

Выраженіе (247) отличается отъ (248) только знакомъ, слѣдовательно, числитель (245) равенъ $-R_{1i}$.

Далѣе, легко замѣтить, что детерминантъ, стоящій въ знаменателѣ (245), получится изъ (246), если мы въ послѣднемъ зачеркнемъ 1-ую строку и первую вертикаль. Т. к. $(-1)^{1+1} = +1$, то послѣдній детерминантъ будетъ миноромъ R_{11}. Такимъ образомъ, искомый коэффиціентъ b_{1i} получаетъ чрезвычайно простой видъ:

$$(249) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad b_{1i} = -\frac{R_{1i}}{R_{11}},$$

а слѣдовательно, въ силу условныхъ равенствъ (243), ур-іе регрессіи будетъ:

$$(250) \quad . \quad . \quad X_1 = -\frac{R_{12}}{R_{11}} \frac{\sigma_1}{\sigma_2} x_2 - \frac{R_{13}}{R_{11}} \frac{\sigma_1}{\sigma_3} x_3 - \ldots - \frac{R_{1n}}{R_{11}} \frac{\sigma_1}{\sigma_n} x_n,$$

или, если вмѣсто отклоненій отъ средняго ариѳметическаго подставить самыя величины:

$$(251) \ldots X_1 - h_1 = -\frac{R_{12}}{R_{11}}\frac{\sigma_1}{\sigma_2}(x_2 - h_2) - \frac{R_{13}}{R_{11}}\frac{\sigma_1}{\sigma_3}(x_3 - h_3) - \\ - \ldots\ldots - \frac{R_{1n}}{R_{11}}\frac{\sigma_1}{\sigma_n}(x_n - h_n).$$

Это будетъ ур-іе, пользуясь которымъ для опредѣленія x_1-ого по $x_2, x_3, \ldots . x_n$-му, мы, какъ выше (§ 23) было доказано, сдѣлаемъ наименьшую ошибку (т. е. рядъ ошибокъ, сумма квадратовъ которыхъ будетъ наименьшей).

Представляетъ также интересъ найти общую формулу для средней квадратичной ошибки. Для этого мы должны въ выраженіе (240) подставить найденныя выше значенія $a_{11}, a_{12}, \ldots . a_{1n}$, раздѣлить все на N и извлечь затѣмъ квадратный корень. Мы получимъ:

$$(252) \ldots N\Sigma_1^2 = \sum\left\{x_1 + \frac{R_{12}}{R_{11}}\frac{\sigma_1}{\sigma_2}x_2 + \frac{R_{13}}{R_{11}}\frac{\sigma_1}{\sigma_3}x_3 + \ldots + \frac{R_{1n}}{R_{11}}\frac{\sigma_1}{\sigma_n}x_n\right\}^2 =$$

$$= \frac{1}{R_{11}^2}\sum\left\{R_{11}x_1 + R_{12}\frac{\sigma_1}{\sigma_2}x_2 + R_{13}\frac{\sigma_1}{\sigma_3}x_3 + \ldots + R_{1n}\frac{\sigma_1}{\sigma_n}x_n\right\}^2.$$

Возвышая въ квадратъ и затѣмъ суммируя, найдемъ:

$$N\Sigma_1^2 = \frac{1}{R_{11}^2}\sum\left\{R_{11}^2x_1^2 + R_{12}^2\frac{\sigma_1^2}{\sigma_2^2}x_2^2 + R_{13}^2\frac{\sigma_1^2}{\sigma_3^2}x_3^2 + \ldots + R_{1n}^2\frac{\sigma_1^2}{\sigma_n^2}x_n^2 + \right.$$

$$+ 2R_{11}R_{12}\frac{\sigma_1}{\sigma_2}x_1x_2 + 2R_{11}R_{13}\frac{\sigma_1}{\sigma_3}x_1x_3 + \ldots +$$

$$+ 2R_{11}R_{1n}\frac{\sigma_1}{\sigma_n}x_1x_n + 2R_{12}R_{13}\frac{\sigma_1^2}{\sigma_2\sigma_3}x_2x_3 + \ldots +$$

$$\left. + 2R_{12}R_{1n}\frac{\sigma_1^2}{\sigma_2\sigma_n}x_2x_n + \text{и т. д.}\right\} =$$

$$= \frac{1}{R_{11}^2}\left\{R_{11}^2 . N\sigma_1^2 + R_{12}^2\frac{\sigma_1^2}{\sigma_2^2} . N\sigma_2^2 + \ldots + R_{1n}^2\frac{\sigma_1^2}{\sigma_n^2} . N\sigma_n^2 + \right.$$

$$+ 2R_{11}R_{12}\frac{\sigma_1}{\sigma_2} . N\sigma_1\sigma_2r_{12} + 2R_{11}R_{13}\frac{\sigma_1}{\sigma_3} . N\sigma_1\sigma_3r_{13} + \ldots +$$

$$+ 2R_{11}R_{1n}\frac{\sigma_1}{\sigma_n} . N\sigma_1\sigma_nr_{1n} + 2R_{12}R_{13}\frac{\sigma_1^2}{\sigma_2\sigma_3} . N\sigma_2\sigma_3r_{23} + \ldots +$$

$$\left. + 2R_{12}R_{1n}\frac{\sigma_1^2}{\sigma_2\sigma_n} . N\sigma_2\sigma_nr_{2n} + \text{и т. д.}\right\}.$$

Раздѣливъ обѣ части на N и произведя всѣ возможныя сокращенія въ отдѣльныхъ членахъ, вынесемъ за скобку σ_1^2 и будемъ имѣть:

$$\Sigma_1^2 = \frac{\sigma_1^2}{R_{11}^2}\left\{R_{11}^2 + R_{12}^2 + R_{13}^2 + R_{14}^2 + \ldots\ldots\ldots + R_{1n}^2 + \right.$$

$$+ 2R_{11}R_{12}r_{12} + 2R_{11}R_{13}r_{13} + 2R_{11}R_{14}r_{14} + \ldots + 2R_{11}R_{1n}r_{1n} +$$

$$+ 2R_{12}R_{13}r_{23} + 2R_{12}R_{14}r_{24} + \ldots + 2R_{12}R_{1n}r_{2n} +$$

$$+ 2R_{13}R_{14}r_{34} + \ldots + 2R_{13}R_{1n}r_{3n} +$$

$$\left. + \text{и т. д.}\right\}.$$

Въ каждый изъ членовъ, кромѣ членовъ первой строки, входитъ удвоенное произведеніе типа $2R_{1i}R_{1j}r_{ij}$. Каждый такой членъ можно представить въ видѣ суммы $(R_{1i}R_{1j}r_{ij} + R_{1i}R_{1j}r_{ij})$, одну часть которой мы напишемъ рядомъ съ R_{1i}^2, другую—рядомъ съ R_{1j}^2.

Такимъ образомъ, мы получимъ:

$$(253) \ldots \Sigma_1^2 = \frac{\sigma_1^2}{R_{11}^2}\left\{R_{11}^2 + R_{11}R_{12}r_{12} + R_{11}R_{13}r_{13} + \ldots + R_{11}R_{1n}r_{1n} + \right.$$

$$+ R_{12}R_{11}r_{21} + R_{12}^2 + R_{12}R_{13}r_{23} + \ldots\ldots + R_{12}R_{1n}r_{2n} +$$

$$+ R_{13}R_{11}r_{31} + R_{13}R_{12}r_{32} + R_{13}^2 + \ldots\ldots + R_{13}R_{1n}r_{3n} +$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\left. + R_{1n}R_{11}r_{n1} + R_{1n}R_{12}r_{n2} + R_{1n}R_{13}r_{n3} + \ldots + R_{1n}^2\right\} =$$

$$= \frac{\sigma_1^2}{R_{11}^2}\left\{R_{11}[R_{11} + R_{12}r_{12} + R_{13}r_{13} + \ldots\ldots + R_{1n}r_{1n}] + \right.$$

$$+ R_{12}[R_{11}r_{21} + R_{12} + R_{13}r_{23} + \ldots\ldots + R_{1n}r_{2n}] +$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\left. + R_{1n}[R_{11}r_{n1} + R_{12}r_{n2} + R_{13}r_{n3} + \ldots\ldots + R_{1n}]\right\}.$$

Изъ теоріи детерминантовъ извѣстно, что каждый детерминантъ можетъ быть представленъ въ видѣ суммы произведеній элементовъ какой нибудь строки на соотвѣтствующіе миноры, слѣдовательно:

$$R_{11} + R_{12}r_{12} + R_{13}r_{13} + \ldots + R_{1n}r_{1n} = R.$$

Легко сообразить, какому детерминанту равно каждое выраженіе въ остальныхъ квадратныхъ скобкахъ. Такъ, напримѣръ, согласно тому-же правилу, очевидно, что

$$R_{11} r_{21} + R_{12} + R_{13} r_{23} + \ldots + R_{1n} r_{2n} =$$

$$= \begin{vmatrix} r_{21} & 1 & r_{23} & r_{24} & \ldots & r_{2n} \\ r_{21} & 1 & r_{23} & r_{24} & \ldots & r_{2n} \\ r_{31} & r_{32} & 1 & r_{34} & \ldots & r_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{n1} & r_{n2} & r_{n3} & r_{n4} & \ldots & 1 \end{vmatrix}.$$

Послѣдній детерминантъ получится изъ R, если первую строку замѣнить 2-ой. Но детерминантъ, у котораго двѣ строки тождественны, равенъ нулю (чтобы въ этомъ убѣдиться стоитъ только изъ первой строки вычесть вторую, что не измѣняетъ величины детерминанта. Тогда всѣ элементы первой строки обратятся въ нули). Подобнымъ образомъ можно доказать, что всѣ выраженія въ квадратныхъ скобкахъ въ формулѣ (253), кромѣ перваго, равны нулю, и мы получимъ (сокращая на R_{11} и извлекая квадратный корень):

$$(254) \ldots\ldots \quad \Sigma_1 = \sigma_1 \sqrt{\frac{R}{R_{11}}},$$

— чрезвычайно простое и очень важное выраженіе.

§ 28. *Случай четырехъ перемѣнныхъ.*

Выраженія (250) и (254) заключаютъ въ себѣ всѣ частные формулы, выведенныя выше для двухъ и трехъ перемѣнныхъ. Напримѣръ, въ случаѣ корреляціи между двумя величинами:

$$R = \begin{vmatrix} 1 & r_{12} \\ r_{12} & 1 \end{vmatrix} = 1 - r_{12}^2; \quad R_{11} = 1; \quad R_{12} = -r_{12}; \quad \text{откуда}$$

ур-іе регрессіи:

$$X_1 - h_1 = -\frac{R_{12}}{R_{11}} \frac{\sigma_1}{\sigma_2} (x_2 - h_2) = r_{12} \frac{\sigma_1}{\sigma_2} (x_2 - h_2),$$

а средняя квадратичная ошибка опредѣленія x_1-го по x_2-му:

$$\Sigma_1 = \sigma_1 \sqrt{\frac{R}{R_{11}}} = \sigma_1 \sqrt{1 - r_{12}^2}.$$

Мы воспользуемся общими формулами для того, чтобы вывести ур-ія, относящіяся къ случаю четырехъ перемѣнныхъ.

Примѣняя введенныя выше обозначенія, будемъ имѣть:

$$X_1 - h_1 = -\frac{R_{12}}{R_{11}}\,\frac{\sigma_1}{\sigma_2}(x_2 - h_2) - \frac{R_{13}}{R_{11}}\,\frac{\sigma_1}{\sigma_3}(x_3 - h_3) -$$
$$-\frac{R_{14}}{R_{11}}\,\frac{\sigma_1}{\sigma_4}(x_4 - h_4), \text{ и}$$
$$\Sigma_1 = \sigma_1\sqrt{\frac{R}{R_{11}}}.$$

Здѣсь

$$(255)\;.\;.\;.\;.\;.\; R = \begin{vmatrix} 1 & r_{12} & r_{13} & r_{14} \\ r_{21} & 1 & r_{23} & r_{24} \\ r_{31} & r_{32} & 1 & r_{34} \\ r_{41} & r_{42} & r_{43} & 1 \end{vmatrix} =$$
$$= 1 - (r_{12}^2 + r_{13}^2 + r_{14}^2 + r_{23}^2 + r_{24}^2 + r_{34}^2) +$$
$$+ (r_{12}^2 r_{34}^2 + r_{23}^2 r_{14}^2 + r_{13}^2 r_{24}^2) +$$
$$+ 2(r_{23}r_{24}r_{34} + r_{34}r_{14}r_{13} + r_{12}r_{14}r_{24} + r_{12}r_{13}r_{23}) -$$
$$- 2(r_{12}r_{14}r_{23}r_{34} + r_{14}r_{13}r_{23}r_{24} + r_{12}r_{13}r_{24}r_{34}).$$

Величину R, конечно, удобнѣе вычислять не по этой формулѣ, а черезъ миноры R_{11}, R_{12}, R_{13} и R_{14}, которые во всякомъ случаѣ должны быть найдены для того, чтобы можно было написать ур-іе регрессіи.

Черезъ свои миноры R выражается, какъ извѣстно, слѣд. обр.:

$$R = R_{11} + R_{12}r_{12} + R_{13}r_{13} + R_{14}r_{14}.$$

Выписывать всѣ миноры нѣтъ надобности, т. к. ихъ выраженія легко получить другъ изъ друга перестановкой значковъ. Покажемъ, какъ найти R_{11} и R_{12}.

Зачеркнувъ въ R первую строку и первую вертикаль, будемъ имѣть:

$$(256)\;.\;.\; R_{11} = + \begin{vmatrix} 1 & r_{23} & r_{24} \\ r_{32} & 1 & r_{34} \\ r_{42} & r_{43} & 1 \end{vmatrix} = 1 - r_{23}^2 - r_{24}^2 - r_{34}^2 + 2r_{23}r_{24}r_{34}.$$

Затѣмъ, зачеркнувъ первую строку и вторую вертикаль и замѣтивъ, что оставшійся детерминантъ нужно взять съ минусомъ, ибо сумма индексовъ $(1+2)$ — число нечетное, будемъ имѣть:

$$(257')\;.\;.\;.\;.\;.\; R_{12} = -\begin{vmatrix} r_{21} & r_{23} & r_{24} \\ r_{31} & 1 & r_{34} \\ r_{41} & r_{43} & 1 \end{vmatrix} =$$
$$= -\{r_{12}(1 - r_{34}^2) - r_{13}r_{23} - r_{14}r_{24} + r_{34}(r_{13}r_{24} + r_{14}r_{23})\}.$$

Аналогично:

$$(257'') \quad . \; . \; R_{13} = -\{ r_{13}(1 - r_{24}^2) - r_{12} r_{32} - r_{14} r_{34} + r_{24}(r_{12} r_{34} + r_{14} r_{32}) \} .$$

$$(257''') \quad . \; . \; R_{14} = -\{ r_{14}(1 - r_{23}^2) - r_{12} r_{42} - r_{13} r_{43} + r_{23}(r_{12} r_{43} + r_{13} r_{42}) \} .$$

Въ качествѣ иллюстраціи разсмотримъ вопросъ о реконструкціи роста человѣка по величинѣ отдѣльныхъ его органовъ (W. R. Macdonell, On Criminal Anthropometry, Biometrica, Vol. I). [1]) Слѣдующія величины относятся къ тюремному населенію главнѣйшихъ тюремъ Англіи и Уэльса и основаны на 3000 формуляровъ, взятыхъ изъ антропометрическаго бюро Scotland Yard'а. Англійская практика дѣлаетъ различіе между „привычными" (habituals) преступниками и „случайными" (non-habituals), преступленія которыхъ и наказанія, коимъ они подвергнуты, сравнительно незначительны. Данныя Macdonell'я относятся къ этой послѣдней категоріи.

Ограничиваясь 4 размѣрами (Macdonell принимаетъ во вниманіе 7), приведемъ слѣдующія числа: [2])

3000 преступниковъ.

	Среднее отклоненіе.	Средняя ариѳметич.
Лѣвый средній палецъ	0,5479 см.	11,5474 см.
Лѣвый локоть	1,9627 „	45,0586 „
Лѣвая ступня	1,1792 „	25,6877 „
Ростъ	6,4540 „	166,4572 „

Коэффиціенты корреляціи для 3000 преступниковъ.

	Палецъ.	Локоть.	Ступня.	Ростъ.
Палецъ	1	0,84638	0,75871	0,66084
Локоть	0,84638	1	0,79699	0,79986
Ступня	0,75871	0,79699	1	0,73636
Ростъ	0,66084	0,79986	0,73636	1

[1]) Другая важная работа по данному вопросу: K. Pearson, On the Reconstruction of the Stature of Prehistoric Races, Phil. Trans., Vol. 192 A., p. 169—244.

[2]) Macdonell, op. cit., p. 202.

Нужно замѣтить, что данныя антропометріи достаточно близко подчиняются въ своемъ распредѣленіи закону Гаусса. Поэтому, какъ доказано будетъ ниже, среднее отклоненіе для каждой величины одинаково во всѣхъ ея строяхъ и средняя квадратичная ошибка всѣхъ опредѣленій равна среднему отклоненію отдѣльнаго строя, равна $\sigma_{1(234)}$.

Вѣроятная ошибка, дающая предѣлы, въ которыхъ должна лежать половина отклоненій, равна, слѣдовательно, $E_{1(234)} = 0{,}67449\,\sigma_{1(234)}$.

Пользуясь вышеприведенными данными, мы по извѣстнымъ формуламъ можемъ найти ур-ія реконструкціи, полученныя Macdonell'емъ (l. cit., p. 210—211).

(1) Реконструкція роста по лѣвому среднему пальцу:

$$H = 166{,}4572 + \frac{6{,}4540}{0{,}5479} \times 0{,}66084 \ (\text{палецъ} - 11{,}5474) =$$
$$= 166{,}4572 + 7{,}7849 \ (\text{палецъ} - 11{,}5474).$$

$$\text{Вѣроятная ошибка} = 0{,}67449 \times 6{,}4540 \sqrt{1 - (0{,}66084)^2} =$$
$$= 3{,}267 \text{ см.}$$

Аналогично:

(2) Реконструкція роста по лѣвому локтю:

$$H = 166{,}4572 + 2{,}6301 \ (\text{локоть} - 45{,}0586).$$

Вѣроятная ошибка $= 2{,}613$ см.

(3) Реконструкція роста по лѣвой ступнѣ:

$$H = 166{,}4572 + 4{,}0301 \ (\text{ступня} - 25{,}6877).$$

Вѣроятная ошибка $= 2{,}945$ см.

(4) Реконструкція роста по пальцу и локтю:

Въ этомъ случаѣ:

$$R = \begin{vmatrix} 1 & 0{,}66084 & 0{,}79986 \\ 0{,}66084 & 1 & 0{,}84638 \\ 0{,}79986 & 0{,}84638 & 1 \end{vmatrix} = 0{,}101914,$$

$$R_{11} = 0{,}28364, \quad R_{12} = 0{,}01614, \quad R_{13} = -0{,}24054,$$
$$\sigma_2 = 0{,}5479, \quad \sigma_3 = 1{,}9627,$$
$$h_2 = 11{,}5474, \quad h_3 = 45{,}0586.$$

Слѣдовательно, ур-іе для H будетъ:

$$H = 166{,}4572 - 0{,}6703 \ (\text{палецъ} - 11{,}5474) + 2{,}7886 \ (\text{локоть} - 45{,}0586).$$

$$\text{Вѣроятная ошибка} = 0{,}67449 \times 6{,}4540 \sqrt{\frac{0{,}101914}{0{,}2836}} =$$
$$= 2{,}609 \text{ см.}$$

Послѣднее ур-іе для реконструкціи роста по пальцу и локтю интересно въ томъ отношеніи, что оно обнаруживаетъ оригинальную зависимость между размѣрами. Именно, изъ субъектовъ съ *данной* длиной локтя тотъ будетъ имѣть большій ростъ, у котораго лѣвый средній палецъ короче, и наоборотъ, чѣмъ длиннѣе палецъ, тѣмъ меньше ростъ.

(5) Реконструкція роста по пальцу и ступнѣ:

$$H = 166{,}4572 + 2{,}8360 \ (\text{палецъ} - 11{,}5474) + 3{,}0304 \ (\text{ступня} - 25{,}6877).$$

Вѣроятная ошибка = 2,865 см.

(6) Реконструкція роста по локтю и ступнѣ:

$$H = 166{,}4572 + 1{,}9198 \ (\text{локоть} - 45{,}0586) + 1{,}4834 \ (\text{ступня} - 25{,}6877).$$

Вѣроятная ошибка = 2,5136 см.

(7) Реконструкція роста по пальцу, локтю и ступнѣ. Ур-іе регрессіи можно написать безъ труда по слѣдующимъ даннымъ (индексъ „1“ относится къ пальцу, „2“ къ локтю, „3“ къ ступнѣ и „4“ къ росту):

$$R = 0{,}031586, \quad R_{44} = 0{,}096392, \quad R_{14} = -0{,}014211,$$

$$R_{24} = -0{,}065693, \ R_{34} = +0{,}029405, \ \sqrt{\frac{R}{R_{44}}} = 0{,}572436.$$ [1]

Вѣроятная ошибка будетъ = 2,492 см.

(8) Если для опредѣленія роста воспользоваться еще тремя величинами: длиной головы, шириной головы и шириной лица, то вѣроятная ошибка такого опредѣленія будетъ = 2,448 см.

Въ слѣдующей таблицѣ сведены всѣ вѣроятныя ошибки:

Реконструкція роста по:	Вѣроятная ошибка.
Пальцу	3,267
Ступнѣ	2,945
Локтю	2,613
Пальцу и ступнѣ	2,865
Пальцу и локтю	2,609
Локтю и ступнѣ	2,514
Пальцу, локтю и ступнѣ	2,492
Пальцу, локтю, ступнѣ, длинѣ головы, ширинѣ головы и ширинѣ лица	2,448

[1]) Macdonell, l. c., p. 209—210.

Эта таблица показываетъ намъ, что увеличеніе числа независимыхъ перемѣнныхъ само по себѣ еще очень мало улучшаетъ точность опредѣленія. Гораздо важнѣе надлежащій выборъ перемѣнныхъ. При реконструкціи роста по одному локтю вѣроятная ошибка равна 2,613 см.; при реконструкціи по локтю и ступнѣ вѣроятная ошибка замѣтно уменьшается, доходя до 2,514 см.; прибавленіе пальца увеличиваетъ точность очень мало, уменьшая вѣроятную ошибку всего на 0,2 мм. Мы можемъ взять еще три органа, но улучшеніе реконструкціи—вѣроятная ошибка уменьшится на 0,4 мм.—совершенно уже незначительно по сравненію съ неимовѣрно возростающей при этомъ сложностью вычисленій.

Характеръ и величину ошибокъ, возникающихъ на практикѣ при употребленіи приведенныхъ формулъ, Macdonell иллюстрируетъ слѣдующей таблицей:

Дѣйств. ростъ.	Ростъ вычисленный по:						Среднее.
	Пальцу.	Локтю.	Ступнѣ.	Пальцу и локтю.	Пальцу и ступнѣ.	Ступнѣ и локтю.	
160,3	+0,3	−2,4	+0,6	−2,4	−0,2	−2,1	−1,03
168,3	−1,4	−2,3	−6,2	−2,3	−5,0	−3,8	−3,50
170,2	−2,6	−4,2	−3,3	−4,3	−3,0	−3,9	−3,55
158,4	+5,4	+1,3	−2,0	+1,2	−0,5	−0,5	+0,82
161,9	+1,1	+3,1	+2,6	+3,3	+1,8	+2,8	+2,45
165,7	+0,3	−0,4	+3,2	−0,5	+2,5	+0,8	+0,98
167,6	+1,6	+0,5	+5,8	+0,4	+5,0	+2,6	+2,65
160,0	+4,5	+1,0	+3,7	+0,9	+3,7	+1,5	+2,55
161,9	+8,1	+6,2	−0,6	+6,0	+1,9	+3,9	+4,25
169,5	−5,0	−1,1	−4,2	−0,8	−4,6	−2,0	−2,95
Среднее	3,03	2,25	3,22	2,21	2,82	2,39	2,47
Вѣр. ош.	±3,3	±2,6	±2,9	±2,6	±2,9	±2,5	±2,8

Здѣсь въ первой графѣ указывается дѣйствительный ростъ 10-ти преступниковъ, выбранныхъ на удачу изъ имѣвшагося матеріала.

Въ остальныхъ графахъ приведены разницы между вычисленнымъ разными способами и дѣйствительнымъ ростомъ. [1])

Результаты реконструкціи только приблизительны, да иными они и не могутъ быть, т. к. формулы регрессіи даютъ не индивидуальный ростъ, а средній—для группы индивидуумовъ съ даннымъ размѣромъ другихъ органовъ. Какъ было доказано въ § 23, эти формулы даютъ наименьшую среднюю квадратичную ошибку, по сравненію со всѣми возможными иными формулами линейной зависимости. Такимъ образомъ, неточности опредѣленія имѣютъ своимъ корнемъ присущую индивидуумамъ измѣнчивость и неполную корреляцію между органами. Для многихъ практическихъ (напримѣръ, полицейскихъ и судебно-медицинскихъ) цѣлей такая точность однако уже болѣе или менѣе достаточна.

§ 29. Нормальная корреляція. Ур-іе распредѣленія.

Теорія корреляціи, изложенная выше, независима отъ какого-либо спеціальнаго закона распредѣленія, и приложимость формулъ регрессіи обусловливается исключительно лишь линейностью распредѣленія эмпирическаго матеріала.

Въ своемъ первоначальномъ видѣ однако теорія корреляціи являлась лишь расширеніемъ закона Гаусса на случай многихъ перемѣнныхъ. Какъ чрезвычайно важный частный случай, она вполнѣ заслуживаетъ хотя бы краткаго изложенія. [2])

Пусть $x_1, x_2, x_3, \ldots x_n$ будутъ отклоненія какимъ либо образомъ сопряженныхъ другъ съ другомъ величинъ отъ ихъ среднихъ ариѳметическихъ. Мы допустимъ, что значенія этихъ величинъ опредѣляются множествомъ причинъ, число которыхъ пусть будетъ m, относительно котораго мы предположимъ, что оно значительно больше n. Отклоненія интенсивностей причинъ отъ ихъ среднихъ значеній пусть будутъ $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots \varepsilon_m$, такъ что при среднемъ значеніи интенсивностей причинъ, т. е. при $\varepsilon_1 = 0$, $\varepsilon_2 = 0, \ldots \varepsilon_n = 0$, и наши величины $x_1, x_2, \ldots x_n$ также равны нулю.

Основныя наши допущенія заключаются, во 1-хъ, въ томъ, что измѣненія въ интенсивности каждой отдѣльной причины на-

---

[1]) Op. cit., p. 213.

[2]) K. Pearson, Regression, Heredity, and Panmixia, Phil. Trans., Vol. 187 A, p. 261 и сл.

столько незначительны, что квадратами ихъ можно пренебречь, во 2-хъ, въ томъ, что распредѣленіе ихъ подчиняется закону Гаусса и, въ 3-хъ, въ томъ, что вѣроятности различныхъ ихъ значеній независимы другъ отъ друга.

Въ силу перваго допущенія зависимость любого x_i отъ $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_m$ будетъ приближенно выражаться линейной функціей (мы здѣсь допускаемъ, что эти функціи можно разложить по ряду Тейлора; тогда квадраты и высшія степени по условію можно отбросить). Такимъ образомъ:

$$(258) \ldots \begin{cases} x_1 = \alpha_{11}\varepsilon_1 + \alpha_{12}\varepsilon_2 + \ldots + \alpha_{1m}\varepsilon_m, \\ x_2 = \alpha_{21}\varepsilon_1 + \alpha_{22}\varepsilon_2 + \ldots + \alpha_{2m}\varepsilon_m, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = \alpha_{n1}\varepsilon_1 + \alpha_{n2}\varepsilon_2 + \ldots + \alpha_{nm}\varepsilon_m. \end{cases}$$

Въ силу 2-го допущенія вѣроятность того, что величина уклоненія интенсивности i-ой причины отъ своего средняго будетъ заключаться между предѣлами ε_i и $\varepsilon_i + \delta\varepsilon_i$ равна

$$C_i e^{-\frac{1}{2}\frac{\varepsilon_i^2}{k_i^2}} . \delta\varepsilon_i,$$

гдѣ k_i есть среднее отклоненіе.

Такъ какъ по допущенію 3-му подобныя вѣроятности независимы другъ отъ друга, то вѣроятность того, что одновременныя отклоненія интенсивностей причинъ будутъ заключаться между ε_1 и $\varepsilon_1 + \delta\varepsilon_1$; ε_2 и $\varepsilon_2 + \delta\varepsilon_2$; $\ldots$ ε_m и $\varepsilon_m + \delta\varepsilon_m$, равна произведенію соотвѣтствующихъ вѣроятностей. Мы имѣемъ, слѣдовательно:

$$(259) \ldots P = Ce^{-\frac{1}{2}\left(\frac{\varepsilon_1^2}{k_1^2} + \frac{\varepsilon_2^2}{k_2^2} + \ldots + \frac{\varepsilon_m^2}{k_m^2}\right)} . \delta\varepsilon_1 \delta\varepsilon_2 \delta\varepsilon_3 \ldots \delta\varepsilon_m.$$

При помощи ур-ій (258) мы можемъ любыя n (напримѣръ первыя n) величинъ изъ числа $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_m$ выразить черезъ n величинъ $x_1, x_2, \ldots x_n$ и остальныя $\varepsilon_{n+1}, \varepsilon_{n+2}, \ldots \varepsilon_m$; получимъ для каждаго ε_i дробное выраженіе съ линейнымъ относительно всѣхъ перемѣнныхъ числителемъ; знаменатели будутъ одинаковы во всѣхъ выраженіяхъ и будутъ содержать только постоянныя. Подставивъ найденныя значенія $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_n$ въ ур-іе (259), мы — для вѣроятности того, что x_1 будетъ заключаться между x_1 и $x_1 + \delta x_1$; x_2 между x_2 и $x_2 + \delta x_2$; $\ldots$ x_n между x_n и $x_n + \delta x_n$;

ε_{n+1} между ε_{n+1} и $\varepsilon_{n+1} + \delta\varepsilon_{n+1}$ и т. д., — будемъ имѣть слѣдующее выраженіе:

$$(260) \ldots P = \text{постоянная} \times e^{-\frac{1}{2}(U+V+W)} \times$$
$$\times \delta x_1 \, \delta x_2 \ldots \delta x_n \, . \, \delta\varepsilon_{n+1} \, \delta\varepsilon_{n+2} \ldots \delta\varepsilon_m ,$$

гдѣ U есть квадратичная функція перемѣнныхъ x_1, x_2, x_n, т. е. сумма квадратовъ и произведеній этихъ величинъ по двѣ съ постоянными коэффиціентами.

V есть такая-же квадратичная функція величинъ ε_{n+1}, $\varepsilon_{n+2}, \ldots \varepsilon_m$.

W есть сумма членовъ, содержащихъ произведенія попарно одного изъ x-овъ и одного изъ ε.

Каково значеніе выраженія (260)? Выраженіе (259), изъ котораго оно получилось, представляетъ изъ себя вѣроятность опредѣленной комбинаціи причинъ, величины интенсивностей которыхъ заключены въ извѣстныхъ границахъ.

Но изъ ур-ій (258) видно, что каждой опредѣленной комбинаціи причинъ соотвѣтствуетъ опредѣленная совокупность значеній $x_1, x_2, \ldots x_n$. Такъ какъ однако число причинъ больше числа величинъ $x_1, x_2, \ldots x_n$, то заданной комбинаціи этихъ послѣднихъ можетъ соотвѣтствовать не одна, а множество комбинацій причинъ. Если однако извѣстны n величинъ $x_1, x_2, \ldots x_n$ и $m - n$ величинъ $\varepsilon_{n+1}, \varepsilon_{n+2}, \ldots \varepsilon_m$, то изъ n ур-ій (258) можно опредѣлить значенія остальныхъ n причинъ $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_n$. Произведенное нами алгебраическое преобразованіе ведетъ за собою измѣненіе точки зрѣнія, такъ что формула (260) даетъ намъ отвѣтъ на вопросъ: какова вѣроятность комбинаціи, состоящей изъ n величинъ $x_1, x_2, \ldots x_n$ и $m - n$ причинъ.

Но послѣднія насъ не интересуютъ, ибо суждены остаться для насъ неизвѣстными; то что намъ нужно знать—это вѣроятность опредѣленной комбинаціи $x_1, x_2, \ldots x_n$ при всѣхъ какихъ угодно значеніяхъ $\varepsilon_{n+1}, \varepsilon_{n+2}, \ldots \varepsilon_m$. Эти величины мы и должны теперь исключить изъ (260). Исключимъ сначала ε_{n+1}. При опредѣленной величинѣ $x_1, x_2, \ldots x_n$, ε_{n+1} можетъ имѣть значенія: отъ ε_{n+1} до $\varepsilon_{n+1} + \delta\varepsilon_{n+1}$, отъ $\varepsilon_{n+1} + \delta\varepsilon_{n+1}$ до $\varepsilon_{n+1} + 2\delta\varepsilon_{n+1}$ и т. д. Вѣроятность того, что мы будемъ имѣть или одинъ, или другой или третій и т. д. изъ этихъ случаевъ равна суммѣ соотвѣтствующихъ вѣроятностей. Это суммированіе производится при помощи операціи интегрированія выраженія (260) по ε_{n+1} отъ $-\infty$ до $+\infty$. Предѣлы опредѣляются на основаніи того соображенія, что мы хотимъ знать вѣроятность комбинаціи $x_1, x_2, \ldots x_n$ при всѣхъ какихъ угодно значеніяхъ ε_{n+1}-го.

Интегрируя по ε_{n+1}, мы оставляемъ безъ измѣненія характеръ величины U въ ур-іи (260), въ выраженіяхъ V и W исчезаетъ ε_{n+1} и въ постоянный членъ входитъ новый множитель, несодержащій ни одной изъ величинъ x и ни одной изъ величинъ ε. [1])

Повторяя аналогичное разсужденіе относительно всѣхъ величинъ $\varepsilon_{n+2}, \varepsilon_{n+3}, \ldots \varepsilon_m$, мы послѣ $m-n$ интегрированій получимъ выраженіе прежняго типа относительно $x_1, x_2, \ldots x_n$, но несодержащее величинъ $\varepsilon_{n+1}, \varepsilon_{n+2}, \ldots \varepsilon_m$.

Такимъ образомъ, вѣроятность того, что величины $x_1, x_2, x_3, \ldots x_n$ будутъ заключаться въ предѣлахъ отъ x_1 до $x_1 + \delta x_1$, отъ x_2 до $x_2 + \delta x_2, \ldots$ отъ x_n до $x_n + \delta x_n$, равна

$$(261) \ldots P = Ce^{-\frac{1}{2}(a_{11}x_1^2 + a_{22}x_2^2 + \ldots + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \ldots)} . \delta x_1 \delta x_2 \delta x_3 \ldots \delta x_n .$$

На основаніи закона большихъ чиселъ частота явленія при большомъ числѣ повтореній приближается къ его вѣроятности. Такимъ образомъ (261) представляетъ изъ себя также и выраженіе для *частоты* явленія, указывая число случаевъ, когда величины $x_1, x_2, x_3, \ldots x_n$ одновременно имѣютъ значенія, заключенныя въ интервалахъ: x_1 до $x_1 + \delta x_1$ и т. д. Все число возможныхъ слу-

[1]) Для читателей, знакомыхъ хорошо съ интегральнымъ исчисленіемъ, сказаннаго, я думаю, достаточно. Для читателей, менѣе освоившихся съ этимъ отдѣломъ математики, не представитъ большого труда выполнить въ качествѣ упражненія описанное преобразованіе. Для этого нужно всѣ члены, несодержащіе ε_{n+1}, вынести изъ подъ знака интеграла, а члены, содержащіе ε_{n+1}, представить въ видѣ полнаго квадрата. Окончательно будемъ имѣть $\int_{-\infty}^{+\infty} e^{-\frac{1}{2}a_{n+1}(\varepsilon_{n+1}+L)^2} . \delta\varepsilon_{n+1}$, гдѣ L — выраженіе, несодержащее ε_{n+1}. Полагая $\frac{1}{2}a_{n+1}(\varepsilon_{n+1}+L)^2 = \xi^2$, будемъ имѣть: $\xi = \sqrt{a_{n+1}/2}(\varepsilon_{n+1}+L)$; $\delta\varepsilon_{n+1} = \sqrt{2/a_{n+1}}\delta\xi$, и интегралъ преобразуется къ виду: $\sqrt{\frac{2}{a}} \int_{-\infty}^{+\infty} e^{-\xi^2} \delta\xi = \sqrt{\frac{2}{a}} . \sqrt{\pi}$ (см. выше стр. 43). Такимъ образомъ, прибавляется новый постоянный множитель, а типъ функціи въ показателѣ остается безъ измѣненія.

чаевъ здѣсь принято за единицу. Если это число равно N, то, умноживъ обѣ части на эту величину и опустивъ дифференціалы, получимъ число случаевъ на единицу интервала каждаго „признака", т. е. частоту распредѣленія:

$$(262) \;.\,.\; Z = Ce^{-\frac{1}{2}(a_{11}x_1^2 + a_{22}x_2^2 + \ldots + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \ldots)}.$$

Величина, стоящая здѣсь справа, есть *нормальная функція распредѣленія n коррелятивно связанныхъ между собою величинъ.*

§ 30. Основныя свойства нормальной функціи распредѣленія. Теорема Edgeworth'а.

Пусть $x_2, x_3, x_4, \ldots . x_n$ имѣютъ опредѣленныя значенія. x_1 при этомъ можетъ имѣть значенія различныя, но, какъ мы увидимъ, средняя ариѳметическая всѣхъ ихъ будетъ функціей $x_2, x_3 \ldots . x_n$. Въ самомъ дѣлѣ, выдѣливъ въ (262) всѣ члены, содержащіе x_1, мы получимъ:

$$Z = Ce^{-\frac{1}{2}\left\{a_{11}\left(x_1^2 + 2\frac{a_{12}}{a_{11}}x_1x_2 + 2\frac{a_{13}}{a_{11}}x_1x_3 + \ldots + 2\frac{a_{1n}}{a_{11}}x_1x_n\right) + V\right\}},$$

гдѣ V есть квадратичная функція остальныхъ величинъ кромѣ x_1 и по условію есть теперь величина постоянная. Члены съ x_1 представимъ въ видѣ полнаго квадрата, а члены, несодержащіе x_1-го, обозначимъ буквой W. Тогда:

$$Z = Ce^{-\frac{1}{2}a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3 \;\ldots\; + \frac{a_{1n}}{a_{11}}x_n\right)^2 + W} =$$
$$= C.e^W.e^{-\frac{1}{2}a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots .\right)^2}.$$

Но W не содержитъ x_1 и при данныхъ $x_2, x_3, \ldots . x_n$ является, слѣд., величиной постоянной. Такимъ образомъ, распредѣленіе значеній x_1-го при указанныхъ условіяхъ подчиняется закону Гаусса и выражается функціей:

$$(263) \;.\,.\; Z = Ce^{-\frac{1}{2}a_{11}\left[x_1 + \left(\frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3 + \ldots + \frac{a_{1n}}{a_{11}}x_n\right)\right]^2}.$$

По доказанному въ § 9 I ч. средняя ариѳметическая всѣхъ значеній x_1-го (при данныхъ $x_2, x_3 \ldots x_n$) будетъ:

$$(264) \quad . \ . \ . \ . \ X = -\frac{a_{12}}{a_{11}} x_2 - \frac{a_{13}}{a_{11}} x_3 - \ . \ . \ . \ - \frac{a_{1n}}{a_{11}} x_n ,$$

а среднее отклоненіе

$$(265) \quad . \ . \ . \ . \ . \ . \ . \ . \ \sigma_{1m} = \frac{1}{\sqrt{a_{11}}} .$$

Мы получаемъ сразу важный результатъ:

(*A*) При *нормальномъ распредѣленіи регрессія носитъ строго линейный характеръ*, и

(*B*) частное *среднее отклоненіе во всѣхъ строяхъ* одинаково.

Выше было доказано (см. стр. 83), что при этихъ условіяхъ средняя ошибка опредѣленія одного перемѣннаго по другимъ равна частному среднему отклоненію. Так. образомъ, мы имѣемъ:

$$(265') \quad . \ . \ . \ . \ . \ \Sigma_1 = \sigma_{1m} = \frac{1}{\sqrt{a_{11}}} .$$

Воспользуемся полученными результатами, чтобы вывести теорему Эджворса, которая устанавливаетъ зависимость между коэффиціентами показательной функціи въ ур-іи нормальнаго распредѣленія, съ одной стороны, и коэффиціентами корреляціи и средними отклоненіями, съ другой.

Мы доказали (§ 27), что всякое ур-іе линейной регрессіи приводится къ виду:

$$X = -\frac{R_{12}}{R_{11}} \frac{\sigma_1}{\sigma_2} x_2 - \frac{R_{13}}{R_{11}} \frac{\sigma_1}{\sigma_3} x_3 - \ . \ . \ . \ - \frac{R_{1n}}{R_{11}} \frac{\sigma_1}{\sigma_n} x_n .$$

Сравнивая это ур-іе съ ур-іемъ линейной регрессіи при нормальномъ распредѣленіи (264), мы найдемъ, что

$$\frac{R_{12}}{R_{11}} \frac{\sigma_1}{\sigma_2} = \frac{a_{12}}{a_{11}}, \qquad \frac{R_{13}}{R_{11}} \frac{\sigma_1}{\sigma_3} = \frac{a_{13}}{a_{11}} \text{ и т. д.,}$$

а сравнивая выраженіе средней ошибки, найденное нами въ § 27 (254), съ только что полученнымъ—(265′), будемъ имѣть:

$$\frac{1}{a_{11}} = \sigma_1^2 \frac{R}{R_{11}} .$$

Отсюда легко находимъ:

$$(266) \quad . \ . \ a_{11} = \frac{R_{11}}{R} \frac{1}{\sigma_1^2} ; \ a_{12} = \frac{R_{12}}{R} \frac{1}{\sigma_1 \sigma_2} ; \ a_{13} = \frac{R_{13}}{R} \frac{1}{\sigma_1 \sigma_3} \text{ и т. д.,}$$

и по аналогіи вообще:

$$(267) \quad . \ . \ . \ . \ a_{ii} = \frac{R_{ii}}{R} \cdot \frac{1}{\sigma_i^2} ; \ a_{ij} = \frac{R_{ij}}{R} \frac{1}{\sigma_i \sigma_j} .$$

Подставляя полученныя величины въ ур-іе нормальнаго распредѣленія (262), мы приведемъ его къ элегантной формѣ:

$$(268) \quad . \; . \; Z = Ce^{-\frac{1}{2}\left(\frac{R_{11}}{R}\frac{x_1^2}{\sigma_1^2} + \frac{R_{22}}{R}\frac{x_2^2}{\sigma_2^2} + \ldots + 2\frac{R_{12}}{R}\frac{x_1 x_2}{\sigma_1 \sigma_2} + \ldots\right)}.$$

Въ этомъ и заключается теорема Edgeworth'a.

Остается найти C.

Въ случаѣ двухъ перемѣнныхъ $R = 1 - r^2$, $R_{11} = 1$, $R_{12} = -r$, и мы имѣемъ:

$$(269) \quad . \; . \; Z = Ce^{-\frac{1}{2(1-r^2)}\left(\frac{x_1^2}{\sigma_1^2} - \frac{2x_1 x_2}{\sigma_1 \sigma_2} r + \frac{x_2^2}{\sigma_2^2}\right)}.$$

Все число случаевъ равно N. Слѣдовательно, если мы проинтегрируемъ выраженіе (269) и по x_1 и по x_2 отъ $-\infty$ до $+\infty$, то мы должны будемъ получить величину N. Такимъ образомъ,

$$N = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} Z dx_1 dx_2 =$$

$$= C\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{1}{2(1-r^2)}\left(\frac{x_1^2}{\sigma_1^2} - \frac{2x_1}{\sigma_1}\cdot r\frac{x_2}{\sigma_2} + r^2\frac{x_2^2}{\sigma_2^2} - r^2\frac{x_2^2}{\sigma_2^2} + \frac{x_2^2}{\sigma_2^2}\right)} . dx_1 dx_2 =$$

$$= C\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{1}{2(1-r^2)}\left(\frac{x_1}{\sigma_1} - r\frac{x_2}{\sigma_2}\right)^2} . e^{-\frac{1}{2}\frac{x_2^2}{\sigma_2^2}} . dx_1 dx_2 =$$

$$= C\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\frac{x_2^2}{\sigma_2^2}} . dx_2 \int_{-\infty}^{+\infty} e^{-\frac{1}{2(1-r^2)}\left(\frac{x_1}{\sigma_1} - r\frac{x_2}{\sigma_2}\right)^2} . dx_1 .$$

Интегрируемъ сначала по x_1. Для этого положимъ

$$\frac{1}{\sqrt{2(1-r^2)}}\left(\frac{x_1}{\sigma_1} - r\frac{x_2}{\sigma_2}\right) = \xi, \text{ откуда}$$

$$dx_1 = \sigma_1 \sqrt{2(1-r^2)}\, d\xi .$$

Слѣдовательно,

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2(1-r^2)}\left(\frac{x_1}{\sigma_1}-r\frac{x_2}{\sigma_2}\right)^2} . dx_1 =$$

$$= \sigma_1 \sqrt{2(1-r^2)} \int_{-\infty}^{+\infty} e^{-\xi^2} d\xi = \sigma_1 \sqrt{2\pi(1-r^2)} \cdot$$

Второй интегралъ найдется подобнымъ-же образомъ:

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\frac{x_2^2}{\sigma_2^2}} . dx_2 = \sqrt{2}\,\sigma_2 \int_{-\infty}^{+\infty} e^{-\left(\frac{x_2}{\sqrt{2}\,\sigma_2}\right)^2} d\left(\frac{x_2}{\sqrt{2}\,\sigma_2}\right) = \sigma_2\sqrt{2\pi}.$$

Такимъ образомъ, весь двойной интегралъ равенъ:

$$C\sigma_1\sigma_2 . 2\pi\sqrt{1-r^2} = N, \text{ откуда}$$

$$C = \frac{N}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}},$$

и ур-іе распредѣленія для двухъ коррелятивно связанныхъ величинъ будетъ:

$$(270) \;.\;.\; Z = \frac{N}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left(\frac{x_1^2}{\sigma_1^2} - \frac{2rx_1x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right)}.$$

Замѣчая, что $\sqrt{1-r^2} = \sqrt{R}$, мы догадываемся, что въ общемъ случаѣ постоянная

$$C = \frac{N}{(2\pi)^{\frac{1}{2}n}\sigma_1\sigma_2 \ldots \sigma_n\sqrt{R}}.$$

Это положеніе справедливо для $n = 2$. Мы докажемъ, что если оно справедливо для числа перемѣнныхъ, равнаго n, то оно будетъ справедливо и для числа ихъ равнаго $n+1$. Такимъ образомъ, это положеніе будетъ доказано вообще.

Обозначимъ извѣстный намъ детерминантъ, составленный изъ коэффиціентовъ корреляціи, черезъ $_nR$ для случая n перемѣнныхъ и черезъ $_{n+1}R$ для случая $n+1$ перемѣннаго.

Пусть для случая $n+1$ перемѣннаго

$$Z_{n+1} = C\, e^{-\frac{1}{2\,_{n+1}R}\left(_{n+1}R_{11}\frac{x_1^2}{\sigma_1^2} + \cdots + _{n+1}R_{n+1,\,n+1}\frac{x_{n+1}^2}{\sigma_{n+1}^2} + 2\,_{n+1}R_{12}\frac{x_1x_2}{\sigma_1\sigma_2} + \cdots\right)}.$$

Это выраженіе нужно проинтегрировать по x_{n+1} между $-\infty$ и $+\infty$, тогда получимъ функцію распредѣленія только для n перемѣнныхъ.

Выраженіе въ показателѣ можно написать такъ:

$$\frac{{}_{n+1}R_{n+1,\,n+1}}{2\,{}_{n+1}R\,\sigma_{n+1}^2}\left(x_{n+1}+U\right)^2-\frac{1}{2}\,V,$$

гдѣ U и V суть многочлены, не содержащіе x_{n+1}, но въ то-же время не содержащіе и членовъ, независимыхъ отъ остальныхъ перемѣнныхъ.

Интегрируя, найдемъ функцію распредѣленія для n перемѣнныхъ:

$$Z_n=\int\limits_{-\infty}^{+\infty} Z_{n+1}\,dx_{n+1}=Ce^{-\frac{1}{2}V}\cdot\int\limits_{-\infty}^{+\infty} e^{\frac{{}_{n+1}R_{n+1,\,n+1}}{2\,{}_{n+1}R\,\sigma_{n+1}^2}(x_{n+1}+U)^2}\cdot dx_{n+1}=$$

$$=C\sqrt{\frac{2\,{}_{n+1}R\,\sigma_{n+1}^2}{{}_{n+1}R_{n+1,\,n+1}}}\cdot\sqrt{\pi}\cdot e^{-\frac{1}{2}V}.$$

Легко замѣтить, что миноръ ${}_{n+1}R_{n+1,\,n+1}=\begin{vmatrix} 1 & r_{12} & r_{13} & \ldots & r_{1n} \\ r_{21} & 1 & r_{23} & \ldots & r_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{n1} & r_{n2} & r_{n3} & \ldots & 1 \end{vmatrix}$

и есть не что иное, какъ ${}_nR$. По допущенію постоянный множитель функціи распредѣленія для n перемѣнныхъ слѣдуетъ предположенному закону. Слѣд.,

$$C\sqrt{\frac{2\pi\,{}_{n+1}R}{{}_nR}}\,\sigma_{n+1}=\frac{N}{(2\pi)^{1/2n}\sigma_1\sigma_2\ldots\sigma_n\sqrt{{}_nR}},\ \text{откуда}$$

$$C=\frac{N}{(2\pi)^{1/2(n+1)}\sigma_1\sigma_2\ldots\sigma_{n+1}\sqrt{{}_{n+1}R}},$$

что и требовалось доказать.

Такимъ образомъ, функція нормальнаго распредѣленія въ окончательномъ видѣ выражается ур-іемъ:

$$(271)\ldots Z=\frac{N}{(2\pi)^{1/2n}\sigma_1\sigma_2\ldots\sigma_n\sqrt{R}}\,e^{-\frac{1}{2}\left\{\sum\frac{R_{ii}}{R}\frac{x_i^2}{\sigma_i^2}+2\sum\frac{R_{ij}}{R}\frac{x_i\,x_j}{\sigma_i\,\sigma_j}\right\}}.$$

§ 31. *О вѣроятности системы коррелятивно связанныхъ между собою отклоненій* [1]).

Въ этомъ параграфѣ я хочу изложить одно изъ наиболѣе блестящихъ и въ то-же время важныхъ приложеній теоріи нормальной корреляціи.

Пусть мы имѣемъ систему величинъ, коррелятивно связанныхъ между собою, со средними ариѳметическими $h_1, h_2, \ldots h_n$, средними отклоненіями $\sigma_1, \sigma_2, \ldots \sigma_n$ и отклоненіями отъ среднихъ значеній $x_1, x_2, \ldots x_n$. Пусть, какъ это во многихъ случаяхъ наблюдается, отклоненія эти будутъ слѣдовать закону Гаусса, а, слѣдовательно, распредѣленіе ихъ будетъ подчиняться нормальному ур-ію распредѣленія (271).

Различныя комбинаціи этихъ величинъ встрѣчаются съ различной частотой, обладаютъ, иначе говоря, различной вѣроятностью. Величина этой вѣроятности зависитъ, очевидно, отъ величины функціи, которая стоитъ въ показателѣ правой части ур-ія (271). Если эта функція остается постоянной, то, какъ бы ни мѣнялись значенія отдѣльныхъ величинъ $x_1, x_2, \ldots x_n$, вѣроятность ихъ комбинаціи отъ этого не измѣнится. Такимъ образомъ ур-іе:

$$(272) \ldots \chi^2 = \sum \frac{R_{ii}}{R} \frac{x_i^2}{\sigma_i^2} + 2 \sum \frac{R_{ij}}{R} \frac{x_i x_j}{\sigma_i \sigma_j}$$

даетъ при постоянномъ χ^2 всѣ возможныя равновѣроятныя значенія нашихъ отклоненій [2]).

Пользуясь для сокращенія выраженіемъ (272), мы ур-іе распредѣленія (271) можемъ короче написать такъ:

$$(273) \ldots\ldots Z = Z_0 e^{-\frac{1}{2}\chi^2}.$$

Вспомнимъ значеніе Z. Это есть выраженіе для той *частоты*, съ которой встрѣчаются во всей совокупности комбинаціи вели-

---

[1]) K. Pearson, On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling, Philosophical Magazine, Vol. 50, 1900, p. 157—175.

[2]) Геометрически (272) представляетъ собою ур-іе обобщеннаго эллипсоида равновѣроятности въ пространствѣ n измѣреній. Чтобы для вообраafженія получить точку опоры, достаточно ограничиться случаемъ трехъ перемѣнныхъ и представлять себѣ дальше при всѣхъ разсужденіяхъ обыкновенный эллипсоидъ. Значенія x_1, x_2, x_3 (вообще $x_1, x_2, \ldots x_n$), соотвѣтствующія любой точкѣ поверхности эллипсоида, — равновѣроятны.

чинъ $x_1, x_2, \ldots x_n$, удовлетворяющія ур-ію (272), т. е. всѣ равновѣроятныя (одинаково частыя) комбинаціи, характеризуемыя опредѣленнымъ значеніемъ величины χ^2. Слѣдовательно, *число* случаевъ, когда первая величина заключается между x_1 и x_1+dx_1, вторая между x_2 и x_2+dx_2 и т. д., равняется

$$(274) \; . \; . \; . \; . \; . \; dN = Z_0 e^{-\frac{1}{2}\chi^2} dx_1 dx_2 \ldots dx_n.$$

Поставимъ вопросъ: какова вѣроятность того, чтобы x_1, $x_2, \ldots x_n$ имѣли какія угодно значенія столь и еще менѣе вѣроятныя, чѣмъ тѣ, какія они въ данномъ случаѣ имѣютъ. Этотъ вопросъ сводится къ другому: какова вѣроятность всѣхъ комбинацій значеній $x_1, x_2, \ldots x_n$, при которыхъ χ^2 имѣетъ значенія равныя и большія даннаго.

Вѣроятность событія равна суммѣ благопріятствующихъ статочностей, дѣленной на всю сумму статочностей. Искомая вѣроятность равна, слѣд., числу всѣхъ случаевъ, когда χ^2 равно и больше данной своей величины, дѣленному на число всѣхъ случаевъ въ совокупности. Первое число найдемъ, если просуммируемъ выраженія, подобныя (274), для всѣхъ значеній $x_1, x_2, \ldots x_n$, начиная съ тѣхъ, при которыхъ χ равно данной величинѣ, и переходя ко всѣмъ тѣмъ, когда χ больше этой величины, т. е. отъ χ даннаго до $\chi=\infty$. Второе число найдемъ, если просуммируемъ выраженія, подобныя (274), для всѣхъ возможныхъ комбинацій значеній $x_1, x_2, \ldots x_n$, начиная съ тѣхъ, при которыхъ $\chi=0$, кончая тѣми, при которыхъ $\chi=\infty$. Такимъ образомъ, искомая вѣроятность будетъ равняться:

$$(275) \; . \; . \; . \; P = \frac{\left[\iiint \ldots \int e^{-\frac{1}{2}\chi^2} dx_1 dx_2 \ldots dx_n\right]_{\chi}^{\infty}}{\left[\iiint \ldots \int e^{-\frac{1}{2}\chi^2} dx_1 dx_2 \ldots dx_n\right]_{0}^{\infty}}.$$

Чтобы упростить это выраженіе, обратимся къ геометрическому представленію (см. прим. 2-е на стр. 181). $dx_1 dx_2 \ldots dx_n$ есть элементарный объемъ, его нужно помножить на $e^{-\frac{1}{2}\chi^2}$ (скажемъ, на плотность) и взять сумму такихъ выраженій, начиная съ поверхности эллипсоида χ и до безконечности. Это суммированіе можно произвести въ такомъ порядкѣ, чтобы сначала найти сумму подобныхъ выраженій для тонкаго эллипсоидальнаго слоя между эллипсоидомъ χ и эллипсоидомъ $\chi+d\chi$. Назовемъ объемъ этого слоя dV. Тогда, очевидно, масса слоя будетъ равна $e^{-\frac{1}{2}\chi^2}dV$, и все выраженіе (275) сведется къ простымъ интеграламъ, т. что мы будемъ имѣть:

$$(276) \quad . \; . \; . \; . \; . \; . \; . \; . \quad P=\frac{\int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2}dV}{\int_{0}^{\infty} e^{-\frac{1}{2}\chi^2}dV}.$$

dV легко найти изъ слѣдующихъ соображеній: если (272) раздѣлить на χ^2, то эта величина войдетъ въ знаменатель каждаго члена каждой суммы. Слѣдовательно, каждый *линейный* размѣръ эллипсоида пропорціоналенъ χ, а объемъ пропорціоналенъ χ^n и равняется $V=C\chi^n$. Отсюда $dV=Cn\chi^{n-1}d\chi$. Подставляя въ выраженіе (276) и замѣчая, что Cn войдетъ въ числителя и въ знаменателя, слѣдовательно, сократится, получимъ:

$$(277) \quad . \; . \; . \; . \; . \; . \; . \; . \quad P=\frac{\int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2}\chi^{n-1}d\chi}{\int_{0}^{\infty} e^{-\frac{1}{2}\chi^2}\chi^{n-1}d\chi}.$$

Такова вѣроятность того, что вслѣдствіе случайныхъ обстоятельствъ можетъ имѣть мѣсто система отклоненій столь-же или еще менѣе вѣроятная, чѣмъ данная.

Выраженіе (277) можно упростить. Интегрируя по частямъ, получимъ:

$$(278) \; . \; . \; . \int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2}\chi^{n-1}d\chi=$$

$$=[\chi^{n-2}+(n-2)\chi^{n-4}+(n-2)(n-4)\chi^{n-6}+\ldots$$

$$+(n-2)(n-4)(n-6)\ldots(n-\overline{2r-2})\chi^{n-2r}]e^{-\frac{1}{2}\chi^2}+$$

$$+(n-2)(n-4)(n-6)\ldots(n-2r)\int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2}\chi^{n-2r-1}.d\chi=$$

$$=(n-2)(n-4)(n-6)\ldots(n-2r)\left\{\int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2}\chi^{n-2r-1}.d\chi+\right.$$

$$+e^{-\frac{1}{2}\chi^2}\left[\frac{\chi^{n-2r}}{n-2r}+\frac{\chi^{n-2r+2}}{(n-2r)(n-2r+2)}+\right.$$

$$+\frac{\chi^{n-2r+4}}{(n-2r).(n-2r+2)(n-2r+4)}+\ldots$$

$$\left.\left.+\frac{\chi^{n-2}}{(n-2r)(n-2r+2)\ldots\ldots(n-2)}\right]\right\}.$$

Далѣе:

$$(279)\ .\ .\ .\int_0^\infty e^{-\frac{1}{2}\chi^2}\chi^{n-1}d\chi=$$

$$=(n-2)(n-4)(n-6)\ldots(n-2r)\int_0^\infty e^{-\frac{1}{2}\chi^2}\chi^{n-2r-1}d\chi.$$

Величина n можетъ быть четной или нечетной.

Случай I, n нечетное. Положивъ $r=\frac{n-1}{2}$, получимъ:

$$P=\frac{\int_\chi^\infty e^{-\frac{1}{2}\chi^2}d\chi+e^{-\frac{1}{2}\chi^2}\left\{\frac{\chi}{1}+\frac{\chi^3}{1.3}+\frac{\chi^5}{1.3.5}+\ldots+\frac{\chi^{n-2}}{1.3.5\ldots(n-2)}\right\}}{\int_0^\infty e^{-\frac{1}{2}\chi^2}d\chi}.$$

Но, такъ какъ

$$\int_0^\infty e^{-\frac{1}{2}\chi^2}d\chi=\sqrt{\frac{\pi}{2}},$$

то

$$(280)\ .\ .\ .\ P=\sqrt{\frac{2}{\pi}}\int_\chi^\infty e^{-\frac{1}{2}\chi^2}d\chi+$$

$$+\sqrt{\frac{2}{\pi}}e^{-\frac{1}{2}\chi^2}\left(\frac{\chi}{1}+\frac{\chi^3}{1.3}+\frac{\chi^5}{1.3.5}+\ldots+\frac{\chi^{n-2}}{1.3.5\ldots(n-2)}\right).$$

Первый членъ этой формулы можетъ быть найденъ изъ таблицъ интеграла вѣроятностей. Въ обозначеніи таблицъ Шеппарда (Леонтовичъ ч. III) онъ равняется $(1-\alpha)$. Таблицы даютъ непосредственно $\frac{1}{2}(1+\alpha)$, откуда, вычитая изъ единицы табличное число, находимъ $\frac{1}{2}(1-\alpha)$, и умножая на два, получаемъ искомую величину. Чтобы найти эту-же величину изъ обычныхъ таблицъ, напр. тѣхъ, что даны у А. А. Чупрова въ Очеркахъ по теоріи статистики или у Акад. Маркова въ Исчисленіи вѣроятностей, нужно сначала раздѣлить χ на $\sqrt{2}$, затѣмъ по полученному числу найти табличную вѣроятность и вычесть ее изъ единицы. Если χ велико и не содержится въ таблицахъ, то вычисленіе можно произвести, воспользовавшись слѣдующимъ рядомъ, небольшое число членовъ котораго даетъ при большомъ χ достаточное приближеніе:

$$\sqrt{\frac{2}{\pi}}\int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2}d\chi = \sqrt{\frac{2}{\pi}}\, e^{-\frac{1}{2}\chi^2}\left(\frac{1}{\chi}-\frac{1}{\chi^3}+\frac{3}{\chi^5}-\frac{3.5}{\chi^7}+\frac{3.5.7}{\chi^9}-\ldots\right).$$

Случай II, n четное. Положивъ въ формулахъ (278) и (279) $r=\frac{1}{2}n-1$, получимъ:

$$P=\frac{\int_{\chi}^{\infty} e^{-\frac{1}{2}\chi^2}\chi d\chi + e^{-\frac{1}{2}\chi^2}\left\{\frac{\chi^2}{2}+\frac{\chi^4}{2.4}+\ldots+\frac{\chi^{n-2}}{2.4\ldots(n-2)}\right\}}{\int_0^{\infty} e^{-\frac{1}{2}\chi^2}\chi d\chi},$$

откуда:

$$(281)\quad \ldots \quad P=e^{-\frac{1}{2}\chi^2}\left(1+\frac{\chi^2}{2}+\frac{\chi^4}{2.4}+\frac{\chi^6}{2.4.6}+\ldots+\right.$$
$$\left.+\frac{\chi^{n-2}}{2.4.6\ldots(n-2)}\right).$$

Вычисленіе формулъ (280) и (281) не представляетъ послѣ сказаннаго никакихъ затрудненій кромѣ чисто ариѳметическихъ, но и ихъ въ большинствѣ случаевъ можно избѣжать, воспользовавшись таблицами, составленными P. Elderton'омъ [1]), въ которыхъ даны значенія P съ достаточной точностью и для большинства приложеній вполнѣ достаточной полнотой. Таблицы вычислены, впрочемъ, при нѣсколько иныхъ предпосылкахъ. Именно, мы до сихъ поръ принимали, что всѣ n величинъ, хотя коррелятивно связаны, но — въ строго функціональномъ смыслѣ — независимы другъ отъ друга. Таблицы-же Elderton'a составлены при предположеніи, что имѣются n' величинъ, связанныхъ однимъ ур-іемъ, слѣдовательно, изъ нихъ независимыхъ будетъ всего $n'-1$. Если наши n величинъ независимы, то, найдя χ^2 по формулѣ (272), мы должны взять $n'=n+1$ и искать въ таблицахъ Elderton'a величину вѣроятности, соотвѣтствующую χ^2 и n'.

---

[1]) Biometrica, Vol. I, p. 155—163. Перепечатаны у Леонтовича, op. cit., ч. I, табл. VII.

§ 32. *Критерій соотвѣтствія теоретическаго распредѣленія эмпирическому* [1]).

Пусть мы имѣемъ нѣкоторую совокупность, и пусть индивидуумы ея по величинѣ какого нибудь признака будутъ распредѣлены на $n+1$ группъ. Изъ каждыхъ N индивидуумовъ пусть на каждую группу приходится въ среднемъ

$$\mu_1,\ \mu_2,\ \mu_3,\ \ldots\ \mu_{n+1}$$

индивидуумовъ. Если мы однако возьмемъ изъ общей (генеральной) совокупности пробную группу въ N индивидуумовъ, то численности отдѣльныхъ группъ окажутся вообще иными. Пусть они будутъ

$$m_1,\ m_2,\ m_3,\ \ldots\ m_{n+1}.$$

Въ среднемъ для множества пробныхъ группъ окажется, что средняя всѣхъ m_i равна μ_i, но въ отдѣльномъ случаѣ мы должны получить ошибки:

$$e_1 = m_1 - \mu_1,\ e_2 = m_2 - \mu_2,\ \ldots\ e_{n+1} = m_{n+1} - \mu_{n+1}.$$

Изъ этихъ ошибокъ только n будутъ независимы, т. к. очевидно, что

$$e_1 + e_2 + \ldots + e_n + e_{n+1} = 0,$$

т. что, зная n ошибокъ, $(n+1)$-ую можно опредѣлить. Поэтому, примѣняя формулы предыдущаго параграфа, нужно считаться только съ n перемѣнными.

Не трудно доказать, что среднее отклоненіе величины e_i при случайныхъ варіаціяхъ, испытываемыхъ m_i, будетъ равняться:

$$(282)\ \ldots\ldots\ \sigma_i = \sqrt{N \cdot \frac{\mu_i}{N} \cdot \left(1 - \frac{\mu_i}{N}\right)},\ ^{2)}$$

---

[1]) K. Pearson, On the Criterion etc., l. cit., p. 160 и сл.

[2]) Вѣроятность извлечь изъ генеральной совокупности индивидуума i-ой группы равна $p = \frac{\mu_i}{N}$, вѣроятность извлечь иного индивидуума — $q = 1 - p = 1 - \frac{\mu_i}{N}$. Вѣроятность того, что *число* индивидуумовъ i-ой группы отклонится отъ наивѣроятнѣйшаго значенія своего μ_i, опредѣляется интеграломъ Лапласа при модулѣ $M = \sqrt{2npq}$ (см. А. А. Чупровъ, Очерки по теоріи статистики, 1909, стр. 211—220; у Чупрова, впрочемъ,

и что коэффиціентъ корреляціи между ошибками опредѣлится по формулѣ:

$$(283) \dots\dots\dots \sigma_i \sigma_j r_{ij} = -\frac{\mu_i \mu_j}{N}. \text{ [1]}$$

Теперь нужно подставить величины $r_{12}, r_{13}, \dots r_{ij} \dots$, опредѣленныя формулами (282) и (283), въ выраженіе для детерминанта R (246), привести его къ болѣе простому виду и опредѣлить всѣ его миноры R_{ii} и R_{ij}. Затѣмъ найденныя величины нужно подставить въ выраженіе для χ^2 (см. формулу 272). Эта работа была выполнена Пирсономъ и привела къ чрезвычайно простому результату [2]). Оказалось, что для данной задачи:

$$(284) \dots\dots\dots \chi^2 = \sum \left(\frac{e^2}{\mu}\right),$$

данъ не модуль абсолютной численности, а модуль относительной частоты $\left(\sqrt{\frac{2pq}{n}}\right)$. Раздѣливъ модуль на $\sqrt{2}$, получаемъ среднее отклоненіе $\sqrt{npq}$, что даетъ формулу (282).

[1]) По основной формулѣ теоріи корреляціи $N\sigma_i\sigma_j r_{ij} = \Sigma\delta\mu_i\delta\mu_j$. Но если отклоненія случайны, то избыточное число индивидуумовъ, попавшее въ i-ую группу *въ среднемъ* должно распредѣлиться пропорціонально между остальными группами. Поэтому будемъ имѣть $\delta\mu_j = -\delta\mu_i \cdot \frac{\mu_j}{N-\mu_i}$ (равенство приближенное, основанное на незначительности отклоненій). Слѣд.,

$$\sigma_i \sigma_j r_{ij} = -\frac{1}{N}\sum\left(\delta\mu_i^2 \cdot \frac{\mu_j}{N-\mu_i}\right) = -\sigma_i^2 \cdot \frac{\mu_j}{N-\mu_i} =$$

$$= -\mu_i\left(1-\frac{\mu_i}{N}\right) \cdot \frac{\mu_j}{N-\mu_i} = -\frac{\mu_i\mu_j}{N}.$$

[2]) Для вычисленія детерминанта R и его миноровъ Пирсонъ употребилъ подстановку: $\frac{\mu_i}{N} = Sin^2\beta_i$, гдѣ β_i—вспомогательная величина. Тогда $\sigma_i = \sqrt{N}\, Sin\beta_i Cos\beta_i$, $r_{ij} = -tg\beta_i tg\beta_j$. Легко найти, что при этомъ условіи $R = (-1)^n tg^2\beta_1 tg^2\beta_2 \dots tg^2\beta_n \cdot J$, гдѣ J есть детерминантъ, діагональная строка котораго будетъ $-ctg^2\beta_1, -ctg^2\beta_2 \dots -ctg^2\beta_n$, а всѣ остальные элементы равны единицѣ. Задача сводится такимъ образомъ къ вычисленію детерминанта J и его миноровъ J_{11}, J_{12} и т. д. Полагая $\eta_q = ctg^2\beta_q = \frac{N}{\mu_q} - 1$, Пирсонъ находитъ $J_{ij} = (-1)^{n-1}\frac{\lambda}{(\eta_i+1)(\eta_j+1)}$, $J = (-1)^n \lambda \frac{\mu_{n+1}}{N}$,

гдѣ суммированіе нужно распространить на всѣ $(n+1)$ ошибокъ. Если обозначимъ число группъ черезъ n', то величины χ^2 и n' дадутъ возможность найти искомую вѣроятность изъ таблицы P. Elderton'a. Если величина χ^2 или n' выходитъ за предѣлы таблицы, то нужно найти число *независимыхъ* ошибокъ $n=n'-1$, и тогда P найдется по одной изъ формулъ (280) или (281).

Примѣры. (1) Такъ называемый законъ большихъ чиселъ неоднократно провѣрялся различными опытами. Такъ напр., проф. Вельдонъ (Weldon) приводитъ слѣдующія данныя. 12 игральныхъ костей были брошены 26306 разъ. Каждый разъ считали сколько изъ 12 костей показали 5 *или* 6 очковъ. Результаты сведены въ въ слѣдующей таблицѣ [1]):

Число костей съ 5-ю или 6-ю очками при одномъ бросаніи.	Теоретическое число случаевъ, μ.	Наблюдавшееся число случаевъ, m.	Отклоненіе эмпирическаго числа отъ теоретическ., $e=m-\mu$.
0	203	185	-18
1	1217	1149	-68
2	3345	3265	-80
3	5576	5475	-101
4	6273	6114	-159
5	5018	5194	$+176$
6	2927	3067	$+140$
7	1254	1331	$+77$
8	392	403	$+11$
9	87	105	$+18$
10	13	14	$+1$
11	1	4	$+3$
12	0	0	0
Всего	26306	26306	

$$J_{ii}=(-1)^{n-1}\frac{\lambda}{1+\eta_i}\left(\frac{\mu_i}{N}+\frac{\mu_{n+1}}{N}\right),\text{ гдѣ } \lambda=(\eta_1+1)(\eta_2+1)\ldots(\eta_n+1).$$

Затѣмъ уже нетрудно найти $\frac{R_{ii}}{R}\frac{1}{\sigma_i^2}=\frac{1}{\mu_i}+\frac{1}{\mu_{n+1}}$ и $\frac{R_{ij}}{R}\frac{1}{\sigma_i\sigma_j}=\frac{1}{\mu_{n+1}}$.

Теперь нахожденіе χ^2 не представляетъ уже никакого труда.

[1]) Цитирую по Пирсону, On the Criterion etc., l. cit., p. 167.

Имѣя такія результаты, спросимъ себя, насколько велико соотвѣтствіе опыта съ теоріей. Можно-ли объяснить полученныя отклоненія одними случайными причинами? Это такой вопросъ, который не поддается глазомѣрному рѣшенію и требуетъ примѣненія критерія Пирсона.

Вычисленіе χ^2 расположено въ слѣдующей таблицѣ (замѣчу, что число знаковъ, взятыхъ Пирсономъ, судя по результату, очевидно, больше, чѣмъ можетъ требоваться при такихъ вычисленіяхъ. Обыкновенно достаточно вычислять χ^2 съ двумя, самое большее съ тремя десятичными знаками).

Группа.	e^2.	e^2/μ.
0	324	1,59606
1	4624	3,79951
2	6400	1,91330
3	10201	1,82945
4	25281	4,03013
5	30976	6,17298
6	19600	6,69628
7	5929	4.72807
8	121	0,30903
9	324	3,72414
10	1	0,07346
11	9	9,00000
12	0	0,00000
Всего		43,87241

Такимъ обр., $\chi^2 = 43{,}87241$ и $\chi = 6{,}623.625$.

Т. к. здѣсь 13 группъ, то независимыхъ перемѣнныхъ у насъ 12, и мы должны примѣнить формулу (281):

$$P = e^{-\frac{1}{2}\chi^2}\left(1 + \frac{\chi^2}{2} + \frac{\chi^4}{2.4} + \frac{\chi^6}{2.4.6} + \frac{\chi^8}{2.4.6.8} + \frac{\chi^{10}}{2.4.6.8.10}\right),$$

которая даетъ намъ:

$$P = 0{,}000016.$$

Это значитъ, что вѣроятность системы отклоненій не болѣе вѣроятной, чѣмъ данная, чрезвычайно ничтожна. Если бы опытъ былъ повторенъ очень много разъ при идеальныхъ условіяхъ, то 62499 разъ мы имѣли бы меньшія отклоненія и только 1 разъ такія или большія. Итакъ можно, поставивъ 62499 противъ 1, т. е. почти съ абсолютной увѣренностью, утверждать, что кости, съ которыми продѣлывался опытъ, сдѣланы неточно, т. что выходъ каждаго очка неодинаково вѣроятенъ. Читатель видитъ, что „провѣрка" теоремы теоріи вѣроятности превратилась въ испытаніе физическихъ свойствъ употреблявшихся при этомъ опытѣ игральныхъ костей, испытаніе, которое вѣроятно потребовало бы меньше времени, если бы его произвести при помощи методовъ измѣрительной физики.

(2) Выше не однократно упоминалось, что различные размѣры человѣческаго тѣла даютъ въ своемъ распредѣленіи картину обыкновенно достаточно близкую къ теоретическому нормальному распредѣленію („законъ" Гаусса). Я хочу поэтому привести изъ этой области одинъ примѣръ.

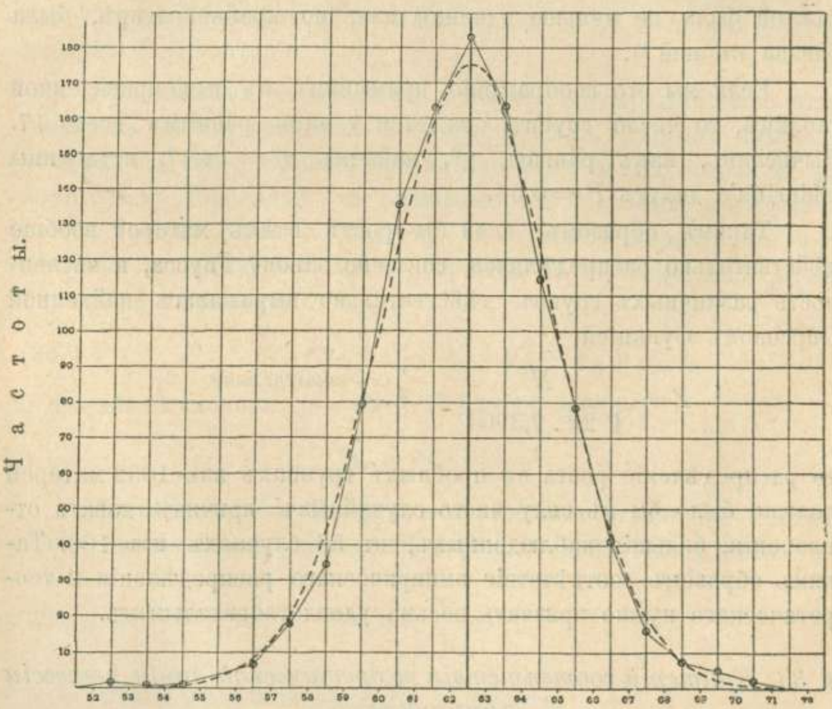
Чер. 30 показываетъ распредѣленіе по росту 1052 матерей по даннымъ Пирсона [1]). Кружками и непрерывной чертой отмѣчено фактическое распредѣленіе, пунктиромъ—теоретическая кривая Гаусса. Слѣдующая таблица содержитъ всѣ данныя, необходимыя для вычисленія критерія соотвѣтствія.

Ростъ въ дюймахъ.	Фактич. численность.	Теоретич. численность.	Ростъ въ дюймахъ.	Фактич. численность.	Теоретич. численность.
52—53	1,5	0,9	62—63	183	174,3
53—54	0,5		63—64	163	159,4
54—55	1		64—65	114,5	122,8
55—56	2	2,6	65—66	78,5	79,5
56—57	6,5	7,9	66—67	41	43,2
57—58	18	20,9	67—68	16	20,1
58—59	34,5	44,5	68—69	7,5	7,7
59—60	79,5	80,8	69—70	4,5	2.5
60—61	135,5	124,1	70—71	2	0,8
61—62	163	160,3			

[1]) Относительно чертежа и таблицы см. K. Pearson and A. Lee, On the Laws of Inheritance in Man. Biometrica, Vol. II, p. 364—365.

Къ этой таблицѣ нужно сдѣлать одно замѣчаніе. Всякая теоретическая кривая частоты представляетъ собою *непрерывную* функцію. Этимъ самымъ она уклоняется отъ дѣйствительности, ибо число дѣйствительныхъ случаевъ въ любой группѣ не можетъ быть меньше единицы, и только условно, когда по своему размѣру индивидуумъ приходится какъ разъ на линіи раздѣла двухъ группъ, мы разбиваемъ его пополамъ и засчитываемъ въ каждую группу по половинѣ. Теоретическая-же численность группы можетъ быть и $^1/_{10}$ и $^1/_{100}$ и $^1/_{1000}$ и меньше.

Чер. 30. Распредѣленіе англійскихъ матерей по росту.



Ростъ матерей въ дюймахъ.

Эту свою слабую сторону теоретическія кривыя и обнаруживаютъ въ крайнихъ группахъ всякаго распредѣленія. Поэтому теоретическія частоты въ рядѣ крайнихъ группъ, вродѣ 0,5, 0,3, 0,2, нужно интерпретировать т. обр., что въ первомъ изъ этихъ подраздѣленій (0,5) въ рядѣ пробныхъ группъ, подобныхъ данной, одинаково часто должно встрѣчаться или 1, или 0 индивидуумовъ, во второмъ подраздѣленіи (0,3) на 10 такихъ-же пробныхъ группъ—только 3 индивидуума и т. д. А всего на каждую пробную группу должно въ среднемъ приходиться по одному индивидууму $(0,5 + 0,3 + 0,2)$ на всѣ три крайнихъ подраздѣленія.

Вслѣдствіе такого чисто условнаго значенія дробныхъ величинъ въ крайнихъ группахъ было бы грубой ошибкой прилагать и къ нимъ критерій Пирсона, считаясь съ каждой изъ нихъ въ отдѣльности. Очевидно, что для того, чтобы обезпечить сравнимость теоретическаго распредѣленія съ эмпирическимъ, необходимо — для примѣненія указаннаго критерія — объединять индивидуумовъ въ такія группы, чтобы теоретическая численность каждой была не меньше единицы или, по крайней мѣрѣ, была близка къ ней [1]).

Если мы это соображеніе примѣнимъ къ вышеприведенной таблицѣ, то число группъ окажется у насъ равнымъ всего 17. Вычисливъ, какъ раньше, χ^2, найдемъ $\chi^2 = 14,47$, и таблицы Elderton'а дадутъ $P = 0,56$.

Такимъ образомъ, если бы ростъ всѣхъ матерей вообще дѣйствительно распредѣлялся согласно закону Гаусса, и численность различныхъ группъ дѣйствительно выражалась найденной Пирсономъ функціей

$$Z = \frac{N}{\sqrt{2\pi} \,.\, 2,3904}\, e^{-\frac{1}{2}(x - 62,484)^2/(2,3904)^2}, \quad ^{2})$$

то распредѣленіе роста въ пробныхъ группахъ изъ 1052 матерей должно было бы въ силу чисто случайныхъ причинъ давать отклоненія, бо́льшія наблюденныхъ, въ 56 случаяхъ изъ 100. Такимъ образомъ соотвѣтствіе эмпирическаго распредѣленія и теоретическаго нужно признать весьма удовлетворительнымъ.

§ 33. *Критерій соотвѣтствія теоретической линіи регрессіи эмпирической.*

Если величины не находятся между собой въ корреляціи, то всѣ коэффиціенты корреляціи равны нулю, детерминантъ R и

---

[1]) См. K. Pearson, On the Criterion etc., l. cit., p. 164.

[2]) Средній ариѳметическій ростъ равнялся 62,484 д., и среднее отклоненіе—2,3904 д.

его миноры типа R_{ii} равны единицѣ, а миноры типа R_{ij} равны нулю. Ур-іе распредѣленія въ этомъ случаѣ будетъ

$$(285) \quad \ldots\ldots\ldots Z = Z_0 e^{-\frac{1}{2}\sum\frac{x_i^2}{\sigma_i^2}},$$

а слѣдовательно

$$(286) \quad \ldots\ldots\ldots \chi^2 = \sum \frac{x_i^2}{\sigma_i^2}.$$

Вѣроятность всѣхъ, не болѣе, чѣмъ данная, вѣроятныхъ системъ отклоненій, опредѣлится по этому значенію χ^2 и по n изъ тѣхъ-же формулъ (280) и (281) или по χ^2 и $n' = n + 1$ изъ таблицъ P. Elderton'a.

Приведенныя соображенія имѣютъ непосредственное отношеніе къ кривой регрессіи. Въ самомъ дѣлѣ, Пирсонъ доказалъ, что ошибки въ среднемъ ариѳметическомъ одного строя не находятся въ корреляціи съ ошибками въ среднемъ ариѳметическомъ другого строя [1]). Поэтому, чтобы найти вѣроятность соотвѣтствія теоретической линіи регрессіи съ эмпирической, нужно вычислить отклоненія

$$(287) \quad \ldots\ldots e_1 = y_{x_1} - Y_1,\ e_2 = y_{x_2} - Y_2 \text{ и т. д.}$$

и среднія отклоненія y-а въ отдѣльныхъ строяхъ:

$$\sigma_{n_{x_1}},\ \sigma_{n_{x_2}} \text{ и т. д.}$$

Среднее отклоненіе e_i-го есть среднее отклоненіе средней ариѳметической и по извѣстной формулѣ (153) будетъ равно:

$$(288) \quad \ldots\ldots\ldots \sigma_{e_i} = \frac{\sigma_{n_{x_i}}}{\sqrt{n_{x_i}}}.$$

Отсюда величина критерія χ^2 будетъ:

$$(289) \quad \ldots\ldots\ldots \chi^2 = \sum \frac{n_{x_i} e_i^2}{\sigma_{n_{x_i}}^2}.$$

Теперь мы видимъ, что рѣшеніе, которое мы дали выше въ § 18 для задачи нахожденія кривой регрессіи, теоретически наи-

---

[1]) K. Pearson, On the General Theory of Skew Correlation etc., p. 13.

болѣе правильно. Въ самомъ дѣлѣ, изъ всѣхъ кривыхъ даннаго типа та будетъ наиболѣе вѣроятной кривой регрессіи, для которой χ^2 (289) будетъ имѣть наименьшую величину, т. е. та, которая удовлетворяетъ условію:

$$(290) \quad . \; . \; . \; . \; . \; . \; . \quad \sum \frac{n_{x_i}}{\sigma^2_{n_{x_i}}} (y_{x_i} - Y)^2 = \text{minimum}.$$

Но это и есть то условіе, на основаніи котораго мы предложили находить коэффиціенты кривой.

Дополнительныя замѣчанія.

I.

О терминологіи.

Русская терминологія въ области изложенныхъ ученій не можетъ еще почитаться установившейся. Вотъ почему не будутъ, можетъ быть, безполезными нѣкоторыя замѣчанія относительно отдѣльныхъ терминовъ.

Прежде всего слѣдуетъ обсудить наименованіе тѣхъ статистическихъ кривыхъ, которыя служатъ для изображенія *распредѣленія* статистическаго матеріала по величинѣ какого-либо признака. Въ англійской литературѣ за ними укрѣпилось названіе – frequency curves. Это выраженіе на русскій языкъ переводится различно: (a) кривыя распредѣленія частотъ (Орженцкій)[1], (b) кривыя частоты (Леонтовичъ)[2], (c) кривыя частости (Некрасовъ)[3], (d) кривыя повтореній (Лахтинъ)[4], (e) кривыя варіацій явленій (Леонтовичъ)[5]. Что касается, прежде всего, послѣдняго выраженія, то оно, не говоря уже о его длинѣ, не точно, т. к. не передаетъ основной идеи понятія. Въ выраженіи, употребляемомъ г. Лахтинымъ слишкомъ отчетливо звучитъ динамическій оттѣнокъ, между тѣмъ какъ frequency curves вполнѣ пригодны и для изображенія статики явленій. Въ самомъ дѣлѣ, со словомъ „повтореніе" слишкомъ сильно связывается представленіе о послѣдовательности *во времени*, чтобы можно было признать этотъ терминъ достаточно адэкватнымъ соотвѣтствующему понятію. Названіе этихъ кривыхъ „кривыми частости" не подходитъ потому, что въ русской литературѣ „*частостью*" принято называть отвлеченное число, аналогичное вѣроятности и равняющееся отношенію числа выходовъ явленія къ числу всѣхъ испытаній[6]). Между тѣмъ frequency curves строятся

---

[1]) Op. cit., стр. 215.

[2]) Op. cit., ч. II, стр. 8.

[3]) Op. cit., стр. 499.

[4]) Л. К. Лахтинъ, О методѣ Пирсона въ приложеніяхъ теоріи вѣроятностей къ задачамъ статистики и біологіи. Математ. Сборникъ, издаваемый Моск. Матем. Обществомъ, т. XXIV, 1903 г., стр. 481—500.

[5]) Op. cit., ч. II, стр. 14.

[6]) См. А. А. Чупровъ, Очерки по теоріи статистики, напр., стр. 208 и П. А. Некрасовъ, Теорія вѣроятностей, стр. 453.

болѣе правильно. Въ самомъ дѣлѣ, изъ всѣхъ кривыхъ даннаго типа та будетъ наиболѣе вѣроятной кривой регрессіи, для которой χ^2 (289) будетъ имѣть наименьшую величину, т. е. та, которая удовлетворяетъ условію:

$$(290) \quad \sum \frac{n_{x_i}}{\sigma_{n_{x_i}}^2}(y_{x_i} - Y)^2 = \text{minimum}.$$

Но это и есть то условіе, на основаніи котораго мы предложили находить коэффиціенты кривой.

Дополнительныя замѣчанія.

I.

О терминологіи.

Русская терминологія въ области изложенныхъ ученій не можетъ еще почитаться установившейся. Вотъ почему не будутъ, можетъ быть, безполезными нѣкоторыя замѣчанія относительно отдѣльныхъ терминовъ.

Прежде всего слѣдуетъ обсудить наименованіе тѣхъ статистическихъ кривыхъ, которыя служатъ для изображенія *распредѣленія* статистическаго матеріала по величинѣ какого-либо признака. Въ англійской литературѣ за ними укрѣпилось названіе – frequency curves. Это выраженіе на русскій языкъ переводится различно: (a) кривыя распредѣленія частотъ (Орженцкій)[1]), (b) кривыя частоты (Леонтовичъ)[2]), (c) кривыя частости (Некрасовъ)[3]), (d) кривыя повтореній (Лахтинъ)[4]), (e) кривыя варіацій явленій (Леонтовичъ)[5]). Что касается, прежде всего, послѣдняго выраженія, то оно, не говоря уже о его длинѣ, не точно, т. к. не передаетъ основной идеи понятія. Въ выраженіи, употребляемомъ г. Лахтинымъ слишкомъ отчетливо звучитъ динамическій оттѣнокъ, между тѣмъ какъ frequency curves вполнѣ пригодны и для изображенія статики явленій. Въ самомъ дѣлѣ, со словомъ „повтореніе" слишкомъ сильно связывается представленіе о послѣдовательности *во времени,* чтобы можно было признать этотъ терминъ достаточно адэкватнымъ соотвѣтствующему понятію. Названіе этихъ кривыхъ „кривыми частости" не подходитъ потому, что въ русской литературѣ „*частостью*" принято называть отвлеченное число, аналогичное вѣроятности и равняющееся отношенію числа выходовъ явленія къ числу всѣхъ испытаній[6]). Между тѣмъ frequency curves строятся

---

[1]) Op. cit., стр. 215.

[2]) Op. cit., ч. II, стр. 8.

[3]) Op. cit., стр. 499.

[4]) Л. К. Лахтинъ, О методѣ Пирсона въ приложеніяхъ теоріи вѣроятностей къ задачамъ статистики и біологіи. Математ. Сборникъ, издаваемый Моск. Матем. Обществомъ, т. XXIV, 1903 г., стр. 481—500.

[5]) Op. cit., ч. II, стр. 14.

[6]) См. А. А. Чупровъ, Очерки по теоріи статистики, напр., стр. 208 и П. А. Некрасовъ, Теорія вѣроятностей, стр. 453.

на абсолютныхъ числахъ, и только знаніе этихъ послѣднихъ даетъ возможность находить вѣроятныя ошибки постоянныхъ распредѣленія.

Наиболѣе точнымъ переводомъ англійскаго термина является, повидимому, выраженіе — „кривая частоты". Однако болѣе внимательное отношеніе къ англійскому словоупотребленію вскрываетъ довольно существенное различіе между русскимъ и англійскимъ терминами. Такъ напр., англичане говорятъ о frequencies различныхъ группъ, подразумѣвая подъ этимъ *численности* ихъ, т. е. числа элементовъ приходящихся на каждую группу. Тотъ-же оттѣнокъ заключается и въ выраженіи frequency distribution, или distribution of frequencies. Отдѣльныя части площади кривой изображаютъ собою по англійской терминологіи frequencies отдѣльныхъ группъ, что нельзя передать на русскій языкъ словомъ „частоты" безъ слишкомъ большой натяжки. Въ большемъ соотвѣтствіи съ духомъ русскаго языка будетъ находиться, какъ мнѣ кажется, терминологія, предлагаемая въ настоящей работѣ и различающая: *численности* группъ, изображаемыя площадями кривой и *частоты*, изображаемыя ординатами той-же кривой. Такъ какъ кривая показываетъ распредѣленіе общаго числа элементовъ совокупности по отдѣльнымъ группамъ сообразно съ размѣрами какого-либо признака, то вполнѣ логичнымъ представляется усвоить этой кривой сокращенное наименованіе *кривой распредѣленія*. Одновременно можно удержать и терминъ „кривая частоты", понимая его въ вышеизъясненномъ смыслѣ. [1])

Попытка П. А. Некрасова (Теорія вѣроятностей, ч. III) пересадить на русскую почву понятіе *корреляціи*, не употребляя этого термина, не представляется мнѣ заслуживающей подражанія. Съ понятіемъ функціональной зависимости обычно связывается представленіе лишь о строгой („крѣпостной") функціональной зависимости. Ввести въ общее сознаніе образованной публики понятіе *нестрогой* зависимости представляется автору весьма важной задачей въ виду того значенія, которое имѣетъ эта еще мало распространенная логическая концепція для правильнаго пониманія характера безчисленнаго множества соціальныхъ, психологическихъ и біологическихъ явленій. Лишнее однако распространяться о томъ, какъ важно для цѣлей популяризаціи и пропаганды связать новый комплексъ идей съ новымъ удачнымъ терминомъ. Если кромѣ того принять во вниманіе, что этотъ терминъ уже началъ примѣняться въ нашей литературѣ, то придется признать, что нѣтъ основаній отказываться отъ его употребленія.

---

[1]) Терминъ „частота", въ смыслѣ англійскаго frequency, употребляется Р. Орженцкимъ (см. op. cit., напр., стр. 211 и сл.). Не говоря уже о нѣкоторомъ несоотвѣтствіи такого словоупотребленія духу нашего языка, я думаю, что предложенную терминологію слѣдовало бы предпочесть и по соображеніямъ экономіи, т. к. въ спеціальной литературѣ всегда чувствуется недостатокъ пригодныхъ словъ для выраженія нужныхъ понятій. Различая *частость*, *численность* и *частоту*, мы удовлетворяемъ и этой немаловажной потребности.

Терминъ „регрессія“ вызываетъ уже нѣкоторыя сомнѣнія. Прежде всего, онъ звучитъ нѣсколько странно для русскаго уха, да и въ самой Англіи онъ не можетъ не казаться нѣсколько искусственнымъ ввиду случайности его происхожденія изъ соображеній біологическаго характера. Я не рѣшился однако на ломку, установившейся въ Англіи терминологіи и, слѣдуя примѣру А. В. Леонтовича, счелъ за лучшее удержать этотъ терминъ и по русски. Будущее покажетъ, не удастся-ли кому замѣнить его достаточно удачнымъ русскимъ выраженіемъ.

II.

О методѣ моментовъ.

Обоснованіе метода моментовъ, данное Пирсономъ, врядъ-ли можетъ быть признано вполнѣ строгимъ. Не говоря уже о томъ, что малость отбрасываемыхъ членовъ въ формулахъ (26) скорѣе декретируется, чѣмъ доказывается, сомнѣнія вызываетъ и самый методъ доказательства, основанный на примѣненіи ряда Тейлора (Маклорена).

Поэтому представляетъ интересъ попытка болѣе строгого обоснованія метода моментовъ, принадлежащая перу русскаго математика Л. К. Лахтина.

Чтобы удовлетворить съ большимъ или меньшимъ приближеніемъ основное ур-іе метода наименьшихъ квадратовъ

$$\int (y - Y)^2\, dx = \text{minimum}$$

(см. выше стр. 21), Лахтинъ разлагаетъ функціи y и Y въ безконечные ряды по Лежандровымъ полиномамъ (т. е. по сферическимъ функціямъ). Это разложеніе отличается болѣе общимъ характеромъ, чѣмъ разложеніе по ряду Тейлора, т. к. для того, чтобы Лежандровы ряды сходились, достаточно, чтобы функціи y и Y въ указанныхъ предѣлахъ были непрерывны и не имѣли безконечнаго числа maxima и minima. Этимъ условіямъ статистическія кривыя обыкновенно удовлетворяютъ.

Приравнивая нулю n первыхъ членовъ разложенія, Лахтинъ получаетъ:

$$\mu_0 = \mu_0', \; \mu_1 = \mu_1', \; \mu_2 = \mu_2', \; \ldots\ldots \; \mu_{n-1} = \mu'_{n-1},$$

т. е. основныя равенства метода моментовъ. Остальными членами разложенія позволительно пренебречь въ силу доказанной сходимости рядовъ. [1])

---

[1]) Л. К. Лахтинъ, О методѣ Пирсона, l. cit., стр. 483—488.

ПРИЛОЖЕНІЕ.

(*A*) Таблицы I—VI даютъ корреляцію среднихъ мѣсячныхъ цѣнъ ржи: въ Москвѣ (x_1), въ Ельцѣ (x_2), въ Самарѣ (x_3) и въ Самарѣ-же за предыдущій мѣсяцъ (x_4). Данныя наши обнимаютъ періодъ 1893—1903 гг. (11 лѣтъ), но относятся только къ 124 мѣсяцамъ изъ 132, т. к., вслѣдствіе пробѣловъ въ нашемъ источникѣ, 8 мѣсяцевъ пришлось опустить. (Напр., цѣна въ Самарѣ въ декабрѣ 1892 г. намъ неизвѣстна, слѣдовательно, x_4 мы имѣемъ съ января 1893 г., а x_1, x_2, x_3—только съ февраля). Источникомъ послужилъ „Сводъ товарныхъ цѣнъ на главныхъ русскихъ и иностранныхъ рынкахъ" за соотвѣтствующіе годы.

(*B*) Таблицы VII и VIII составлены по даннымъ Б. Веселовскаго, приведеннымъ въ его „Исторіи земства за сорокъ лѣтъ", т. I, Спб. 1909. Расходы уѣздныхъ земствъ на содержаніе земскаго управленія и на народное образованіе взяты изъ Приложенія XV § II и § V, стр. 674—682. За торгово-промышленный доходъ земства принята сумма доходовъ, (1), съ документовъ на право торговли и промысловъ и, (2), съ завод., фабр. и торгово-промышленныхъ помѣщеній. См. тамъ-же Приложеніе XII § II и § IV, стр. 659—667. Всѣ данныя относятся къ 1901 г. [1]).

(*C*) Таблица IX составлена по матеріаламъ, сгруппированнымъ въ книгѣ O. Schmitz'a, Die Bewegung der Warenpreise in Deutschland von 1851 bis 1902, Berlin 1903, стр. 65—69 и 217—221.

---

[1]) Считаю своимъ пріятнымъ долгомъ выразить благодарность г. Добрыдену за помощь при перенесеніи данныхъ изъ таблицъ Веселовскаго на карточки.

Таблица I.

Корреляція между средними мѣсячными цѣнами ржи въ Москвѣ (x_1) и въ Ельцѣ (x_2) за 1893—1903 г.г.

Москва („овинная“).

Елецъ („тяжелая“).

Коп. за пудъ	35	40	45	50	55	60	65	70	75	80	85	90	95	Всего
75													1	1
70								3½	2½		1	1½		8½
65							5	7			½			12½
60						3	13	6						22
55					9	4	8	1	2	1				25
50				7	6			4	1	1				19
45		½	3	2										5½
40		2½	9½	1	1	3	1	1						19
35	2½	6	1											9½
30	2													2
Всего	4½	9	13½	10	16	10	27	22½	5½	2	1½	1½	1	124

$h_1 = 59{,}40 \pm 0{,}77$ коп. $\sigma_1 = 12{,}64 \pm 0{,}54$ коп.

$h_2 = 52{,}64 \pm 0{,}65$ „ $\sigma_2 = 10{,}74 \pm 0{,}46$ „

$$r_{12} = 0{,}79215 \pm 0{,}023$$

Таблица II.

Корреляція между средними мѣсячными цѣнами ржи въ Ельцѣ (x_2) и въ Самарѣ (x_3) за 1893—1903 г.г.

Елецъ („тяжелая").

Самара (безъ обозначенія сорта).

Коп. за пудъ	25	30	35	40	45	50	55	60	65	70	75	80	Всего
80													
75										2			2
70								1	2	2	1		6
65							3	2½	3	3			11½
60							2	6½	3	1			12½
55							3	6					9
50						4	12	5	4½	½			26
45						9	5	1					15
40					1	3							4
35				4	1	3							8
30			6½	12	2½								21
25		2	3	3	1								9
20													
Всего		2	9½	19	5½	19	25	22	12½	8½	1		124

$h_3 = 47{,}04 \pm 0{,}84$ коп. $\sigma_3 = 13{,}84 \pm 0{,}59$ коп.

$r_{23} = 0{,}87796 \pm 0{,}014$

Таблица III.

Корреляція между средними мѣсячными цѣнами ржи въ Москвѣ (x_1) и въ Самарѣ (x_3) за 1893—1903 г.г.

Москва („овинная").

Самара (безъ обозначенія сорта).

	35	40	45	50	55	60	65	70	75	80	85	90	95	Всего
75								2						2
70							1	2	1		1/2	1/2	1	6
65							5	5	1 1/2					11 1/2
60							5	6 1/2				1		12 1/2
55						2	6	1						9
50					9	4	9	1	1	1	1			26
45				2	5	1		4	2	1				15
40				3	1									4
35			1	4		1	1	1						8
30	1	7 1/2	9 1/2	1		2								21
25	3 1/2	1 1/2	3		1									9
Всего	4 1/2	9	13 1/2	10	16	10	27	22 1/2	5 1/2	2	1 1/2	1 1/2	1	124

$$r_{13} = 0,76818 \pm 0,025$$

Таблица IV.

Корреляція между средней мѣсячной цѣной ржи въ Москвѣ (x_1) и средней цѣной ржи за предыдущій мѣсяцъ въ Самарѣ (x_4) за время съ 1893—1903 г.

Самара (предыдущій мѣсяцъ).

Москва.

Коп. за пудъ	20	25	30	35	40	45	50	55	60	65	70	75	
95											1		1
90									1		$\frac{1}{2}$		$1\frac{1}{2}$
85									1		$\frac{1}{2}$		$1\frac{1}{2}$
80						1	1						2
75						1	2			$\frac{1}{2}$	2		$5\frac{1}{2}$
70						5	1	1	$5\frac{1}{2}$	6	2	2	$22\frac{1}{2}$
65				1		1	9	6	4	6			27
60			2	2			3	2	1				10
55				1	2	4	9						16
50		1	1	3	2	3							10
45		2	$10\frac{1}{2}$	1									$13\frac{1}{2}$
40		2	7										9
35		4	$\frac{1}{2}$										$4\frac{1}{2}$
		9	21	8	4	15	25	9	$12\frac{1}{2}$	$12\frac{1}{2}$	6	2	124

$h_4 = 47{,}16 \pm 0{,}84$ коп. $\sigma_4 = 13{,}93 \pm 0{,}60$ коп.

$r_{14} = 0{,}78706 \pm 0{,}023$

Таблица III.

Корреляція между средними мѣсячными цѣнами ржи въ Москвѣ (x_1) и въ Самарѣ (x_3) за 1893—1903 г.г.

Москва („овинная“).

Самара (безъ обозначенія сорта).

	35	40	45	50	55	60	65	70	75	80	85	90	95	Всего
75								2						2
70							1	2	1		$1/2$	$1/2$	1	6
65							5	5	$1^1/_2$					$11^1/_2$
60							5	$6^1/_2$				1		$12^1/_2$
55						2	6	1						9
50					9	4	9	1	1	1	1			26
45				2	5	1		4	2	1				15
40				3	1									4
35			1	4		1	1	1						8
30	1	$7^1/_2$	$9^1/_2$	1		2								21
25	$3^1/_2$	$1^1/_2$	3		1									9
Всего	$4^1/_2$	9	$13^1/_2$	10	16	10	27	$22^1/_2$	$5^1/_2$	2	$1^1/_2$	$1^1/_2$	1	124

$$r_{13} = 0{,}76818 \pm 0{,}025$$

Таблица IV.

Корреляція между средней мѣсячной цѣной ржи въ Москвѣ (x_1) и средней цѣной ржи за предыдущій мѣсяцъ въ Самарѣ (x_4) за время съ 1893—1903 г.

Самара (предыдущій мѣсяцъ).

Москва.

Коп. за пудъ	20	25	30	35	40	45	50	55	60	65	70	75	
95											1		1
90									1		1/2		1 1/2
85									1		1/2		1 1/2
80						1	1						2
75						1	2			1/2	2		5 1/2
70						5	1	1	5 1/2	6	2	2	22 1/2
65				1		1	9	6	4	6			27
60			2	2			3	2	1				10
55				1	2	4	9						16
50		1	1	3	2	3							10
45		2	10 1/2	1									13 1/2
40		2	7										9
35		4	1/2										4 1/2
		9	21	8	4	15	25	9	12 1/2	12 1/2	6	2	124

$h_4 = 47{,}16 \pm 0{,}84$ коп. $\sigma_4 = 13{,}93 \pm 0{,}60$ коп.

$r_{14} = 0{,}78706 \pm 0{,}023$

Таблица V.

Корреляція между средней мѣсячной цѣной ржи въ Ельцѣ (x_2) и такой-же цѣной въ Самарѣ за предыдущій мѣсяцъ (x_4) за время съ 1893—1903 г.

Самара (предыдущій мѣсяцъ).

Елецъ.

Коп. за пудъ	25	30	35	40	45	50	55	60	65	70	75	Всего
75										1		1
70						1		1½	1	4	1	8½
65					1	3	1	2½	3	1	1	12½
60			1			5	4	5½	6½			22
55					3	13	4	3	2			25
50		1	2	4	9	3						19
45	1½	2	1		1							5½
40	3½	10½	4		1							19
35	2	7½										9½
30	2											2
25												
Всего	9	21	8	4	15	25	9	12½	12½	6	2	124

$$r_{24} = 0{,}85593 \pm 0{,}016$$

Таблица VI.

Корреляція между средней мѣсячной цѣной ржи въ Самарѣ (x_3) и средней мѣсячной цѣной ржи въ Самарѣ-же, но за предыдущій мѣсяцъ (x_4) за время съ 1893—1903 г.

Самара (предыдущій мѣсяцъ).

Самара.

Коп. за пудъ	25	30	35	40	45	50	55	60	65	70	75	Всего
75										1	1	2
70									1	4	1	6
65						1	1	$3^1/_2$	5	1		$11^1/_2$
60							2	$5^1/_2$	5			$12^1/_2$
55						3	2	$2^1/_2$	$1^1/_2$			9
50					2	19	4	1				26
45			1	2	10	2						15
40			1	2	1							4
35		3	3		2							8
30	6	13	2									21
25	3	5	1									9
Всего	9	21	8	4	15	25	9	$12^1/_2$	$12^1/_2$	6	2	124

$$r_{34} = 0{,}93292 \pm 0{,}008.$$

Таблица VII.

Расходъ на содержаніе уѣздной земской управы въ %%-хъ всего расходнаго бюджета (1901 г.) (x).

Расходъ на народное образованіе въ %%-хъ всей расходной смѣты уѣзднаго земства (1901 г.) (y).

%%	2—4	4—6	6—8	8—10	10—12	12—14	14—16	16—18	18—20	20—22	22—24	Всего уѣздн. земст.
$43^3/_4$—$41^1/_4$						1						1
$41^1/_4$—$38^3/_4$		2										2
$38^3/_4$—$36^1/_4$		$^1/_2$	$2^1/_2$	4								7
$36^1/_4$—$33^3/_4$		1	5	3	2	1	2					14
$33^3/_4$—$31^1/_4$			7	3	1	2						13
$31^1/_4$—$28^3/_4$		4	7	3	7							21
$28^3/_4$—$26^1/_4$		2	5	7	$3^1/_2$	$3^1/_2$			1			22
$26^1/_4$—$23^3/_4$		2	$11^1/_2$	$12^1/_2$	8	1	2	1	1			39
$23^3/_4$—$21^1/_4$		7	8	9	6	4	1	1				36
$21^1/_4$—$18^3/_4$		6	$12^1/_2$	$12^1/_2$	$8^1/_2$	$5^1/_2$	1	1			1	48
$18^3/_4$—$16^1/_4$		6	17	9	2	2						36
$16^1/_4$—$13^3/_4$		$6^1/_2$	18	16	$5^1/_2$	4	1					51
$13^3/_4$—$11^1/_4$	$^1/_2$	$5^1/_2$	$10^1/_2$	$8^1/_2$	6	3		1				35
$11^1/_4$— $8^3/_4$		4	7	7	2	1						21
$8^3/_4$— $6^1/_4$	1	1	1	$2^1/_2$	$1^1/_2$	1	1	1				10
$6^1/_4$— $3^3/_4$			$1^1/_2$	$1^1/_2$								3
Всего уѣздн. земствъ	$1^1/_2$	$47^1/_2$	$113^1/_2$	$98^1/_2$	53	29	8	5	2	0	1	359

$$h_x = 8{,}77159\,\% \qquad \sigma_x = 2{,}971\,\%$$

$$h_y = 20{,}64067\,\% \qquad \sigma_y = 7{,}685\,\%$$

$$r_{xy} = +\,0{,}054 \pm 0{,}035$$

Таблица VIII.

Корреляція между торгово-промышленнымъ доходомъ уѣздныхъ земствъ и расходомъ ихъ на народное образованіе (въ %% къ смѣтѣ) въ 1901 г.

Торгово-промышленный доходъ (x).

Расходъ на народное образованіе (y).

%%	0—5	5—10	10—15	15—20	20—25	25—30	30—35	35—40	40—45	45—50	50—55	55—60	60—65	65—70	70—75	75—80	80—85	Всего
43,75—41,25						1												1
41,25—38,75		1	1															2
38,75—36,25	1	1	1			2			1		1							7
36,25—33,75	2	3	2	1	1		1	2	1	1								11
33,75—31,25	1	1	3	1		1	1	1		2	1			1				13
31,25—28,75	3	6	7		1		1		1	1	1							21
28,75 26,25	3	4	2	2	1	6	1	1	1				1					2
26,25—23,75	6	18	7	1	1	2	1			1	2							39
23,75—21,25	6	17	5	3	1	2				1							1	36
21,25—18,75	11	16½	7½	5	4	2		2										48
18,75—16,25	6½	17½	6	2	1	1	2											36
16,25—13,75	9	23	12	3		2		1				1						51
13,75—11,25	7	17½	5½	3	1	1												35
11,25—8,75	3	9	8	½	½													21
8,75—6,25	4	3	1	2														10
6,25—3,75	1	1		1														3
Всего	63½	138½	68	24½	11½	20	7	7	4	6	5	1	1	1	0	0	1	359

Таблица IX.

Корреляція между средними мѣсячными цѣнами чугуна и ржи въ Германіи за 1879—1900 г.

Цѣна чугуна (x).

Цѣна ржи (y).

Марки за 1000 К⁰	45	50	55	60	65	70	75	80	85	90	95	Всего
234		1	3	1								5
222		1										1
210		2	2	5½	2							11½
198		4		½	2	1						7½
186			1½	4½	1	1						8
174			2	5	1	3	½	2½	1½	3	1½	20
162				4	3	2	1½	3	½			14
150			4½	21½	10	6	1	5	1	4	1	54
138		2½	12½	25	4	3			1	6	1	55
126	1	21	15½	9								46½
114		15½	7½	13½	1							37½
102				1								1
Всего	1	47	48½	90½	24	16	3	10½	4	13	3½	261

$$h_x = 61{,}609 \pm 0{,}46, \qquad \sigma_x = 11{,}092 \pm 0{,}33,$$
$$h_y = 147{,}333 \pm 1{,}19, \qquad \sigma_y = 28{,}572 \pm 0{,}84,$$
$$r_{xy} = +0{,}19162 \pm 0{,}040.$$

ДИМИТРІЙ ГРАВЕ.
Профессоръ Университета Св. Владиміра.

МАТЕМАТИКА СТРАХОВОГО ДѢЛА.

Курсъ Кіевскаго Коммерческаго Института.

КІЕВЪ.
1912.

D. GRAVÉ.
Prof. de l'Université de St Vladimira Kieff.

Théorie mathématique d'assurances.

Cours d'études de l'Institut Commercial de Kieff.

Kieff.
1912.

Предисловіе.

Настоящая книга предназначается въ качествѣ пособія для студентовъ страхового подотдѣла Кіевскаго Коммерческаго Института, которымъ читается курсъ страховой математики.

Въ книгѣ я предполагаю извѣстными основныя понятія теоріи вѣроятностей и анализа безконечно малыхъ.

Для удобства приводятся въ разныхъ мѣстахъ книги ссылки на соотвѣтственныя мѣста моей книги „Энциклопедія Математики", гдѣ читатель найдетъ необходимыя свѣдѣнія.

Профессоръ Д. Граве.

31 іюня 1912 г.

ОГЛАВЛЕНІЕ.

Математика страхового дѣла.

I.

Основныя положенія.

§ 1. Страховая математика имѣетъ своимъ предметомъ приложеніе математической теоріи вѣроятностей къ страховому дѣлу [1]). Эти приложенія основываются на нѣкоторыхъ допущеніяхъ, имѣющихъ характеръ постулатовъ, или аксіомъ.

Перечислимъ эти допущенія:

Первая гипотеза. Предполагается существованіе нѣкоторой вѣроятности человѣку возраста x дожить до возраста $x+m$. Мы будемъ обозначать эту вѣроятность: $p(x, x+m)$.

Вторая гипотеза. Вѣроятности $p(x, x+m)$, относящіяся къ различнымъ людямъ, независимы между собою.

Эта гипотеза необходима для примѣненія теоремы объ умноженіи вѣроятностей (см. Энцикл. матем. гл. XIV, § 15).

Третья гипотеза. Предполагается существованіе нѣкотораго класса людей, находящихся въ одинаковыхъ жизненныхъ условіяхъ, причемъ для каждаго индивидуума этого класса существуетъ одна, опредѣленная и общая для всѣхъ вѣроятность $p(x, x+m)$, зависящая только отъ x и $x+m$.

Очевидно, что выводы, получаемые черезъ примѣненіе вышеописанныхъ трехъ гипотезъ къ разсмотрѣнію совокупностей людей, будутъ тѣмъ ближе къ дѣйствительности, чѣмъ разсматриваемая совокупность ближе подходитъ къ понятію о классѣ.

§ 2. Мы вводимъ въ разсмотрѣніе нѣкоторую функцію l_x отъ независимаго перемѣннаго x, которую мы будемъ называть *числомъ живущихъ возраста* x.

Данное нами названіе функціи l_x не относится къ разсмотрѣнію какихъ либо опредѣленныхъ людей.

---

[1]) Предметомъ нашего курса будетъ разсмотрѣніе страхованія, связаннаго съ различными обстоятельствами жизни и смерти людей.

Функцію l_x мы будемъ опредѣлять слѣдующими свойствами:

1) Эта функція задана съ точностью до постояннаго множителя.

2) Функція l_x не возрастающая.

3) „ l_x не отрицательная.

4) Существуетъ равенство: $p(x, x+m) = \frac{l_{x+m}}{l_x}$.

Если будутъ указаны пріемы нахожденія значенія вѣроятностей $p(x, x+m)$ при всякихъ x и m на основаніи статистическихъ наблюденій надъ жизнью людей извѣстнаго класса, то на основаніи свойства (4) можно будетъ вычислить функцію l_x для всякаго значенія x. Такъ напримѣръ, полагая x равнымъ 0 и взявъ произвольное l_0 (число новорожденныхъ), мы получаемъ для всякаго m:

$$l_m = l_0 p(0, m).$$

Такимъ образомъ у насъ получается рядъ значеній l_x для ряда численныхъ значеній независимаго перемѣннаго x. Обыкновенно разсматриваются промежутки времени, равные одному году, и мы вычисляемъ функцію l_x для значеній:

$$x = 0;\ x = 1;\ x = 2;\ x = 3;\ \ldots\ldots \text{ и т. д.}$$

Получается таблица:

$$l_0,\ l_1,\ l_2,\ l_3,\ l_4\ \ldots\ldots\ldots \text{ и т. д.},$$

носящая названіе таблицы смертности.

Наблюденія приводятъ насъ къ убѣжденію въ существованіи предѣльнаго возраста ω, при которомъ $l_\omega = 0$ и $l_{\omega+\varepsilon} = 0$.

II.

Принципы, на которыхъ основано составленіе таблицы смертности.

§ 1. Нахожденіе вѣроятностей $p(x, x+m)$ могло бы быть основано на производствѣ наблюденій надъ нѣкоторымъ большимъ числомъ L людей возраста x. Если мы замѣтимъ, что по прошествіи m лѣтъ остается въ живыхъ изъ общаго числа людей L, подлежавшихъ нашему наблюденію, нѣкоторое число L_1 людей, то дробь $\frac{L_1}{L}$ можетъ быть приближенно принята за значеніе искомой вѣроятности $p(x, x+m)$. Законъ большихъ чиселъ (Э. м. XIV, § 27), показываетъ, что дробь $\frac{L_1}{L}$ будетъ тѣмъ ближе къ искомой вѣроятности, чѣмъ большее число наблюденій произведено, т. е. чѣмъ больше число L.

§ 2. Таблицы смертности раздѣляются на двѣ главныя категоріи: 1) таблицы, относящіяся къ общему населенію (города, провинціи или страны) и 2) таблицы, составленныя по индивидуальнымъ наблюденіямъ надъ людьми извѣстнаго класса. Къ этой категоріи принадлежатъ таблицы, составленныя по наблюденіямъ надъ кліентами страховыхъ обществъ. Хотя для страхового дѣла являются наиболѣе важными эти послѣднія таблицы, тѣмъ не менѣе мы скажемъ нѣсколько словъ о составленіи таблицъ общаго народонаселенія. Такъ какъ общее народонаселеніе никоимъ образомъ не можетъ быть подведено подъ понятіе объ одномъ классѣ людей, то такія таблицы могутъ имѣть значеніе при разсмотрѣніи очень большого числа людей, ибо только въ этомъ случаѣ можно допустить, что законъ большихъ чиселъ можетъ проявить свое нивеллирующее значеніе; при этомъ случайныя уклоненія въ ту и другую сторону потонутъ въ большомъ числѣ среднихъ цифръ. Та часть статистики, которая относится къ составленію таблицъ смертности, носитъ названіе *формальной теоріи народонаселенія*. Въ послѣднее время этой теоріи придаютъ математическое изложеніе.

Мы должны здѣсь перечислить наиболѣе важныя сочиненія по разсматриваемой теоріи, а именно:

Knapp. Sterblichkeit in Sachsen. Leipzig. 1869.

Zeuner. Abhandlungen aus der mathematischen Statistik. Leipzig. 1869.

Lexis. Theorie des Bevölkerungswechsels. Braunschweig. 1874.

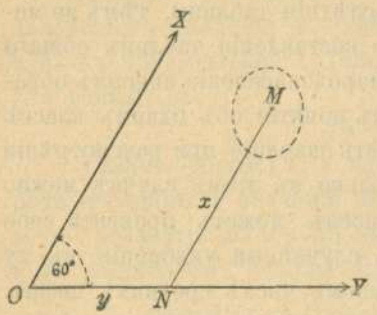
„ Einleitung in die Theorie der Bevölkerungsstatistik. Strassburg. 1875.

III.

Геометрическая теорія Lexis'a.

§ 1. Въ виду практическаго неудобства и даже почти невозможности при разсмотрѣніи общаго народонаселенія прослѣдить за группой достаточно большого числа людей, имѣвшихъ въ данный моментъ времени опредѣленный возрастъ, приходится измѣнять объектъ наблюденія и получать затѣмъ необходимыя вѣроятности, исходя изъ соображеній формальной теоріи народонаселенія. Кромѣ того такое прямое наблюденіе, если бы оно даже и было возможно, требовало бы для составленія полной таблицы смертности слишкомъ большого времени, около 100 лѣтъ. Ибо предѣльный возрастъ ω близокъ къ числу 100.

Lexis предлагаетъ слѣдующую схему: наблюдаемыя смерти людей изображаются (черт. 1) точками на плоскости, въ которой взяты координатныя оси (см. Э. м. II, § 60), образующія уголъ въ 60^0. При этомъ на оси OY откладывается дата рожденія y умершаго человѣка, т. е. другими словами координата y обозначаетъ промежутокъ времени между моментомъ разсматриваемаго рожденія и какимъ нибудь началомъ счета временъ, напримѣръ началомъ христіанскаго лѣтосчисленія. Другую координату $x = NM$ мы беремъ пропорціональной возрасту смерти. Тогда всякая зарегистрированная смерть человѣка, опредѣляемая двумя данными y и x, изображается на плоскости нѣкоторой точкой M. Прямая NM называется *линіей жизни* человѣка, смерть котораго зарегистрирована точкой M. Мы имѣемъ полное основаніе предполагать, что при большомъ числѣ наблюденій точки смертей заполнятъ нѣкоторую часть плоскости равномѣрно, подобно тому, какъ падающій дождь смачиваетъ равномѣрно ту площадь, на которую онъ падаетъ. Конечно, гипотеза такого равномѣрнаго распредѣленія тѣмъ болѣе допустима, чѣмъ больше въ численномъ отношеніи составъ народонаселенія. Если мы разсматриваемъ небольшой раіонъ площади, не болѣе одного года, во всѣ стороны, то такое допущеніе равномѣрности совершенно основательно. Указанное допущеніе равномѣрности распредѣленія смертей на плоскости есть единственный въ настоящее время способъ, дающій возможность при составленіи таблицы смертности общаго народонаселенія замѣнять элементы, неудобные для статистическаго учета, другими элементами, болѣе удобными. Геометрически это приводится къ тому, что число



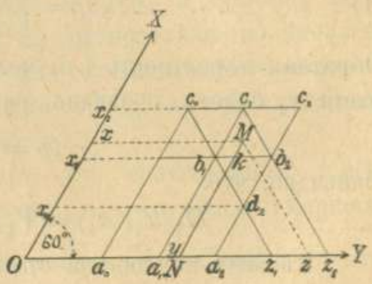
Черт. 1.

смертей, находящихся внутри нѣкоторой площади, можно считать пропорціональнымъ этой площади (вслѣдствіе вышеуказанной равномѣрности) и мы можемъ, слѣдовательно, замѣнять однѣ фигуры другими, равновеликими имъ по площади. Является удобнымъ, кромѣ двухъ указанныхъ координатъ, y и x, ввести еще третью перемѣнную z, опредѣляемую равенствомъ:

$$x + y = z;$$

очевидно, что z будетъ не чѣмъ инымъ, какъ величиной, опредѣляющей моментъ смерти (моментъ рожденія плюсъ продолжительность жизни).

Будемъ откладывать на оси Y (черт. 2) моменты a_0, a_1, a_2, соотвѣтствующіе началамъ промежутковъ времени, принятыхъ за единицу счета времени. Обыкновенно въ страховой математикѣ за единицу времени принимается годъ, и слѣдовательно точкамъ a_0, a_1, a_2 могутъ соотвѣтствовать начала послѣдовательныхъ годовъ. Разсмотримъ промежутокъ времени между точками a_1 и a_2



Черт. 2.

и построимъ линіи жизни всѣхъ людей подлежащей нашему наблюденію совокупности и родившихся въ промежуткѣ (a_1, a_2). Будемъ разсматривать также на другой координатной оси точки x_0, x_1, x_2 . . . . и т. д., соотвѣтствующія моментамъ времени, когда разсматриваемый человѣкъ имѣетъ возрастъ цѣлаго числа лѣтъ. Такъ напримѣръ, если x_0 соотвѣтствуетъ возрасту 19 лѣтъ, то x_1 будетъ соотвѣтствовать возрасту 20 лѣтъ, x_2 — возрасту 21 годъ и т. д. Разсмотримъ человѣка, родившагося въ промежутокъ времени (a_1, a_2) и умершаго въ возрастѣ между x_1 и x_2 лѣтъ. Тогда точка M его смерти попадаетъ въ параллелограммъ $b_1c_1c_2b_2$. Разсмотримъ точки k пересѣченія линіи жизни людей, родившихся въ промежуткѣ (a_1, a_2), съ линіей (b_1, b_2) параллелограмма. Тогда число такихъ точекъ k будетъ, очевидно, представлять число людей данной совокупности, дожившихъ до возраста x_1 лѣтъ. Это число мы будемъ называть *первой главной совокупностью живущихъ* и будемъ обозначать знакомъ $V(x_1, a_1)$. Въ этомъ знакѣ первое число (x_1) обозначаетъ возрастъ живущихъ, а второе число (a_1) начало того промежутка времени (a_1, a_2), въ которомъ родились всѣ разсматриваемые люди. Далѣе, число всѣхъ смертей въ параллелограммѣ $b_1c_1c_2b_2$ мы назовемъ *первой главной совокупностью умершихъ* и будемъ обозначать ее знакомъ $M_1(x_1, a_1)$. Этотъ знакъ будетъ обозначать, слѣдовательно, число умершихъ въ возрастѣ отъ x_1 до x_2 лѣтъ изъ числа родившихся въ промежуткѣ (a_1, a_2). Очевидно, что отношеніе $\frac{M_1(x_1, a_1)}{V_1(x_1, a_1)}$ будетъ представлять приближенно вѣроятность человѣку возраста x_1 лѣтъ не дожить до возраста x_2 лѣтъ. Будемъ обозначать эту вѣроятность черезъ q. Итакъ:

$$(1) \quad \ldots\ldots\ldots\ldots \quad q = \frac{M_1(x_1, a_1)}{V_1(x_1, a_1)}.$$

Обратная вѣроятность (p) человѣку возраста x_1 дожить до возраста x_2 будетъ, очевидно, равняться:

$$p = 1 - q. \quad \text{(См. Э. м. XIV, § 14).}$$

Очевидно, что:

$$M_1(x_1, a_1) = V_1(x_1, a_1) - V_1(x_2, a_1).$$

Въ виду неудобства статистическаго подсчета первыхъ главныхъ совокупностей живущихъ и умершихъ вводятся новыя понятія.

Мы будемъ называть *второй главной совокупностью живущихъ* величину $V_2(a_1, z_1)$ число людей родившихся въ промежуткѣ (a_1, a_2) и доживших до нѣкотораго момента времени $z_1 = a_1 + x_1$. Мы можемъ наглядно геометрически представить себѣ величину $V_2(a_1, z_1)$. Проводимъ изъ точки b_1, имѣющей координаты (a_1, x_1), третью сторону b_1z_1 равносторонняго треугольника $a_1b_1z_1$, тогда очевидно, что $V_2(a_1, z_1)$ будетъ представлять число точекъ встрѣчи линіи жизни подлежащихъ наблюденію людей съ отрѣзкомъ b_1d_2, соотвѣтствующее прямой b_1z_1, дающей моментъ времени z_1 [1]). Будемъ называть *второй главной совокупностью умершихъ* число смертей M_2, попадающихъ внутрь параллелограмма $b_1c_1b_2d_2$, и будемъ обозначать эту величину $M_2(a_1, z_1)$. Очевидно, что это вторая главная совокупность умершихъ представляетъ число людей. родившихся въ промежутокъ (a_1, a_2) и умершихъ въ промежуткѣ (z_1, z_2). Очевидно также, что

---

[1]) Вслѣдствіе равносторонности треугольника $a_1b_1z_1$ точка z_1 соотвѣтствуетъ такому моменту времени, при которомъ $z_1 = a_1 + x_1$.

$$M_2(a_1, z_1) = V_2(a_1, z_1) - V_2(a_1, z_2).$$

Наконецъ, введемъ въ разсмотрѣніе *третью главную совокупность умершихъ*: $M_3(x_1, z_1)$, которая представляетъ собою число смертей, падающихъ внутри параллелограмма $c_0c_1b_2b_1$; эта главная совокупность удобна для статистическихъ наблюденій, потому что она представляетъ изъ себя число умершихъ въ промежутокъ времени отъ z_1 до z_2 въ возрастѣ отъ x_1 до x_2 лѣтъ.

Для вычисленія вѣроятности q_2 человѣку въ возрастѣ x_1 лѣтъ не дожить до возраста x_2 лѣтъ, мы можемъ вмѣсто первой главной совокупности $M_1(a_1, x_1)$, взять третью главную совокупность $M_3(x_1, z_1)$, если допустимъ высказанную нами раньше гипотезу равномѣрнаго распредѣленія смертей по плоскости. Въ самомъ дѣлѣ, параллелограммы $b_1c_1c_2b_2$ и $c_0c_1b_2b_1$, соотвѣтствующіе совокупностямъ M_1 и M_3 равновелики по площади. Подобнымъ же образомъ удобнѣе для статистическаго подсчета вмѣсто первой главной совокупности $V_1(x_1, a_1)$ разсматривать вторую главную совокупность $V_2(a_1, z_1)$.

Разсмотримъ число умершихъ въ треугольникѣ $b_1c_1b_2$. Съ одной стороны мы можемъ сказать, что это число умершихъ будетъ $= \frac{1}{2} M_3(x_1, z_1)$, потому что этотъ треугольникъ по площади равняется половинѣ параллелограмма $c_0c_1b_2b_1$; съ другой стороны, очевидно, что то же число смертей можетъ быть представлено по формулѣ: $V_1(x_1, a_1) - V_2(a_1, z_2)$. Итакъ, мы получаемъ равенство: $\frac{1}{2} M_3(x_1, z_1) = V_1(x_1, a_1) - V_2(a_1, z_2)$,

отсюда:

$$V_1(x_1, a_1) = V_2(a_1, z_2) + \frac{1}{2} M_3(x_1, z_1).$$

Подставляя въ формулу (1) вмѣсто первыхъ главныхъ совокупностей ихъ выраженія черезъ вторыя и третьи совокупности, получаемъ окончательно слѣдующую формулу, удобную для статистическихъ приложеній:

$$q = \frac{M_3(x_1, z_1)}{V_2(a_1, z_2) + \frac{1}{2} M_3(x_1, z_1)}.$$

§ 2. Въ предыдущемъ параграфѣ мы показали, какъ при помощи статистическихъ наблюденій получить для общаго народонаселенія данной мѣстности вѣроятность q человѣку возраста x

не дожить до возраста x_2. Послѣ того, какъ вычислена вѣроятность q, можно будетъ вычислить вѣроятность $p = 1 - q$, человѣку возраста x_1 лѣтъ дожить до слѣдующаго возраста x_2 лѣтъ. Если промежутокъ между x_1 и x_2 равняется одному году, то вѣроятность p, вычисленная такимъ образомъ, будетъ представлять изъ себя не что иное, какъ ту вѣроятность, которая соотвѣтственно обозначенію § 1 должна быть обозначена такъ: $p(x_1, x+1)$. Очевидно, что общая вѣроятность $p(x, x+m)$ по теоремѣ объ умноженіи вѣроятностей будетъ выражаться такъ:

$$(1) \quad . \; . \; p(x, x+m) = p(x, x+1)\, p(x+1, x+2)\, p(x+2, x+3) \; . \; . \; . \; . \; . \; . \; . \; . \; p(x+m-1, x+m).$$

Такъ какъ, по только что сказанному, мы можемъ вывести изъ статистическихъ наблюденій величины всѣхъ множителей: $p(x, x+1)$, $p(x+1, x+2)$, $p(x+2, x+3)$ и т. д..., выражающихъ вѣроятности прожить одинъ годъ, то, перемножая эти множители, мы получимъ величину $p(x, x+m)$, а слѣдовательно по правиламъ § 2, главы I, составимъ таблицу смертности.

IV.

Составленіе таблицъ смертности по индивидуальнымъ наблюденіямъ.

§ 1. Таблица смертности, составленная по правиламъ, изложеннымъ въ предыдущей главѣ, изъ статистическихъ наблюденій надъ общимъ народонаселеніемъ, обладаетъ главнымъ недостаткомъ, состоящимъ въ томъ, что общее народонаселеніе данной мѣстности никоимъ образомъ не можетъ подойти подъ понятіе класса людей въ указанномъ нами въ главѣ первой смыслѣ, между тѣмъ, какъ приложенія теоріи вѣроятностей могутъ быть тѣмъ болѣе надежны, чѣмъ дѣло идетъ ближе къ разсмотрѣнію нѣкотораго класса людей, находящихся по отношенію къ смертности приблизительно въ одинаковыхъ жизненныхъ условіяхъ. Практика страховыхъ учрежденій, особенно въ Англіи, гдѣ страховое дѣло уже давно пользуется большимъ распространеніемъ, привела къ необходимости для страховыхъ учрежденій вести собственную статистику смертности среди своихъ кліентовъ. Мотивами къ введенію такой статистики служили слѣдующія соображенія:

1) Съ одной стороны является труднымъ точно установить понятіе о классѣ людей относительно смертности, съ другой стороны кліенты страхового учрежденія являются естественно весьма важнымъ объектомъ для наблюденій самого этого страхового учрежденія.

2) Страховое учрежденіе имѣетъ возможность слѣдить за жизнью своихъ кліентовъ и, слѣдовательно, имѣетъ возможность непосредственно получать вѣроятности дожитія не прибѣгая ни къ какимъ добавочнымъ гипотезамъ вродѣ равномѣрности распредѣленія смертей, о которой мы упоминали при разсмотрѣніи общаго народонаселенія.

§ 2. Итакъ, страховое учрежденіе можетъ взять извѣстное число L своихъ кліентовъ, находящихся въ какомъ нибудь возрастѣ x и произвести наблюденіе, какое число D изъ этого числа кліентовъ умретъ, не достигнувъ возраста $x+1$. Тогда вѣроятность q получится непосредственно по формулѣ:

$$(1) \quad \ldots\ldots\ldots\ldots \quad q=\frac{D}{L},$$

гдѣ D число умершихъ изъ числа совокупности L въ возрастѣ $(x, x+1)$.

Практика страховыхъ учрежденій обнаруживаетъ однако необходимость нѣкоторой поправки въ формулѣ (1), происходящей отъ иммиграціи новыхъ кліентовъ и возможнаго ухода—эмиграціи раньше бывшихъ кліентовъ. Понятіе объ эмиграціи надо принимать въ самомъ широкомъ смыслѣ слова. Такъ напримѣръ, если дѣло идетъ о статистикѣ смертности холостыхъ людей, то вступленіе въ бракъ должно разсматриваться, какъ эмиграція. Окончательная формула, въ которой приняты въ расчетъ эмиграція и иммиграція, обыкновенно употребляется въ такомъ видѣ:

$$(2) \quad \ldots\ldots\ldots\ldots \quad q=\frac{D}{L+\frac{I-E}{2}}.$$

Здѣсь L есть число людей возраста x, подлежавшихъ въ началѣ года наблюденію. Число D есть число зарегистрованныхъ обществомъ смертей въ промежуткѣ отъ x до $x+1$ лѣтъ, I—число иммигрантовъ, а E—число эмигрантовъ за тотъ же промежутокъ времени $(x, x+1)$ лѣтъ. Въ виду частаго употребленія формулы (2) полезно обратить вниманіе на математическіе доводы въ пользу этой формулы.

Выводъ этой формулы основывается на слѣдующихъ предположеніяхъ:

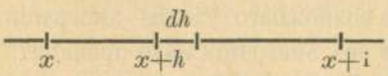
I. Эмиграція и иммиграція распредѣляются равномѣрно въ продолженіи цѣлаго года.

II. Всѣ функціи l_x, $p(x, y)$, $q(x, y)$, таковы, что въ промежуткѣ отъ x до $x+1$ они измѣняются прямолинейно (см. Э. м. XI, § 4). Такъ напримѣръ, если мы будемъ подразумѣвать подъ q вѣроятность, которую мы обозначаемъ $q(x, x+1)$, т. е. вѣроятность человѣку возраста x не дожить до возраста $x+1$, то вѣроятность $q(x+h, x+1)$ человѣку возраста $x+h$ не дожить до возраста $x+1$ должна быть меньше вѣроятности q во столько разъ, во сколько промежутокъ $1-h$ меньше промежутка 1. Другими словами, мы предполагаемъ существованіе формулы:

(3) . . . . . . $q(x+h, x+1) = (1-h)q.$ [1])

III. Вѣроятность $q(x+k, x+h)$ зависитъ отъ чиселъ $k<1$ и $h<1$, а не зависитъ отъ того, было ли данное лицо въ промежутокъ времени $(x+k, x+h)$ подъ наблюденіемъ страховыхъ учрежденій, или же оно не принадлежало къ числу объектовъ наблюденія.

Предположимъ, что среди года разсматривается безконечно малый промежутокъ времени dh (чер. 3), начало котораго пусть будетъ $x+h$, а конецъ $x+h+dh$.



Черт. 3.

Будемъ называть этотъ промежутокъ ε. Предполагая равномѣрность иммиграціи, мы можемъ сказать, что если за цѣлый годъ будетъ I иммигрантовъ, то въ промежутокъ ε удалятся изъ сферы наблюденія $I.dh$. Этому числу соотвѣтствуетъ:

(4) . . . . $\dfrac{Idh}{p(x, x+h)}$

число людей возраста x, которые должны были бы входить въ расчетъ при наблюденіяхъ, если бы они уже въ началѣ года вошли въ составъ кліентовъ и за вычетомъ умершихъ изъ нихъ, представляли бы для промежутка ε какъ разъ число Idh.

---

[1]) Конечно, мы предполагаемъ $h<1$

Очевидно, что если лица въ числѣ (4) подлежали бы наблюденію страхового учрежденія съ начала года, то страховое учрежденіе зарегистрировало бы число смертей:

$$(5) \ldots \quad \frac{Idh}{p(x, x+h)} - Idh = Idh\left[\frac{1}{p(x, x+h)} - 1\right].$$

Итакъ, послѣдняя формула (5) указываетъ число незарегистрированныхъ смертей, потому что эти смерти произошли до поступленія людей въ число кліентовъ общества. Совершенно подобное же происходитъ съ эмиграціей. Очевидно, что эмигрантовъ въ промежутокъ ε будетъ Edh; изъ этого числа доживутъ до конца года очевидно: $Edh \, . \, p(x+h, \; x+1)$. Значитъ число не зарегистрированныхъ смертей, потому что они послѣдовали послѣ выхода этихъ людей изъ подъ наблюденія, будетъ:

$$(6) \ldots \quad Edh - Edhp(x+h, x+1) = Edh\,[1 - p(x+h, \; x+1)].$$

Чтобы составить себѣ окончательное понятіе о полномъ числѣ не зарегистрированныхъ смертей, придется просуммировать за всѣ промежутки ε, заполняющіе цѣлый годъ; другими словами, придется проинтегрировать формулы (5) и (6) по всему промежутку года. Интегрируя формулу (5), мы имѣемъ:

$$(7) \ldots \ldots \quad I\left[\int_0^1 \frac{dh}{p(x, x+h)} - 1\right].$$

Совершенно подобнымъ же образомъ интегрируя формулу (6), получаемъ:

$$(8) \ldots \ldots \quad E\int_0^1 q(x+h, \; x+1)dh.$$

Итакъ полное число смертей, какъ зарегистрированныхъ обществомъ, такъ и не зарегистрированныхъ, очевидно будетъ:

$$(9) \ldots \quad D + I\left[\int_0^1 \frac{dh}{p(x, x+h)} - 1\right] + E\int_0^1 q(x+h \;\; x+1)dh,$$

число же всѣхъ людей, къ началу года имѣвшихъ возрастъ x лѣтъ, должно состоять изъ L людей, дѣйствительно подлежавшихъ наблюденію въ началѣ года, и изъ числа тѣхъ людей, которые за вычетомъ умершихъ дали иммигрантовъ, т. е. другими словами надо проинтегрировать формулу (4). Итакъ, все число людей возраста x будетъ выражаться по формулѣ:

$$(10) \ldots\ldots\ldots L+I\int_0^1 \frac{dh}{p(x, x+h)}.$$

Итакъ, мы получаемъ окончательную формулу для вѣроятности q:

$$(11) \ldots q = \frac{D+I\left[\int_0^1 \frac{dh}{p(x, x+h)} - 1\right] + E\int_0^1 q(x+h, x+1)dh}{L+I\int_0^1 \frac{dh}{p(x, x+h)}}.$$

На основаніи формулы (3) можно выраженіе (11) упростить. Въ самомъ дѣлѣ, имѣемъ:

$$(12) \ldots \int_0^1 q(x+h, x+1)dh = q\int_0^1 (1-h)qdh = q\int_0^1 (1-h)dh =$$

$$= q\left[\int_0^1 dh - \int_0^1 hdh\right] = q\left(1-\frac{1}{2}\right) = \frac{q}{2}.$$

Итакъ, мы можемъ преобразовать формулу (12) такъ:

$$(13) \ldots\ldots q\left(L+I\int_0^1 \frac{dh}{p(x, x+h)} - \mathrm{E}\,\frac{1}{2}\right) =$$

$$= D+I\left[\int_0^1 \frac{dh}{p(x, x+h)} - 1\right].$$

На основаніи таблицы смертности имѣемъ:

$$p(x, x+h) = \frac{l_{x+h}}{l_x},$$

отсюда

$$(14) \ldots\ldots \int_0^1 \frac{dh}{p(x, x+h)} = l_x\int_0^1 \frac{dh}{l_{x+h}}.$$

Обозначая черезъ $k=\int_0^1 \frac{dh}{l_{x+h}}$, мы перепишемъ формулу (12) въ такомъ видѣ:

$$(15) \ldots\ldots q\left(L+Il_xk - E\frac{1}{2}\right) = D+I\left(l_xk-1\right).$$

Послѣднюю формулу можно будетъ переписать такъ:

$$(16) \ldots\ldots q\left(L+Il_xk - \frac{E}{2} - I\frac{l_xk-1}{q}\right) = D.$$

Но $q=1-p(x, x+1)$. Не трудно показать, что коэфиціентъ при I въ скобкахъ лѣвой части уравненія (16) равенъ $\frac{1}{2}$; въ самомъ дѣлѣ, этотъ коэффиціентъ равенъ:

$$l_x k-\frac{l_x k-1}{q}=\frac{1-l_x k(1-q)}{q}=\frac{1-l_x k p(x, x+1)}{q}=$$

$$=\frac{1-k l_{x+1}}{q}=\frac{1-\int_0^1 \frac{l_{x+1}}{l_{x+h}} dh}{q}=\frac{1-\int_0^1 p(x+h, x+1) dh}{q}=$$

$$=\frac{\int_0^1 dh-\int_0^1 p(x+h, x+1) dh}{q}=\frac{\int_0^1\left[1-p(x+h, x+1)\right] dh}{q}=$$

$$=\frac{\int_0^1 q(x+h, x+1) dh}{q};$$

послѣдняя же дробь на основаніи формулы (12) можетъ быть написана такъ: $\frac{\frac{q}{2}}{q}=\frac{1}{2}$. Значитъ, формула (16) обращается въ:

$$q\left(L+I \frac{1}{2}-E \frac{1}{2}\right)=D$$

и слѣдовательно получается формула (2), справедливость которой мы и хотѣли показать.

V.

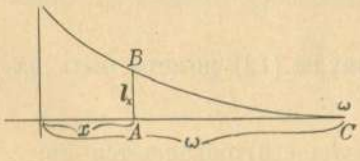
Введеніе новыхъ біометрическихъ функцій.

§ 1. Въ предыдущемъ изложеніи мы познакомились со свойствами функціи l_x, частыя значенія которой составляютъ таблицу смертности. На конгрессахъ актуаріевъ приняты слѣдующія обозначенія: $l_x-l_{x+1}=d_x$. При этомъ величина d_x выражаетъ фиктивно взятое изъ таблицы смертности число умершихъ въ возрастѣ $(x, x+1)$. Вѣроятность человѣку возраста x лѣтъ прожить еще одинъ годъ, принято обозначать p_x. Вѣроятность умереть въ ближайшемъ году обозначается q_x.

Очевидно, что: $p_x = \frac{l_{x+1}}{l_x}$, $q_x = \frac{d_x}{l_x}$. Кромѣ того:

$$q_x = \frac{d_x}{l_x} = \frac{l_x - l_{x+1}}{l_x} = 1 - \frac{l_{x+1}}{l_x} = 1 - p_x \,.$$

Разсмотримъ еще рядъ функцій, связанныхъ съ вѣроятностью человѣческой жизни и часто употребляющихся въ приложеніяхъ страховой математики. Прежде всего будемъ разсматривать l_x, какъ непрерывную функцію, значенія которой для цѣлаго числа лѣтъ даны въ таблицѣ смертности, а для долей года мы предполагаемъ установленными нѣкоторыя правила интерполированія. При этомъ мы будемъ считать функцію l_x непрерывной функціей отъ x, соотвѣтствующаго начальному возрасту таблицы смертности, до предѣльнаго возраста ω. Функція l_x будучи не возрастающей, представится нѣкоторой падающей кривой линіей, достигающей оси абсциссъ въ точкѣ, соотвѣтствующей предѣльному возрасту ω (см. черт. 4). Наша функція l_x, будучи непрерывной, будетъ очевидно функціей интегрируемой (см. Э. м. III, § 149).



Черт. 4.

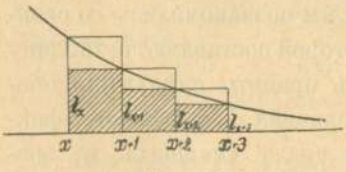
Разсмотримъ интегралъ:

$$\int_x^\omega l_x dx \,.$$

Очевидно, что этотъ интегралъ, есть не что иное, какъ площадь ABC. Этотъ интегралъ оказывается нѣкоторой функціей отъ нижняго предѣла x. Эта функція убывающая, ибо съ возрастаніемъ x, площадь убываетъ до 0.

Функція $E_x = \frac{1}{l_x}\int_x^\omega l_x dx$ называется *среднею продолжительностью жизни*, или *ожиданіемъ жизни* лица возраста x.

Построимъ ординаты l_x, l_{x+1}, l_{x+2}, ... соотвѣтствующія цѣлымъ годамъ (черт. 5): x, $x+1$, $x+2$, ... и т. д., тогда криволинейная площадь, опредѣляемая $\int_x^\omega l_x d_x$, будетъ больше, чѣмъ сумма площадей вписанныхъ (на чертежѣ заштрихованныхъ) прямоугольниковъ.



Черт. 5.

Такъ какъ основанія этихъ прямоугольниковъ суть единицы, а высоты будутъ l_{x+1}, l_{x+2}, l_{x+3} . . . . . , то сумма этихъ площадей будетъ очевидно равна:

$$l_{x+1}\cdot 1 + l_{x+2}\cdot 1 + l_{x+3}\cdot 1 + \ldots\ldots \text{ и т. д.}$$

и мы получаемъ, очевидно, неравенство:

$$\int_x^\omega l_x dx > l_{x+1} + l_{x+2} + l_{x+3} + \ldots\ldots$$

Если бы мы взяли прямоугольники выходящіе, то, очевидно, получили бы неравенство:

$$\int_x^\omega l_x dx < l_x + l_{x+1} + l_{x+2} + l_{x+3} + \ldots\ldots$$

Обозначимъ черезъ

$$(1) \quad \ldots\ldots\ldots \quad e_x = \frac{l_{x+1} + l_{x+2} + l_{x+3} + \ldots\ldots}{l_x},$$

тогда, раздѣляя неравенства:

$$l_{x+1} + l_{x+2} + \ldots\ldots < \int_x^\omega l_x dx < l_x + l_{x+1} + l_{x+2} + \ldots\ldots$$

на l_x мы получимъ:

$$e_x < E_x < 1 + e_x.$$

Величина l_x вычисляется непосредственно по таблицамъ смертности, что видно изъ формулы (1). Что же касается ожиданія жизни E_x, то вычисленіе его представляетъ затрудненіе въ томъ смыслѣ, что предварительно требуется при помощи пріемовъ интерполированія составить непрерывную функцію l_x. Поэтому на практикѣ для упрощенія принимаютъ обыкновенно:

$$E_x = e_x + \frac{1}{2}\cdot$$

Въ англійской литературѣ величина $e_x + \frac{1}{2}$ называется обыкновенно *полнымъ ожиданіемъ жизни* [1]).

§ 2. Часто разсматривается понятіе о такъ называемой *вѣроятной продолжительности жизни лица возраста* x, которая опре-

---

[1]) $e_x + \frac{1}{2}$ — complte expectation of life

l_x — curtate capecctation of life.

дѣляется такимъ образомъ: подъ вѣроятной продолжительностью жизни w_x разумѣется величина, удовлетворяющая уравненію:

$$l_{x+w_x} = \frac{1}{2} l_x .$$

Другими словами w_x выражается число лѣтъ, послѣ котораго функція l_x становится равной $\frac{1}{2}$ первоначальнаго ея значенія, соотвѣтствующаго разсматриваемому возрасту x лѣтъ. Такимъ образомъ, каждое лицо возраста x лѣтъ имѣетъ одинаковую вѣроятность умереть, какъ до истеченія вѣроятной продолжительности жизни, такъ и по истеченіи ея.

§ 3. Обращаемся теперь къ опредѣленію такъ называемаго *коэффиціента смертности* [1]). Для этой цѣли мы предположимъ функцію l_x непрерывной отъ перемѣнной x и введемъ въ разсмотрѣніе новую функцію μ_x, опредѣляемую слѣдующимъ уравненіемъ:

$$\frac{l_x - l_{x+dx}}{l_x} = \mu_x dx .$$

Слѣдовательно, у насъ получается вѣроятность лицу возраста x умереть въ безконечно маломъ промежуткѣ времени $(x, x+dx)$; очевидно, что:

$$(1) \quad . \; . \; . \; . \; . \; . \; . \; . \quad \mu_x = -\frac{dl_x}{l_x dx} = -\frac{dlgl_x}{dx} .$$

Величина μ_x и есть то, что называютъ коэффиціентомъ смертности. Существенное различіе коэффиціента смертности отъ вѣроятности заключается въ томъ, что вѣроятность измѣряетъ смертность въ нѣкоторый конечный промежутокъ времени, а коэффиціентъ смертности — въ извѣстный моментъ жизни. Интегрируя (1) мы получаемъ:

$$(2) \quad . \; . \; . \; . \; . \quad lgl_x = -\int \mu_x dx; \text{ откуда: } l_x = e^{-\int \mu_x dx},$$

такъ что коэффиціентъ смертности равняется функціи, замѣняющей функцію l_x. Двѣ функціи l_x и μ_x находятся въ такомъ соотношеніи, что, зная одну, мы находимъ другую.

---

[1]) Taux de mortabilité instantané.
Sterblichkeits Intensität oder Sterblichkeits Kraft
Force of mortality.

VI.

Различные виды таблицъ смертности.

§ 1. Основнымъ требованіемъ для составленія таблицы смертности мы ставили понятіе о классѣ людей; въ виду этого мы должны допустить возможность измѣненія вѣроятностей жизни при переходѣ отъ одного класса людей къ другому; также является важнымъ вопросъ о возможности измѣненія смертности людей извѣстнаго класса въ зависимости отъ измѣненія культурныхъ условій жизни. Статистическій опытъ страховыхъ учрежденій, а также статистическія изслѣдованія надъ общимъ составомъ народонаселенія не даютъ въ настоящее время какихъ либо категорическихъ указаній относительно этихъ вопросовъ, но все-таки слѣдуетъ подчеркнуть небольшое число фактовъ, установленныхъ съ полной несомнѣнностью. Первый фактъ такого рода состоитъ въ несомнѣнномъ различіи законовъ смертности въ зависимости отъ пола. При этомъ, относительно смертности женщинъ по сравненію съ таковой же для мужчинъ приходится отмѣтить слѣдующіе факты: въ первые года жизни до 10 лѣтъ смертность меньше; между 10 и 15-ю годами, а также 27—35-ю годами, замѣчается рѣзкое ухудшеніе. Въ дальнѣйшіе годы, особенно послѣ 50-ти лѣтъ, женская смертность отклоняется въ благопріятномъ смыслѣ отъ мужской.

Высказывавшіяся предположенія о томъ, что съ развитіемъ культуры смертность падаетъ, не подтверждаются съ достаточной ясностью наблюденіями надъ общимъ народонаселеніемъ. Одинъ фактъ только несомнѣненъ, это паденіе дѣтской смертности въ Европѣ за XIX столѣтіе. Что же касается большихъ возрастовъ, начиная съ 50-ти лѣтъ, то замѣчается скорѣе ухудшеніе, чѣмъ улучшеніе.

Иначе обстоитъ дѣло съ таблицей смертности, составленной страховыми учрежденіями по непосредственнымъ наблюденіямъ надъ ихъ кліентами. Такъ, напримѣръ, опытъ англійскихъ страховыхъ учрежденій даетъ замѣтное улучшеніе смертности.

§ 2. Практика составленія таблицъ страховыми учрежденіями привела къ цѣлому ряду весьма важныхъ фактовъ. Такъ, напримѣръ, оказалось, что смертность различна при различныхъ видахъ страхованія, поэтому въ настоящее время является важнымъ имѣть особенныя таблицы, выведенныя на основаніи статистическаго опыта, для различныхъ видовъ страхованія. Напримѣръ, если мы разсмотримъ два вида страхованій: 1) страхованіе на случай смерти и 2) страхованіе на случай дожитія, то причина возможныхъ различныхъ вѣроятностей жизни является болѣе или менѣе очевидной, ибо самъ характеръ страхованія соотвѣтствуетъ извѣстному подбору кліентовъ. Такъ, страхованія на случай смерти заключаютъ съ бóльшей охотой люди слабаго здоровья, тогда какъ страхованія на случай дожитія нравятся людямъ болѣе здоровымъ. Въ заключеніе мы должны указать на весьма важный фактъ, замѣченный Higham'омъ въ 1851 году относительно англійскихъ страховыхъ обществъ, а именно, что на смертность по стольку, по скольку она выражается въ составленныхъ таблицахъ смертности, вліяетъ также продолжительность страхового договора. Оказывается, что съ теченіемъ времени нейтрализуется вліяніе первоначальнаго подбора, зависящаго отъ того или другого характера страхованія и таблица смертности принимаетъ болѣе и болѣе общій характеръ. Особенно замѣтно повышеніе смертности застрахованныхъ противъ, такъ сказать, нормальной нормы, въ первые 3—5 лѣтъ послѣ заключенія страхованія.

Въ виду сказаннаго, въ послѣднее время начинаютъ распространяться таблицы смертности, въ которыхъ принимается въ расчетъ также и истекшая продолжительность страхового договора [1]).

VII.

Выравниваніе таблицъ. Законъ Gompertz-Makeham'а.

§ 1. Всякая таблица смертности, употребляемая въ настоящее время въ страховой математикѣ, представляетъ изъ себя рядъ чиселъ, соотвѣтствующихъ цѣлымъ годамъ. Геометрически это соотвѣтствуетъ ряду отдѣльныхъ ординатъ l_0, l_1 . . . . (черт. 6), разстояніе между которыми по оси x-овъ, изображающей время, равняется одному году. Получаются такимъ образомъ отдѣльно стоящія точки: M_0, M_1, M_1, M_3 . . . . и т. д.

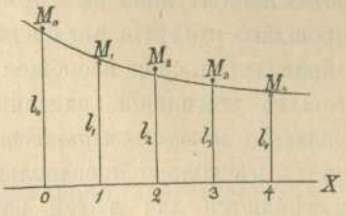
Подъ выравниваніемъ таблицы смертности понимается проведеніе плавной кривой линіи, которая по возможности ближе под-

---

[1]) Doppelt abgestufte Sterbetafel.

ходила бы къ ряду точекъ, дающихся таблицей. Мы приходимъ такимъ образомъ къ задачамъ интерполированія (Э. м. XI, § 4).

Въ виду неопредѣленности указанной задачи существуетъ цѣлый рядъ пріемовъ такого выравниванія. Отсылая для болѣе подробнаго знакомства съ этими пріемами къ большимъ трактатамъ страховой математики [1]), мы тѣмъ не менѣе укажемъ на нѣкоторые основные принципы такого выравниванія. Существуютъ *геометрическіе* пріемы выравниванія, при помощи которыхъ даются правила для чертежника, нанесшаго на бумагу точки M_0, M_1, M_2, ... , проведенія искомой плавной линіи. Другіе пріемы носятъ *аналитическій* характеръ. Тутъ ищется непрерывная функція l_x такая, чтобы она по возможности мало уклонялась для цѣлыхъ значеній годовъ отъ соотвѣтствующихъ по таблицѣ смертности величинъ.



Черт. 6.

При этомъ можетъ быть употребленъ способъ *наименьшихъ квадратовъ* (Э. м. IX, § 6).

Мы разсмотримъ болѣе подробно способъ выравниванія таблицъ смертности, основанный на подысканіи такой аналитической функціи, которая выражала бы по возможности близко законъ смертности, характеризуемый таблицей.

Попытки нахожденія такихъ *формулъ смертности* начались сравнительно давно. Конечно, всѣ такія формулы могутъ имѣть только теоретическій характеръ и они тѣмъ болѣе заслуживаютъ практическаго вниманія, чѣмъ ближе соотвѣтствуютъ таблицамъ смертности.

Первая гипотеза была сдѣлана Moivres'омъ. Moivres на основаніи Halley'евской таблицы [2]) пришелъ къ заключенію, что съ достаточнымъ для того времени приближеніемъ можно считать число доживающихъ падающимъ по прямой линіи. Поэтому Moivres далъ такую гипотетическую формулу: $l_x = a\,(86 - x)$.

---

[1]) Для первоначальнаго знакомства съ этимъ вопросомъ можно рекомендовать въ русской литературѣ: С. Е. Савичъ. Элементарная теорія страхованія жизни и трудоспособности. С.-Петербургъ. 1909 г. Въ иностранной: Hugo Broggi. Versicherungs Mathematik. Leipzig. 1911 г.

[2]) Эта таблица составлена по наблюденіямъ надъ населеніемъ города Бреславля за время 1686—1691 г. г.

Здѣсь a есть нѣкоторый постоянный коэффиціентъ, число 86 представляетъ предѣльный возрастъ Halley'евской таблицы, такъ что $l_{86} = 0$.

Не перечисляя дальнѣйшихъ предложенныхъ формулъ, мы остановимся лишь на формулѣ, предложенной въ 30-хъ годахъ прошлаго столѣтія англійскимъ актуаріемъ Gompertz'омъ, которая обратила на себя всеобщее вниманіе. Эта формула, послѣ нѣкоторыхъ измѣненій, внесенныхъ въ нее Makeham'омъ, получила названіе *закона смертности Gompertz - Makeham'a*. Она заключаетъ нѣсколько произвольныхъ постоянныхъ, значенія которыхъ измѣняются для людей различныхъ классовъ. При извѣстномъ выборѣ этихъ постоянныхъ, формула, кромѣ дѣтскихъ возрастовъ, давала во всѣхъ случаяхъ, когда были сдѣланы сравненія, результаты, близко совпадающіе съ наблюденіями. Оба автора исходятъ изъ той мысли, что смертность зависитъ отъ двухъ главныхъ комплексовъ причинъ: 1) отъ внѣшнихъ и случайныхъ причинъ смерти, не имѣющихъ ничего общаго съ физическимъ состояніемъ, а слѣдовательно и съ возрастомъ каждаго отдѣльнаго индивидуума, и 2) причинъ, зависящихъ исключительно отъ возраста.

Gompertz предполагаетъ коэфиціентъ смертности возрастающимъ въ геометрической прогрессіи съ возрастомъ. Makeham добавляетъ еще одинъ постоянный коэффиціентъ, долженствующій выразить вліяніе случайныхъ причинъ.

Итакъ, формула Gompertz-Makeham'а будетъ такая:

$$(1) \ldots\ldots\ldots\ldots \mu_x = A + Bc^x,$$

гдѣ A, B и c нѣкоторыя постоянныя величины, подлежащія опредѣленію изъ наблюденій. На основаніи формулы (2) V гл. § 3 мы получаемъ:

$$(2) \ldots lgl_x = -\int \left(A + Bc^x\right) dx = -Ax - \frac{B}{lgc} c^x + lgk,$$

гдѣ k есть постоянная величина, вводимая интегрированіемъ.

Обозначая: $-A = lgs$; $-\frac{B}{lgc} = lgg$, получимъ: $lgl_x = xlgs + c^x lgg + lgk$, откуда:

$$(3) \ldots\ldots\ldots\ldots l_x = ks^x g^{c^x}.$$

Въ послѣднемъ видѣ формула употребляется англійскими писателями. Удобнѣе однако перейти къ обычному виду показательныхъ функцій, взятыхъ при Неперовскомъ основаніи e.

Въ самомъ дѣлѣ, положимъ въ формулѣ (2):

$$A=\alpha;\ \frac{B}{lgc}=\beta;\ k=C;\ c=e^{\gamma};$$

тогда, очевидно, получимъ:

$$(4)\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ l_x=Ce^{-\alpha x-\beta e^{\gamma x}}.$$

Постоянная $k=C$, конечно, совершенно произвольна, что слѣдуетъ на основаніи свойства 1-го функціи l_x, указаннаго въ главѣ I, § 2.

Вычислимъ по формулѣ (4) вѣроятность $p\ (x, x+n)$. Такъ какъ $p(x,\ x+n)$ равняется $\frac{l_{x+n}}{l_x}$, то, прилагая формулу (4), получаемъ:

$$(5)\ .\ .\ .\ p(x,\ x+n)=\frac{Ce^{-\alpha(x+n)-\beta e^{\gamma(x+n)}}}{Ce^{-\alpha x-\beta e^{\gamma x}}}=e^{-\alpha n-\beta(c^n-1)c^x},$$

такъ какъ

$$c=e^{\gamma}.$$

Разсмотримъ r индивидуумовъ одного или различныхъ возрастовъ: (x), (y), (z) . . . . . . Вѣроятность, что они всѣ проживутъ еще n лѣтъ, по теоріи объ умноженіи вѣроятностей будетъ такая:

$$e^{-\alpha n-\beta(c^n-1)c^x}\,.\,e^{-\alpha n-\beta(c^n-1)c^y}\,.\,e^{-\alpha n-\beta(c^n-1)c^z}\ .\ .\ .\ .\ .=$$

$$=e^{-\alpha rn-\beta(c^n-1)(c^x+c^y+c^z+\ .\ .\ .\ .\ .)}.$$

Совершенно же подобнымъ образомъ, если мы имѣемъ группу ρ индивидуумовъ одного и того же возраста ξ, то вѣроятность имъ всѣмъ прожить еще n лѣтъ будетъ выражаться формулой:

$$e^{-\alpha\rho n-\beta(c^n-1)(c^\xi+c^\xi+c^\xi+\ .\ .\ .\ .\ .)}=e^{-\alpha\rho n-\beta\rho(c^n-1)c^\xi}.$$

Рѣшимъ задачу, когда двѣ группы: $[x,\ y,\ z\ .\ .\ .\ .\ .]_r$ и $[\xi, \xi, \xi,\ .\ .\ .\ .\ ]_\rho$ будутъ имѣть общую вѣроятность прожить n лѣтъ. Мы приходимъ, очевидно, къ равенству:

$$(6)\ .\ -\alpha\rho n-\beta(c^n-1)\rho c^\xi=-\alpha rn-\beta(c^n-1)\,[c^x+c^y+c^z+\ .\ .\ .].$$

Разсмотримъ сначала случай формулы Gompertz'а, когда $\alpha=0$, тогда формула (6) даетъ:

$$-\beta(c^n-1)\rho c^\xi = -\beta(c^n-1)\,[c^x+c^y+c^z+\ldots]$$

или, полагая здѣсь $\rho=1$, получимъ:

$$(7) \ldots\ldots\ldots c^\xi = c^x + c^y + c^z \ldots$$

Эта формула носитъ названіе *свойства Gompertz'а* и показываетъ, что можно свести вычисленіе вѣроятности для группы людей $(x, y, z, \ldots)$ прожить еще n лѣтъ на вычисленіе вѣроятности одному лицу прожить еще n лѣтъ. Возрастъ ξ этого фиктивнаго лица вычисляется при законѣ Gompertz'а по формулѣ (7). Свойство Gompertz'а перестаетъ имѣть мѣсто, если α не равно 0, т. е. если мы приходимъ къ гипотезѣ Makeham'а; но тогда можно получить изъ уравненія (6) формулу подобную, а именно, если положить $\rho=r$. Тогда получимъ:

$$(8) \ldots\ldots\ldots r c^\xi = c^x + c^y + c^z + \ldots\ldots$$

Формула (8) выражаетъ такъ называемое *свойство Makeham'а*. По этой формулѣ вычисленіе вѣроятности дожитія группы r индивидуумовъ $(x, y, z \ldots)$ приводится къ вычисленію вѣроятности такого же дожитія фиктивной группы того же числа r индивидуумовъ возраста ξ, опредѣляемаго по формулѣ (8).

VIII.

Выводъ нѣкоторыхъ формулъ политической ариѳметики.

§ 1. Обозначимъ буквой i [1]) доходъ, приносимый въ годъ единицей капитала и назовемъ его *нормой годового роста капитала* или просто *ростомъ капитала*. Обыкновенно ростъ выражается по формулѣ:

$$(1) \ldots\ldots\ldots\ldots i = \frac{k}{100},$$

причемъ число k называется процентомъ, и тогда говорятъ, что капиталъ приноситъ k процентовъ въ годъ, что записываютъ такъ: $k\,^0/_0$.

[1]) Начальная буква слова intérêt.

Проценты, какъ извѣстно, бываютъ простые и сложные. Формула роста единицы капитала по простымъ процентамъ такая:

(2) . . . . $c'_n = 1 + ni$,

гдѣ n число лѣтъ, въ которое капиталъ приноситъ проценты. При сложныхъ процентахъ формула прироста капитала будетъ такая:

(3) . . . . $c_n = (1+i)^n$ [1]).

Сравнивая формулы (2) и (3) мы замѣчаемъ, что ростъ капитала по сложнымъ процентамъ совершается быстрѣе, ибо раскладывая по биному Ньютона формулу (3), мы получаемъ:

$$c_n = 1 + ni + \frac{n(n-1)}{1.2} i^2 + \frac{n(n-1)(n-2)}{1.2.3.} i^3 + \ldots ,$$

откуда

$$c_n > 1 + ni \quad \text{или} \quad c_n > c'_n.$$

§ 2. Если мы годъ раздѣлимъ на m частей, напримѣръ на 12 мѣсяцевъ, и примемъ за единицу времени эту долю года, то, если мы будемъ сложные проценты разсматривать съ ростомъ, равнымъ $\frac{i}{m}$ единицы капитала въ каждую изъ этихъ долей года, то по истеченіи года единица капитала обратится въ

(4) . . . . $\left(1+\frac{i}{m}\right)^m$.

Легко убѣдиться, что число (4) будетъ больше числа $1+i$. Ибо, раскладывая по формулѣ бинома, мы получимъ:

$$\left(1+\frac{i}{m}\right)^m = 1 + m\frac{i}{m} + \frac{m(m-1)}{1.2}\frac{1^2}{m^2} + \frac{m(m-1)(m-2)}{1.2.3}\frac{1^3}{m^3}\cdots$$

или

$$\left(1+\frac{i}{m}\right)^m > 1 + m\frac{i}{m}$$
$$> 1 + i.$$

Послѣдняя формула учитъ, что, если мы раздѣлимъ годъ на нѣкоторое число m частей и каждой этой части сопоставимъ ростъ, равный $\frac{i}{m}$ первоначальнаго годового роста, то нароста-

---

[1]) Начальная алгебра Киселева.

ніе капитала по сложнымъ процентамъ пойдетъ быстрѣе. Такъ, напримѣръ, если мы разсмотримъ съ одной стороны приращеніе капитала изъ 6 сложныхъ годовыхъ процентовъ, а съ другой стороны изъ 3 сложныхъ процентовъ полугодовыхъ, то во второмъ случаѣ капиталъ будетъ расти быстрѣе.

§ 3. Въ теоріи долгосрочныхъ финансовыхъ операцій часто разсматриваютъ непрерывное наростаніе капитала, т. е., другими словами, берется нѣкоторый фиктивный годовой ростъ капитала j и годъ дѣлится на m равныхъ частей, причемъ каждой части сопоставляется ростъ $\frac{j}{m}$. Тогда въ продолженіи года, какъ было сказано выше, капиталъ обратится въ такой:

$$(1) \ldots\ldots\ldots \left(1+\frac{j}{m}\right)^m.$$

Подъ непрерывнымъ ростомъ капитала мы будемъ понимать то видоизмѣненіе, которое будетъ претерпѣвать формула (1) при увеличеніи числа m до безконечности. Какъ извѣстно изъ дифференціальнаго исчисленія (Э. м. гл. III, § 29), мы будемъ имѣть:

$$lim\left(1+\frac{j}{m}\right)^m = e^j, \qquad \text{при } m=\infty.$$

Итакъ, годовой приростъ капитала, соотвѣтствующій непрерывному наростанію процентовъ, найдется по формулѣ:

$$1+i=e^j, \quad \text{т. е.} \quad i=e^j-1.$$

§ 4. На основаніи всего вышесказаннаго мы видимъ, что капиталъ C по истеченіи n лѣтъ обратится въ нѣкоторый новый капиталъ M, причемъ этотъ новый капиталъ будетъ вычисляться при простыхъ процентахъ по формулѣ:

$$(1) \ldots\ldots\ldots M=C(1+ni).$$

При сложныхъ процентахъ при годовомъ расчетѣ роста по формулѣ:

$$(2) \ldots\ldots\ldots M=C(1+i)^n.$$

И наконецъ по сложнымъ процентамъ, но при наростаніи процентовъ въ каждую долю $\frac{i}{m}$ года по формулѣ:

$$(3) \ldots\ldots\ldots M=C\left(1+\frac{i}{m}\right)^{mn}.$$

Если мы въ формулахъ (1), (2), (3) положимъ $M=1$, то получаемъ:

$$(1) \ldots C=\frac{1}{1+ni}; \quad (2) \ C=\frac{1}{(1+i)^n}; \quad (3) \ C=\frac{1}{\left(1+\frac{i}{m}\right)^{mn}}.$$

Такимъ образомъ мы получили формулы, выражающія число C, т. е. *современную стоимость капитала, равнаго единицѣ*, причемъ эта единица капитала подлежитъ уплатѣ черезъ n лѣтъ. Разница $1-C$ носитъ названіе *дисконта*. Во всемъ дальнѣйшемъ мы будемъ разсматривать годовой дисконтъ и обозначать его такъ:

$$d=1-V, \qquad \text{гдѣ} \qquad V=\frac{1}{1+i}.$$

При этомъ, если не будетъ оговорокъ, то мы будемъ употреблять формулу (2) обыкновенныхъ годовыхъ сложныхъ процентовъ.

Итакъ мы видимъ, что, если мы желаемъ найти современную стоимость капитала M, уплачиваемаго черезъ n лѣтъ, или, какъ говорятъ, будемъ дисконтировать будущую уплату капитала M къ настоящему моменту, то придется согласно формулѣ (2) капиталъ M умножить на выраженіе:

$$V^n=\frac{1}{(1+i)^n},$$

такъ что современная стоимость капитала будетъ MV^n. Множитель V^n носитъ названіе дисконтирующаго множителя.

Ренты.

§ 5. Подъ *рентой* разумѣется періодически производимая уплата денегъ. Уплачиваемыя суммы при рентахъ могутъ и не быть одинаковыми. Мы оставимъ однако такіе случаи въ сторонѣ и обратимъ наше вниманіе на ренты, въ которыхъ періодически уплачивается одна и таже сумма. Рентъ, по характеру производимыхъ уплатъ, различается нѣсколько сортовъ. Прежде всего мы отличаемъ ренты временныя отъ рентъ вѣчныхъ.

Временной рентой называется такая, которая прекращается послѣ извѣстнаго числа платежей.

Рента называется *вѣчной*, если предполагается, что платежи ея никогда не прекращаются. Само собой разумѣется, что понятіе о вѣчной рентѣ есть понятіе фиктивное, и такая рента фактически не можетъ быть осуществлена.

Далѣе ренты раздѣляются на годовыя, полугодовыя, уплачиваемыя въ четверть года, мѣсячныя и т. д. соотвѣтственно промежутку времени, черезъ который происходятъ платежи.

Ренты раздѣляются также на *ренты praenumerando,* когда уплаты происходятъ въ началѣ каждаго промежутка времени, и на *ренты postnumerando,* когда уплата происходитъ въ концѣ каждаго промежутка времени.

Рента называется *немедленной,* если промежутки времени, соотвѣтствующіе уплатамъ, начинаются съ момента договора.

Рента называется *отсроченной*, если начало счета промежутковъ временъ отсрачивается на нѣкоторое число m лѣтъ отъ момента заключенія договора.

Указанныя выше различныя ренты соотвѣтствуютъ различнымъ математическимъ обстоятельстваыъ, связаннымъ съ вычисленіями, относящимися къ этимъ рентамъ. Мы оставимъ поэтому въ сторонѣ различныя подраздѣленія рентъ по признакамъ не математическаго характера, такъ, напримѣръ, существуетъ названіе ренты погашенія, когда цѣлью ренты является погашеніе долга и ренты сбереженія, когда цѣлью ренты является накопленіе къ нѣкоторому моменту времени извѣстнаго капитала.

§ 6. Мы будемъ разсматривать ренты только съ одной точки зрѣнія, а именно, разсматривать современную дисконтированную стоимость всѣхъ платежей ренты, т. е., другими словами, будемъ разсматривать стоимость всей ренты въ моментъ заключенія сдѣлки. При этомъ обнаруживается весьма важное явленіе, что современная стоимость ренты есть всегда число конечное, будетъ ли рента временная или вѣчная.

Разсмотримъ прежде всего *временную немедленную ренту postnumerando.* Очевидно, что мы можемъ разсматривать ренту, состоящую въ уплатѣ въ каждый срокъ единицы капитала, потому что, если рента будетъ состоять въ уплатѣ a единицъ капитала, то придется просто умножить на a тѣ числа, которыя мы получимъ. Итакъ, пусть произведено n платежей равнаго числа единицъ капитала въ концѣ 1-го, 2-го, ... n-го годовъ. Тогда, если мы обозначимъ стоимость такой ренты черезъ V_n, то, очевидно, мы будемъ имѣть:

$$(1) \ldots V_n = V + V^2 + V^3 + \ldots\ldots + V^n = V\frac{1 - V^n}{1 - V}.$$

Если мы примемъ въ послѣдней формулѣ, что $V = \frac{1}{1+i}$, то формулу (1) можно будетъ написать такъ:

$$(2) \ldots V_n = \frac{1}{1+i} + \left(\frac{1}{1+i}\right)^2 + \ldots\ldots + \left(\frac{1}{1+i}\right)^n =$$

$$= \frac{(1+i)^{n-1} + (1+i)^{n-2} + \ldots(1+i) + 1}{(1+i)^n} = \frac{C_n}{(1+i)^n}.$$

Здѣсь

$$C_n = (1+i)^{n-1} + (1+i)^{n-2} + \ldots(1+i) + 1.$$

Легко видѣть, что C_n представлять изъ себя капиталъ, накопленный отъ платежей ренты къ моменту послѣдняго взноса ренты. Въ самомъ дѣлѣ, къ моменту послѣдняго взноса ренты, этотъ послѣднiй взносъ равняется единицѣ. Предпослѣднiй взносъ ренты, бывшiй одинъ годъ въ ростѣ, дастъ $(1+i)$. Третiй съ конца взносъ, бывшiй два года въ ростѣ, дастъ $(1+i)^2$ и т. д. Очевидно, что капиталъ, накопившiйся отъ всѣхъ взносовъ, будетъ равенъ $1 + (1+i) + (1+i)^2 + \ldots$, т. е. какъ разъ величинѣ C_n. Формула (2) даетъ намъ, очевидно, теорему:

Капитализированная стоимость ренты V_n равняется капиталу C_n, накопленному изъ отдѣльныхъ платежей ренты къ моменту послѣдняго взноса и учтенному [раздѣленному на $(1+i)^n$] къ началу перваго года счета ренты.

§ 7. Разсмотримъ теперь ренту praenumerando, уплачиваемую въ размѣрѣ одного рубля, причемъ производится n взносовъ. Очевидно, что дисконтъ къ началу перваго года всѣхъ этихъ взносовъ дастъ числа: 1, V, V^2, ... V^{n-1}.

Отсюда, если мы обозначимъ стоимость такой ренты черезъ V'_n, то получимъ:

$$(1) \ldots V'_n = 1 + V + V^2 + \ldots\ldots + V^{n-1} = \frac{1 - V^n}{1 - V}.$$

Сравнивая послѣднюю формулу съ формулой (1), мы получаемъ:

$$(2) \ldots\ldots V'_n = \frac{V_n}{V} = V_n(1+i);$$

Мы приходимъ къ теоремѣ:

Капитализированная стоимость ренты praenumerando равняется стоимости ренты postnumerando, наращенной процентами за одинъ годъ.

§ 8. Предполагая n равнымъ безконечности въ формулѣ: $V_n = V\frac{1-V^n}{1-V}$, получимъ стоимость вѣчной ренты:

$$V_\infty = \frac{V}{1-V} = \frac{\frac{1}{1+i}}{1-\frac{1}{1+i}} = \frac{1}{i}.$$

Получается теорема: *Стоимость вѣчной ренты есть величина обратная годовому росту.*

Отсюда слѣдуетъ, что съ увеличеніемъ процента капитализированная величина ренты убываетъ. Легко убѣдиться, что подобныя же обстоятельства имѣютъ мѣсто и для временной ренты, потому что мы имѣемъ:

$$V_n = V + V^2 + \ldots\ldots + V^n.$$

Но величина $V = \frac{1}{1+i}$, очевидно, убываетъ при возрастаніи процента i, а значитъ убываютъ всѣ числа V, V^2, V^3, ... V^n, а значитъ и число V_n.

§ 9. Обращаемая теперь къ разсмотрѣнію *отсроченной ренты.* Мы будемъ употреблять знакъ:

$$_m|V_n$$

для выраженія капитализированной стоимости временной ренты, уплачиваемый postnumerando въ теченіе n лѣтъ. Лѣвый значекъ m, отдѣленный отъ V вертикальной чертой, указываетъ, что начало платежей ренты отсрочено на m лѣтъ. Стоимость отдѣльныхъ платежей въ этомъ случаѣ будетъ, очевидно:

$$V^{m+1},\ V^{m+2},\ \ldots\ldots\ V^{m+n},$$

откуда стоимость всей ренты выразится:

$$_m|V_n = V^{m+1} + V^{m+2} + \ldots + V^{m+n} = V^m(V + V^2 + \ldots + V^n),$$

или другими словами:

$$_m|V_n = V^m V_n.$$

Получаемъ теорему: *Капитализированная стоимость $_m|V_n$ временной въ теченіи n лѣтъ ренты, отсроченной на m лѣтъ,* *равна капитализированной стоимости немедленной ренты, дисконтированной назадъ за промежутокъ времени отсрочки.*

§ 10. Совершенно подобнымъ образомъ получимъ для отсроченной ренты praenumerando формулу:

$$_{m|}V'_n = v^m V'_n.$$

§ 11. Теорія рентъ находится въ тѣсной связи съ теоріей такъ называемыхъ *срочныхъ уплатъ*. Срочными уплатами называются платежи, вносимые въ постоянномъ размѣрѣ и черезъ одинаковые промежутки времени, необходимые и достаточные для погашенія занятаго капитала съ процентами. Такъ, напримѣръ, если долгъ въ A рублей долженъ быть погашенъ срочными уплатами въ размѣрѣ x_n рублей, производимыми разъ въ году въ концѣ его, причемъ долгъ долженъ быть уплаченъ въ n лѣтъ, тогда мы получаемъ, какъ извѣстно изъ элементарной математики, для нахожденія срочной уплаты x_n формулу:

(1) . . $A(1+i)^n = x_n[(1+i)^{n-1} + (1+i)^{n-2} + \ldots (1+i) + 1]$,

т. е.

(2) . . . . $A = x_n V_n$.

Имѣемъ теорему: *Занятый капиталъ представляетъ капитализированную стоимость ренты, уплачиваемую въ размѣрѣ, равномъ срочной уплатѣ, и въ одинаковые съ нею сроки.*

IX.

Основные принципы безобиднаго страхованія.

§ 1. Приступая къ изложенію основныхъ положеній математической теоріи страхованія жизни, „начнемъ съ конкретнаго примѣра. Пусть нѣкоторое лицо A въ возрастѣ m лѣтъ обращается въ страховое общество B, причемъ страхуетъ въ этомъ обществѣ свою жизнь. Другими словами, оно заключаетъ съ обществомъ такую сдѣлку: общество B обязано въ случаѣ смерти A уплатить немедленно нѣкоторый капиталъ a наслѣдникамъ умершаго, съ другой стороны лицо A обязывается вносить пожизненно въ общество нѣкоторую ежегодную уплату α.

Лицо A, страхующее жизнь, мы будемъ называть *страхователемъ*, общество B — *страховщикомъ*. Если страхователь страхуетъ не свою жизнь, а жизнь нѣкотораго третьяго лица, то это лицо мы будемъ называть *застрахованнымъ*. Ежегодная уплата α страхователя страховщику носитъ названіе *страховой преміи*. Въ удостовѣреніе заключеннаго договора страхователь получаетъ отъ страховщика бумагу, называемую *страховымъ полисомъ*. По предъявленіи этого полиса наслѣдники получаютъ застрахованный капиталъ a.

Указанный нами договоръ называется страхованіемъ *на случай смерти*. Другая форма страхованія, носящая названіе *страхованія на дожитіе*, состоитъ въ томъ, что въ случаѣ достиженія страхователемъ нѣкотораго опредѣленнаго возраста n лѣтъ страхователь получаетъ самъ застрахованный имъ капиталъ.

Кромѣ этихъ двухъ главныхъ формъ страхованія потребности жизни выработали цѣлый рядъ комбинацій, связанныхъ съ различными возможными обстоятельствами жизни. Сюда относятся самыя разнообразныя пенсіонныя кассы для вдовъ и сиротъ, а также на случай старости и утраты трудоспособности.

Разсматривая приведенную выше сдѣлку страхователя A со страховщикомъ B мы прежде всего замѣчаемъ, что эта сдѣлка подходитъ вполнѣ подъ опредѣленіе понятія объ игрѣ, данное въ § 30 (гл. XIV, Э. м.).

Въ самомъ дѣлѣ, уплата каждой преміи α является событіемъ недостовѣрнымъ, такъ какъ эта уплата происходитъ только въ томъ случаѣ, если страхователь въ моментъ уплаты живъ. Чѣмъ страхователь старше, тѣмъ вѣроятность уплаты преміи дѣлается меньше. Съ другой стороны, уплата обществомъ застрахованнаго капитала a совершается послѣ смерти страхователя; точный моментъ смерти неизвѣстенъ ни страхователю, ни страховщику. Итакъ, переходъ денежныхъ суммъ отъ страхователя къ страховщику и обратно совершается при обстоятельствахъ недостовѣрныхъ, а слѣдовательно у насъ имѣется наличность нѣкоторой игры.

Отсюда очевидно, что при установленіи математическихъ принциповъ страховыхъ операцій необходимо придерживаться правила математической безобидности. Въ самомъ дѣлѣ, если мы допустимъ отрицательное математическое ожиданіе для страхового общества, то при достаточно большомъ числѣ операцій такое общество придетъ къ банкротству. Если, съ другой стороны, допустить положительное математическое ожиданіе для общества, то это равносильно допущенію обогащенія этого общества въ ущербъ интересамъ остального населенія". [1])

§ 2. Для установленія математической безобидности страховой сдѣлки мы установимъ понятіе о т. наз. *современномъ ожиданіи платежа a*, который долженъ послѣдовать черезъ m лѣтъ при наступленіи нѣкотораго недостовѣрнаго событія A, имѣющаго вѣроятность p. Въ такихъ случаяхъ мы будемъ всегда разсматривать математическое ожиданіе ap этого платежа и полученную величину дисконтировать къ начальному моменту, такъ что получимъ окончательное выраженіе

(1) . . . . apV^m.

Выраженіе (1) мы будемъ называть *современнымъ ожиданіемъ платежа a*. Если платежъ a будетъ имѣть мѣсто непремѣнно, т. е. другими словами, если событіе A достовѣрно (p равно единицѣ), то современное ожиданіе платежа обращается въ современную стоимость платежа, о которой мы говорили въ § 4 предыдущей главы.

Ясное дѣло, что для математической безобидности страховой сдѣлки необходимо, чтобы сумма современныхъ ожиданій всѣхъ возможныхъ платежей страхователя равнялась суммѣ современныхъ ожиданій всѣхъ возможныхъ платежей страховщика. Получается такимъ образомъ уравненіе, дающее возможность установить размѣръ подлежащей къ уплатѣ страховой преміи. Такая премія, вычисленная на основаніи принципа безобидности, носитъ названіе netto-преміи.

Страховое общество не можетъ, однако, удовольствоваться netto-преміями, такъ какъ оно несетъ расходы, связанные съ администраціей общества, расходы на жалованіе служащихъ, на наемъ помѣщеній и т. д. Поэтому къ вычисленной netto-преміи прибавляется нѣкоторая надбавка, дающая окончательную премію, которую и платитъ фактически страхователь. Эта премія носитъ названіе brutto-преміи.

Размѣры brutto-премій, или, другими словами, дѣйствительные *тарифы* страхового учрежденія устанавливаются до извѣстной степени произвольно на основаніи общихъ законовъ конкуренціи и соотвѣтствія между спросомъ и предложеніемъ. Если

---

[1]) См. Э. м., стр. 558, 559.

подъ вліяніемъ конкурренціи страховое общество желаетъ возможно болѣе понизить тарифы, то оно прежде всего должно знать точный размѣръ netto-премій, чтобы не назначить тарифы ниже этихъ netto-премій. Для этой цѣли необходимо имѣть возможно совершенныя таблицы смертности, чтобы вычисленныя по нимъ вѣроятности соотвѣтствовали обстоятельствамъ, имѣющимъ на самомъ дѣлѣ мѣсто въ средѣ кліентовъ общества.

У насъ въ Россіи вслѣдствіе отсутствія надежныхъ таблицъ смертности, когда приходится пользоваться нѣмецкими таблицами, вся дѣятельность страховыхъ обществъ происходитъ до нѣкоторой степени въ темную.

§ 3. Ограничиваясь въ нашей математической теоріи безобиднымъ страхованіемъ, мы должны однако упомянуть о существующемъ обобщеніи понятія страхованія, о такъ называемомъ соціальномъ страхованіи [1]).

Нѣкоторые авторы понятіе о страхованіи на столько широко обобщаютъ, что разумѣютъ подъ нимъ всякую организованную взаимопомощь. При этомъ предполагается уклоненіе въ широкой мѣрѣ отъ принципа безобидности. Всѣ такого рода видоизмѣненія понятія о страхованіи можно характеризовать желаніемъ перейти отъ примѣненія понятія математическаго ожиданія къ примѣненію понятія нравственнаго ожиданія Даніила Бернулли, о которомъ упоминается въ Э. м. гл. XIV, § 31.

§ 4. Формы страхованія жизни могутъ быть очень разнообразными, и довольно трудно установить какую-нибудь опредѣленную ихъ классификацію. Прежде всего можно подраздѣлить формы страхованія по числу людей застрахованныхъ; при этомъ мы получаемъ двѣ главныхъ категоріи страхованій:

1) страхованіе одной жизни и

2) страхованіе на нѣсколько жизней.

Далѣе, какъ мы сказали выше, можно установить слѣдующихъ два главныхъ вида страхованій:

1) страхованіе на дожитіе, 2) страхованіе на случай смерти. Впрочемъ, надо оговориться, что указанныя категоріи могутъ быть болѣе или менѣе точно разграничены только въ случаѣ страхованія одного лица.

---

[1]) Соціальное страхованіе. Др. Виндоршкъ.

Въ случаѣ же страхованія нѣсколькихъ жизней указанное различіе теряется, и одна и та же сдѣлка можетъ имѣть характеръ какъ сдѣлки на дожитіе, такъ и сдѣлки на случай смерти, смотря по тому, на котораго изъ застрахованныхъ лицъ мы обращаемъ вниманіе.

§ 5. Гораздо важнѣе подраздѣленіе задачъ страховой математики на двѣ слѣдующія категоріи: 1) задачи, въ которыхъ вычисленіе netto-преміи совершается на основаніи вполнѣ опредѣленныхъ данныхъ условія сдѣлки и 2) задачи, для математическаго рѣшенія которыхъ требуются добавочныя гипотезы.

Мы будемъ называть задачи, не требующія добавочныхъ гипотезъ, задачами 1-го класса, и задачи, требующія такихъ гипотезъ, задачами 2-го класса. Такъ, напримѣръ, всѣ задачи, гдѣ приходится разсматривать вѣроятности дожитія или смерти, относящіяся къ дробнымъ годамъ, будутъ считаться во второмъ классѣ, потому что таблица смертности даетъ числа только относящіяся къ цѣлымъ годамъ. Какъ мы видѣли выше, для полученія вѣроятностей, относящихся къ дробнымъ годамъ, приходится употреблять пріемы интерполированія, причемъ дѣлать добавочную гипотезу о непрерывности функціи l_x.

§ 6. Во всемъ дальнѣйшемъ мы ограничимся разсмотрѣніемъ задачъ страхованія только по отношенію къ одной жизни, ибо на этихъ задачахъ выясняются основныя положенія страховой математики. Обобщеніе же задачъ на случай нѣсколькихъ жизней является сравнительно простымъ и нетребующимъ новыхъ принциповъ.

§ 7. Первая задача, которую мы разсмотримъ, будетъ состоять въ вычисленіи современной стоимости математическаго ожиданія платежей одной изъ сторонъ, заключающихъ страховой договоръ, безразлично которой — страхователя или страховщика. Пусть

$$(1) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; a_1, \; a_2, \; . \; . \; . \; . \; a_n$$

будутъ платежи, которые должны быть произведены въ концѣ перваго, второго и т. д. n-го годовъ, причемъ эти платежи будутъ имѣть мѣсто въ зависимости отъ нѣкоторыхъ недостовѣрныхъ событій, напримѣръ въ предположеніи, что лицо, долженствующее произвести платежъ, живо въ моментъ этой уплаты.

Пусть вѣроятности платежей (1) будутъ:

$$\pi_1,\ \pi_2,\ \ldots\ldots\ \pi_n,$$

тогда современное ожиданіе платежей (1), очевидно, выразится по формулѣ:

$$(2)\ \ldots\ldots\ a_1\pi_1 V + a_2\pi_2 V^2 + \ldots\ldots + a_n\pi_n V^n;$$

для сокращенія обозначимъ выраженіе (2) черезъ:

$$(3)\ \ldots\ldots\ldots\ \sum_{i=1}^{i=n} a_i\pi_i V^i.$$

Предположимъ, что въ тѣ же самые сроки противуположная сторона должна произвести уплаты:

$$b_1,\ b_2,\ \ldots\ldots\ b_n,$$

причемъ вѣроятность этихъ уплатъ будетъ:

$$\rho_1,\ \rho_2,\ \ldots\ldots\ \rho_n;$$

тогда получаемъ формулу:

$$(4)\ \ldots\ldots\ldots\ \sum_{i=1}^{i=n} b_i\rho_i V^i,$$

выражающую современное ожиданіе уплатъ другой стороны.

Условіе безобидности сдѣлки мы получимъ въ такомъ видѣ:

$$(5)\ \ldots\ldots\ldots\ \sum_{i=1}^{i=n} a_i\pi_i V^i = \sum_{i=1}^{i=n} b_i\rho_i V^i.$$

Пусть формула (3) относится къ платежамъ страховщика, а формула (4) къ платежамъ страхователя. Если мы обозначимъ для сокращенія:

$$a_i\pi_i V^i - b_i\rho_i V^i = \delta_i,$$

то условіе безобидности принимаетъ видъ:

$$(6)\ \ldots\ldots\ldots\ \sum_{i=1}^{i=n} \delta_i = 0.$$

Установленное для момента заключенія сдѣлки равенство ожиданій платежей обѣихъ сторонъ не сохраняется уже въ дальнѣйшемъ теченіи сдѣлки. Въ самомъ дѣлѣ, если бы мы захотѣли разсмотрѣть соотношеніе между обязательствами двухъ сторонъ въ началѣ второго года, то съ одной стороны замѣтили бы, что ожиданіе платежей страхового учрежденія къ началу второго года было бы:

$$a_2\pi_2 V + a_3\pi_3 V^2 + \ldots\ldots + a_n\pi_n V^{n-1}$$

или, что одно и то же,

$$(7) \quad \ldots\ldots\ldots\ldots \quad V^{-1} \sum_{i=2}^{i=n} a_i\pi_i V^i .$$

Совершенно подобнымъ же образомъ, въ тотъ же моментъ начала второго года, получимъ ожиданіе платежей страхователя въ такомъ видѣ:

$$(8) \quad \ldots\ldots\ldots\ldots \quad V^{-1} \sum_{i=1}^{i=n} a_i\pi_i V^i .$$

Не трудно убѣдиться, что выраженія (7) и (8), вообще говоря, неравны между собою, потому что, если мы вычтемъ изъ выраженія (7) выраженіе (8), мы получимъ:

$$(9) \quad \ldots\ldots\ldots\ldots \quad V^{-1} \sum_{i=2}^{i=n} \delta_i$$

представляя равенство (6) въ видѣ

$$\delta_1 + \sum_{i=2}^{i=n} \delta_i = 0,$$

мы получимъ:

$$(10) \quad \ldots\ldots \quad -V^{-1}\delta_1 = V^{-1} \sum_{i=2}^{i=n} \delta_i ;$$

послѣдняя формула показываетъ, что дѣйствительно выраженіе (9) отлично отъ 0, если δ_1 отлично отъ 0.

Выраженіе (9) называется *резервомъ* страхового учрежденія по разсматриваемой сдѣлкѣ, относящимся къ началу второго года. Очевидно, что резервъ страхового учрежденія въ началѣ третьяго года сдѣлки будетъ выражаться по формулѣ:

$$(11) \quad \ldots\ldots\ldots \quad V^{-2} \sum_{i=3}^{i=n} \delta_i .$$

Вслѣдствіе того, что формулу (6) можно будетъ переписать такъ:

$$\sum_{i=1}^{i=2} \delta_i + \sum_{i=3}^{i=n} \delta_i = 0,$$

мы получимъ:

$$(12) \quad . \; . \; . \; . \; . \quad V^{-2} \sum_{i=3}^{i=n} \delta_i = - V^{-2} \sum_{i=1}^{i=2} \delta_i .$$

Вообще говоря, къ началу нѣкотораго года k резервъ будетъ имѣть видъ:

$$(13) \quad . \; . \; . \; . \quad v^{-(k-1)} \sum_{i=k}^{i=n} \delta_i = - v^{-(k-1)} \sum_{i=1}^{i=k-1} \delta_i .$$

Формула (13) даетъ возможность двояко вычислять резервы. Если мы будемъ этотъ резервъ вычислять по лѣвой части уравненія (13), то получаемъ методъ, называемый *прямымъ вычисленіемъ резервовъ*. По этому способу вычисленія мы принимаемъ въ соображеніе будущіе платежи обѣихъ сторонъ. Правая часть уравненія (13) даетъ основаніе для такъ называемаго *обратнаго* (ретроспективнаго) *вычисленія резервовъ*. По этой новой методѣ мы принимаемъ въ расчетъ уже сдѣланные къ моменту вычисленія резерва платежи обѣихъ сторонъ. Знакъ минусъ, стоящій во второй части уравненія (13) показываетъ, что при вычисленіи резерва по обратному способу надо будетъ вычитать изъ ожиданій платежей страхователя ожиданія платежей страховщика, а не такъ, какъ это было при прямомъ методѣ, гдѣ мы вычитали изъ ожиданій платежей страхового учрежденія ожиданія платежей страхователя.

§ 8. Въ предыдущихъ параграфахъ мы установили основные принципы главнѣйшихъ задачъ страховой математики. Такъ называемыя netto-преміи вычисляются на основаніи уравненія (6), выражающаго безобидность сдѣлки. Съ другой стороны, мы упомянули о вычисленіи резервовъ.

Задача вычисленія резервовъ имѣетъ громадную важность, ибо страховое учрежденіе обязано въ каждый моментъ времени имѣть наготовѣ сумму денегъ, равную резерву этого момента. При неисполненіи этого требованія, т. е. при недостаточности резервовъ, является для страхового учрежденія серьезная опасность не имѣть возможности выполнить своевременно обязательства сдѣлки.

Основныя вѣроятности страхованія жизни.

§ 9. Итакъ, обращаемся къ страхованію жизни одного лица.

Тутъ на практикѣ обыкновенно встрѣчаются три слѣдующія вѣроятности:

I. Что лицо возраста x лѣтъ (x число цѣлое) доживетъ до возраста $x+n$ (n также цѣлое число).

II. Что лицо (x) умретъ въ промежутокъ $(x, x+n)$.

III. Что лицо (x) умретъ въ промежутокъ $(x+n, x+n+1)$.

Особенно важное значеніе имѣютъ вѣроятности I и III.

Въ страховой математикѣ принято вѣроятность I-ую обозначать знакомъ: ${}_np_x$, причемъ вѣроятность ${}_1p_x$ принято писать просто такъ: p_x, такъ что p_x обозначаетъ вѣроятность лицу возраста x прожить еще одинъ годъ.

На основаніи всего предыдущаго мы можемъ написать:

$$ {}_np_x = \frac{l_{x+n}}{l_x}, \quad p_x = \frac{l_{x+1}}{l_x}. $$

Вѣроятность II-ую принято обозначать знакомъ: ${}_nq_x$, съ тѣмъ же самымъ условіемъ: ${}_1q_x = q_x$. Очевидно, что:

$$ {}_nq_x = 1 - {}_np_x. $$

Что же касается вѣроятности III-й, то ее принято обозначать знакомъ: ${}_{n|}q_x$. Если мы введемъ въ разсмотрѣніе величину $l_x - l_{x+1} = d_x$, то получимъ:

$$ {}_{n|}q_x = \frac{d_{x+n}}{l_x}. $$

Получаемъ также слѣдующія формулы, которыя полезно запомнить. (Гл. III, § 2).

$$ (1) \quad {}_np_x = p_x \cdot p_{x+1} \cdot p_{x+2} \cdot \ldots \cdot p_{x+n-1}. $$

$$ (2) \quad {}_nq_x = q_x + {}_{1|}q_x + {}_{2|}q_x + \ldots + {}_{n-1|}q_x. $$

$$ (3) \quad {}_{n|}q_x = \frac{d_{x+n}}{l_x} = \frac{l_{x+n} - l_{x+n+1}}{l_x} = {}_np_x - {}_{n+1}p_x. $$

$$ (4) \quad {}_{n|}q_x = \frac{l_{x+n} - l_{x+n+1}}{l_x} = \frac{l_{x+n}}{l_x} \cdot \frac{l_{x+n} - l_{x+n+1}}{l_{x+n}} = {}_np_x \cdot q_{x+n}. $$

Коммутаціонныя числа.

§ 10. Въ различныхъ задачахъ, относящихся къ страхованію жизни, входятъ въ вычисленіе нѣкоторыя характернаго вида суммы, которыя называются коммутаціонными числами. Оказывается полезнымъ имѣть таблицы для вычисленія этихъ суммъ; тогда значительно сокращается процессъ вычисленія.

Основное коммутаціонное число D_x [1]) представляетъ изъ себя произведеніе l_x числа живущихъ возраста x на дисконтирующій множитель V^x. Такъ что:

$$(1) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad D_x = l_x V^x .$$

Число D_x называется *дисконтированнымъ числомъ доживающихъ до возраста* x.

Относительно другихъ коммутаціонныхъ чиселъ существуетъ извѣстное различіе въ ихъ опредѣленіи у различныхъ авторовъ. Наиболѣе распространенными являются обозначенія англійскихъ актуаріевъ. Мы ихъ приведемъ:

$$(2) \; . \; . \quad N_x = \sum_{x=1}^{\omega} D_x = D_{x+1} + D_{x+2} + \; . \; . \; . \; . + D_\omega =$$
$$= l_{x+1} V^{x+1} + l_{x+2} V^{x+2} + \; . \; . \; . \; . + l_\omega V^\omega .$$

$$(3) \; . \; . \quad S_n = \sum_x N_x = N_x + N_{x+1} + \; . \; . \; . \; . + N_{\omega-1} = \qquad (N_\omega = 0)$$
$$\begin{aligned} = D_{x+1} + D_{x+2} + D_{x+3} + \; . \; . \; . \; . + D_\omega + \\ D_{x+2} + D_{x+3} + \; . \; . \; . \; . + D_\omega + \\ + D_{x+3} + \; . \; . \; . \; . + D_\omega + \\ . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \\ + D_{\omega-1} + D_\omega + \\ + D_\omega = \end{aligned}$$
$$= D_{x+1} + 2 D_{x+2} + 3 D_{x+3} + \; . \; . \; . + (\omega - x) D_\omega .$$

Буквы D, N, S избраны потому, что въ простѣйшихъ формулахъ страхованія (пожизненной ренты) числа D играютъ роль

---

[1]) Величина D_x была введена въ первый разъ: Tetens „Einleitung zur Berechnung der Leibrenten". Leipzig. 1785—1786 г.

знаменателя (dénominateur), числа N — числителя (numérateur), а числа S представляютъ сумму чиселъ N. Кромѣ того, вводятся въ разсмотрѣніе еще слѣдующія коммутаціонныя числа:

$$(4) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad C_x = d_x V^{x+1} ;$$

это число называется дисконтированнымъ числомъ умирающихъ возраста x. Существуетъ слѣдующая простая зависимость между C_x и D_x:

$$C_x = d_x V^{x+1} = (l_x - l_{x+1}) V^{x+1} = V l_x V^x - l_{x+1} V^{x+1} = V D_x - D_{x+1} .$$

Слѣдующее, пятое коммутаціонное число получимъ, если возьмемъ сумму дисконтированныхъ чиселъ умирающихъ въ возрастѣ x и выше, т. е.

$$(5) \quad . \; . \; . \; . \quad M_x = C_x + C_{x+1} + \; . \; . \; . \; . + C_\omega = \sum_x^\omega C_x ;$$

наконецъ, вводится въ разсмотрѣніе еще шестое коммутаціонное число:

$$(6) \quad . \; . \; . \; . \; . \; . \quad R_x = M_x + M_{x+1} + \; . \; . \; . \; . \; M_\omega .$$

Обозначенія C, M, R взяты, какъ ближайшія буквы къ вышеуказаннымъ буквамъ D, N, S.

Нѣкоторые писатели, особенно французскіе и американскіе, находили неудобнымъ, что суммы M и R начинаются со знака x а сумма N съ $x+1$, и поэтому они вводили для однообразія вмѣсто формулы (2) формулу:

$$(7) \quad . \; . \; . \; . \; . \; . \; . \quad N_x = D_x + D_{x+1} + \; . \; . \; . \; .$$

§ 11. При страхованіи на случай смерти является извѣстное принципіальное затрудненіе при вычисленіи, состоящее въ томъ, что неизвѣстенъ моментъ смерти застрахованнаго. Изъ этого затрудненія можно выйти, конечно, если при заключеніи договора установить правило, что застрахованный капиталъ выдается въ концѣ года смерти застрахованнаго. Если же предположить, что застрахованный капиталъ уплачивается немедленно по представленіи полиса послѣ смерти застрахованнаго, то, особенно при большихъ застрахованныхъ капиталахъ, можетъ быть сравнительно большая разница въ дисконтѣ, судя по тому, въ какой части года умеръ застрахованный, въ началѣ года, или въ концѣ. Существуетъ правило, которое часто употребляется, состоящее въ томъ, чтобы считать моментъ смерти приходящимся какъ разъ въ серединѣ года. Въ этомъ случаѣ приходится всѣ дисконтируемые платежи, относящіеся къ цѣлымъ годамъ, умножать на дисконтирующій множитель за полгода, т. е. на выраженіе:

$$V^{-\frac{1}{2}} (1+i)^{\frac{1}{2}} = \sqrt{1+i}\,.$$

Поэтому часто вводятся вмѣсто обозначеній (4), (5), (6) величины, которыя отличаются отъ нихъ на указанный множитель $V^{\frac{1}{2}}$.

На второмъ международномъ конгрессѣ актуаріевъ, имѣвшемъ мѣсто въ Брюсселѣ въ 1902 году, были предложены слѣдующія новыя обозначенія:

$$\overline{C}_x = V^{-\frac{1}{2}} C_x, \quad \overline{M}_x = V^{-\frac{1}{2}} M_x, \quad \overline{R}_x = V^{-\frac{1}{2}} R_x;$$

кромѣ того, на томъ же конгрессѣ принято: сохранить въ употребленіи коммутаціонное число (2) N; число же (7) обозначать жирнымъ шрифтомъ, а именно:

$$N_x = D_{x+1} + D_{x+2} + \ldots\ldots$$

$$\mathbf{N}_x = D_x + D_{x+1} + \ldots\ldots,$$

такъ что:

$$N_x = \mathbf{N}_{x+1}\,.$$

Въ концѣ книги мы приводимъ таблицу смертности двадцати британскихъ обществъ вмѣстѣ съ таблицей коммутаціонныхъ чиселъ, вычисленныхъ при предположеніи $3^1/_2\,^0/_0$ роста.

X.

Вычисленіе стоимостей важнѣйшихъ видовъ платежей, встрѣчающихся въ страхованіи жизни.

§ 1. Будемъ разсматривать современную стоимость платежа, не обращая вниманія сначала на то, кто производитъ этотъ платежъ,—страховщикъ или страхователь.

При страхованіи на дожитіе (Erlebensversicherung) будемъ обозначать знакомъ $_nE_x$ современную стоимость единицы капитала, уплачиваемаго черезъ n лѣтъ въ томъ случаѣ, если лицо, имѣющее въ данный моментъ возрастъ x, доживетъ до возраста $x+n$ лѣтъ. Очевидно, что мы будемъ имѣть формулу:

$$(1) \quad {}_nE_x = V^n \cdot {}_np_x \,.$$

Выраженіе (1) можно будетъ слѣдующимъ образомъ преобразовать и свести на вычисленіе коммутаціонныхъ чиселъ:

$$(2) \quad {}_nE_x = V^n \frac{l_{x+n}}{l_x} = \frac{V_{x+n} l_{x+n}}{V^x l_x} = \frac{D_{x+n}}{D_x} \,.$$

Необходимо обратить вниманіе на слѣдующую формулу:

$$(3) \quad {}_{n+m}E_x = \frac{D_{x+n+m}}{D_x} = \frac{D_{x+n}}{D_x} \cdot \frac{D_{x+n+m}}{D_{x+n}} \,,$$

откуда получаемъ:

$${}_{n+m}E_x = {}_nE_x \,.\, {}_mE_{x+n} \,.$$

Послѣднее равенство (3) находится въ полной аналогіи съ соотвѣтствующимъ равенствомъ для вѣроятности дожитія:

$${}_{n+m}p_x = {}_np_x \cdot {}_mp_{x+n} \,.$$

Необходимость вычисленія выраженія ${}_nE_x$, относящагося къ страхованію на доживіе, можетъ встрѣтиться при самыхъ разнообразныхъ формахъ страхованія. Такъ, напримѣръ, это выраженіе является при вычисленіи современной стоимости уплаты, равной 1 рублю, страховымъ обществомъ, въ случаѣ страхованія на доживіе, т. е. когда застрахованное лицо возраста x достигнетъ возраста $x+n$. Но и при другихъ формахъ страхованія можетъ встрѣтиться надобность въ выраженіи ${}_nE_x$; такъ, напримѣръ, при страхованіи на случай смерти лица возраста x намъ придется разсматривать современную стоимость уплаты страхователемъ страховой премія, соотвѣтствующей возрасту $x+n$.

§ 2. *Пожизненная годовая рента или пенсія* (Leibrenten).

I. *Postnumerando*. Предположимъ, что нѣкоторому лицу возраста x уплачивается ежегодно postnumerando пожизненная пенсія въ размѣрѣ одного рубля. Спрашивается, чему равна современная стоимость этой пенсіи. Эта современная стоимость въ страховой математикѣ обозначается знакомъ a_x. Очевидно, что вычисленіе a_x приводитъ къ вычисленіямъ предыдущей задачи § 1-го. Получаемъ:

$$a_x = {}_1E_x + {}_2E_x + {}_3E_x + \ldots + {}_{(\omega-x)}E_x \,.$$

отсюда:

$$(1) \ldots\ldots a_x = \frac{D_{x+1} + D_{x+2} + \ldots + D_\omega}{D_x} = \frac{N_x}{D_x}.$$

Относительно пожизненной пенсіи a_x можно указать формулу послѣдовательнаго ея вычисленія для послѣдовательныхъ цѣлыхъ чиселъ лѣтъ, а именно:

$$a_x = \frac{N_x}{D_x} = \frac{D_{x+1} + N_{x+1}}{D_x} = \frac{D_{x+1}}{D_x}\left(1 + \frac{N_{x+1}}{D_{x+1}}\right),$$

но

$$\frac{D_{x+1}}{D_x} = \frac{V^{x+1} l_{x+1}}{V^x l_x} = V p_x; \quad \frac{N_{x+1}}{D_{x+1}} = a_{x+1},$$

откуда:

$$(2) \ldots\ldots\ldots a_x = V p_x (1 + a_{x+1});$$

очевидно, $a_\omega = 0$.

II. *Praenumerando.* Стоимость ежегодной пенсіи, уплачиваемой praenumerando въ размѣрѣ одного рубля, обозначается въ страховой математикѣ $\mathbf{a}_x$. Получаемъ, очевидно:

$$\mathbf{a}_x = 1 + {}_1E_x + {}_2E_x + \ldots = \frac{D_x + D_{x+1} + \ldots}{D_x}$$

или:

$$(1) \ldots\ldots\ldots\ldots \mathbf{a}_x = \frac{\mathbf{N}_x}{D_x};$$

зависимость между пенсіями praenumerando и postnumerando получается такая:

$$(2) \ldots\ldots\ldots\ldots \mathbf{a}_x = 1 + a_x.$$

§ 3. *Немедленная временная годичная рента.*

I. *Postnumerando.* Если лицо возраста x лѣтъ начинаетъ получать годичную ренту въ размѣрѣ одного рубля postnumerando, причемъ получаетъ ее только n разъ до возраста $x + n$, то современная стоимость такой ренты обозначается знакомъ

$$|_n a_x.$$

Очевидно, что мы приходимъ къ формулѣ:

$$|_n a_x = \frac{D_{x+1} + D_{x+2} + \ldots + D_{x+n}}{D_x}$$

или окончательно:

$$(1) \ldots\ldots\ldots |_n a_x = \frac{N_x - N_{x+n}}{D_x}.$$

II. *Praenumerando.* Получаемъ формулу:

$$(1) \ldots\ldots\ldots \quad |_{n}\mathbf{a}_x = \frac{\mathbf{N}_x - \mathbf{N}_{x+n}}{D_x}.$$

§ 4. *Отсроченная пожизненная рента.*

I. *Postnumerando.* Если лицо возраста x будетъ получать годовую пожизненную ренту postnumerando въ размѣрѣ одного рубля, начиная съ возраста $x+m$ лѣтъ, то такая рента называется отсроченной на m лѣтъ пожизненной рентой postnumerando стоимость такой ренты мы будемъ обозначать знакомъ:

$$_{m}|\,a_x.$$

Получаемъ, очевидно, формулу:

$$_{m}|\,a_x = \frac{D_{x+m+1} + D_{x+m+2} + \ldots + D_{\omega}}{D_x}$$

или окончательно:

$$(1) \ldots\ldots\ldots \quad _{m}|\,a_x = \frac{N_{x+m}}{D_x}.$$

II. *Praenumerando.* Для отсрочки на m лѣтъ ренты praenumerando получаемъ формулу:

$$_{m}|\,\mathbf{a}_x = \frac{\mathbf{N}_{x+m}}{D_x}.$$

Сравнивая немедленную пожизненную ренту съ рентами временной и отложенной, получаемъ очевидныя формулы:

$$a_x = |_{n}a_x + {}_{n}|\,a_x,$$

$$\mathbf{a}_x = |_{n}\mathbf{a}_x + {}_{n}|\,\mathbf{a}_x.$$

§ 5. *Отсроченная временная рента.*

I. *Postnumerando.* Положимъ, что лицо возраста x будетъ получать, начиная съ возраста $x+m$ лѣтъ, ренту postnumerando въ размѣрѣ одного рубля до возраста $x+n$ (очевидно $n>m$). Стоимость такой отложенной временной ренты мы будемъ обозначать знакомъ:

$$_{m}|_{n-m}a_x;$$

въ этомъ знакѣ лѣвая отъ вертикальной черты буква m напоминаетъ, что рента отсрочена на m лѣтъ, правая же отъ вертикальной черты выраженіе $n-m$, показываетъ, что рента будетъ уплочена ровно $n-m$ разъ. Получаемъ, очевидно, формулу:

$$_{m\,|\,n-m}a_x = \frac{D_{x+m+1} + D_{x\ m+2} + \ \ldots\ D_{x+n}}{D_x},$$

что можно написать окончательно въ такомъ видѣ:

$$(1)\ \ldots\ldots\ _{m\,|\,n-m}a_x = \frac{N_{x+m} - N_{x+n}}{D_x}.$$

II. *Praenumerando.* Для отсроченной временной ренты praenumerando получается формула:

$$_{m\,|\,n-m}\mathbf{a}_x = \frac{\mathbf{N}_{x+m} - \mathbf{N}_{x+n}}{D_x}.$$

§ 6. *Непрерывная рента.*

I. *Postnumerando.* Дѣлимъ годъ на m равныхъ частей и предположимъ, что въ концѣ каждой этой части (postnumerando) уплачивается рента въ размѣрѣ доли $\frac{1}{m}$ рубля; предположимъ, что уплата этой ренты совершается немедленно и пожизненно. Обозначимъ современную ея стоимость знакомъ: $a_x^{(m)}$.

Очевидно, что разсматриваемая нами задача принадлежитъ къ задачамъ второго класса, гдѣ требуются добавочныя гипотезы. Мы поставимъ гипотезу о существованіи непрерывности функціи l_x. Очевидно, мы получимъ слѣдующую формулу:

$$(1)\ \ldots\ a_x^{(m)} = V^{\frac{1}{m}} \cdot \frac{l_{x+\frac{1}{m}}}{l_x} \cdot \frac{1}{m} + V^{\frac{2}{m}} \cdot \frac{l_{x+\frac{2}{m}}}{l_x} \cdot \frac{1}{m} +$$

$$+ V^{\frac{3}{m}} \cdot \frac{l_{x+\frac{3}{m}}}{l_x} \cdot \frac{1}{m} + \ldots$$

Если мы будемъ увеличивать число m до безконечности, то мы замѣтимъ, что перемѣнная величина $a_x^{(m)}$, въ которой x остается безъ перемѣны, а m возрастаетъ до безконечности, стремится къ нѣкоторому предѣлу, который легко вычислить, и который мы будемъ обозначать знакомъ $\overline{a}_x$. Предѣлъ

$$\overline{a}_x = \lim_{m=\infty} a_x^{(m)}$$

называется пожизненной непрерывной рентой лица возраста x, соотвѣтствующей единицѣ капитала. Обращаемся теперь къ вычисленію сказаннаго предѣла. Формулу (1) можно будетъ переписать такъ:

$$(2) \ldots a_x^{(m)} = \frac{V^{x+\frac{1}{m}} l_{x+\frac{1}{m}} \cdot \frac{1}{m} + V^{x+\frac{2}{m}} l_{x+\frac{2}{m}} \cdot \frac{1}{m} + \ldots}{V^x l_x}$$

$$\frac{\ldots + V^{x+\frac{k}{m}} l_{x+\frac{k}{m}} \cdot \frac{1}{m} + \ldots}{V^x l_x}.$$

Если мы будемъ считать число m безконечно большимъ, тогда $\frac{1}{m}$ будетъ безконечно малая величина. И если мы обозначимъ $x + \frac{k}{m} = y$, гдѣ k пробѣгаетъ цѣлыя значенія, то при увеличеніи k на единицу y получаетъ приращеніе на $\frac{1}{m}$. Поэтому можно положить:

$$\frac{1}{m} = dy.$$

Значитъ сумма, стоящая въ числителѣ, имѣетъ своимъ предѣломъ интегралъ:

$$\int V^y l_y dy.$$

Что касается предѣловъ этого интеграла, то, очевидно, что придется интегрировать отъ начальнаго возраста до предѣльнаго возраста ω. Можно интегрировать до безконечности, потому что все равно послѣ предѣльнаго возраста подинтегральная функція l_x равна 0.

Итакъ, мы получаемъ формулу:

$$(3) \ldots\ldots\ldots \bar{a}_x = \frac{\int_x^\infty V^y l_y dy}{V^x l_x}.$$

Если бы мы дѣлили промежутокъ года также на m частей, но ренту увеличивали бы въ началѣ каждаго промежутка—praenumerando, то мы, очевидно, для стоимости $\mathbf{a}_x^{(m)}$ такой новой ренты получили бы число, отличающееся отъ числа $a_x^{(m)}$ только на первый взносъ $\frac{1}{m}$, совершающійся въ настоящій моментъ. И мы имѣемъ:

$$\mathrm{a}_x^{(m)} = \frac{1}{m} + a_x^{(m)},$$

отсюда:

$$\lim_{m=\infty} \mathrm{a}_x^{(m)} = \lim_{m=\infty} \frac{1}{m} + \lim_{m=\infty} a_x^{(m)} = 0 + \bar{a}_x = \bar{a}_x.$$

Если ростъ i равенъ нулю, то $V = 1$, и непрерывная рента обращается въ выраженіе:

$$\frac{1}{l_x}\int_x^\infty l_y dy,$$

которое мы назвали средней продолжительностью жизни (§ 1, гл. V).

Итакъ мы видимъ, что непрерывная рента, выводимая изъ ренты praenumerando, оказывается тѣмъ же самымъ числомъ, что, конечно, можно было ожидать a priori, потому что при уменьшеніи промежутковъ времени концы этихъ промежутковъ сближаются, и, значитъ, должны сближаться стоимости двухъ рентъ postnumerando и praenumerando.

§ 7. *Непрерывная временная рента*, уплачиваемая лицу возраста x до возраста $x+n$, очевидно, выразится по формулѣ:

$$(1) \quad \ldots\ldots\ldots \quad {}_{|n}\bar{a}_x = \frac{\int_x^{x+n} V^y l_y dy}{V^x l_x}.$$

Непрерывная, отсроченная на m лѣтъ рента лица возраста x выразится по формулѣ:

$$(2) \quad \ldots\ldots\ldots \quad {}_{m|}\bar{a}_x = \frac{\int_{x+m}^{\infty} V^y l_y dy}{V^x l_x}.$$

Непрерывная отсроченная, временная рента выразится по формулѣ:

$$(3) \quad \ldots\ldots \quad {}_{m|n-m}\bar{a}_x = \frac{\int_{x+m}^{x+n} V^y l_y dy}{V^x l_x}.$$

При разсмотрѣнiи непрерывной ренты въ формулахъ (1), (2), (3), можно предполагать числа m и n какими угодно положительными, какъ цѣлыми, такъ и дробными.

§ 8. *Страхованiе на случай смерти.* Положимъ, что страховое учрежденiе обязалось уплатить наслѣдникамъ страхователя капиталъ, равный одному рублю въ концѣ того года, въ которомъ произойдетъ смерть страхователя. Требуется опредѣлить современную стоимость обязательства страхового общества. Такъ какъ неизвѣстно, на которомъ году жизни умретъ страхователь, то надо перечислить всѣ возможныя комбинацiи уплаты застрахованнаго капитала. Если страхователь въ моментъ заключенiя договора имѣетъ x лѣтъ, тогда очевидно, что уплата застрахованнаго капитала можетъ произойти въ тѣ моменты времени, которые будутъ соотвѣтствовать возрастамъ:

$$x+1,\ x+2,\ x+3,\ \ldots\ldots;$$

этотъ рядъ продолжается вплоть до предѣльнаго возраста ω. Итакъ, страхованiе на случай смерти можно разсматривать, какъ совокупность слѣдующихъ частичныхъ страхованiй: 1) обязательство общества уплатить застрахованный капиталъ въ моментъ $x+1$, если смерть послѣдуетъ въ промежутокъ времени $(x, x+1)$; 2) обязательство общества уплатить застрахованный капиталъ въ моментъ $x+2$, если смерть послѣдуетъ въ промежутокъ времени $(x+1, x+2)$; 3) подобное же обязательство для момента $x+3$ и промежутка $(x+2, x+3)$ и т. д. . . . .

Очевидно, что математическое ожиданiе полученiя застрахованнаго капитала должно равняться суммѣ математическихъ ожиданiй, относящихся ко всѣмъ вышеприведеннымъ частичнымъ страхованiямъ.

Итакъ, опредѣлимъ прежде всего современную стоимость частичнаго страхованiя, соотвѣтствующаго промежутку времени $(x+n-1, x+n)$. Вѣроятность смерти въ этомъ промежуткѣ, какъ извѣстно, опредѣляется по формулѣ:

$$_{n-1|}q_x = \frac{d_{x+n-1}}{l_x};$$

по условiю страхованiя, капиталъ въ одинъ рубль долженъ быть уплоченъ страховымъ обществомъ въ концѣ промежутка $x+n$; тогда, очевидно, придется дисконтировать къ начальному моменту при помощи множителя V^n, и мы получимъ современную стоимость указаннаго частичнаго страхованія въ такомъ видѣ:

$$V^n {}_{n-1}|q_x = \frac{V^{x+n} d_{x+n-1}}{V^x l_x} = \frac{C_{x+n-1}}{D_x}.$$

Итакъ, общая современная стоимость уплаты обществомъ застрахованнаго капитала, которую мы обозначимъ A_x (assurance), выразится суммой:

$$\sum \frac{C_{x+n-1}}{D_x},$$

распространенной на всѣ значенія n, отъ единицы до предѣльнаго возраста. Итакъ, мы получаемъ слѣдующую формулу стоимости страхованія на случай смерти:

$$A_x = \frac{C_x + C_{x+1} + \ldots\ldots C_\omega}{D_x} = \frac{M_x}{D_x}.$$

Необходимо обратить вниманіе на важное соотношеніе между стоимостью страхованія на случай смерти и рентами. Въ самомъ дѣлѣ, мы имѣемъ:

$$\begin{aligned}(1)\ldots M_x &= C_x + C_{x+1} + C_{x+2} + C_{x+3} + \ldots\ldots = \\ &= (VD_x - D_{x+1}) + (VD_{x+1} - D_{x+2}) + (VD_{x+2} - D_{x+3}) + \ldots = \\ &= V(D_x + D_{x+1} + D_{x+2} + \ldots) - (D_{x+1} + D_{x+2} + D_{x+3} \ldots) = \\ &= VN_{x-1} - N_x.\end{aligned}$$

Итакъ, на основаніи формулы (1) мы получаемъ:

$$(2)\ldots A_x = \frac{VN_{x-1} - N_x}{D_x} = V\frac{\mathbf{N}_x}{D_x} - \frac{N_x}{D_x} = V\mathbf{a}_x - a_x.$$

Отсюда мы можемъ выразить стоимость страхованія на случай смерти черезъ ту или другую изъ рентъ a_x или $\mathbf{a}_x$. Напримѣръ, выражая черезъ ренту praenumerando получимъ:

$$(3)\ldots A_x = V\mathbf{a}_x - (\mathbf{a}_x - 1) = 1 - \mathbf{a}_x(1 - V) = 1 - d\mathbf{a}_x;$$

$$\text{если } i = 0, \text{ то } V = 1,\ d = 0,$$

и мы получаемъ $A_x = 1$, что можно было бы ожидать, ибо $\sum_{n-1}|q_x = 1$. Послѣднее уравненіе выражаетъ не что иное, какъ достовѣрность смерти.

§ 9. Относительно страхованія на случай смерти существуетъ тоже самое подраздѣленіе страхованія на 1) немедленное временное, 2) отсроченное и 3) отсроченное временное.

Временнымъ немедленнымъ страхованіемъ называется такое, въ которомъ уплата капитала послѣ смерти лица (x) должна послѣдовать только въ случаѣ, если смерть послѣдуетъ въ теченіе опредѣленнаго срока, напримѣръ до достиженія возраста $(x+n)$ лѣтъ. Современная стоимость временнаго на n лѣтъ страхованія единицы капитала обозначается:

$$|_n A_x .$$

Очевидно, мы получимъ формулу:

$$(1) \ldots |_n A_x = \frac{C_x + C_{x+1} + \ldots C_{x+n-1}}{D_x} = \frac{M_x - M_{x+n}}{D_x} .$$

Отсроченнымъ на m лѣтъ называется такое страхованіе, въ которомъ уплата капитала производится только въ томъ случаѣ, если лицо возраста (x) умретъ послѣ достиженія возраста $(x+m)$ лѣтъ. Современная стоимость единицы капитала при такомъ страхованіи обозначается черезъ:

$$_m| A_x .$$

Получается, очевидно, формула:

$$(2) \ldots \, _m| A_x = \frac{C_{x+m} + C_{x+m+1} + \ldots C_\omega}{D_x} = \frac{M_{x+m}}{D_x} .$$

Наконецъ, мы получаемъ формулу для современной стоимости временнаго на $n-m$ лѣтъ и отсроченнаго на m лѣтъ страхованія смерти:

$$(3) \ldots \, _m|_{n-m} A_x = \frac{C_{x+m} + C_{x+m+1} + \ldots + C_{x+n-1}}{D_x} =$$

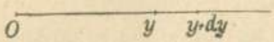
$$= \frac{M_{x+m} - M_{x+n}}{D_x} .$$

§ 10. *Непрерывное страхованіе на случай смерти.* Подобно тому, какъ мы разсмотрѣли переходъ отъ ренты, уплачиваемой черезъ извѣстный срокъ, къ рентѣ непрерывной, точно такъ же можетъ быть введено въ разсмотрѣніе *непрерывное страхованіе* на случай смерти. При этомъ мы будемъ обобщать вышеприведенную формулу:

$$(1)\ .\ .\ A_x = \sum_{n=1}^{n=\omega} V^n \frac{d_{x+n-1}}{l_x} = \frac{1}{V^x l_x} \cdot \sum_{n=1}^{n=\omega} V^{x+n} d_{x+n-1} =$$

$$= \frac{1}{D_x} \sum_{n=1}^{n=\omega} V^{x+n} (l_{x+n-1} - l_{x+n})\,;$$

это обобщеніе мы произведемъ такъ: предположимъ, что годъ раздѣленъ на безконечное число разныхъ промежутковъ.



Черт. 7.

Если мы обозначимъ: $dl_y = l_{y+dy} - l_y$ (черт. 7), тогда мы можемъ какъ предѣлъ формулы (1) разсматривать слѣдующую формулу непрерывнаго страхованія:

$$(2)\ .\ .\ .\ .\ .\ .\ .\ .\ \overline{A}_x = \frac{-1}{D_x} \int_x^{\infty} V^y dl_y\,;$$

интегрируя по частямъ, можно преобразовать эту формулу такимъ образомъ:

$$\int_x^{\infty} V^y dl_y = \Big[V^y l_y \Big]_x^{\infty} - \int_x^{\infty} l_y d(V^y) = - V^x l_x - lg\, V \int_x^{\infty} V^y l_y dy\,;$$

итакъ получаемъ:

$$(3)\ .\ .\ .\ .\ .\ .\ .\ .\ \overline{A}_x = 1 - \frac{lg V}{D_x} \int_x^{\infty} V^y l_y dy\,.$$

§ 11. Не трудно вывести понятіе объ отсроченномъ, временномъ и отсроченномъ временномъ непрерывномъ страхованіи. Мы не будемъ эти формулы подробно выписывать, потому что все дѣло приводится къ установленію предѣловъ интегрированія. Подобное установленіе предѣловъ нами было подробно указано при разсмотрѣніи непрерывныхъ рентъ (см. § 6).

§ 12. Разсмотрѣнное непрерывное страхованіе можетъ имѣть особенное значеніе въ вопросахъ государственнаго страхованія, гдѣ вопросы о точномъ установленіи netto-преміи имѣютъ большое значеніе, и гдѣ государство имѣетъ возможность вести широкую статистику смертности; при этомъ таблицы смертности могутъ быть представлены въ достаточно совершенномъ видѣ и давать вѣрную картину риска страховщика. Непрерывное страхованіе даетъ возможность вычислять netto-преміи при страхованіи на случай смерти, при условіи немедленной послѣ смерти уплаты капитала; при этомъ не требуется никакихъ добавочныхъ условій, вродѣ уплаты капитала въ концѣ года смерти, а также не требуется никакихъ произвольныхъ способовъ вычисленія, вродѣ умноженія на дисконтирующій множитель $V^{-\frac{1}{2}}$ съ цѣлью приведенія момента смерти къ серединѣ года.

XI.

Вычисленіе netto-премій.

§ 1. Вышеприведенныя формулы даютъ возможность вычислять netto-преміи для разнообразныхъ видовъ страхованія одной жизни. Мы оставимъ въ сторонѣ задачи такого страхованія, при которомъ netto-премія уплачивается страхователемъ только одинъ разъ при заключеніи договора. Очевидно, что эта netto-премія должна равняться современной стоимости всѣхъ обязательствъ страховщика, а потому полученіе такой netto-преміи производится непосредственно. Напримѣръ, если капиталъ K застрахованъ на дожитіе, то netto-премія будетъ равняться K_nE_x. Въ страхованіи на случай смерти мы получимъ KA_x. И наконецъ, страхованіе ренты въ размѣрѣ K рублей будетъ сопровождаться единовременной преміей въ размѣрѣ Ka_x или $K\mathfrak{a}_x$. Понятно, что при этомъ, какъ страхованіе на случай смерти, такъ и страхованіе рентъ могутъ быть немедленныя, отсроченныя и временныя.

§ 2. Въ обычной практикѣ однако уплата netto-преміи производится равными частями, черезъ равные промежутки времени и обыкновенно praenumerando, такъ что первый взносъ преміи страхователь производитъ въ моментъ заключенія договора. Премія, такимъ образомъ, представляетъ изъ себя ренту praenumerando. Премія можетъ уплачиваться или пожизненно, или же извѣстное число разъ. Въ послѣднемъ случаѣ страхованіе носитъ названіе *страхованія съ временнымъ взносомъ премій*. Во всемъ дальнѣйшимъ мы будемъ обозначать знакомъ P пожизненную годовую премію; знакомъ же ${}_nP$ будемъ обозначать годовую премію, подлежащую n разъ уплатѣ. Очевидно, что знакъ ${}_1P$ будетъ выражать разобранную въ предыдущемъ параграфѣ единовременную премію. Знакомъ $P^{(m)}$, ${}_nP^{(m)}$ будемъ обозначать премію, выплачиваемую m разъ въ годъ.

Обыкновенно въ обозначеніе преміи вводится знакъ, указывающій на планъ страхованія, причемъ этотъ знакъ ставится въ скобкахъ послѣ знака преміи; напримѣръ, знакъ $P({}_nE_x)$ ставится при страхованіи на дожитіе; $P(\,|\,{}_nA_x)$ представляетъ собою премію при временномъ страхованіи на случай смерти; знакъ ${}_nP({}_m\,|\,A_x)$ обозначаетъ временную n разъ вносимую годовую премію при отсроченномъ на m лѣтъ страхованіи на случай смерти.

Покажемъ вычисленіе netto-преміи на рядѣ задачъ.

Задача 1.

§ 3. Разсматривается страхованіе единицы капитала на случай смерти [1]), требуется вычислить пожизненно уплачиваемую (praenumerando) годовую премію лица возраста x.

Согласно нашимъ условіямъ искомая премія должна быть обозначена знакомъ:

$$P(A_x).$$

Такъ какъ единица капитала, уплачиваемая ежегодно praenumerando въ видѣ пожизненной преміи, даетъ современную стоимость $\mathbf{a}_x$, то очевидно, что современная стоимость уплаты искомой ренты будетъ $P(A_x)\,\mathbf{a}_x$; съ другой стороны, современная стоимость обязательства общества будетъ A_x. Отсюда получаемъ уравненіе:

$$P(A_x)\mathbf{a}_x = A_x\,.$$

Искомая премія выразится по формулѣ:

$$(1)\;.\;.\;.\;\;P(A_x) = \frac{A_x}{\mathbf{a}_x} = \frac{M_x}{D_x} : \frac{\mathbf{N}_x}{D_x} = \frac{M_x}{\mathbf{N}_x} = \frac{M_x}{N_{x-1}};$$

премія же для капитала K будетъ:

$$(2)\;.\;.\;.\;.\;.\;.\;.\;.\;\;P(A_x) = K\,\frac{M_x}{N_{x-1}}\,.$$

Численный примѣръ. Лицо возраста 41 годъ застраховало жизнь на случай смерти въ размѣрѣ 10000 рублей. Требуется вычислить размѣръ годовой netto-преміи.

---

[1]) Съ уплатой въ концѣ года смерти.

Получаемъ искомую премію: $P(A_{41})=10000\cdot\frac{M_{41}}{N_{40}}=$

$$=10000\cdot\frac{8560{,}74\,^{1)}}{355429}=240\text{ р. }85\text{ к.}$$

Задача II.

§ 4. Страхованіе капитала на случай смерти съ временнымъ взносомъ премій. Искомая netto-премія можетъ быть обозначена такъ:

$${}_nP(A_x),$$

гдѣ n указываетъ число взносовъ преміи, послѣ которыхъ прекращается уплата преміи. Очевидно, что, если застрахована единица капитала, то мы получаемъ:

$${}_nP(A_x)\cdot|_n\mathbf{a}_x=A_x,$$

откуда:

$$(1)\;\ldots\;{}_nP(A_x)=\frac{A_x}{|_n\mathbf{a}_x}=\frac{M_x}{D_x}:\frac{\mathbf{N}_x-\mathbf{N}_{x+n}}{D_x}=\frac{M_x}{\mathbf{N}_x-\mathbf{N}_{x+n}}=$$

$$=\frac{M_x}{N_{x-1}-N_{x+n-1}}.$$

Сравнивая эту формулу съ формулой (1) предыдущаго параграфа, мы получаемъ ${}_nP>P$, т. е., другими словами, при одной и той же формѣ страхованія размѣръ временной преміи будетъ выше размѣра пожизненной.

Задача III.

§ 5. *Страхованіе на случай смерти съ немедленной послѣ смерти уплатой застрахованнаго капитала.* Очевидно, что все отличіе задачъ I и II будетъ состоять въ томъ, что вмѣсто A_x придется взять $A_x\sqrt{1+i}$. Если мы находимъ предположеніе момента смерти, относящагося къ серединѣ года, не достаточно точнымъ, то можно будетъ разсматривать непрерывное страхованіе на случай смерти

$$\overline{A}_x$$

и придется вычислять интегралъ, что не представляетъ особаго затрудненія, въ виду не особенно большой точности, которая обыкновенно требуется при такомъ вычисленіи.

---

[1]) См. таблицу въ концѣ книги.

Задача IV.

§ 6. *Временное страхованіе капитала на случай смерти съ пожизненной уплатой премій.* Подъ временнымъ страхованіемъ капитала на случай смерти разумѣется такая форма страхованія, когда страховщикъ обязуется уплатить капиталъ послѣ смерти застрахованнаго, если только эта смерть послѣдуетъ въ теченіе опредѣленнаго числа n лѣтъ отъ начала страхованія.

Допустимъ, что заключено временное на n лѣтъ страхованіе капитала въ 1 руб. на случай смерти лица возраста x. Мы будемъ предполагать, какъ въ этой задачѣ, такъ и въ дальнѣйшихъ, что капиталъ уплачивается въ концѣ года смерти застрахованнаго.

Итакъ, современная стоимость обязательствъ страховщика будетъ:

$$|_n A_x .$$

Обыкновенно при такомъ страхованіи расчеты страхователя и страховщика оканчиваются или смертью застрахованнаго, или истеченіемъ срока, на который страхованіе заключено (n лѣтъ). Поэтому, по разсматриваемому плану договоры съ пожизненной уплатой преміи не заключаются. Обыкновенно премія вносится при жизни застрахованнаго, но не далѣе срока самого страхованія. Если остановиться на предположеніи, что премія вносится не далѣе срока страхованія, то мы приходимъ къ разсмотрѣнію ${}_nP(|_n A_x)$ и, слѣдовательно, получаемъ:

$${}_nP(|_n A_x) \cdot |_n \mathrm{a}_x = |_n A_x ,$$

т. е. $${}_nP(|_n A_x) = |_n A_x : |_n \mathrm{a}_x = \frac{M_x - M_{x+n}}{D_x} : \frac{N_{x-1} - N_{x+n-1}}{D_x} =$$

$$= \frac{M_x - M_{x+n}}{N_{x-1} - N_{x+n-1}} .$$

Численный примѣръ. Застрахована жизнь лица возраста 35 лѣтъ на случай смерти въ размѣрѣ 45000 рублей, причемъ страхованіе временное на 25 лѣтъ. Требуется найти netto-премію $P(|_{25} A_{35})$. Получаемъ для искомой преміи выраженіе:

$$45000 \cdot \frac{M_{35} - M_{60}}{N_{34} - N_{59}} = 45000 \cdot \frac{9805{,}72 - 4735{,}38}{501101 - 88791{,}3} = 553{,}39.$$

Задача V.

§ 7. *Временное въ продолженіи n лѣтъ страхованіе на случай смерти съ временнымъ k разъ взносомъ премій.* При этомъ, конечно, $k < n$. Получаемъ:

$${}_kP(\,|_nA_x)\,.\,|_k\mathrm{a}_x = |_nA_x\,,$$

откуда:

$${}_kP(\,|_nA_x) = \frac{M_x - M_{x+n}}{N_{x-1} - N_{x+k-1}}\,.$$

Численный примѣръ. Лицо возраста 20 лѣтъ застраховало свою жизнь въ 30000 рублей на случай смерти временнымъ въ продолженіи 20 лѣтъ страхованіемъ, подъ условіемъ уплаты преміи въ продолженіи 15 лѣтъ. Требуется найти netto-премію ${}_{15}P(\,|_{20}A_{20})$. Получаемъ:

$$30000 \cdot \frac{M_{20} - M_{40}}{N_{19} - N_{34}} = 30000 \cdot \frac{13594{,}03 - 8761{,}58}{1075856 - 501101} = 252{,}23.$$

Задача VI.

§ 8. *Отсроченное страхованіе на случай смерти съ пожизненной уплатой преміи.* Netto-премія такого страхованія можетъ быть обозначена знакомъ:

$$P(_n|\,A_x)$$

и, очевидно, вычисляется по формулѣ:

$$P(_n|\,A_x) = \frac{_n|\,A_x}{\mathrm{a}_x} = \frac{M_{x+n}}{N_{x-1}}\,.$$

Численный примѣръ. Лицо возраста 32 года застраховало на случай смерти капиталъ въ 1000 рублей, съ условіемъ, что этотъ капиталъ будетъ уплоченъ лишь въ томъ случаѣ, когда смерть наступитъ не ранѣе пять лѣтъ послѣ заключенія договора. Получаемъ:

$$P(_5|\,A_{32}) = \frac{M_{37}}{N_{31}} \cdot 1000 = \frac{9378{,}96}{589246} \cdot 1000 = 15{,}91.$$

Задача VII.

§ 9. *Отсроченное страхованіе на случай смерти съ временнымъ взносомъ премій.* Получаемъ формулу:

$$ {}_kP\,({}_n|\,A) = \frac{{}_n|\,A}{|_k\mathbf{a}_x} = \frac{M_{x+n}}{N_{x-1} - N_{x+k-1}}. $$

Численный примѣръ. Лицо возраста 25 лѣтъ застраховало жизнь въ 5000 рублей съ отсроченнымъ на 5 лѣтъ страхованіемъ и съ временной въ продолженіи 10 лѣтъ уплатой преміи. Получаемъ:

$$ 5000 \cdot \frac{M_{30}}{N_{24} - N_{34}} = 5000 \cdot \frac{10946{,}14}{843679 - 501101} = 159{,}76. $$

Задача VIII.

§ 10. *Отсроченное временное страхованіе на случай смерти.* Въ случаѣ пожизненнаго взноса преміи, получаемъ:

$$ P({}_m|\,{}_{n-m}A_x) = \frac{{}_m|\,{}_{n-m}A_x}{\mathbf{a}_x} = \frac{M_{x+m} - M_{x+n}}{N_{x-1}}. $$

При временномъ же взносѣ премій получаемъ:

$$ {}_kP({}_m|\,{}_{n-m}A_x) = \frac{M_{x+m} - M_{x+n}}{N_{x-1} - N_{x+k-1}}. $$

Задача IX.

§ 11. *Отсроченное страхованіе съ условіемъ возврата внесенныхъ премій въ случаѣ смерти застрахованнаго до истеченія періода отсрочки.* Обыкновенно подлежатъ возврату brutto-преміи, т. е. дѣйствительные взносы страхователя, а не netto-преміи. Подобный планъ страхованія принятъ русскими Государственными сберегательными кассами на основаніи закона 1904 года.

Обозначимъ черезъ P_x искомую netto-премію указаннаго страхованія, а черезъ P brutto-премію, такъ что $P = P_x(1+\alpha)$, гдѣ α есть обыкновенно правильная дробь.

Если смерть застрахованнаго лица возраста x произойдетъ въ промежутокъ времени $(x+k-1,\ x+k)$, вѣроятность чего будетъ ${}_{k-1}|\,q_x$, то въ этомъ случаѣ подлежитъ возврату $k \cdot P$.

Итакъ, современная стоимость всѣхъ brutto-премій, подлежащихъ возврату, будетъ:

$$ (1) \ldots\ldots\ldots W = \sum_{k=1}^{k=n} {}_{k-1}|\,q_x \cdot V^k \cdot k \cdot P, $$

гдѣ n указываетъ періодъ отсрочки страхованія. При этомъ, очевидно, при $k=1$, мы имѣемъ: $q_x = {}_0|q_x$ . . . .

Итакъ, искомая премія P_x при отсроченномъ на n лѣтъ страхованіи съ возвратомъ brutto-преміи, будетъ вычисляться по формулѣ:

$$(2) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; P_x \mathbf{a}_x = W + {}_n|A_x.$$

Приведемъ P_x къ вычисленію коммутаціонныхъ чиселъ:

$$W = q_x VP + {}_1|q_x V^2 2P + \ldots + {}_{n-1}|q_x V^n . nP =$$

$$= P\frac{C_x + 2C_{x+1} + 3C_{x+2} + \ldots \; nC_{x+n-1}}{D_x} =$$

$$= P\frac{R_x - R_{x+n} - nM_{x+n}}{D_x}, \; {}^{1)}$$

отсюда формула (2) перепишется такъ:

$$P_x \frac{\mathbf{N}_x}{D_x} = P_x(1+\alpha)\frac{R_x - R_{x+n} - n.M_{x+n}}{D_x} + \frac{M_{x+n}}{D_x},$$

откуда получаемъ:

$$P_x = \frac{M_{x+n}}{N_{x-1} - (1+\alpha)\left[R_x - R_{x+n} - nM_{x+n}\right]}.$$

Численный примѣръ: $x=30$, $n=7$, $\alpha=0{,}1$.

Получаемъ:

$$P_{30} = \frac{M_{37}}{N_{29} - 1{,}1\left[R_{30} - R_{37} - 7M_{37}\right]} =$$

$$= \frac{9378{,}96}{654523 - 1{,}1\left[294665{,}43 - 222867{,}11 - 7.9378{,}96\right]} = 0{,}0144.$$

---

1) $R_x = C_x + 2C_{x+1} + \ldots + n.C_{x+n-1} + (n+1)C_{x+n} +$
$+ (n+2)C_{x+n+1} + \ldots\ldots$

$R_{x+n} = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots . C_{x+n} +$
$+ 2C_{x+n+1} + \ldots\ldots;$

отсюда получаемъ черезъ вычитаніе:

$$R_x - R_{x+n} = C_x + 2C_{x+1} + \ldots + n.C_{x+n-1} + $$
$$+ n\left[C_{x+n} + C_{x+n+1} + \ldots\right],$$

но выраженіе въ послѣднихъ скобкахъ есть не что иное, какъ M_{x+n}.

Задача X.

§ 12. *Страхованіе на опредѣленный срокъ.* Подъ страхованіемъ на опредѣленный срокъ разумѣютъ такой видъ страхованія, когда страховщикъ уплачиваетъ капиталъ по истеченіи указаннаго срока, независимо отъ того, остается ли застрахованное лицо въ живыхъ, или нѣтъ. Если подлежитъ уплатѣ черезъ n лѣтъ единица капитала, то современная стоимость обязательствъ страховщика будетъ равняться V^n, ибо вѣроятность уплаты есть единица.

Netto-премія ${}_nP(V^n)$ по такому страхованію вычислится по формулѣ:

$$ {}_nP(V^n)\,.\,|{}_n\mathbf{a}_x = V^n $$

или

$$ {}_nP(V^n) = V^n \frac{D_x}{\mathbf{N}_x - \mathbf{N}_{x+n}}\,. $$

Задача XI.

§ 13. *Смѣшанное страхованіе нѣкотораго капитала.* Подъ смѣшаннымъ страхованіемъ капитала мы разумѣемъ такое, когда страховщикъ обязуется уплатить застрахованный капиталъ какъ въ случаѣ смерти застрахованнаго лица, если она наступитъ въ теченіе опредѣленнаго періода времени, такъ и въ случаѣ дожитія застрахованнаго въ концѣ этого періода. Мы видимъ слѣдовательно, что смѣшанное страхованіе есть комбинація двухъ страхованій: временнаго страхованія на случай смерти и страхованія на доживіе до конца дѣйствія перваго страхованія.

Пусть лицо возраста x лѣтъ заключаетъ смѣшанное страхованіе единицы капитала на срокъ n лѣтъ, тогда очевидно, что современная стоимость платежей страховщика будетъ состоять: 1) изъ стоимости $|{}_nA_x$ страхованія на случай смерти, происходящей во время указаннаго періода страхованія, и изъ современной стоимости ${}_nE_x$ страхованія на доживіе до конца періода. Полная стоимость обязательствъ страховщика по этому страхованію будетъ, очевидно, ${}_nE_x + |{}_nA_x$. Такъ какъ годовая netto-премія такого страхованія не можетъ быть уплачена болѣе n разъ, то мы ее обозначимъ:

$$ {}_nP({}_nE_x + |{}_nA_x). $$

Получаемъ рѣшеніе задачи по слѣдующей формулѣ:

$$ {}_nP({}_nE_x + |{}_nA_x) = \frac{{}_nE_x + |{}_nA_x}{|{}_n\mathbf{a}_x} = \frac{D_{x+n}}{\mathbf{N}_x - \mathbf{N}_{x+n}} + \frac{M_x - M_{x+n}}{\mathbf{N}_x - \mathbf{N}_{x+n}} . $$

Задача XII.

§ 14. *Страхованіе на дожитіе.* Страхованіе на дожитіе заключается въ томъ, что страховщикъ обязуется уплатить условленную сумму по истеченіи опредѣленнаго числа n лѣтъ въ томъ случаѣ, если застрахованное лицо останется къ этому времени въ живыхъ.

Очевидно, что ежегодная премія по этому страхованію вычислится по формулѣ:

$$ {}_nP({}_nE_x) = \frac{{}_nE_x}{|{}_n\mathbf{a}_x} = \frac{D_{x+n}}{\mathbf{N}_x - \mathbf{N}_{x+n}} . $$

Если взносъ преміи по страхованію на дожитіе совершается только въ теченіе m лѣтъ, причемъ $m < n$, то премія будетъ вычисляться по формулѣ:

$$ {}_mP({}_nE_x) = \frac{{}_nE_x}{|{}_m\mathbf{a}_x} = \frac{D_{x+n}}{\mathbf{N}_x - \mathbf{N}_{x+m}} . $$

Въ случаѣ единовременной преміи мы получаемъ:

$$ {}_1P({}_nE_x) = \frac{{}_nE_x}{|{}_1\mathbf{a}_x} = \frac{D_{x+n}}{D_x} = {}_nE_x . $$

Задача XIII.

§ 15. *Страхованіе на дожитіе съ возвратомъ въ случаѣ смерти внесенныхъ премій.* Соображенія для рѣшенія этой задачи будутъ отличаться отъ тѣхъ, которыя мы имѣли въ задачѣ IX, только тѣмъ, что вмѣсто отсроченнаго страхованія смерти ${}_n|A_x$ придется взять страхованіе на дожитіе ${}_nE_x$ и вмѣсто пожизненной ренты $\mathbf{a}_x$ надо будетъ написать временную ренту $|{}_n\mathbf{a}_x$, такъ что вмѣсто формулы (2) § 11 придется написать такую:

(1) . . . . $P_x|{}_n\mathbf{a}_x = W + {}_nE_x ,$

получаемъ:

$$ P_x \frac{N_{x-1} - N_{x+n-1}}{D_x} = P_x(1+\alpha)\frac{R_x - R_{x+n} - nM_{x+n}}{D_x} + \frac{D_{x+n}}{D_x} , $$

откуда:

$$P_x = \frac{D_{x+n}}{N_{x-1} - N_{x+n-1} - (1+\alpha)\left[R_x - R_{x+n} - nM_{x+n}\right]}.$$

Задача XIV.

§ 16. *Страхованіе пенсій.* Подъ страхованіемъ пенсій разумѣются различнаго вида задачи, гдѣ предполагается уплата извѣстнаго числа разъ премій, причемъ страховщикъ обязуется черезъ извѣстный срокъ послѣ окончанія уплаты премій уплачивать страхователю періодическую пенсію.

Очевидно, что эта пенсія можетъ разсматриваться, какъ рента и потому формула вычисленія netto-преміи такого страхованія будетъ, очевидно, такая:

$$P \,|_k\mathbf{a} = \mathfrak{a}\,,$$

гдѣ k выражаетъ число разъ уплаты преміи, а $\mathfrak{a}$ представляетъ изъ себя современную стоимость пенсіи, которую будетъ получать страхователь. Очевидно, что эта современная стоимость пенсіи будетъ вычислятся по формуламъ, даннымъ въ § 2 главы X; напримѣръ, если пенсія единовременная, то k равняется единицѣ, $|_k\mathbf{a}$ тоже равняется единицѣ, и мы получаемъ: $P = \mathfrak{a}$. Если ежегодная пенсія будетъ равняться K единицъ капитала и будетъ пожизненно уплачиваться postnumerando, то получится

$$P = Ka_x.$$

Если пенсія будетъ отсрочена на m лѣтъ и будетъ уплачиваться praenumerando, то получается формула:

$$P = K\,{}_m|\mathbf{a}_x.$$

Подобнымъ же образомъ мы можемъ вычислить netto-премію въ случаѣ отсроченной временной пенсіи.

XII.

Вычисленіе резервовъ.

§ 1. Въ § 7 гл. IX мы дали понятіе о резервѣ (Dackungs-kapital, Prämienreserve). Не трудно видѣть, что резервъ, какъ излишекъ современной стоимости обязательствъ страховщика, надъ современной стоимостью обязательствъ застрахованнаго, представляетъ изъ себя какъ бы математическую мѣру риска страховщика, связанную съ даннымъ страхованіемъ. Поэтому практическое значеніе резерва состоитъ въ томъ, что страховщикъ долженъ изъ полученныхъ имъ премій оставить въ своемъ запасѣ неизрасходованную сумму денегъ, равную резерву страхователя. Практическая достаточность такого резерва основывается, конечно, на законѣ большихъ чиселъ. Если, съ одной стороны, при вычисленіи резервовъ употребляются хорошо составленныя таблицы смертности, и правильно установленъ техническій процентъ роста капитала, и если, съ другой стороны, страховое учрежденіе имѣетъ очень большое число кліентовъ, причемъ дѣйствительная смертность этихъ кліентовъ не отклоняется значительно въ неблагопріятную сторону отъ теоретической смертности, указываемой въ таблицѣ смертности, то можно почти съ достовѣрностью утверждать, что, пользуясь скопленными резервами, страховое учрежденіе будетъ въ состояніи своевременно произвести уплату по всѣмъ своимъ обязательствамъ. Въ виду сказаннаго, правильная постановка вычисленія резервовъ является кардинальнымъ вопросомъ страхового дѣла.

§ 2. На основаніи разъясненій § 7 главы IX вычисленіе резервовъ не представляетъ никакого затрудненія. Въ виду этого мы ограничимся болѣе подробнымъ указаніемъ пріемовъ вычисленія на частныхъ задачахъ.

Мы будемъ во всемъ дальнѣйшемъ, обозначать резервы такъ: V (valeur, Valuation). Направо въ скобкахъ мы будемъ указывать также, какъ и при netto-преміяхъ, къ какому плану страхованія резервъ относится. Значкомъ слѣва внизу мы будемъ обозначать къ концу какого года отъ начала страхованія резервъ вычисляется. Такъ, знакъ ${}_nV(A_x)$ будетъ обозначать резервъ въ концѣ n-го года отъ начала страхованія при страхованіи на случай смерти.

Если происходитъ временная уплата премій, то мы будемъ употреблять знакъ: ${}^{(m)}_{\ n}V$, который укажетъ, что резервъ относится къ концу n-го года отъ начала страхованія, причемъ страхованіе совершенно съ временнымъ m разъ взносомъ преміи. При единовременной уплатѣ преміи знакъ резерва будетъ ${}^{(1)}_{\ n}V$.

Если премія уплачивается p разъ въ году, то употребляется знакъ ${}^{(m)}_{\ n}V^{(p)}$.

§ 3. Разсмотримъ страхованіе единицы капитала на случай смерти. Тогда, очевидно, резервъ опредѣлится по слѣдующей формулѣ:

$$(1) \ldots\ldots\ldots {}_nV(A_x) = A_{x+n} - P(A_x)\mathbf{a}_{x+n},$$

ибо черезъ n лѣтъ послѣ начала страхованія стоимость обязательствъ страховщика будетъ A_{x+n}, а стоимость всѣхъ премій, предстоящихъ къ уплатѣ страхователемъ будетъ $P(A_x)\mathbf{a}_{x+n}$.

Въ § 3 гл. XI мы видѣли, что netto-премія выражается по формулѣ:

$$(2) \ldots\ldots\ldots\ldots P(A_x) = \frac{A_x}{\mathbf{a}_x};$$

отсюда получаемъ для вычисленія резерва такую формулу:

$$(3) \ldots {}_nV(A_x) = A_{x+n} - A_x \frac{\mathbf{a}_{x+n}}{\mathbf{a}_x} = \frac{M_{x+n}}{D_{x+n}} - \frac{M_x}{D_x} \cdot \frac{\mathbf{N}_{x+n} \cdot D_x}{\mathbf{N}_x \cdot D_{x+n}} =$$

$$= \frac{M_{x+n}\mathbf{N}_x - \mathbf{N}_{x+n}M_x}{D_{x+n}\mathbf{N}_x}.$$

§ 4. Если премія уплачивается временно въ теченіе m лѣтъ, тогда при вычисленіи резерва ${}_n^{(m)}V(A_x)$ нужно разобрать два случая:

1) $n \geqslant m$. Въ этомъ случаѣ всѣ преміи уже выплачены къ моменту резерва, слѣдовательно стоимость обязательствъ страхователя равняются нулю. Резервъ состоитъ цѣликомъ изъ стоимости обязательствъ страховщика, и мы получаемъ:

$${}_n^{(m)}V(A_x) = A_{x+n}.$$

2) Если же $n < m$, то остается еще $m - n$ уплатъ премій, и тогда мы, конечно, придемъ къ формулѣ:

$${}_n^{(m)}V(A_x) = A_{x+n} - P(A_x) \cdot |_{m-n}\mathbf{a}_{x+n};$$

въ правой части послѣдняго равенства всѣ выраженія по формуламъ, указаннымъ раньше, можно выразить черезъ коммутаціонныя числа.

Очевидно, что мы сумѣемъ очень просто вычислить резервы, относящіеся ко всѣмъ видамъ страхованій, о которыхъ мы говорили въ предыдущей главѣ.

Напримѣръ, если мы желаемъ вычислить резервъ

$${}_n^{(m)}V(|_mA_x),$$

который относится къ временному на m лѣтъ страхованію жизни, причемъ, конечно, мы предполагаемъ $m > n$, ибо черезъ m лѣтъ сдѣлка прекращается, то получаемъ:

$$(1) \quad \ldots \quad {}^{(m)}_{n}V(|_{m}A_x) = |_{m-n}A_{x+n} - {}_{m}P(|_{m}A_x) \cdot |_{m-n}\mathsf{a}_{x+n} \,.$$

§ 5. Резервъ по смѣшанному страхованію на n лѣтъ очевидно вычислится по формулѣ:

$$ {}^{(m)}_{n}V(|_{m}A_x + {}_{m}E_x) =$$
$$= |_{m-n}A_{x+n} + {}_{m-n}E_{x+n} - {}_{m}P(|_{m}A_x + {}_{m}E_x) \cdot |_{m-n}\mathsf{a}_{x+n}.$$

§ 6. Въ случаѣ страхованія на дожитіе, резервъ вычислится по формулѣ:

$$ {}^{(m)}_{n}V({}_{m}E_x) = {}_{m-n}E_{x+n} - {}_{m}P({}_{m}E_x) \, |_{m-n}\mathsf{a}_{x+n} \,.$$

§ 7. Мы не будемъ останавливаться на вычисленіи резервовъ по обратному методу, ибо такое вычисленіе приводитъ въ концѣ концовъ къ тѣмъ же формуламъ, что и вышеприведенный прямой способъ.

§ 8. *Послѣдовательное вычисленіе резервовъ.* Практика страховыхъ учрежденій обратила вниманіе на удобство т. наз. послѣдовательнаго вычисленія резервовъ. Это послѣдовательное вычисленіе резервовъ указываетъ правила, по которому, зная резервъ ${}_{n}V$ лѣтъ, можно вычислить резервъ на ${}_{n+1}V$ лѣтъ. Практическое удобство этого новаго способа состоитъ въ томъ, что мы не вычисляемъ въ каждомъ отдѣльномъ году резервъ заново, а пользуемся величиною уже вычисленнаго въ прошломъ году резерва. Послѣдній способъ еще удобенъ въ томъ отношеніи, что даетъ извѣстный контроль правильности произведенныхъ вычисленій и дѣлаетъ невозможными особенно грубыя ошибки.

Мы ограничимся разъясненіемъ способа послѣдовательнаго вычисленія резервовъ лишь на одномъ примѣрѣ; возьмемъ страхованіе на случай смерти.

Тогда, по формулѣ (3) § III-го мы получаемъ:

$$(1) \ldots\ldots\ldots {}_{n}V = \frac{M_{x+n}\,\mathbf{N}_x - M_x\,\mathbf{N}_{x+n}}{D_{x+n}\,\mathbf{N}_x} \,;$$

увеличивая n на единицу, мы получимъ:

$$(2) \ldots\ldots\ldots {}_{n+1}V = \frac{M_{x+n+1}\,\mathbf{N}_x - M_x\,\mathbf{N}_{x+n+1}}{D_{x+n+1}\,\mathbf{N}_x} \,;$$

равенства (1) и (2) можно будетъ переписать такъ:

$$(3) \ldots\ldots \quad D_{x+n\,.\,n}V = M_{x+n} - \frac{M_x \mathbf{N}_{x+n}}{\mathbf{N}_x};$$

$$(4) \ldots\ldots \quad D_{x+n+1\,.\,n+1}V = M_{x+n+1} - \frac{M_x \mathbf{N}_{x+n+1}}{\mathbf{N}_x};$$

вычитая формулу (3) изъ формулы (4) получимъ:

$$(5) \ldots \quad D_{x+n+1\,.\,n+1}V - D_{x+n\,.\,n}V = -C_{x+n} + \frac{M_x}{\mathbf{N}_x} D_{x+n}\ ^{1)} =$$

$$= -C_{x+n} + P_x D_{x+n}.$$

Формула (5) даетъ простой способъ по резерву $_nV$ найти резервъ $_{n+1}V$.

Уменьшенные резервы.

§ 9. Приведенные выше способы вычисленія резерва имѣютъ характеръ чисто теоретическій, ибо по этимъ пріемамъ вычисленія разсматривалось нѣкоторое невозможное на практикѣ страховое учрежденіе, не имѣющее никакихъ собственныхъ расходовъ и не желающее получать никакихъ прибылей, причемъ мы вычисляемъ для такого страхового учрежденія теоретическую величину риска по данному страхованію въ предположеніи, что смертность среди кліентовъ этого учрежденія соотвѣтствуетъ теоретическимъ предположеніямъ таблицы смертности и извѣстному, вполнѣ опредѣленному техническому проценту. Практическая дѣятельность страховыхъ учрежденій въ значительной мѣрѣ уклоняется отъ такой теоретической схемы. Страховое учрежденіе несетъ свои расходы, оно не можетъ поручиться за то, что дѣйствительная смертность его кліентовъ не будетъ значительно выше теоретической. Въ виду этого страховое учрежденіе не можетъ довольствоваться netto-преміями и устанавливаетъ дѣйствительный тарифъ выше теоретическихъ премій, не говоря уже о томъ, что на практикѣ страховыя учрежденія стремятся къ выгодному помѣщенію ихъ капиталовъ, чтобы тѣмъ увеличить доходность предпріятія.

---

1) Въ самомъ дѣлѣ:

$$M_{x+n+1} - M_{x+n} = -C_{x+n}; \text{ а } \mathbf{N}_{x+n+1} - \mathbf{N}_{x+n} = -D_{x+n}.$$

Въ виду указаннаго обстоятельства, взиманія страховымъ учрежденіемъ brutto-преміи, а не netto-преміи, является на практикѣ коммерческихъ страховыхъ обществъ возможность откладывать резервы въ меньшемъ размѣрѣ. Наиболѣе элементарной формой является *вычисленіе резервовъ по методу brutto-преміи.*

Вычисленіе по этому методу производится по формуламъ, указаннымъ выше въ §§ 3—8, съ тою лишь разницей, что вмѣсто netto-преміи ставится въ этихъ формулахъ brutto-премія. Такъ какъ по прямому методу вычисленія резервовъ, премія находится въ вычитаемомъ (т. е. въ членѣ со знакомъ минусъ), то очевидно, что резервъ по методу brutto-премій выйдетъ меньше, чѣмъ резервъ по способу netto-премій. Методъ brutto-премій имѣетъ однако большія неудобства. При такомъ способѣ страховое учрежденіе въ текущихъ своихъ поступленіяхъ не имѣетъ источника для покрытія ежегодныхъ расходовъ по управленію дѣлами учрежденія, для обезпеченія дохода по основному капиталу, для составленія запаснаго капитала, для покрытія убытковъ, превышающихъ теоретическіе расчеты, и т. д.

Въ виду этого чаще на практикѣ при вычисленіи резервовъ берется часть brutto-преміи, превышающая однако netto-премію.

Всѣ такіе пріемы вычисленія резервовъ носятъ названіе *пріемовъ гипотетическихъ премій,* такъ какъ при этомъ вмѣсто дѣйствительныхъ brutto или netto-премій берутся нѣкоторыя произвольныя числа. Къ числу такихъ гипотетическихъ пріемовъ относится пріемъ, получившій большое распространеніе подъ названіемъ *цильмерованія.* Первоначальная идея этого метода принадлежитъ англійскому актуарію Spraag'у, но методъ этотъ подробно развитъ и примѣненъ на практикѣ Zillmer'омъ въ его книгѣ: Beitrage zûr Teorie der Premienreserven. Stettin. 1863 г.

§ 10. Мы введемъ въ разсмотрѣніе новый терминъ: *достаточная премія.*

Положимъ, что при страхованіи единицы капитала по нѣкоторому опредѣленному плану этого страхованія причитается netto-премія P_x. Очевидно, что это страхованіе, какъ и всякое другое, сопряжено съ извѣстными расходами для страхового общества. Эти расходы можно разбить на двѣ части: 1) *единовременные* или *первые расходы* при заключеніи договора и 2) *текущіе расходы,* повторяющіеся ежегодно.

Къ числу первоначальныхъ расходовъ можно отнести вознагражденіе агентовъ за привлеченіе новыхъ страхователей, далѣе—жалованіе служащихъ, которымъ приходится принимать участіе при заключенія страхованій, печатаніе объявленій и рекламъ, гонораръ за врачебное освидѣтельствованіе, правительственныя пошлины и т. п.

Къ текущимъ, повторяющимся ежегодно расходамъ можно причислить вознагражденіе служащихъ, наемъ помѣщенія, расходы по корреспонденціи, ежегодныя правительственныя пошлины, вознагражденіе агентовъ при incasso премій и т. д.

Перечисленные расходы носятъ названіе вычислительныхъ величинъ второго порядка, въ отличіе отъ вычислительныхъ величинъ перваго порядка, къ которымъ принадлежитъ техническій процентъ и даваемыя таблицей смертности величины. Весьма важно замѣтить, что величины второго порядка могутъ быть точнѣе установлены, чѣмъ величины перваго порядка.

Подъ достаточными преміями разумѣются преміи, превышающія netto-преміи и достаточныя для покрытія всѣхъ расходовъ страхованія, какъ первыхъ, такъ и текущихъ. Обозначимъ черезъ P'_x такую достаточную премію по разсматриваемой нами единицѣ капитала. Часть этой достаточной преміи, идущая на покрытіе текущихъ годовыхъ расходовъ, можетъ быть обозначена черезъ $\gamma P'_x$, гдѣ γ есть положительная правильная дробь. Если черезъ δ мы обозначимъ первый расходъ страхованія, то очевидно, что современная стоимость $P'_x\mathbf{a}$ уплаты достаточныхъ премій въ продолженіи всего страхованія должна состоять изъ трехъ частей: а) изъ современной стоимости $P_x\mathbf{a}$ netto-премій, б) изъ современной стоимости $\gamma P_x\mathbf{a}$ - текущихъ расходовъ и в) изъ единовременнаго расхода δ.

Мы получаемъ:

$$P'_x\mathbf{a} = P_x\mathbf{a} + \gamma P'_x\mathbf{a} + \delta,$$

или

$$(1) \quad . \; . \; . \; . \; . \; . \; . \quad P'_x = P_x + \gamma P'_x + \frac{\delta}{\mathbf{a}};$$

отсюда

$$P'_x(1-\gamma) = P_x + \frac{\delta}{\mathbf{a}} =$$

$$(2) \quad . \; . \; . \; . \; . \; . \; . \quad P'_x = P_x \frac{1}{1-\gamma} + \frac{\delta}{\mathbf{a}(1-\gamma)}.$$

Формула (2) показываетъ, какъ вычислять достаточную премію P'_x, если извѣстна netto-премія P_x.

§ 11. Предполагая, что при вычисленіи резерва вмѣсто netto-преміи употребляется достаточная премія, получаемъ тотъ уменьшенный резервъ, который былъ предложенъ Zillmer'омъ.

Если мы будемъ обозначать черезъ ${}_mV$ резервъ netto-премій, вычисленный къ концу m-го года, то соотвѣтствующій Zillmer'овскій резервъ будемъ обозначать ${}_mV'$.

Для того, чтобы сдѣлать наши разсужденія совершенно общими, обозначимъ черезъ $A_{(n)}$ [1]) стоимость обязательствъ страхового общества къ моменту вычисленія резерва, совершенно независимо отъ того, какой планъ страхованія разсматривается. Подобнымъ же образомъ обозначимъ знакомъ $\mathbf{a}_{(n)}$ стоимость единицы преміи даннаго страхованія, относящуюся къ моменту вычисленія резерва. Подобнымъ же образомъ черезъ $P_{(n)}$ мы обозначимъ netto-премію, которую пришлось бы платить страхователю по страхованію того же вида, если бы страхованіе началось въ моментъ вычисленія резерва, т. е. черезъ n лѣтъ послѣ дѣйствительнаго начала страхованія. Имѣемъ

$$(1) \quad \ldots\ldots\ldots\ldots \quad P_{(n)}\mathbf{a}_{(n)} = A_{(n)},$$

или

$$P_{(n)} = \frac{A_{(n)}}{\mathbf{a}_{(n)}}.$$

§ 12. Очевидно, что Zillmer'овскій резервъ долженъ быть вычисленъ по формулѣ:

$${}_nV' = A_{(n)} + \gamma \cdot P'_x \cdot \mathbf{a}_{(n)} - P'_x \mathbf{a}_{(n)}.$$

Здѣсь въ первой части прибавляется второй членъ, представляющій современную стоимость всѣхъ слѣдующихъ за годомъ резерва текущихъ расходовъ по страхованію; въ членъ же со знакомъ минусъ, какъ было сказано выше, вставлена достаточная премія вмѣсто netto-преміи.

Примѣняя формулу (1) § 10, мы получимъ:

$${}_nV' = A_{(n)} + \gamma P'_x \mathbf{a}_{(n)} - \left(P_x + \gamma P'_x + \frac{\delta}{\mathbf{a}}\right) \mathbf{a}_{(n)},$$

---

[1]) Такъ напримѣръ подъ знакомъ $A_{(n)}$ можно разумѣть одинъ изъ знаковъ ${}_{m-n}E_{x+n}$, A_{x+n}, ${}_{m-n}E_{x+n} + A_{x+n}$ и т. д., судя по тому, какой планъ страхованія разсматривается.

или окончательно:

$$(1) \ldots\ldots\ldots \quad {}_nV' = A_{(n)} - \left(P_x + \frac{\delta}{\mathbf{a}}\right)\mathbf{a}_{(n)}\,;$$

послѣдняя формула учитъ, что Zillmer'овскій резервъ вычисляется точно по тому же правилу, что и резервъ netto-преміи, съ тою лишь разницей, что вмѣсто netto-преміи берется величина $P_x + \frac{\delta}{\mathbf{a}}$, которая носитъ названіе: *Zillmer'овская резервная премія*. Легко составить себѣ ясное представленіе объ этой преміи. Въ самомъ дѣлѣ, если мы представимъ себѣ, что первые расходы страхованія δ представляютъ изъ себя современную стоимость нѣкоторой годовой ренты α, то мы будемъ имѣть:

$$\delta = \alpha\mathbf{a}\,,$$

откуда:

$$\alpha = \frac{\delta}{\mathbf{a}}\,.$$

Итакъ, Zillmer'овская премія оказывается не чѣмъ инымъ, какъ величиной $P_x + \alpha$.

Мы приходимъ къ слѣдующей окончательной характеристикѣ Zillmer'овской преміи: Zillmer предлагаетъ погасить расходъ δ не единовременно, а разсрочить его въ видѣ пожизненной ренты съ ежегоднымъ взносомъ α и прибавлять этотъ взносъ α къ netto-преміи.

§ 13. Принимая во вниманіе формулу (1) § 11-го, мы получимъ изъ формулы (1) § 12-го слѣдующую:

$${}_nV' = \left(P_{(n)} - P_x - \frac{\delta}{\mathbf{a}}\right)\mathbf{a}_{(n)}\,.$$

Если первоначальный расходъ δ будетъ очень значительный, то резервъ (1) можетъ выйти числомъ отрицательнымъ. Такъ какъ по самому понятію резерва, резервъ всегда возрастаетъ съ теченіемъ времени, то, если резервъ въ концѣ перваго года былъ не отрицательный, онъ будетъ числомъ положительнымъ и дальше.

Итакъ, условіемъ необходимымъ и достаточнымъ для отсутствія отрицательнаго резерва является неравенство:

$$P_{(1)} - P_x - \frac{\delta}{\mathbf{a}} \geqq 0\,,$$

или:

$$\delta \leqq \mathbf{a}(P_{(1)} - P_x)\,.$$

Величина $\mathfrak{a}(P_{(1)} - P_x)$ называется *Zillmer'овскимъ максимумомъ первоначальныхъ расходовъ.*

Zillmer рекомендуетъ во избѣжаніе отрицательныхъ резервовъ стараться, чтобы первоначальные расходы δ на единицу застрахованнаго капитала не превзошли числа $\mathfrak{a}(P_{x+1} - P_x)$, ибо, очевидно, $P_{(1)}$ есть не что иное, какъ P_{x+1}. При этомъ это число P_{x+1} надо понимать согласно сказанному выше, какъ такую премію, которую бы пришлось назначить, если бы отложить на годъ заключеніе сдѣлки, не нарушая всѣхъ остальныхъ условій этой сдѣлки.

Если мы предположимъ, что первые расходы равняются въ точности максимуму Zillmer'а, тогда $\frac{\delta}{\mathfrak{a}} = P_{x+1} - P_x$, и, подставляя это выраженіе въ формулу (1) § 12, получимъ:

$$_nV' = A_{(n)} - (P_x + P_{x+1} - P_x)\mathfrak{a}_{(n)},$$

или окончательно

$$_nV' = A_{(n)} - P_{x+1}\mathfrak{a}_{(n)}.$$

На основаніи послѣдней формулы часто называютъ способъ вычисленія резерва въ предположеніи Zillmer'овскаго максимума „*методой* $x + 1$“.

§ 14. Хотя цильмерованіе приводитъ къ меньшему значенію резерва, чѣмъ по способу netto-премій, однако при правильно выбранномъ процентѣ и хорошей таблицѣ смертности получается резервъ вполнѣ достаточный.

Если страховое общество молодое и не обладаетъ достаточно большимъ, внесеннымъ учредителями, капиталомъ, то первые расходы по заключенію страхованій будутъ на столько велики, что едвали можно будетъ обойтись безъ цильмерованія.

Въ старыхъ обществахъ, имѣющихъ достаточное число прежнихъ страхованій, имѣются извѣстныя сбереженія, дающія возможность покрывать первые расходы новыхъ страхованій безъ цильмерованія.

Въ пользу цильмерованія говоритъ также слѣдующее обстоятельство. Если вычисляется резервъ по методѣ netto-премій, то въ случаѣ неожиданно большого наплыва новыхъ страхованій, первые по нимъ расходы могутъ быть настолько значительны, что могутъ повести не только къ уменьшенію дивидендовъ, но даже къ дефициту.

Правительственные органы надзора не особенно однако относятся благосклонно къ методу Zillmer'a. Такъ напримѣръ, по Германскому закону цильмерованіе допускается только въ размѣрѣ не превышающемъ $12^1/_2$ на 1000 застрахованнаго капитала, т. е. другими словами допускаетъ только $\delta \leqq 0{,}0125$.

XIII.

Выкупъ и редуцированіе полисовъ.

§ 1. При заключеніи страхового договора обыкновенно принимается, что страховщикъ не имѣетъ права отказаться отъ договора въ теченіе времени его дѣйствія, если только страхователь и застрахованный точно выполняютъ всѣ полисныя условія. Страхователь же наоборотъ, имѣетъ почти всегда право прекратить дѣйствіе договора, когда ему угодно, подъ угрозой, конечно, потерять внесенные платежи.

При заключеніи страхового договора, въ нѣкоторыхъ случаяхъ, довольно часто, страхователю представляется право при добровольномъ оставленіи страхованія получать обратно извѣстную долю изъ внесенныхъ имъ страховщику страховыхъ суммъ.

Такая операція носитъ названіе *выкупа страхованія*. Размѣръ выкупныхъ суммъ опредѣляется обыкновенно полисными условіями, т. е. соглашеніемъ страховщика и страхователя, а иногда закономъ.

§ 2. Теоретически разсуждая, съ точки зрѣнія математической безобидности, страхователь при выкупѣ имѣетъ право получить весь резервъ по его страхованію, который страховщику, очевидно, не нуженъ, вслѣдствіе прекращенія договора. На практикѣ, однако, страховыя общества въ своихъ договорахъ со страхователями устанавливаютъ выкупную сумму въ размѣрѣ меньшемъ причитающагося резерва, чтобы такимъ путемъ повліять на сокращеніе числа лицъ, желающихъ прекратить страхованіе, ибо большое число прекращеній страхованій можетъ отразится весьма неблагопріятно на дѣлахъ страхового учрежденія.

Ни одно страховое учрежденіе не допускаетъ выкупа страхованія, если оно заключено на дожитіе безъ возврата премій.

§ 3. Иногда допускается договоромъ страхователя и страховщика такъ называемое *редуцированіе страхованія*.

Подъ этимъ разумѣютъ уменьшенiе размѣра страховой суммы въ соотвѣтствiи съ уменьшенiемъ и даже прекращенiемъ уплаты премiи.

XIV.

Балансъ.

§ 1. Страховое общество ежегодно должно подводитъ балансъ своего актива и пассива.

Способы сведенiя баланса претерпѣваютъ извѣстныя видоизмѣненiя въ частности, но въ общемъ установленъ рядъ принциповъ, принятыхъ въ большинствѣ страховыхъ учрежденiй.

Мы сообщимъ основныя понятiя, касающiяся подведенiя баланса, имѣя въ виду главнымъ образомъ практику германскихъ страховыхъ учрежденiй. Въ Германiи правила подведенiя баланса установлены предписанiями правительственнаго органа, называемаго das Kaiserliche Aufsichtsamt für Privatversicherungen.

Примѣръ баланса, составленнаго согласно этимъ предписанiямъ, можно найти въ Veröffentlichungen des Kais. Aufsichtsamtes f. Privatv. Jahrg. 1902; s. 38.

§ 2. Основную статью пассива образуютъ резервы по всѣмъ страхованiямъ; ихъ совокупность образуетъ такъ называемый *общiй резервъ* общества.

При этомъ надо обратить вниманiе на то обстоятельство, что къ моменту составленiя баланса не истекаетъ обыкновенно полнаго числа лѣтъ различныхъ страхованiй, ибо вступленiе новыхъ страхователей происходитъ въ разные времена года.

Пусть ко дню баланса продолжительность нѣкотораго страхованiя будетъ

$$m+\alpha,$$

гдѣ m цѣлое число, а α правильная дробь. Обозначая резервъ единицы этого страхованiя въ день баланса черезъ

$$(1) \quad \dots\dots\dots\dots \quad {}_{m+\alpha}V_x$$

мы замѣчаемъ, что для точнаго вычисленiя резерва (1) пришлось бы примѣнять правила непрерывнаго вычисленiя страхованiя, что привело бы къ довольно сложнымъ формуламъ интегральнаго исчисленiя.

На практикѣ, однако, достаточно приближенное рѣшенiе, основанное на слѣдующихъ соображенiяхъ.

Въ концѣ m-аго года резервъ будетъ ${}_mV_x$, въ началѣ же $m+1$-аго года резервъ увеличится на сумму преміи P_x, такъ что будетъ равенъ

$$ {}_mV_x + P_x . $$

Въ справедливости сказаннаго легко убѣдиться изъ того соображенія, что стоимость уплатъ страхователя измѣняется въ безконечно малый промежутокъ времени

$$ (m - \varepsilon,\ m + \varepsilon), $$

гдѣ ε безконечно малая величина, на сумму P_x преміи, уплата которой приходится на средину m промежутка. Съ другой стороны, вслѣдствіе безконечной малости промежутка измѣненіе стоимости обязательствъ страховщика ничтожно мало.

Итакъ, резервъ страхованія измѣняется въ продолженіе $m+1$-аго года, отъ значенія ${}_mV_x + P_x$ до значенія ${}_{m+1}V_x$. Это измѣненіе можно предполагать равномѣрнымъ, т. е. предполагать, что измѣненіе резерва за долю года α будетъ равно части α всего годового измѣненія

$$ {}_{m+1}V_x - {}_mV_x - P_x , $$

такъ что

$$ {}_{m+\alpha}V_x = {}_mV_x + P_x + \alpha\,({}_{m+1}V_x - {}_mV_x - P_x) . $$

Часто на практикѣ допускаютъ еще большее упрощеніе, полагая $\alpha = \frac{1}{2}$, т. е. приводятъ вычисленіе резерва къ срединѣ года

$$ {}_{m+\frac{1}{2}}V_x = \frac{{}_mV_x + {}_{m+1}V_x}{2} + \frac{P_x}{2} . $$

Часто подъ счетомъ резервовъ пишутъ только часть

$$ \frac{1}{2}\,({}_mV_x + {}_{m+1}V_x) $$

резерва, остальную же часть $\frac{P_x}{2}$ записываютъ въ пассивъ подъ особеннымъ счетомъ *переноса премій* (Prämienübertrag).

Многія общества пишутъ въ счетѣ переноса премій brutto-преміи, или же достаточныя преміи, если при вычисленіи резервовъ употреблялась метода Zillmer'a.

Въ пассивъ входятъ два слѣдующихъ важныхъ счета: 1) резервъ убытковъ (Schaden-und Unkostenreserve), 2) резервъ риска (Risikoreserve).

Въ составъ резерва убытковъ можно помѣстить всѣ подлежащія уплатѣ страховымъ обществомъ застрахованныя суммы, назначенныя къ выдачѣ до дня баланса и оставшіяся по какимъ либо причинамъ не выплаченными къ этому дню. Въ этомъ счету надо показать стоимость подлежащихъ уплатѣ страхованій на опредѣленный срокъ, если ко дню баланса всѣ преміи договоровъ уже выплачены. Вообще говоря, тутъ надо показать всѣ расходы, могущіе быть по страхованіямъ свободнымъ отъ уплаты премій. Резервъ риска составляетъ экстраординарныя надбавки къ преміямъ въ случаяхъ особеннаго риска страховщика (слабость здоровья застрахованнаго, опасныя профессіи и т. д.).

Кромѣ указанныхъ болѣе крупныхъ счетовъ могутъ быть другіе болѣе мелкіе въ зависимости отъ устава страхового предпріятія.

Къ пассиву должны быть причислены всѣ экстренные фонды, не составляющіе имущества общества, какъ то: недополученный страхователями дивидендъ, имущество третьихъ лицъ, преждевременно уплаченныя преміи и т. д.

Пассиву противупоставляется все имущество страхового общества, какъ активъ: чистыя деньги, процентныя бумаги и т. д.

Необходимо обратить вниманіе на то обстоятельство, что къ активу должны принадлежать счета *ссудъ подъ полисы* и *отсроченныя преміи*.

§ 3. Разница между величиной актива и пассива составляетъ *чистую прибыль*, которая вносится при подведеніи баланса въ пассивъ.

Назначеніе и распредѣленіе чистой прибыли совершается по уставу учрежденія согласно общимъ законамъ страны.

Страховыя общества раздѣляются на двѣ категоріи: 1) акціонерныя, 2) взаимныя.

Въ акціонерныхъ обществахъ изъ чистой прибыли уплачиваются дивиденды акціонеровъ. Въ обществахъ взаимныхъ страхователи участвуютъ въ прибыляхъ общества. Часть дивиденда причитающагося каждому страхователю или выдается ему наличными, или списывается съ его преміи.

§ 4. Если мы проанализируемъ источники прибылей страховаго учрежденія, то мы можемъ замѣтить слѣдующіе главнѣйшіе изъ этихъ источниковъ:

1) прибыль отъ благопріятной смертности;

2) прибыль отъ выгоднаго помѣщенія капиталовъ общества;

3) случайные источники прибыли.

Что касается до прибыли отъ благопріятной смертности, то можетъ случиться явленіе обратное, а именно, смертность можетъ оказаться очень неблагопріятной для страхового учрежденія, и тогда вмѣсто прибыли можетъ получиться значительный убытокъ.

Въ виду возможности очень неблагопріятной смертности страховое учрежденіе является всегда связаннымъ съ извѣстнымъ рискомъ. Въ послѣднее время составлена математическая теорія риска страхового учрежденія; мы не будемъ на ней останавливаться въ нашемъ элементарномъ изложеніи, отсылая читателя къ слѣдующимъ сочиненіямъ:

Landré. Aperçu succinct des théories du plein de l'assurance Transact. of the sec. internat. actuarialcongress. London, 1899, S. 110.

Radtke. Die Stabilität der Lebensversicherungsanstalten Zeitschrift für d. ges. Versicherungswissenschaft Bd. 3. 399—459 (1903).

Bohlmann. Die Theorie des mittleren Risikos in der Lebens versicherung. Sechster internationaler Kongress f. Versicherungswissenschaft. Bd. 1. S. 593 (1909).

Hugo Broggi. Versicherungsmathematik. 1911.

Практически страховыя общества обезпечиваютъ свой рискъ перестраховкой убытковъ въ другихъ обществахъ.

Прибыль отъ выгоднаго помѣщенія капиталовъ состоитъ въ томъ, что дѣйствительные проценты, получаемые обществомъ на его капиталы, выше того техническаго процента, при которомъ производятся разсчеты премій.

§ 5. Не останавливаясь подробно на способахъ распредѣленія дивиденда между страхователями, мы упомянемъ о такъ называемомъ „*тонтинномъ*“ распредѣленіи.

Подъ названіемъ *тонтины* извѣстна финансовая операція, предложенная въ XVII столѣтіи неаполитанскимъ врачемъ Lorenzo Tonti. Эта операція состоитъ въ томъ что нѣсколько лицъ платятъ извѣстныя суммы для образованія капитала, проценты котораго распредѣляются между пережившими другихъ. Если вся группа вымираетъ, то капиталъ переходитъ въ казну.

Тонтины дѣйствовали сначала въ итальянскихъ и нѣмецкихъ городахъ, позднѣе (1686, 1696, 1706 и пять разъ въ періодъ

1731—1759) французское и англійское правительства дѣлали попытки установленія государственной тонтины.

Тонтину необходимо разсматривать какъ первообразъ современнаго страхового учрежденія.

XV.

Страхованіе на нѣсколько жизней.

§ 1. Наиболѣе важнымъ случаемъ страхованія на нѣсколько жизней являетъ страхованіе двухъ жизней. Такого рода страхованіе встрѣчается, напримѣръ, подъ видомъ страхованія *вдовьихъ* и *сиротскихъ пенсій*.

Мы ограничимся случаемъ разсмотрѣнія двухъ лицъ, причемъ для опредѣленности рѣчи будемъ разсматривать двухъ супруговъ, мужа возраста x лѣтъ и жену возраста y лѣтъ. Вмѣсто супруговъ можно разсматривать x-лѣтняго отца и y-лѣтняго сына, x-лѣтняго брата и y-лѣтнюю сестру и т. д.

Вѣроятности жизни мужа будемъ разсматривать въ зависимости отъ значеній функцій

$$l_x,$$

а вѣроятности жизни жены опредѣлимъ функціей

$$l_y.$$

Обѣ функціи l_x и l_y мы можемъ брать изъ одной таблицы смертности, или же для смертности женщинъ составлять особенную таблицу l_y.

§ 2. Современная стоимость капитала равнаго единицѣ и подлежащаго уплатѣ въ случаѣ дожитья обоихъ супруговъ (x) и (y) до конца n-аго года, есть произведеніе дисконтированной стоимости его V^n на вѣроятность ${}_np_{xy}$ обоимъ супругамъ дожить до указаннаго срока. Если обозначить черезъ ${}_nE_{xy}$ искомую стоимость, то получимъ

$$(1) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad {}_nE_{xy} = {}_np_{xy} V^n;$$

по теоремѣ объ умноженіи вѣроятностей мы будемъ имѣть, очевидно,

$$(2) \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad {}_np_{xy} = \frac{l_{x+n}}{l_x} \cdot \frac{l_{y+n}}{l_y}.$$

§ 3. При разсмотрѣніи страхованія на двѣ жизни вводятъ обыкновенно слѣдующія коммутаціонныя числа (de Morgan)

$$D_{x,y}=V^{\frac{x+y}{2}}\, l_x \cdot l_y\,,$$

$$\sum_{(x+1,\ y+1)} D_{x,y}=N_{x,y}\,,$$

$$VD_{x,y}-D_{x+1,\ y+1}=V^{1+\frac{x+y}{2}}\,(l_x l_y-l_{x+1}l_{y+1})=C_{x,y}\,,$$

$$\sum_{(x,\ y)} C_{x,y}=M_{x,y}\,.$$

§ 4. Вводя въ разсмотрѣніе эти коммутаціонныя числа, мы получимъ по формуламъ (1) и (2) § 2

$${}_nE_{xy}=\frac{l_{x+n}l_{y+n}V^{n+\frac{x+y}{2}}}{l_x l_y V^{\frac{x+y}{2}}}=\frac{D_{x+n,\ y+n}}{D_{x,y}}\,.$$

§ 5. Разсмотримъ теперь современную стоимость a_{xy} ренты, уплачиваемой postnumerando въ размѣрѣ единицы капитала до тѣхъ поръ, пока оба супруга (x) и (y) живы; получаемъ

$$a_{xy}=p_{xy}\,V+{}_2p_{xy}\,V^2+\ .\ .\ .\ ={}_1E_{xy}+{}_2E_{xy}+\ .\ .\ .\ .$$

На основаніи формулы предыдущаго параграфа получаемъ

$$a_{xy}=\frac{D_{x+1,\ y+1}+D_{x+2,\ y+2}+\ .\ .\ .}{D_{x,y}}=\frac{N_{x,y}}{D_{x,y}}\,.$$

Очевидно, что рента praenumerando вычисляется по формуламъ:

$$\mathbf{a}_{xy}=a_{xy}+1\,,\ \ \mathbf{a}_{xy}=\frac{N_{x-1,\ y-1}}{D_{x,\ y}}\,.$$

Совершенно подобнымъ образомъ стоимость временной ренты postnumerando вычислится по формулѣ

$$|_n\,\mathbf{a}_{xy}=\frac{D_{x+1,\ y+1}+\ .\ .\ .\ +D_{x+n,\ y+n}}{D_{x,\ y}}=\frac{N_{x,\ y}-N_{x+n,\ x+n}}{D_{x,\ y}}\,,$$

для ренты praenumerando получаемъ

$$|_n\,\mathbf{a}_{xy}=\frac{N_{x-1,\ y-1}-N_{x+n-1,\ y+n-1}}{D_{x,\ y}}\,.$$

Подобнымъ же образомъ вычисляются отсроченныя ренты

$${}_m|\,a_{xy}\,,\ \ {}_m|\,\mathbf{a}_{xy}\,,\ \ {}_m|\,{}_{n-m}\,a_{xy}\,,\ \ {}_m|\,{}_{n-m}\,\mathbf{a}_{xy}\,.$$

§ 6. *Задача I.* Мужъ возраста x страхуетъ своей y-лѣтней женѣ пожизненную *вдовью пенсію* въ размѣрѣ 1 съ условіемъ, чтобы эта пенсія уплачивалась вдовѣ praenumerando съ начала года слѣдующаго за годомъ смерти мужа.

Обозначимъ современную стоимость этой пенсіи черезъ

$$a_{x \mid y};$$

мы должны будемъ для нахожденія искомой пенсіи разсуждать такъ: если бы уплата пенсіи для y-лѣтней жены началась тотчасъ, то мужъ, очевидно, долженъ былъ бы внести единовременную netto-премію въ размѣрѣ

(1) . . . . $\mathbf{a}_y$.

Но пенсія не уплачивается, пока оба супруга живы, слѣдовательно, изъ суммы $\mathbf{a}_y$ надо вычесть величину

$$\mathbf{a}_{xy}$$

netto-преміи по двойному страхованію до ближайшей смерти одного изъ супруговъ, и мы получаемъ

$$a_{x \mid y} = \mathbf{a}_y - \mathbf{a}_{xy}.$$

§ 7. *Задача II.* Отецъ въ возрастѣ x лѣтъ страхуетъ для своего y лѣтняго сына годовую *сиротскую пенсію* въ размѣрѣ 1 съ условіемъ чтобы эта пенсія выплачивалась сыну въ продолженіе n лѣтъ praenumerando начиная съ года слѣдующаго за годомъ смерти отца.

Очевидно, что эта задача отличается отъ предыдущей лишь тѣмъ, что вмѣсто пожизненной пенсіи разсматривается временная, и мы получаемъ для искомой единовременной netto-преміи выраженіе

$$|_n\mathbf{a}_y - |_n\mathbf{a}_{xy}.$$

§ 8. То, что сказано относительно двухъ жизней можетъ быть обобщено и на случай какого угодно числа жизней.

По мѣрѣ увеличенія числа жизней казуистика усложняется.

Пусть разсматривается группа $[x_1, x_2, \ldots . x_m]$ изъ m индивидуумовъ возрастовъ $x_1, x_2, \ldots . x_m$. Приходится отличать слѣдующія вѣроятности:

Знакомъ

$$_np_{x_1x_2 \ldots x_m}$$

обозначается вѣроятность того, что черезъ n лѣтъ *всѣ лица*, составляющія группу, *еще живы*.

Знакомъ

$${}_nq_{x_1x_2\ldots x_m}$$

обозначается вѣроятность того, что черезъ n лѣтъ вся группа *вымерла*.

Могутъ быть случаи промежуточные, а именно, черезъ n лѣтъ *остались въ живыхъ* только r лицъ группы, вѣроятность чего обозначается знакомъ

$${}_np\frac{[r]}{x_1x_2\ldots x_m} = P_r.$$

Формула

$$(1) \ldots {}_np_{x_1x_2\ldots x_r}\,{}_nq_{x_{r+1}x_{r+2}\ldots x_m} =$$

$$= {}_np_{x_1x_2\ldots x_r}(1 - {}_np_{x_{r+1}})(1 - {}_np_{x_{r+2}}) \ldots (1 - {}_np_{x_m})$$

даетъ вѣроятность того, что *остались въ живыхъ* черезъ n лѣтъ r *опредѣленныхъ лицъ*, а именно

$$(x_1), (x_2), \ldots (x_r),$$

остальные же умерли.

Просуммировавъ

$$\sum {}_np_{x_1x_2\ldots x_r}(1 - {}_np_{x_{r+1}})(1 - {}_np_{x_{r+2}}) \ldots (1 - {}_np_{x_m})$$

формулу (1) на всѣ возможныя комбинаціи изъ m лицъ группы по r, получаемъ вѣроятность P_r дожитія r лицъ безъ указанія, которыхъ именно.

§ 9. Употребляется одно изъ двухъ обозначеній

$$D_{x_1x_2\ldots x_m} = V^{\frac{x_1+x_2+\ldots+x_m}{m}} l_{x_1} l_{x_2} \ldots l_{x_m},$$

$$D_{x_1x_2\ldots x_m} = V^{\xi} l_{x_1} l_{x_2} \ldots l_{x_m},$$

гдѣ ξ наибольшее изъ чиселъ $x_1\, x_2 \ldots x_m$.

§ 10. Въ таблицѣ смертности приведенный въ концѣ книги даны таблицы стоимости рентъ postnumerando

$$a_x,\ a_{xx},\ a_{xxx}$$

для одной жизни, двухъ и трехъ жизней.

XVI.

Государственное страхованіе.

§ 1. Подъ государственнымъ страхованіемъ разумѣется такое, при которомъ въ качествѣ страховщика выступаетъ государство.

Государственное страхованіе можно подраздѣлить на два главныхъ типа: 1) *страхованіе въ собственномъ смыслѣ слова* и 2) *соціальное страхованіе.*

Государственное страхованіе въ собственномъ смыслѣ слова основывается на принципѣ математической безобидности, а потому къ нему примѣнимы всѣ наши вышеизложенныя теоретическія соображенія.

Соціальное страхованіе характеризуется отступленіями отъ принципа строгой математической безобидности.

Принципы, которые лежатъ въ основѣ соціальнаго страхованія, восходятъ къ XVII столѣтію. Buffon, авторъ извѣстнаго сочиненія „Essaid'arithmétique morale", а также Daniel Bernoulli находили правило математической безобидности игръ несправедливымъ съ общежитейской точки зрѣнія и предполагали давать нѣкоторыя преимущества болѣе бѣднымъ игрокамъ. При этомъ вмѣсто *математическаго* ожиданія выгоды они предлагали ввести такъ называемое *нравственное* ожиданіе этой выгоды, при опредѣленіи котораго входило бы въ разсмотрѣніе имущество, которымъ обладаетъ игрокъ, а также могли играть роль и другіе факторы взаимныхъ отношеній между игроками.

§ 2. Остановимся сначала на государственномъ страхованіи въ собственномъ смыслѣ слова. Примѣромъ такого государственнаго страхованія могутъ служить производимыя въ Россіи на основаніи закона 1904 года операціи страхованія въ государственныхъ сберегательныхъ кассахъ.

Въ пользу желательности развитія государственнаго страхованія говорятъ слѣдующіе доводы. Во первыхъ, государство можетъ удешевить страхованіе и сдѣлать его болѣе доступнымъ для всѣхъ классовъ населенія, ибо оно можетъ удовольствоваться самыми незначительными надбавками къ netto-преміямъ. Второй доводъ состоитъ въ фактѣ скопленія въ страховыхъ учрежденіяхъ всего міра колоссальныхъ капиталовъ, которые въ рукахъ государства могли бы быть употреблены на культурныя цѣли.

Защитники частныхъ страховыхъ учрежденій указываютъ на то обстоятельство, что частные предприниматели заинтересованные матеріально въ развитіи страхого дѣла, лучше будутъ способствовать его прогрессу. Они говорятъ, что удешевленіе страхованія можетъ явиться само собой, какъ результатъ свободной конкурренціи.

По развитію государственнаго страхованія всѣхъ типовъ стоитъ впереди Германія, гдѣ болѣе половины всего населенія является кліентами этого страхованія.

Въ послѣднее время представляетъ большой какъ теоретическій, такъ и практическій интересъ законопроектъ, внесенный въ законодательныя учрежденія Италіи, о введеніи государственной монополіи страхового дѣла съ обязательнымъ прекращеніемъ на всей территоріи государства дѣятельности частныхъ страховыхъ учрежденій.

§ 3. Къ типу соціальнаго страхованія принадлежитъ страхованіе рабочихъ. Жизнь показала, что обычныя формы безобиднаго страхованія не достаточны для удовлетворительной постановки обезпеченія рабочаго класса и удовлетворенія главной его потребности, а именно, матеріальной поддержки въ случаѣ утраты трудоспособности. Условія жизни рабочихъ даютъ возможность только немногимъ изъ нихъ обезпечить себѣ эту поддержку по собственной инициативѣ и за счетъ своихъ сбереженій. Государства многихъ странъ признали за собой нравственную обязанность и экономическую необходимость пойти на встрѣчу потребностямъ рабочаго класса. Было признано необходимымъ извѣстное отступленіе отъ принципа безобидности въ сторону принудительнаго участія работодателей въ расходахъ по обезпеченію рабочихъ.

§ 4. Страхованіе трудоспособности производится въ слѣдующихъ формахъ:

1) страхованіе на случай болѣзни,

2) страхованіе отъ несчастныхъ случаевъ,

3) страхованіе на старость,

4) страхованіе на случай утраты трудоспособности.

Страхованіе пенсій на старость не нуждается съ математической точки зрѣнія въ добавочныхъ разъясненіяхъ. Это есть обыкновенное страхованіе пожизненной ренты, отсроченной до извѣстнаго возраста.

Подобнымъ же образомъ очень элементарно представляются съ математической точки зрѣнія вопросы страхованія на случай болѣзни или отъ несчастныхъ случаевъ.

Тутъ обыкновенно дѣло идетъ лишь о раскладкѣ между участниками расходовъ ихъ уплатъ.

Обычной организаціей является учрежденіе обществъ взаимнаго страхованія владѣльцевъ промышленныхъ предпріятій отъ убытковъ, которые на нихъ упадутъ, если несчастный случай въ предпріятіи повлечетъ за собой смерть или увѣчіе рабочаго.

Обязательства такихъ обществъ состоятъ въ уплатѣ пенсій или утратившимъ трудоспособность рабочимъ, или семьямъ умершихъ отъ несчастныхъ случаевъ.

Математическій интересъ представляетъ до извѣстной степени способъ раскладки расходовъ взаимнаго товарищества предпринимателей между его участниками. Существуютъ три способа раскладки:

1) капитализаціонный,

2) раскладочный,

3) способъ среднихъ премій.

По капитализаціонной системѣ подлежатъ раскладкѣ *капитализированныя современныя стоимости* назначенныхъ въ теченіе года пенсій (немедленныхъ для увѣчныхъ, пожизненныхъ для вдовъ и временныхъ для сиротъ). По раскладочной системѣ подлежатъ раскладкѣ только уплаты одиночныхъ пенсій, причитающихся къ году раскладки, независимо отъ того, когда назначена пенсія и сколько разъ она была раньше выплачена.

Способъ среднихъ премій основанъ на гипотетической потребности, введеной статистическимъ подсчетомъ. Такимъ образомъ установливается нѣкоторая гипотетически опредѣленная ежегодная сумма расходовъ. При этой системѣ, конечно, необходимо провѣрять періодически соотвѣтствуетъ ли гипотеза фактамъ дѣйствительности.

§ 5. Переходя къ вопросамъ страхованія утраты трудоспособностей, или *инвалидности*, мы должны подчеркнуть большую разницу между этимъ страхованіемъ и страхованіемъ отъ несчастныхъ случаевъ.

Въ страхованіи отъ несчастныхъ случаевъ размѣръ пенсіи независитъ отъ возраста рабочаго и отъ числа лѣтъ его службы, причемъ самъ рабочій не участвуетъ въ образованіи суммъ, обезпечивающихъ выплату пенсіи.

При страхованіи на случай инвалидности размѣръ пенсіи существеннымъ образомъ зависитъ отъ взносовъ самого рабочаго, размѣръ которыхъ въ свою очередь зависитъ отъ его заработка. въ этомъ страхованіи обыкновенно участвуетъ государство, приплаты котораго увеличиваютъ размѣръ пенсій. Иногда участвуютъ въ при платахъ также предприниматели.

Вопросы страхованія инвалидности требуютъ кромѣ обыкновенной таблицы смертности еще таблицы вѣроятностей потерять трудоспособность.

Два фактора смерть и инвалидность, въ различныхъ комбинаціяхъ ихъ проявленій значительно усложняютъ математическую теорію разсчетовъ.

Для желающихъ познакомится съ теоріей страхованія трудоспособности можно рекомендовать на русскомъ языкѣ книгу *С. Е. Савичъ*: „Элементарная теорія страхованія жизни и трудоспособности". Наиболѣе извѣстными являются слѣдующія таблицы инвалидности: 1) *Zeuner'a* (1869), составленныя по наблюденіямъ надъ рабочими фрейбергскаго горнаго округа въ Саксоніи, 2) *Zimmermann'a* (1886), относящіяся къ служащимъ на дорогахъ германскаго желѣзно дорожнаго союза, 3) *Малешевскаго* (1894), по наблюденіямъ также надъ желѣзнодорожными служащими.

§ 6. Крайній предѣлъ отступленія отъ принципа математической безобидности представляютъ такъ называемыя *государственныя пенсіи*, выплачиваемыя на старости изъ суммъ государственнаго казначейства независимо отъ положенія и безъ какихъ либо предварительныхъ взносовъ пенсіонера.

Эти пенсіи установлены въ 1891 году въ Даніи, въ 1898 въ Новой Зеландіи, въ 1908 въ Австраліи и въ 1909 году въ Англіи.

ТАБЛИЦА

смертности H^M двадцати Британскихъ Страховыхъ Обществъ.

Таблица смертности H^M двадцати

x	l_x	d_x	μ_x	D_x	N_x	S_x
0	127 283	14 358	0,15920	127 283	2 553 055	52 129 621
1	112 925	3 962	,07901	109 110	2 425 772	49 703 849
2	108 963	2 375	,02366	101 720	2 316 662	47 278 077
3	106 588	1 646	,01787	96 137	2 214 942	44 961 415
4	104 942	1 325	,01379	91 451	2 118 805	42 746 473
5	103 617	1 061	,01142	87 243	2 027 354	40 627 668
6	102 556	852	,00925	83 430	1 940 111	38 600 314
7	101 704	683	,00748	79 939	1 856 681	36 660 203
8	101 021	557	,00607	76 717	1 776 742	34 803 522
9	100 464	464	,00502	73 714	1 700 025	33 026 780
10	100 000	408	,00428	70 892	1 626 311	31 326 755
11	99 592	369	,00388	68 215	1 555 419	29 700 444
12	99 223	346	,00359	65 664	1 487 204	28 145 025
13	98 877	337	,00342	63 224	1 421 540	26 657 821
14	98 540	337	,00340	60 877	1 358 316	25 236 281
15	98 203	360	,00353	58 615	1 297 439	23 877 965
16	97 843	384	,00378	56 426	1 238 824	22 580 526
17	97 459	425	,00414	54 304	1 182 398	21 341 702
18	97 034	465	,00458	52 238	1 128 094	20 159 304
19	96 569	508	,00504	50 231	1 075 856	19 031 210
20	96 061	548	,00550	48 277	1 025 625	17 955 354
21	95 513	582	,00592	46 378	977 348	16 929 729
22	94 931	609	,00629	44 537	930 970	15 952 381
23	94 322	631	,00659	42 754	886 433	15 021 411
24	93 691	647	,00682	41 033	843 679	14 134 978
25	93 044	658	,00701	39 371	802 646	13 291 299
26	92 386	664	,00716	37 771	763 275	12 488 653
27	91 722	673	,00729	36 231	725 504	11 725 378
28	91 049	678	,00742	34 750	689 273	10 999 874
29	90 371	686	,00755	33 324	654 523	10 310 601
30	89 685	691	,00768	31 953	621 199	9 656 078
31	88 994	700	,00782	30 634	589 246	9 034 879
32	88 294	709	,00798	29 366	558 612	8 445 633
33	87 585	719	,00815	28 145	529 246	7 887 021
34	86 866	729	,00833	26 970	501 101	7 357 775
35	86 137	742	,00854	25 839	474 131	6 856 674
36	85 395	756	,00876	24 750	448 292	6 382 543
37	84 639	770	,00901	23 702	423 542	5 934 251
38	83 869	786	,00928	22 692	399 840	5 510 709
39	83 083	806	,00957	21 719	377 148	5 110 869
40	82 277	823	,00990	20 781	355 429	4 733 721
41	81 454	846	,01025	19 877	334 648	4 378 292
42	80 608	871	,01064	19 006	314 771	4 043 644
43	79 737	895	,01106	18 165	295 765	3 728 873
44	78 842	924	,01153	17 353	277 600	3 433 108
45	77 918	954	,01204	16 570	260 247	3 155 508
46	76 964	986	,01260	15 814	243 677	2 895 261
47	75 978	1 021	,01321	15 083	227 863	2 651 584
48	74 957	1 061	,01388	14 377	212 780	2 423 721
49	73 896	1 101	,01462	13 694	198 403	2 210 941

Британскихъ Страховыхъ Обществъ.

C_x	M_x	R_x	a_x	a_{xx}	a_{xxx}	A_x	x
13 872	40 948	785 908	20,058	15,079	11,633	0,32171	0
3 698,5	27 075,5	744 959,6	22,233	18,513	15,760	,24861	1
2 142,1	23 377,0	717 884,1	22,775	19,467	17,004	,22982	2
1 434,4	21 234,9	694 507,1	23,039	19,974	17,696	,22088	3
1 115,6	19 800,5	673 272,2	23,169	20,260	18,107	,21652	4
863,14	18 684,94	653 471,69	23,238	20,447	18,393	,21417	5
669,67	17 821,80	634 786,75	23,255	20,546	18,567	,21362	6
518,68	17 152,13	616 964,95	23,225	20,570	18,642	,21456	7
408,70	16 633,45	599 812,82	23,160	20,530	18,633	,21681	8
328,94	16 224,75	583 179,37	23,062	20,439	18,555	,22011	9
279,46	15 895,81	566 954,62	22,940	20,307	18,424	,22423	10
244,20	15 616,35	551 058,81	22,802	20,146	18,257	,22892	11
221,24	15 372,15	535 442,46	22,648	19,964	18,060	,23410	12
208,19	15 150,91	520 070,31	22,484	19.765	17,843	,23964	13
201,15	14 942,72	504 919,40	22,312	19,555	17,612	,24547	14
207,61	14 741,57	489 976,68	22,134	19,337	17,372	,25151	15
213,96	14 533,96	475 235,11	21,955	19,118	17,132	,25758	16
228,80	14 320,00	460 701,15	21,774	18,901	16,895	,26370	17
241,87	14 091,20	446 381,15	21,596	18,690	16,669	,26974	18
255,30	13 849,33	432 289,95	21,418	18,486	16,452	,27571	19
266,09	13 594,03	418 440,62	21,245	18,289	16,248	,28159	20
273,04	13 327,94	404 846,59	21,074	18,101	16,055	,28738	21
276,05	13 054,90	391 518,65	20,903	17,917	15,870	,29313	22
276,35	12 778,85	378 463,75	20,733	17,736	15,691	,29890	23
273,77	12 502,50	365 684,90	20,561	17,555	15,514	,30471	24
269,02	12 228,73	353 182,40	20,387	17,374	15,337	,31061	25
262,29	11 959,71	340 953,67	20,208	17,189	15,159	,31664	26
256,86	11 697,42	328 993,96	20,024	17,000	14,976	,32284	27
250,01	11 440,56	317 296,54	19,835	16,805	14,787	,32924	28
244,41	11 190,55	305 855,98	19,641	16,605	14,593	,33582	29
237,86	10 946,14	294 665,43	19,441	16,399	14,394	,34257	30
232,81	10 708,28	283 719,29	19,235	16,187	14,189	,34954	31
227,83	10 475,47	273 011,01	19,023	15,968	13,978	,35671	32
223,23	10 247,64	262 535,54	18,804	15,744	13,761	,36412	33
218,69	10 024,41	252 287,90	18,580	15,513	13,558	,37167	34
215,06	9 805,72	242 263,49	18,349	15,277	13,309	,37949	35
211,70	9 590,66	232 457,77	18,112	15,035	13,075	,38750	36
208,33	9 378,96	222 867,11	17,870	14,787	12,836	,39571	37
205,47	9 170,63	213 488,15	17,620	14,532	12,590	,40414	38
203,58	8 965,16	204 317,52	17,365	14,272	12,339	,41278	39
200,84	8 761,58	195 352,36	17,103	14,007	12,084	,42161	40
199,47	8 560,74	186 590,78	16,835	13,736	11,824	,43068	41
198,42	8 361,27	178 030,04	16,562	13,459	11,559	,43993	42
196,99	8 162,85	169 668,77	16,282	13,179	11,291	,44938	43
196,49	7 965,86	161 505,92	15,997	12,893	11,018	,45904	44
196,02	7 769,37	153 540,06	15,706	12,603	10,742	,46889	45
195,74	7 573,35	145 770,69	15,409	12,308	10,462	,47892	46
195,83	7 377,61	138 197,34	15,107	12,010	10,180	,48914	47
196,63	7 181,78	130 819,73	14,800	11,709	9,895	,49953	48
197,14	6 985,15	123 637,95	14,488	11,404	9,608	,51008	49

x	l_x	d_x	μ_x	D_x	N_x	S_x
50	72 795	1 144	0,01542	13 034	184 709	2 012 538
51	71 651	1 193	,01631	12 395	171 675	1 827 829
52	70 458	1 243	,01727	11 777	159 280	1 656 154
53	69 215	1 296	,01833	11 178	147 503	1 496 874
54	67 919	1 353	,01950	10 598	136 325	1 349 371
55	66 566	1 414	,02077	10 035	125 727	1 213 046
56	65 152	1 475	,02216	9 490,1	115 691,8	1 087 319,3
57	63 677	1 541	,02369	8 961,5	106 201,7	971 627,5
58	62 136	1 612	,02536	8 448,9	97 240,2	865 425,8
59	60 524	1 682	,02719	7 951,5	88 791,3	768 185,6
60	58 842	1 755	,02920	7 469,1	80 839,8	679 394,3
61	57 087	1 830	,03140	7 001,3	73 370,7	598 554,5
62	55 257	1 906	,03381	6 547,7	66 369,4	525 183,8
63	53 351	1 983	,03645	6 108,0	59 821,7	458 814,4
64	51 368	2 059	,03934	5 682,1	53 713,7	398 992,7
65	49 309	2 133	,04251	5 270,0	48 031,6	345 279,0
66	47 176	2 204	,04599	4 871,5	42 671,6	297 247,4
67	44 972	2 273	,04979	4 486,8	37 890,1	254 485,8
68	42 699	2 334	,05396	4 116,0	33 403,3	216 595,7
69	40 365	2 388	,05853	3 759,5	29 287,3	183 192,4
70	37 977	2 434	,06353	3 417,4	25 527,8	153 905,1
71	35 543	2 468	,06901	3 090,2	22 110,4	128 377,3
72	33 075	2 490	,07502	2 778,4	19 020,2	106 266,9
73	30 585	2 496	,08160	2 482,3	16 241,8	87 246,7
74	28 089	2 487	,08881	2 202,7	13 759,5	71 004,9
75	25 602	2 459	,09671	1 939,7	11 556,8	57 245,4
76	23 143	2 412	,10536	1 694,1	9 617,1	45 688,6
77	20 731	2 343	,11485	1 466,3	7 923,0	36 071,5
78	18 388	2 255	,12523	1 256,6	6 456,7	28 148,5
79	16 133	2 146	,13662	1 065,2	5 200,1	21 691,8
80	13 987	2 018	,14909	892,26	4 134,89	16 491,72
81	11 969	1 873	,16275	737,72	3 242,63	12 356,83
82	10 096	1 712	,17772	601,23	2 504,91	9 114,20
83	8 384	1 540	,19412	482,39	1 903,68	6 609,29
84	6 844	1 361	,21209	380,47	1 421,29	4 705,61
85	5 483	1 180	,23177	294,50	1 040,82	3 284,32
86	4 303	1 002	,25343	223,31	746,32	2 243,50
87	3 301	830	,27697	165,52	523,01	1 497,18
88	2 471	671	,30286	119,71	357,49	974,17
89	1 800	527	,33123	84,252	237,784	616,80
90	1 273	402	,36230	57,571	153,532	378,896
91	871	296	,39635	38,058	95,961	225,364
92	575	209	,43366	24,275	57,903	129,403
93	366	144	,47453	14,929	33,628	71,500
94	222	93	,51930	8,749	18,699	37,872
95	129	58	,56830	4,912	9,950	19,173
96	71	34	,62211	2,612	5,038	9,223
97	37	18	,68100	1,315	2,426	4,185
98	19	10	,74552	0,653	1,111	1,759
99	9	5	,81621	,296	0,458	0,648
100	4	3	,89366	,128	,159	,190
101	1	1	,97851	,031	,031	,031
102	0	...	∞	...	...	...

C_x	M_x	R_x	a_x	a_{xx}	a_{xxx}	A_x	x
197,91	6 788,01	116 652,80	14,172	11,096	9,320	0,52079	50
199,41	6 590,10	109 864,79	13,850	10,785	9,030	,53166	51
200,74	6 390,69	103 274,69	13,524	10,474	8,740	,54265	52
202,22	6 189,95	96 884,00	13,196	10,161	8,451	,55377	53
203,98	5 987,73	90 694,05	12,864	9,847	8,161	,56500	54
205,96	5 783,75	84 706,32	12,528	9,532	7,873	,57634	55
207,58	5 577,79	78 922,57	12,191	9,218	7,587	,58775	56
209,54	5 370,21	73 344,78	11,851	8,904	7,303	,59925	57
211,78	5 160,67	67 974,57	11,509	8,592	7,021	,61082	58
213,51	4 948,89	62 813,90	11,167	8,282	6,742	,62239	59
215,24	4 735,38	57 865,01	10,823	7,973	6,468	,63400	60
216,85	4 520,14	53 129,63	10,480	7,668	6,197	,64561	61
218,21	4 303,29	48 609,49	10,136	7,366	5,931	,65722	62
219,35	4 085,08	44 306,20	9,794	7,068	5,671	,66881	63
220,06	3 865,73	40 221,12	9,453	6,775	5,416	,68033	64
220,26	3 645,67	36 355,39	9,114	6,486	5,167	,69178	65
219,89	3 425,41	32 709,72	8,778	6,204	4,925	,70315	66
219,11	3 205,52	29 284,31	8,445	5,927	4,690	,71443	67
217,38	2 986,41	26 078,79	8,115	5,656	4,462	,72555	68
214,89	2 769,03	23 092,38	7,790	5,393	4,241	,73653	69
211,62	2 554,14	20 323,35	7,470	5,136	4,028	,74738	70
207,32	2 342,52	17 769,21	7,155	4,887	3,823	,75804	71
202,09	2 135,20	15 426,69	6,846	4,646	3,626	,76851	72
195,73	1 933,11	13 291,49	6,543	4,413	3,437	,77874	73
188,43	1 737,38	11 358,38	6,247	4,189	3,256	,78877	74
180,01	1 548,95	9 621,00	5,958	3,973	3,083	,79856	75
170,60	1 368,94	8 072,05	5,677	3,765	2,919	,80802	76
160,11	1 198,34	6 703,11	5,404	3,567	2,764	,81726	77
148,89	1 038,23	5 504,76	5,138	3,377	2,616	,82623	78
136,90	889,34	4 466,54	4,882	3,195	2,476	,83491	79
124,38	752,44	3 577,20	4,634	3,023	2,344	,84330	80
111,54	628,06	2 824,76	4,395	2,859	2,220	,85135	81
98,503	516,521	2 196,695	4,166	2,705	2,104	,85911	82
85,611	418,018	1 680,174	3,946	2,558	1,995	,86656	83
73,102	332,407	1 262,156	3,736	2,420	1,894	,87368	84
61,236	259,305	929,749	3,534	2,290	1,799	,88050	85
50,241	198,069	670,444	3,342	2,168	1,711	,88699	86
40,210	147,828	472,375	3,160	2,055	1,630	,89314	87
31,407	107,618	324,547	2,986	1,948	1,555	,89902	88
23,833	76,211	216,929	2,822	1,849	1,486	,90457	89
17,565	52,378	140,718	2,667	1,756	1,423	,90981	90
12,496	34,813	88,340	2,521	1,671	1,366	,91472	91
8,5251	22,3170	53,5272	2,385	1,594	1,315	,91936	92
5,6751	13,7919	31,2102	2,253	1,517	1,264	,92385	93
3,5412	8,1168	17,4183	2,137	1,455	1,225	,92773	94
2,1338	4,5756	9,3015	2,106	1,394	1,186	,93152	95
1,2086	2,4418	4,7259	1,929	1,347	1,157	,93480	96
0,6183	1,2332	2,2841	1,844	1,321	1,145	,93665	97
,3317	0,6149	1,0509	1,702	1,261	1,112	,94248	98
,1603	,2832	0,4360	1,533	1,202	1,086	,94816	99
,0930	,1229	,1528	1,242	1,060	1,015	,95803	100
,0299	,0299	,0299	...	...	...	,96618	101
...	...	...				...	102

ОПЕЧАТКИ.

		Напечатано:	Должно быть:
Стр. 13	7 снизу	частыя	частныя
Стр. 29	5 сверху	v	V
Стр. 32	1 снизу	Виндоршкъ	Вигдорчикъ
Стр. 35	9 сверху	$V^{-1}\sum_{i=1}^{i=n} a_i \pi_i V^i$	$V^{-1}\sum_{i=2}^{i=n} b_i \rho_i V^i$
Стр. 36	6 сверху	v	V
Стр. 37	1 и 2 снизу	${}_nq_x$	${}_{n\mid}q_x$
Стр. 40	5 сверху	$V^{-\frac{1}{2}}(1+i)^{\frac{1}{2}}$	$V^{-\frac{1}{2}}=(1+i)^{\frac{1}{2}}$

Изданія Кіевскаго Коммерческаго Института:

В. Г. Бажаевъ. Къ вопросу о законахъ аграрной эволюціи.	— р.	15 коп.
И. В. Егоровъ. Техническій анализъ. Кіевъ 1909 г.	2 „	— „
Кэтлэ. Соціальная физика. Т. I. (Переводъ студентовъ Института). Цѣна для студент. 1 р. 50 к., для постороннихъ.	2 „	— „
Труды Общества экономистовъ при Кіевскомъ Коммерческомъ Институтѣ. Кіевъ. Вып. I. 1910г.	—	50 „
Труды Общества Экономистовъ. Кіевъ. Вып. II. 1910 . . .	1 „	— „
И. В. Егоровъ. Объ окиси декаметиленгликоля.	—	10 „
Записка о Кіевскомъ Коммерческомъ Институтѣ изд. 1909 г.	—	15 „
Записка о Кіевскомъ Коммерческомъ Институтѣ изд. 1910 г. .	—	25 „
Отчетъ о музеѣ товаровѣдѣнія при Кіевскомъ Коммерческомъ Институтѣ. Кіевъ 1910 г.	—	25 „
Музей товаровѣдѣнія, 3-й выпускъ, изданія 1911 г.	1 „	25 „
Kurzer Bericht über das Waarenkundemuseum, изд. 1911 г. . .	—	20 „
Notice sur le Musée de Marchandises, изд. 1911 г.	—	20 „
Обозрѣніе преподаванія на 1912—1913 академическій годъ въ Кіевскомъ Коммерческомъ Институтѣ. Кіевъ 1912 г.	—	20 „
М. В. Довнаръ-Запольскій. На зарѣ крестьянской свободы. Кіевъ. 1911	1 „	— „
Его-же. Обзоръ новѣйшей русской исторіи т. I. Кіевъ 1912 г.	2 „	50 „
Д. А. Граве. Энциклопедія математики. Кіевъ 1912 г. . .	3 „	50 „
Е. Е. Слуцкій. Теорія корреляціи и элементы ученія о кривыхъ распредѣленія, Кіевъ. 1912 г.	1 „	25 „

Означенныя книги продаются у кассира Института; у него же продаются:

М. В. Довнаръ-Запольскій. Изъ исторіи общественныхъ теченій въ Россіи, изд. 2-ое. Кіевъ 1910 г.	1 „	50 „
Его-же. Русская Исторія т. I, изд. 2-ое	2 „	— „
т. II	2 „	— „
Его-же. Изслѣдованія (этнографія и соціологія, обычное право и статистика). Кіевъ 1909 г.	3 р.	— „
А. А. Русовъ. О желѣзнодорожной статистикѣ	—	20 „

„Извѣстія Кіевскаго Коммерческаго Института“

выходятъ 4—6 разъ въ годъ, по мѣрѣ накопленія матеріала въ редакціи. Въ „Извѣстіяхъ“, кромѣ оффиціальныхъ свѣдѣній о дѣятельности Института и состоящихъ при немъ учрежденій, помѣщаются научные труды преподавателей Института, работы слушателей, одобренныя къ напечатанію Учебнымъ Комитетомъ, а также труды „Общества Экономистовъ“, состоящаго при К. К. Институтѣ.

Подписная цѣна: 3 руб. въ годъ, съ пересылкой **3** руб. **50** коп. Цѣна для слушателей Института (безъ доставки) **2** руб.

Отдѣльная книжка **75** коп., для слушателей **50** коп.

Адресъ Редакціи: Кіевъ, Бибиковскій Бульваръ № 21, Коммерческій Институтъ.

Редакторъ **А. А. Русовъ.**

Типографія И. И. Чоколова, Кіевъ, Б.-Житомірская № 20, соб. д.