від Е. В. Чекотовського

Р. Оржецкій.

Учебникъ
МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ.

Юридическій
книжный складъ и книгоиздательство „ПРАВО"
С.-ПЕТЕРБУРГЪ,
Литейный просп., № 28.

1914.

Типографія Губернскаго Правленія.

Изложеніе математической статистики имѣется въ курсахъ Ю. Э. Янсона «Теорія статистики» (5 изданіе 1912); А. Н. Анциферова «Курсъ элементарной статистики» (2 изданіе 1911); А. А. Кауфмана «Теорія и методы статистики» (1912); этой части методы посвящены спеціальныя работы: брошюра В. А. Косинскаго «о научной обработкѣ статистическихъ данныхъ», 1889; цѣнный трудъ А. А. Чупрова «Очерки по теоріи статистики» (2 изд. 1913); изслѣдованіе Е. Е. Слуцкаго «Теорія корреляціи», 1912; сводная работа А. Леонтовича «Элементарное пособіе къ примѣненію методовъ Gauss'a и Pearson'a» въ 3 частяхъ, 1909—1911, съ таблицами. Практическое примѣненіе и иллюстрацію пріемовъ на примѣрахъ можно найти въ работахъ М. Б. Гуревича «Примѣненіе нѣкоторыхъ пріемовъ математической статистики», 1912 и «Методъ Оцѣноки городскихъ недвижимыхъ имуществъ», 1913.

Въ иностранной литературѣ пріемы математической статистики въ общихъ курсахъ изложены у Bowley, Elements of Statistics, 1907; Udny Yule, An introduction to the Theory of Statistics, 1911; R. Benini, Principii di Statistica metodologica, 1906; G. Duncker, Die Methode der Variationsstatistik, 1899; A. Gabaglio, Teoria generale della Statistica, 1888; въ справочномъ сборникѣ C. B. Davenport, Statistical Methods, 1904; методы Пирсона въ частности у W. P. Elderton, Frequency Curves and Correlation. Часть работъ W. Lexis'a, имѣвшихъ столь важное значеніе въ развитіи математическаго метода въ статистикѣ, собрана въ книгѣ «Abhandlungen zur Theorie der Bevölkerungs-und Moralstatistik», 1903; многочисленныя цѣнныя работы проф. Борткевича (Bortkiewicz) разсѣяны въ разныхъ экономическихъ, статистическихъ и спеціальныхъ журналахъ и изданіяхъ; здѣсь можно упомянуть лишь объ особо изданномъ изслѣдованіи «Das Gesetz der kleinen Zahlen» von Dr. L. von Bortkewitsch, 1898, и статьѣ: «Kritische Betrachtungen zur theoretischen Statistik, въ журналѣ Jahrbücher für Nationalökonomie und Statistik, 1894, 1895, 1896, III Folge, Band 8, 10, 11. Основныя работы Pearson'a Contributions to the Mathematical Theory of Evolution помѣщены въ Philosophical Transactions of the Royal Society of London A, т. 186 и сл., и въ Draper's Company Research Memoirs; изслѣдованія F. G. Edgeworth въ журналѣ Journal of the Royal Statistical Society. Вопросамъ математи ческой статистики и ея примѣненію посвященъ журналъ Biometrica. Подробныя библіографическія указанія у А. Леонтовича (цит. соч.).

Оглавленіе.

1. Общія понятія.

Статистика, какъ методологическая наука.

Подъ статистикой многіе понимаютъ какъ систематическое изложеніе фактовъ и явленій статистическаго характера и результатовъ ихъ изученія, такъ и особый статистическій способъ наблюденія и изслѣдованія; иначе говоря, статистику считаютъ, съ одной стороны, матеріальной наукой, описывающей извѣстнаго рода явленія и излагающей ихъ законы, и, съ другой, методологической дисциплиной, имѣющей своимъ предметомъ методъ изслѣдованія. Въ качествѣ науки матеріальной, статистика должна была бы имѣть свой особый предметъ изслѣдованія. Однако данныя, получаемыя статистическимъ путемъ, относятся къ явленіямъ, которыя входятъ въ область другихъ наукъ—зоологіи, ботаники, антропологіи, біологіи, политической экономіи, социологіи. Естественно, что и разработка такихъ данныхъ требуетъ соотвѣтствующихъ специальныхъ познаній. Въ тѣхъ случаяхъ, когда статистическіе матеріалы собираются для практическихъ надобностей, использованіе ихъ предполагаетъ специальную освѣдомленность въ прикладныхъ вопросахъ. Отсюда ясно, что статистики, какъ особой матеріальной науки, не существуетъ. Правда, статистикъ—практикъ, владѣющій техникой собираніи статистическихъ матеріаловъ, или теоретикъ, знакомый съ логическими основами статистическаго метода, нерѣдко заимствуютъ изъ разныхъ областей свѣдѣнія, дающія имъ возможность разобраться въ матеріальной природѣ явленій, надъ которыми они оперируютъ; это обстоятельство можетъ сдѣлать такого статистика-практика или теоретика въ лучшемъ случаѣ энциклопедистомъ, въ худшемъ—диллетантомъ, но не можетъ сдѣлать изъ собранія разнородныхъ статистическихъ матеріаловъ предмета особой самостоятельной науки. Совершенно правы поэтому тѣ, кто отрицая существованіе и возможность статистики, какъ самостоятельной области знанія, въ смыслѣ матеріальной науки, признаютъ существованіе лишь статистическаго метода. Этимъ методомъ долженъ пользоваться всякій спеціалистъ, теоретикъ или практикъ—зоологъ, ботаникъ, антропологъ, экономистъ, врачъ, педагогъ, актуарій, поскольку онъ желаетъ распространить свои изслѣдованія на тѣ стороны явленій своей спеціальной области, которыя не поддаются обычному родовому способу изслѣдованія. Изученіе самого метода, напротивъ, представляетъ особую и самостоятельную задачу, не совпадающую съ задачей познанія природы и законовъ какой-либо спеціальной области явленій, и составляетъ предметъ особой и самостоятельной методологической науки—статистики. Статистика въ этомъ смыслѣ распадается на двѣ связанныя другъ съ другомъ дисциплины: практическую и теоретическую; первая имѣетъ своимъ предметомъ техническіе пріемы собиранія и подсчета статистическихъ матеріаловъ, вторая логическое обоснованіе статистическаго метода и вытекающіе изъ логики метода пріемы изслѣдованія матеріала. Эта вторая теоретическая задача составитъ предметъ дальнѣйшаго изложенія.

Предметъ статистическаго метода.

Необходимость особыхъ статистическихъ пріемовъ изученія вызывается существованіемъ своеобразныхъ свойствъ и правильностей явленій, не поддающихся изслѣдованію другими обычными способами. Въ этомъ отношеніи свойства и правильности явленій бываютъ двоякаго рода. Одни свойства таковы, что проявляются одинаково въ каждомъ единичномъ случаѣ опредѣленнаго рода; напримѣръ, вода при разложеніи даетъ всегда два атома водорода на одинъ кислорода. Свойства и правильности такого рода, будучи установлены для опредѣленнаго рода явленій или предметовъ, являются правиломъ или закономъ и для каждаго единичнаго случая того же рода. Такія свойства и правильности мы можемъ назвать родовыми. Въ другихъ же случаяхъ правильность обнаруживается только въ болѣе или менѣе значительной совокупности однородныхъ случаевъ, но незамѣтна или отсутствуетъ въ единичныхъ случаяхъ совокупности. Такъ, напримѣръ, продолжительность жизни отдѣльныхъ индивидовъ опредѣленнаго возраста, національности и профессіи бываетъ различна; но средняя продолжительность жизни значительной совокупности такихъ индивидовъ представляетъ собою величину довольно устойчивую; рожденія мальчиковъ и рожденія дѣвочекъ наступаютъ въ каждомъ отдѣльномъ случаѣ безъ видимого порядка; но совокупность рожденій, взятая за достаточно большой промежутокъ времени, обнаруживаетъ нѣкоторое правильное отношеніе числа рожденій дѣвочекъ къ числу рожденій мальчиковъ.

Сводные признаки и правильности.

Свойства и правильности, которыя обнаруживаются только въ совокупности случаевъ и не могутъ быть открыты въ единичныхъ случаяхъ этой совокупности, можно назвать сводными свойствами и закономѣрностями. Для изслѣдованія ихъ нужны особые пріемы, отличные отъ тѣхъ, какіе примѣняются къ изслѣдованію родовыхъ свойствъ и правильностей. Пріемы изученія сводныхъ признаковъ и сводныхъ закономѣрностей въ тѣхъ случаяхъ, когда эти признаки и закономѣрности могутъ быть выражены въ числовыхъ величинахъ, составляютъ содержаніе статистическаго метода *). Въ частности статистическій методъ состоитъ въ характеристикѣ предметовъ и яв-

*) Особой науки о методѣ изученія сводныхъ признаковъ и правильностей въ тѣхъ случаяхъ, когда они не могутъ быть выражены количественно, не существуетъ. Самый же способъ сводной характеристики несомнѣнно существуетъ; онъ примѣняется всякій разъ, когда изслѣдователь имѣетъ дѣло съ индивидуально

леній *сводными признаками*, т. е. такими признаками, которые относятся къ цѣлой совокупности предметовъ или явленій извѣстнаго рода и неприсущи единичнымъ случаямъ этой совокупности, какъ то : средній ростъ, средняя температура, процентное отношеніе рожденій къ числу населенія и т. п.; въ изученіи *сводной закономѣрности* явленій, т. е. такой закономѣрности, которая проявляется только при сведеніи ряда качественно или количественно отличныхъ другъ отъ друга единичныхъ случаевъ извѣстнаго рода въ болѣе или менѣе значительную совокупность, какъ, напримѣръ: наступленіе опредѣленнаго числа рожденій мальчиковъ и опредѣленнаго числа рожденій дѣвочекъ въ общей массѣ рожденій и т. п.; и, наконецъ, въ изученіи *сводной зависимости*, т. е. такой зависимости, когда правильность количественныхъ измѣненій извѣстнаго рода явленій обнаруживается не въ каждомъ единичномъ случаѣ, а лишь въ достаточно большой ихъ совокупности, какъ, напримѣръ, зависимость средней продолжительности жизни отъ возраста и т. п.

Условія примѣненія. Сводные признаки, закономѣрности и зависимости выражаются въ среднихъ величинахъ, отношеніяхъ и функціяхъ, т. е. имѣютъ количественное выраженіе, поэтому статистическій методъ примѣнимъ лишь въ тѣхъ случаяхъ, когда 1) явленія имѣютъ количественный характеръ и выражаются въ мѣрѣ, величинѣ или числѣ (ростъ, возрастъ, число лицъ въ семьѣ и т. п.); 2) когда качественно различныя явленія могутъ быть сосчитаны (число мужчинъ и женщинъ въ населеніи) или 3) когда качественныя различія поддаются какому-либо количественному сопоставленію (расположеніе учениковъ въ рядъ по ихъ умственному развитію), или мо-

---

различными явленіями, смыслъ которыхъ раскрывается въ ихъ общей совокупности, какъ это бываетъ при характеристикѣ литературныхъ направленій, школъ, историческихъ эпохъ, общественныхъ настроеній и т. п.

гутъ быть условно выражены въ количественномъ видѣ (оцѣнка способностей посредствомъ отмѣтокъ и т. п.).

Область примѣненія статистическаго метода опредѣляется такимъ образомъ двумя условіями: во 1-хъ, своднымъ характеромъ признака, закономѣрности или зависимости и, во 2-хъ, возможностью измѣренія или учета явленій и ихъ признаковъ. Всѣ другія особенности явленій не имѣютъ значенія; поэтому статистическій методъ примѣнимъ одинаково какъ къ явленіямъ механическимъ, такъ и къ явленіямъ органической жизни—физической, психической и соціальной.

Совокупность. Основнымъ понятіемъ статистическаго метода является понятіе совокупности. Совокупность образуется такимъ образомъ, что индивидуально отличные съ точки зрѣнія изучаемаго свойства случаи мы объединяемъ въ одно понятіе по наличности въ нихъ какого либо общаго признака или по принадлежности ихъ къ одному и тому же роду. Такъ напримѣръ, изучая продолжительность жизни умершихъ, мы объединяемъ большее или меньшее число случаевъ смерти въ различныхъ возрастахъ по принадлежности умершихъ къ опредѣленному мѣсту, времени или къ опредѣленной профессіи и получаемъ такимъ образомъ совокупность умершихъ опредѣленнаго государства, или опредѣленнаго года, или извѣстной профессіи.

Исчерпывающія и выборочныя. Совокупности образуются изъ единичныхъ случаевъ либо такимъ образомъ, что совокупность исчерпываетъ всѣ случаи опредѣленнаго рода въ предѣлахъ извѣстнаго времени и мѣста, напримѣръ, всѣ смертные случаи въ какомъ либо государствѣ за извѣстный годъ; или же совокупность обнимаетъ только часть всѣхъ случаевъ, причемъ эта часть взята изъ общей массы на удачу, случайно или хотя бы и по опредѣленной системѣ, но исключающей возможность подбора случаевъ по какому-либо другому общему признаку, кромѣ того, который долженъ служить для образованія данной совокупности, напримѣръ, когда совокупность составлена посредствомъ отбора каждаго десятаго случая смерти. Совокупности перваго рода называются исчерпывающими, второго—частичными или выборочными.

Общія и частныя.

Отъ указаннаго дѣленія слѣдуетъ отличать дѣленіе совокупностей на общія и частныя. Частная совокупность образуется выдѣленіемъ изъ общей совокупности тѣхъ случаевъ, которые отличаются какимъ-либо дополнительнымъ признакомъ, напримѣръ, выдѣленіемъ изъ общаго числа смертныхъ случаевъ тѣхъ, которые относятся къ городскому населенію; или обратно, общая совокупность получается объединеніемъ частныхъ совокупностей по какому либо общему имъ всѣмъ признаку, напримѣръ, объединеніемъ совокупностей умершихъ отъ каждой причины въ отдѣльности въ общую совокупность умершихъ вообще.

Разница эта существенна въ томъ отношеніи, что частная совокупность представляетъ собою собраніе случаевъ, качественно отличныхъ отъ всѣхъ другихъ случаевъ общей совокупности, тогда какъ частичная совокупность заключаетъ въ себѣ тѣ же случаи, что и исчерпывающая, но лишь въ меньшемъ количествѣ. Поэтому частичная совокупность при извѣстныхъ условіяхъ, можетъ служить представителемъ исчерпывающей совокупности— чего нельзя сказать относительно частной совокупности— и въ этомъ смыслѣ называется также репрезентативной совокупностью.

Практическія затрудненія.

Хотя логическая разница между способомъ образованія выборочной и частной совокупности не вызываетъ сомнѣній, но практически опредѣлить, въ какихъ случаяхъ мы имѣемъ дѣло съ совокупностью того или другого рода, весьма часто бываетъ затруднительно. Происходитъ это отъ того, что случайные по видимости способы отбора явленій часто въ скрытомъ видѣ заключаютъ въ себѣ условія для систематическаго подбора; такъ, при анкетномъ опросѣ отвѣчающими являются, по преимуществу, либо односторонне заинтересованные, либо особенно сознательно относящіеся къ вопросу индивиды;и н тѣ, ни другіе не могутъ служить представителями всей массы опрашиваемыхъ; отборъ смертныхъ случаевъ по опредѣленному дню недѣли или числу мѣсяца могъ бы дать искуственный подборъ фактовъ въ тѣхъ или другихъ специальныхъ случаяхъ, напримѣръ, при изслѣдованіи смертности отъ пьянства, драки, самоубійствъ, въ зависимости отъ отношенія выбраннаго дня къ праздникамъ, времени полученія жалованія, обычному сроку платежей и т. п. Рѣшеніе вопроса, имѣетъ ли мѣсто репрезентативный отборъ или же систематическій подборъ явленій, зависитъ отъ того, связанъ ли признакъ или способъ отбора какимъ-либо причиннымъ образомъ съ изслѣдуемымъ явленіемъ, что весьма часто обнаруживается лишь по результатамъ изслѣдованія.

Отношеніе къ родовымъ понятіямъ и истинамъ.

Совокупность, будетъ ли она исчерпывающей или частичной, общей или частной, во всякомъ случаѣ представляетъ собою понятіе, состоящее изъ общаго признака, объединяющаго единичные случаи, и представленія о нѣкоторомъ числѣ такихъ случаевъ. Этимъ послѣднимъ моментомъ совокупность отличается отъ родового понятія, которое содержитъ въ себѣ только общій признакъ или общіе признаки, объединяющіе всѣ единичные случаи въ одинъ общій родъ. Число единичныхъ случаевъ не входитъ въ содержанія понятія и составляетъ лишь объемъ его. Отсюда явствуетъ, что отъ всякой совокупности можно перейти къ родовому понятію и обратно. Если изъ совокупности исключить представленіе о количествѣ случаевъ и оставить только родовой признакъ, получается родовое понятіе; такъ совокупности «населеніе города N» отвѣчаетъ родовое понятіе «житель города N». Если же въ содержаніе родового понятія включить представленіе о нѣкоторомъ числѣ случаевъ, какіе заключаются въ его объемѣ, получимъ понятіе совокупности. При этомъ, если мы возьмемъ все число случаевъ, какое фактически содержится въ объемѣ понятія въ предѣлахъ опредѣленнаго времени и мѣста, мы получимъ исчерпывающую совокупность; если же включимъ только часть случаевъ,—получимъ частичную репрезентативную или частную совокупность, смотря по способу подбора единичныхъ случаевъ. Если въ содержаніе родового понятія ввести представленіе о всѣхъ возможныхъ единичныхъ случаяхъ, получится совокупность съ неопредѣленнымъ количествомъ случаевъ, отвѣчающимъ всему объему родового понятія.

Такъ, родовому понятію «человѣкъ» отвѣчаетъ совокупность «человѣческій родъ»; совокупности «городское населеніе» отвѣчаетъ родовое понятіе «житель города».

При изслѣдованіи явленій мы могли бы—и въ иныхъ случаяхъ такъ и поступаемъ—объединять ихъ то въ родовое понятіе, то въ совокупность, смотря по тому, приходится ли намъ изслѣдовать въ даннаго рода явленіяхъ ихъ родовыя свойства, присущія каждому случаю явленія въ тождественномъ видѣ, или же свойства индивидуально различныя, получающія одно общее для всей совокупности выраженіе лишь въ видѣ сводныхъ признаковъ и сводной закономѣрности. Такъ, напримѣръ, понятіе «домашняя кошка» имѣетъ родовой характеръ, поскольку зоологъ описываетъ такія свойства этого животнаго, какъ наличность позвоночнаго столба, расположеніе органовъ, характерное устройство пальцевъ и т. п.; и становится понятіемъ совокупности, когда рѣчь идетъ о величинѣ животнаго, его вѣсѣ, окраскѣ и другихъ подобныхъ свойствахъ; точно также, напримѣръ, преступность можетъ быть изучаема и какъ родовое понятіе, и какъ совокупность.

Однако въ большинствѣ случаевъ для изученія тѣхъ свойствъ и правильностей, какія обнаруживаются лишь въ совокупностяхъ, мы объединяемъ единичные случаи по инымъ общимъ признакамъ, чѣмъ при изученіи свойствъ и правильностей родовыхъ. Такая перегруппировка единичныхъ случаевъ для образованія родовыхъ понятій и совокупностей вызывается тѣмъ обстоятельствомъ, что общіе признаки, дающіе весьма цѣлесообразныя родовыя понятія, т. е. понятія, содержащія въ себѣ въ то же время большое количество другихъ общихъ родовыхъ свойствъ и правильностей, даютъ совокупности, бѣдныя сводными признаками и сводными закономѣрностями; и обратно, признаки, весьма цѣлесообразно объединяющіе единичные случаи въ совокупности, т. е. дающіе совокупности съ обильнымъ содержаніемъ сводныхъ свойствъ и закономѣрностей, будучи употреблены для образованія родовыхъ понятій, даютъ понятія, бѣдныя родовымъ содержаніемъ. Такъ, напримѣръ, родовое понятіе «вода», образованное по признаку химическаго состава изъ H_2O, обнаруживаетъ большое количество весьма существенныхъ общихъ химическихъ и физическихъ свойствъ родового характера; но если объединить по тому же признаку всѣ возможные случаи наличности воды въ понятіе совокупности, послѣднее окажется совершенно безплоднымъ для цѣлей изслѣдованія; съ другой стороны, такой признакъ, какъ тридцатилѣтній возрастъ даетъ родовое понятіе съ весьма бѣднымъ содержаніемъ родовыхъ свойствъ, кромѣ тѣхъ, разумѣется, которыя содержатся уже въ понятіи «человѣкъ»; но въ смыслѣ признака, служащаго для образованія совокупности, можетъ дать понятіе, обладающее значительнымъ количествомъ сводныхъ признаковъ и правильностей.

Иначе говоря, обычно бываетъ такъ, что сводные признаки и правильности связаны съ одними общими свойствами явленій, тогда какъ родовые признаки и законы съ другими.

Преувеличенное значеніе, придаваемое указанному здѣсь обстоятельству, привело нѣкоторыхъ авторовъ къ ошибочному утвержденію, будто существуютъ особыя области явленій, допускающія и не допускающія примѣненіе статистическаго метода. Такъ напримѣръ, Rümelin утверждалъ, что понятіе общества допускаетъ и требуетъ статистическаго метода изученія, тогда какъ понятіе государства не допускаетъ таковаго. Иногда утверждаютъ, будто явленія общественной жизни могутъ быть изслѣдуемы только статистическимъ методомъ, вопреки очевидному существованію ряда общественныхъ дисциплинъ (науки права, теоретическая экономія), устанавливающихъ исключительно родовыя свойства и правильности. Въ дѣйствительности нѣтъ явленій, которыя по своей природѣ допускали бы только одинъ какой-либо способъ изученія и не допускали другого; существуютъ лишь въ каждомъ явленіи свойства и правильности, требующія примѣненія статистическаго способа изслѣдованія, и свойства, требующія родовыхъ пріемовъ; сообразно съ этимъ, одни и тѣ же явленія служатъ для образованія различныхъ понятій—родовыхъ и сводныхъ.

Общность родовыхъ истинъ. Родовыя понятія заключаютъ въ своемъ объемѣ все возможное число случаевъ даннаго рода; поэтому всякая истина, имѣющая родовой характеръ, въ силу этого имѣетъ значеніе и силу для всякаго любого случая, не связана съ какими либо опредѣленными конкретными случаями даннаго рода, иначе говоря, имѣетъ общій и абстрактный характеръ. Такъ, если вода при разложеніи даетъ два атома водорода и одинъ кислорода, или если извѣстное количество механической работы эквивалентно опредѣленному количеству тепловой энергіи, то такого рода истины имѣютъ значеніе для всякаго случая разложенія воды или превращенія тепловой энергіи въ работу, совершенно независимо отъ тѣхъ конкретныхъ случаевъ, на которыхъ эти истины установлены и вообще независимо отъ какихъ бы то ни было условій времени и мѣста.

Эмпиризмъ сводныхъ величинъ Понятіе совокупности всегда содержитъ въ себѣ представленіе о нѣкоторомъ числѣ случаевъ. Правда, можно образовать совокупность такимъ образомъ, чтобы въ содержаніе ея входили всѣ возможные и мыслимые случаи даннаго рода; тогда фактически она будетъ относиться къ тому же неопредѣленному числу случаевъ, какое содержится въ объемѣ родового понятія. Такъ, можно образовать понятіе совокупности рожденій вообще. Однако даже и здѣсь, когда мы приступаемъ къ нахожденію сводныхъ признаковъ и правильностей, мы должны имѣть конкретное число случаевъ съ конкретной величиной или конкретнымъ качествомъ каждаго случая, такъ какъ сводные признаки и сводныя правильности состоятъ изъ среднихъ чиселъ, отношеній, численной зависимости, и вычисленіе такихъ сводныхъ величинъ требуетъ конкретныхъ чиселъ. Поэтому всякій статистическій выводъ имѣетъ прежде всего значеніе для той конкретной совокупности случаевъ, на основаніи которой онъ вычисленъ.

Каждый конкретный случай, необходимый для образованія конкретной совокупности, относится къ опредѣленному мѣсту, опредѣленному времени, поэтому и совокупность, образованная изъ конкретныхъ случаевъ, также ограничена предѣлами времени и мѣста. А слѣдовательно, и выводъ, сдѣланный для такой совокупности, связанъ съ рамками времени и мѣста, или, иначе говоря, имѣетъ эмпирическій характеръ. Правда, и родовыя понятія могутъ быть образованы лишь на основаніи наблюденія отдѣльныхъ случаевъ или экземпляровъ рода, взятыхъ въ опредѣленномъ мѣстѣ и времени; какъ равнымъ образомъ всякій абстрактный и общій законъ устанавливается на конкретномъ случаѣ опыта или испытанія. Иначе говоря, всякое родовое понятіе, родовой признакъ и общій законъ точно также имѣютъ эмпирическій источникъ происхожденія.

Однако въ процессѣ обобщенія, съ того момента, какъ понятіе, признакъ или истина получаютъ родовой характеръ, они теряютъ связь съ конкретнымъ источникомъ своего происхожденія и пріобрѣтаютъ общій и абстрактный характеръ, выражающійся въ томъ, что самое представленіе о какомъ либо числѣ случаевъ исчезаетъ изъ ихъ содержанія.

Допускаютъ ли онѣ обобщеніе.

Между тѣмъ въ понятіи совокупности это представленіе о нѣкоторомъ числѣ случаевъ составляетъ необходимую часть содержанія понятія; отсюда возникаетъ вопросъ, не связанъ ли эмпирическій характеръ не только съ процессомъ образованія сводныхъ величинъ, но и вообще съ самой логической природой совокупности, какъ непремѣнное ея условіе, и могутъ ли сводные признаки и закономѣрности пріобрѣтать общее и абстрактное значеніе, подобно родовымъ понятіямъ, признакамъ и законамъ?

Сводная величина не есть обобщеніе.

Прежде чѣмъ отвѣтить на этотъ вопросъ, необходимо устранить возможное недоразумѣніе: можно было бы думать, что сводная величина, сама по себѣ, является уже нѣкоторымъ обобщеніемъ, такъ какъ она представляетъ собою общее выраженіе для совокупности, выведенное изъ единичныхъ случаевъ. Однако въ самомъ процессѣ образованія сводной величины нѣтъ никакого обобщенія, ибо сводная величина лишь измѣняетъ выраженіе признака или соотношенія, приписывая ихъ совокупности въ иномъ видѣ, чѣмъ они присущи отдѣльнымъ случаямъ; но при этомъ число случаевъ не измѣняется: въ видѣ сводной величины признакъ относится къ совокупности тѣхъ же случаевъ, которымъ онъ присущъ въ своемъ первоначальномъ видѣ. Обобщеніе же состоитъ въ томъ, что нѣчто, выведенное на основаніи однихъ случаевъ, приписывается другимъ, новымъ и инымъ случаямъ.

Обобщеніе въ предѣлахъ совокупности.

Что касается обобщенія въ этомъ послѣднемъ смыслѣ, то въ отношеніи сводныхъ величинъ здѣсь слѣдуетъ различать два вопроса. Обобщеніе сводной величины можетъ итти, прежде всего, внутри самой совокупности, отъ извѣстнаго числа наблюденныхъ случаевъ къ бóльшему числу тѣхъ же случаевъ. Совокупность, какъ извѣстно, можетъ быть образована двояко: или изъ всего числа случаевъ даннаго рода, имѣвшихъ мѣсто въ рамкахъ даннаго мѣста и времени, или только изъ части этихъ случаевъ, набранныхъ безъ какого бы то ни было систематическаго подбора; иначе говоря, конкретная совокупность можетъ быть исчерпывающей или частичной. Но даже если совокупность является исчерпывающей для даннаго времени и мѣста, то, расширивъ предѣлы времени и мѣста, мы исчерпывающую совокупность можемъ сдѣлать, въ свою очередь, частичной съ точки зрѣнія совокупности болѣе обширнаго числа случаевъ. Поэтому возникаетъ вопросъ: если выводъ сдѣланъ на основаніи частичной совокупности, можетъ ли онъ быть приписанъ бóльшему числу случаевъ совокупности исчерпывающей или вообще совокупности съ бóльшимъ количествомъ такихъ же случаевъ, не входившихъ въ процессъ образованія сводной величины? Практическій опытъ указываетъ, что для опредѣленія средней температуры, средняго урожая, средняго роста и т. п. нѣтъ надобности измѣрять температуру каждой минуты въ каждой точкѣ территоріи; мы довольствуемся для этого нѣсколькими моментами сутокъ и показаніями сравнительно немногихъ станцій; мы ограничиваемся опросомъ достаточно большого числа хозяевъ относительно урожая; равнымъ образомъ, измѣривъ достаточно большое количество индивидовъ, мы принимаемъ полученную среднюю величину за общее выраженіе средняго роста всей совокупности индивидовъ данной націи и даннаго возраста. Статистическій методъ, какъ мы увидимъ ниже, даетъ доказательство, что такое обобщеніе, при соблюденіи извѣстныхъ условій, допустимо: если частичная совокупность имѣетъ строго репрезентативный характеръ и состоитъ изъ достаточно большого числа случаевъ, то выведенныя на основаніи ея сводныя величины могутъ быть отвлечены отъ нея и приняты за выраженіе сводныхъ признаковъ и закономѣрностей любой достаточно большой совокупности однородныхъ случаевъ въ предѣлахъ того же явленія, въ томъ числѣ и совокупности исчерпывающей. Если предѣлы времени и мѣста мы возьмемъ такъ широко, что они исчерпываютъ все количество случаевъ, какое фактически могло бы быть подведено подъ объемъ родового понятія, то сводныя величины будутъ имѣть такое же значеніе для характеристики рода явленій, какъ и признаки родовые, хотя самая сводная величина будетъ выведена на основаніи репрезентативной части случаевъ. Но, разумѣется, такое обобщеніе возможно лишь при условіи, если бо́льшая совокупность состоитъ изъ случаевъ совершенно однородныхъ съ тѣми, какіе послужили для вывода сводной величины. Чѣмъ болѣе мы расширяемъ рамки совокупности, тѣмъ труднѣе сохранить однородность вновь привходящихъ случаевъ, поэтому распространеніе репрезентативно полученной сводной величины на весь родъ явленій практически возможно въ рѣдкихъ случаяхъ и предполагаетъ крайнюю простоту своднаго признака и большую устойчивость явленія.

Обобщенія этого рода характеризуются слѣдовательно тѣмъ, что совокупность остается качественно неизмѣнной, остается совокупностью того же рода, и измѣняется лишь число случаевъ, въ нее вводимыхъ; увеличенію числа случаевъ совокупности сопутствуетъ распространеніе на нихъ значенія сводной величины, выведенной на основаніи части случаевъ. Обобщеніе послѣдней происходитъ здѣсь внутри самой совокупности и выражается въ томъ, что сводная величина становится независимой отъ конкретной опредѣленности случаевъ и ихъ конкретнаго числа въ предѣлахъ одного и того же явленія. Указанный способъ обобщенія основанъ на т. н. статистическомъ законѣ большихъ чиселъ и имѣетъ специфически статистическій характеръ.

Обобщеніе въ обычномъ смыслѣ.

Въ другомъ способѣ обобщенія мы имѣемъ дѣло съ примѣненіемъ къ совокупностямъ общихъ методовъ индукціи. Здѣсь вопросъ идетъ о томъ, что если какая либо сводная величина или закономѣрность установлена для даннаго времени или мѣста, то можетъ ли она быть распространена и на всѣ сходные случаи, независимо отъ условій времени и мѣста? Если, напримѣръ, для Англіи и Уэльса установлено, что средній ростъ преступниковъ на опредѣленную величину ниже средняго роста всего населенія, то насколько такое соотношеніе можетъ быть распространено на другія страны или вообще на преступное и непреступное населеніе независимо отъ условій пространства и времени.

Здѣсь мы не имѣемъ уже дѣла съ отношеніемъ сводной величины къ бо́льшему или меньшему числу случаевъ въ предѣлахъ одной и той же совокупности или съ отношеніемъ одной части случаевъ къ другой части случаевъ въ предѣлахъ совокупности одного и того же явленія; здѣсь совокупность случаевъ—большая или малая, репрезентативная или исчерпывающая безразлично—разсматривается какъ одно единое явленіе, какъ единичный случай наблюденія и сопоставляется съ другими сходными совокупностями, разсматриваемыми каждая также какъ единичный случай, и вопросъ состоитъ въ томъ, возможно ли сводную величину или закономѣрность, установленную для одного случая, распространить на всѣ случаи и сдѣлать статистическій выводъ независимымъ отъ условій времени и мѣста. Опытъ показываетъ, что преобладающая масса среднихъ цыфръ, отношеній и статистическихъ зависимостей носитъ временный и мѣстный характеръ и не допускаетъ отвлеченія и обобщенія, хотя нерѣдко статистическія цыфры обнаруживаютъ замѣчательную устойчивость. Такъ, напримѣръ, коэффиціентъ брачности представляетъ для многихъ странъ довольно устойчивую величину и лишь въ рѣдкихъ случаяхъ колеблется по годамъ сколько-нибудь значительно; однако мы не можемъ считать его какимъ-либо абстрактнымъ закономъ и допускаемъ возможность, что онъ можетъ измѣниться и весьма значительно. Еще болѣе устойчиво отношеніе числа рожденій мальчиковъ и дѣвочекъ, но все же оно не носитъ столь абстрактно-обязательнаго характера, какъ численное отношеніе химическихъ элементовъ въ сложномъ тѣлѣ. Большинство же статистическихъ величинъ прямо связаны съ опредѣленной страной и временемъ и совершенно не допускаютъ распространенія на другія страны или обобщенія во времени.

Этимъ эмпирическимъ характеромъ сводные признаки, сводныя закономѣрности и сводныя зависимости отличаются отъ родовыхъ признаковъ и родовыхъ или точныхъ законовъ, примѣромъ которыхъ могутъ служить законы физическіе. Коренится ли эта разница въ логической природѣ совокупностей или же въ фактическомъ характерѣ тѣхъ совокупностей, съ которыми намъ обычно приходится имѣть дѣло?

Индукція. Рѣшеніе этого вопроса, очевидно, связано съ вопросомъ о томъ, какимъ образомъ получаютъ свой абстрактный и общій характеръ родовыя понятія и законы, которые также имѣютъ эмпирическое происхожденіе и должны быть обоснованы на конкретныхъ случаяхъ.

Родовыя истины, имѣющія общій и абстрактный характеръ, состоятъ либо въ утвержденіи, что извѣстный признакъ связанъ съ какимъ-нибудь комплексомъ другихъ признаковъ, какъ напримѣръ, атомный вѣсъ связанъ съ комплексомъ признаковъ, образующихъ химическій элементъ; или что какое либо обстоятельство причинно связано съ другимъ, какъ нагрѣваніе съ расширеніемъ тѣлъ; или что одно слѣдуетъ за другимъ, какъ напримѣръ, смерть за рожденіемъ. Вообще всѣ абстрактныя и общія истины состоятъ въ установленіи необходимой связи между двумя моментами. Въ зависимости отъ того, какъ размѣщаются связанные моменты въ пространствѣ (т. е. относятся ли они къ одному и тому же объекту или разнымъ) и во времени (относятся ли они къ одному или разнымъ моментамъ времени), связь моментовъ получаетъ различныя наименованія; въ этомъ отношеніи можно различать слѣдующіе случаи:

1. оба момента относятся къ одному и тому же предмету или явленію въ одно и то же время; въ этомъ случаѣ одинъ моментъ мы разсматриваемъ какъ предметъ или явленіе, другой какъ необходимый признакъ этого предмета или явленія; такъ, мы говоримъ: золото обладаетъ удѣльнымъ вѣсомъ; или оба момента разсматриваемъ, какъ признаки объекта, и тогда говоримъ о необходимой связи двухъ признаковъ, какъ, напримѣръ, связь атомнаго и удѣльнаго вѣса въ золотѣ;

2. оба момента относятся къ одному и тому же явленію, но въ послѣдующіе моменты времени; въ этомъ случаѣ необходимая связь носитъ названіе закона; таковъ законъ паденія, законъ развитія и т. п.;

3. оба момента относятся къ разнымъ объектамъ въ одинъ и тотъ же или разные моменты времени—въ этомъ случаѣ мы говоримъ о причинной связи или причинной зависимости.

Постулатъ обобщенія. Во всѣхъ случаяхъ, когда мы имѣемъ подобный случай связи, съ характеромъ общей и безусловной истины, такого рода истина, если только она получается изъ опыта—наблюденія или испытанія—всегда устанавливается на опредѣленномъ конкретномъ случаѣ или немногихъ конкретныхъ случаяхъ, и поэтому прежде всего имѣетъ значеніе для этихъ именно случаевъ. Общій, отвлеченный отъ конкретнаго случая и безусловный характеръ она получаетъ лишь путемъ дополнительной логической операціи, основанной на дедукціи изъ подразумѣваемой нами при всякомъ отвлеченіи и обобщеніи истины: если что-либо зависѣло отъ какого-либо обстоятельства одинъ разъ, то невозможно, чтобы оно не зависѣло другой разъ, если обстоятельство повторится въ тождественномъ видѣ (т. н. постулатъ постоянства познаваемаго міра).

Изъ характера этой постулативной истины вытекаютъ тѣ условія, при которыхъ опытъ одного случая можетъ быть распространенъ на любой другой случай и слѣдовательно получаетъ характеръ общаго положенія, независимаго отъ конкретныхъ условій времени и мѣста первоначальнаго опыта. А именно, необходимо, чтобы обстоятельство обуславливающее, при повтореніи, повторялось въ тождественномъ видѣ, и, во-вторыхъ, чтобы связь этого обстоятельства съ другимъ была установлена хотя бы въ одномъ случаѣ. Первое условіе можетъ быть удовлетворено лишь въ томъ случаѣ, когда обуславливающее обстоятельство представляетъ собою какой-либо элементарно-простой моментъ, ибо только элементарно-простой моментъ абсолютно тождественъ и неизмѣненъ и, слѣдовательно, если повторяется, то повторяется въ тождественномъ видѣ; всякое же сложное обстоятельство, состоящее изъ стеченія нѣсколькихъ моментовъ, раньше—въ другомъ либо другихъ случаяхъ повторялось въ иномъ составѣ моментовъ—ибо иначе мы не могли бы знать, что оно сложно—а слѣдовательно и впредь можетъ повториться въ иномъ видѣ. Второе условіе—необходимость установленія связи двухъ моментовъ хотя бы въ одномъ случаѣ—подразумѣваетъ, во-первыхъ, необходимость установить, что изслѣдуемый моментъ *связанъ* съ другимъ моментомъ, т. е. находится въ такомъ къ нему отношеніи, что наличность или отсутствіе этого другого момента обуславливаетъ наличность или отсутствіе даннаго момента; и, во-вторыхъ, подразумѣваетъ, что изслѣдуемый моментъ зависитъ указаннымъ способомъ именно отъ даннаго, а не какого-либо иного момента, а это, въ свою очередь, требуетъ исчерпывающаго учета всѣхъ моментовъ явленія, съ цѣлью устраненія сомнѣній, что появленіе и исчезновеніе изслѣдуемаго момента можетъ зависѣть отъ какихъ либо неучтенныхъ обстоятельствъ.

Если имѣются на лицо указанныя условія разложенія явленія на элементарно-простыя обстоятельства и точнаго учета послѣднихъ, то достаточно въ данныя условія опыта ввести новый элементарно-простой моментъ, или, обратно, исключить его, и наступающее измѣненіе представитъ собою то, что связано съ этимъ моментомъ (пріемъ единственной разницы); или, иначе, необходимо измѣнить всѣ обстоятельства опыта, оставивъ только одно, и тогда то общее, что новый опытъ имѣетъ съ прежнимъ, будетъ связано съ оставшимся неизмѣннымъ обстоятельствомъ (пріемъ единственнаго сходства).

Примѣненіе указанныхъ пріемовъ составляетъ сущность точной индукціи; индукція слѣдовательно требуетъ двухъ условій: разложенія наблюдаемаго случая на элементарно-простыя части и точнаго учета послѣднихъ. Вопросъ о томъ, могутъ ли сводныя величины быть отвлечены отъ эмпирически данной совокупности и получить характеръ общаго свойства или закона, сводится такимъ образомъ, къ тому, допускаютъ ли совокупности примѣненіе къ нимъ индукціи, что, въ свою очередь, равносильно вопросу, допускаютъ ли онѣ разложеніе на элементарно-простые моменты и учетъ послѣднихъ?

Въ этомъ отношеніи мы различаемъ въ совокупности, прежде всего, двѣ части: съ одной стороны, сводную величину, подлежащую изслѣдованію, и, съ другой—ту сумму родовыхъ признаковъ (или тотъ признакъ), по которой каждый единичный случай относится къ данной совокупности. Такъ напримѣръ, въ совокупности «шотландцы, мужчины, 30 лѣтъ» мы различаемъ сводную величину, подлежащую изученію—скажемъ, средній ростъ или отношеніе числа холостыхъ къ числу состоящихъ въ бракѣ, и, съ другой стороны, общіе признаки каждаго единичнаго случая, а именно: то, что каждый индивидъ совокупности представляетъ собою «шотландца», «мужчину» и обладаетъ «возрастомъ 30 лѣтъ». Сводный признакъ можетъ быть связанъ съ однимъ изъ родовыхъ признаковъ группы, и индукція можетъ быть направлена прежде всего на отысканіе этой связи. Однако вопросъ не исчерпывается этимъ. Съ признакомъ, по которому мы подбираемъ единичные случаи, можетъ быть связанъ цѣлый рядъ обстоятельствъ, которыя одновременно съ родовымъ признакомъ, выраженнымъ expressis verbis, проникаютъ въ совокупность въ скрытомъ видѣ. Далѣе, если совокупность взята въ рамкахъ опредѣленнаго времени и мѣста, то всѣ обстоятельства этого времени и этого мѣста составляютъ какъ бы среду, въ которой находится совокупность съ ея сводными свойствами и правильностями. Такимъ образомъ задача индуктивнаго изслѣдованія распространяется на всю сумму обстоятельствъ, заключающихся въ самой совокупности, и обстоятельствъ, заключающихъ въ себѣ совокупность. При всемъ фактическомъ обиліи и разнообразіи всѣхъ этихъ обстоятельствъ, они, по своей формальной логической природѣ, могутъ относиться либо къ разряду сводныхъ величинъ, либо къ разряду родовыхъ свойствъ и явленій, и такимъ образомъ задача индуктивнаго изслѣдованія сводится къ отысканію связи изслѣдуемой сводной величины съ какой-либо другой сводной же величиной или съ обстоятельствомъ родового характера.

Что касается обстоятельствъ родового характера, то возможность ихъ разложенія на элементарныя части и учета послѣднихъ съ логической стороны не представляетъ никакихъ специфическихъ особенностей, по сравненію съ тѣми случаями, когда мы имѣемъ дѣло съ приложеніемъ индукціи къ явленіямъ исключительно родового характера: тотъ фактъ, что разложеніе имѣетъ цѣлью сопоставить элементарные моменты разлагаемыхъ обстоятельствъ со сводными величинами, не измѣняетъ логическаго характера ни самого процесса разложенія, ни разлагаемыхъ явленій. Иначе обстоитъ вопросъ съ фактической стороны: множество обстоятельствъ, содержащихся въ самой совокупности, и условій, составляющихъ ея среду, дѣлаютъ простой учетъ этихъ обстоятельствъ и условій мало осуществимымъ практически.

Остается вопросъ о сводныхъ величинахъ; и здѣсь мы должны различать фактическую сторону отъ принципіальной. Если бы сводныя величины были принципіально, по своей логической природѣ, несводимы къ элементарно-простымъ моментамъ, въ такомъ случаѣ ни одна совокупность съ ея сводными правильностями не могла бы быть сведена къ такимъ моментамъ, и, обратно, изъ простыхъ моментовъ нельзя было бы воспроизвести сводныхъ величинъ съ ихъ закономѣрностью. Такое предположеніе не отвѣчаетъ дѣйствительности. Прежде всего теоретически и абстрактно мы можемъ представить себѣ систему элементарно-простыхъ моментовъ, которые дадутъ намъ рядъ сводныхъ закономѣрностей, логически вытекающихъ изъ принятыхъ нами допущеній. Такая именно элементарно-простая система моментовъ, дающая рядъ сводныхъ закономѣрностей, лежитъ въ основѣ математическаго исчисленія вѣроятностей въ теоріи вѣроятности. Мало того, путемъ эксперимента мы можемъ создать условія, дѣйствующія, насколько это допускаетъ техника, согласно апріорнымъ допущеніямъ теоріи вѣроятности и дающія согласные съ ней результаты. Таковы испытанія, предпринимаемыя съ спеціальною цѣлью эмпирической провѣрки дедуктивныхъ выводовъ теоріи, какъ то: опыты съ выниманіемъ наугадъ шаровъ, подбрасываніемъ монетъ, метаніемъ кости; сюда же относятся и наблюденія надъ исходами игръ, построенныхъ на расчетѣ массовой правильности случайныхъ событій, какъ то: рулетки, лотто. Теоретически выведенныя и провѣренныя на опытѣ сводныя правильности, проявляющіяся въ простѣйшихъ совокупностяхъ, служатъ исходной точкой для объясненія тѣхъ сложныхъ закономѣрностей, съ которыми мы имѣемъ дѣло въ эмпирическихъ совокупностяхъ.

Такимъ образомъ, съ точки зрѣнія логической, сводныя величины поддаются сведенію къ дѣйствію элементарныхъ моментовъ, доступныхъ точному учету. Иначе обстоитъ дѣло съ фактической стороны. Тѣ совокупности, съ которыми мы имѣемъ дѣло на практикѣ, представляютъ слишкомъ сложное строеніе, и сводныя правильности являются результатомъ дѣйствія столь большого числа различныхъ и мѣняющихся моментовъ, что фактически мы не въ состояніи выдѣлить ихъ, подвергнуть учету и изслѣдованію. Благодаря именно этой сложности, отсутствуютъ условія, необходимыя для приложенія къ эмпирическимъ совокупностямъ точной индукціи, и мы не можемъ ни свести эмпирическія сводныя величины къ элементарно-простому выраженію, ни установить связь ихъ въ такомъ видѣ съ какими-либо другими простыми моментами и сообщить своднымъ величинамъ и закономѣрностямъ характеръ точнаго закона. Одной изъ задачъ статистическаго изученія является именно разложеніе эмпирически данныхъ совокупностей и сводныхъ величинъ на совокупности возможно простого типа, совокупности такого рода, чтобы въ нихъ, съ одной стороны, опредѣляющіе совокупность родовые признаки имѣли возможно простой характеръ и, съ другой, сводныя величины обладали, по возможности, такими свойствами, какія отвѣчаютъ элементарно-простымъ гипотетическимъ совокупностямъ.

Такого рода изслѣдованія представляютъ большой теоретическій интересъ и, быть можетъ, большую практическую важность. Но и эмпирическій характеръ статистическихъ цыфръ и правильностей нисколько не лишаетъ ихъ глубокаго жизненнаго интереса, такъ какъ задача познанія не ограничивается идеаломъ точнаго описанія конечныхъ простыхъ элементовъ въ ихъ взаимныхъ зависимостяхъ, но обнимаетъ собою и эмпирическое изученіе явленій, какъ въ конкретныхъ условіяхъ опредѣленнаго времени и мѣста, такъ и въ формѣ эмпирически обобщенныхъ правилъ.

Хотя такимъ образомъ приведеніе получаемыхъ въ дѣйствительности сводныхъ величинъ къ ихъ элементарно-простому виду является задачей практически мало осуществимой, но самое понятіе элементарно-простой совокупности является исходной точкой для объясненія закономѣрности сводныхъ величинъ въ эмпирическихъ совокупностяхъ и служитъ обоснованіемъ тѣхъ логическихъ операцій, которыя составляютъ содержаніе статистическаго метода. Логическое построеніе элементарно-простой совокупности показываетъ намъ тѣ условія, при наличности которыхъ могутъ образоваться индивидуально-различные по количеству или качеству случаи, дающіе сводную правильность въ ихъ совокупности. Теоретическое изученіе образованія совокупности съ закономѣрностью ея сводныхъ величинъ мы находимъ въ теоріи вѣроятностей. Теорія вѣроятности изучаетъ законы образованія совокупностей и ея сводныхъ величинъ дедуктивно, исходя изъ нѣкоторыхъ апріорныхъ допущеній. Статистическій методъ примѣняетъ выводъ теоріи вѣроятностей къ изслѣдованію наблюдаемыхъ совокупностей, принимая, что въ послѣднихъ дѣйствуютъ и проявляются условія, до извѣстной степени аналогичныя допущеніямъ теоріи.

2. Значенiе теорiи вѣроятности для статистическаго метода.

Вѣроятность. Представимъ себѣ два (или больше) элементарно-простыхъ момента, скажемъ a и b (или a, b, c,... n), изъ которыхъ каждый, осуществляясь въ воображаемомъ опытѣ, даетъ соотвѣтственно событiя A и B (или A, B.... N); напримѣръ, наличность въ монетѣ двухъ сторонъ—орла и рѣшетки—даетъ при подбрасыванiи монеты паденiе ея орломъ или рѣшеткой кверху. Пусть эти моменты будутъ связаны другъ съ другомъ такимъ образомъ, что въ каждомъ опытѣ непремѣнно осуществляется, въ видѣ соотвѣтствующаго явленiя, одинъ изъ моментовъ, и осуществленiе одного исключаетъ осуществленiе въ томъ же опытѣ другого момента (или другихъ моментовъ)—если монета упала орломъ кверху, то въ томъ же самомъ случаѣ она одновременно не можетъ упасть рѣшеткой кверху; назовемъ связанные такъ другъ съ другомъ моменты комплексно—множественной причиной. Но пусть при этомъ до опыта осуществленiе въ немъ одного изъ моментовъ представляется столь же мыслимымъ, какъ и всякаго другого, или пусть осуществленiе въ опытѣ каждаго изъ моментовъ будетъ равновозможнымъ; такъ напримѣръ, до того, какъ монета упала, паденiе ея той или другой стороной представляется намъ одинаково возможнымъ. Наконецъ, допустимъ, что осуществленiе въ одномъ или многихъ опытахъ того или иного изъ моментовъ не предрѣшаетъ исхода другихъ опытовъ, иначе говоря, представимъ, что моменты другъ отъ друга независимы.

Такимъ образомъ моменты комплексно—множественной причины мыслятся, какъ исчерпывающiе причину, исключающiе другъ друга, равновозможные и независимые, а вслѣдствiе того и соотвѣтствующiя имъ событiя представляются, какъ исчерпывающія, исключающія другъ друга, равновозможныя и независимыя. При указанныхъ допущеніяхъ, если каждый изъ моментовъ принять за единицу, то отношеніе одного момента къ числу всѣхъ моментовъ комплексно-множественной причины называется вѣроятностью соотвѣтствующаго данному моменту событія. Такъ, при двухъ исключающихъ другъ друга, равновозможныхъ и независимыхъ моментахъ, какъ напримѣръ, наличность бѣлаго и чернаго шара, изъ которыхъ одинъ вынимается наугадъ, вѣроятность появленія бѣлаго шара составляетъ $^1/_2$, и вѣроятность появленія чернаго шара тоже $^1/_2$; при трехъ шарахъ различнаго цвѣта вѣроятность появленія каждаго цвѣта составляетъ $^1/_3$. Если нѣсколько индивидуально различныхъ событій объединяется по какому либо общему признаку въ понятіе одного общаго родового событія, то вѣроятность такого событія опредѣляется отношеніемъ суммы моментовъ, соотвѣтствующихъ каждому изъ объединяемыхъ событій, къ числу всѣхъ моментовъ; такъ, при наличности двухъ бѣлыхъ шаровъ и одного чернаго, вѣроятность появленія каждаго шара составляетъ $^1/_3$; если же появленіе одного или другого бѣлаго шара мы объединимъ въ родовое понятіе появленія бѣлаго шара вообще, то вѣроятность такого событія выразится въ видѣ $^2/_3$. Въ этомъ случаѣ говорятъ, что появленію даннаго событія благопріятствуютъ два мыслимыхъ исхода, появленію другого событія—одинъ; поэтому вѣроятность можно опредѣлить, какъ отношеніе числа исходовъ, благопріятствующихъ данному событію, къ числу всѣхъ мыслимыхъ исходовъ. Въ такомъ видѣ опредѣленіе удобно тѣмъ, что его можно приложить къ тѣмъ случаямъ, когда исходъ представляетъ событіе сложное и вызывается дѣйствіемъ не одного простого момента, а сочетаніемъ моментовъ или неизвѣстной причиной; если только мы можемъ установить, что исходы равновозможны, независимы, исключаютъ другъ друга и въ совокупности исчерпываютъ всѣ мыслимые случаи, то отношеніе одного или многихъ исходовъ, принимаемыхъ за данное событіе, къ числу всѣхъ возможныхъ исходовъ, составитъ вѣроятность этого событія. Изъ опредѣленія вѣроятности явствуетъ, что вѣроятность представляетъ собою выраженіе (ожидаемаго) дѣйствія комплексно-множественной причины, отнесенное къ одному случаю.

Численное выраженіе вѣроятности, при принятыхъ допущеніяхъ исчерпывающаго учета, взаимнаго исключенія, равновозможности и независимости комплексно-дѣйствующихъ моментовъ даетъ возможность математическаго исчисленія вѣроятности болѣе сложныхъ исходовъ, состоящихъ изъ различнаго сочетанія простыхъ событій A и B въ рядѣ мыслимыхъ послѣдовательныхъ или одновременныхъ опытовъ.

Вѣроятность сложныхъ исходовъ.

1. Предварительно замѣтимъ, что такъ какъ вѣроятность событія равна суммѣ благопріятствующихъ ему исходовъ, дѣленной на число всѣхъ исходовъ, то сумма вѣроятностей всѣхъ взаимно исключающихъ другъ друга событій равна 1; такъ, если вѣроятность p одного событія A равна $^1/_2$ и другого B—q равна $^1/_2$, то

$$p+q=1.$$

2. Отсюда если вѣроятность наступленія событія равна p, то вѣроятность ненаступленія событія или q составитъ

$$q=1-p.$$

3. При двухъ событіяхъ A и B съ вѣроятностями $p={}^1/_2$ и $q={}^1/_2$, въ двухъ послѣдовательныхъ или одновременныхъ опытахъ мыслимы слѣдующіе исходы:

въ первомъ опытѣ—	A,	во второмъ	A	или:	AA
»	A	»	B	»	AB
»	B	»	A	»	BA
»	B	»	B	»	BB

такъ какъ мыслимыхъ равновозможныхъ исходовъ 4, то вѣроятность каждаго составитъ $^1/_4$; принимая во вниманіе, что вѣроятность каждаго изъ простыхъ событій A и B составляетъ $^1/_2$, видимъ, что

$$^1/_4 = ^1/_2 \cdot ^1/_2 ,$$

или вообще: вѣроятность P сложнаго событія, состоящаго въ повтореніи или совпаденіи простыхъ ообытій съ вѣроятностями p и q, равняется произведенію вѣроятностей этихъ послѣднихъ событій—

$P = pp$ (для сложнаго событія AA)
$P = pq$ (для AB и BA)
$P = qq$ (для BB);

при трехъ опытахъ мыслимы слѣдующіе исходы:

AAA AAB BBA BBB
ABA BAB
BAA ABB

вѣроятность каждаго составляетъ

$$^1/_8 = ^1/_2 \cdot ^1/_2 \cdot ^1/_2 ;$$

указанное правило составляетъ такъ наз. *теорему умноженія.*

4. Если въ послѣднемъ примѣрѣ не обращать вниманія на порядокъ событій A и B въ случаяхъ смѣшаннаго ихъ появленія, а лишь на число повтореній, то исходовъ съ двумя событіями A и однимъ B окажется 3, и вѣроятность появленія событія A два раза и событія B одинъ разъ составитъ $^3/_8$; принимая во вниманіе, что вѣроятность каждаго изъ этихъ трехъ исходовъ въ отдѣльности равна $^1/_8$, видимъ, что

$$^3/_8 = ^1/_8 + ^1/_8 + ^1/_8 ;$$

или вообще, вѣроятность P сложнаго событія, состоящаго въ появленіи одного, безразлично какого именно изъ нѣсколькихъ событій, съ вѣроятностями p_1, p_2, и т. д., выбранныхъ изъ всего числа взаимно исключающихъ другъ друга событій, равняется суммѣ вѣроятностей этихъ нѣсколькихъ событій

$$P = p_1 + p_2 + \ldots$$

Указанный выводъ составляетъ *теорему сложенія* вѣроятностей.

5. Вообще вѣроятность различныхъ сочетаній событій опредѣляется соотвѣтствующими членами разложенія бинома; такъ, если вѣроятность событія A равняется $^1/_2$ и событія B тоже $^1/_2$, то при двухъ послѣдовательныхъ или одновременныхъ опытахъ вѣроятность появленія событій A и B въ разныхъ сочетаніяхъ выражается соотвѣтствующими членами разложенія бинома

$$(^1/_2 + ^1/_2)^2 ,$$

а именно, вѣроятность появленія:

оба раза A — первымъ членомъ	$(^1/_2)^2 = ^1/_4$
одйнъ разъ A, одинъ B — вторымъ	$2 \cdot ^1/_2 \cdot ^1/_2 = ^1/_2$
оба раза B — третьимъ членомъ	$(^1/_2)^2 = ^1/_4$;

при трехъ опытахъ вѣроятность появленія

три раза A	$(^1/_2)^3 = ^1/_8$
два раза A, одинъ разъ B . . .	$3 \cdot (^1/_2)^2 \cdot (^1/_2) = ^3/_8$
» » B » » A . . .	$3 \cdot (^1/_2) \cdot (^1/_2)^2 = ^3/_8$
три раза B	$(^1/_2)^3 = ^1/_8$

Или, при вѣроятности событія A равной $^1/_3$ и B — равной $^2/_3$, при трехъ испытаніяхъ, вѣроятность появленія различнаго числа разъ A и B (независимо отъ порядка) составитъ:

AAA	$(^1/_3)^3 = ^1/_{27}$
AAB	$3 \cdot (^1/_3)^2 \cdot (^2/_3) = ^6/_{27}$
ABB	$3 \cdot (^1/_3) \cdot (^2/_3)^2 = ^{12}/_{27}$
BBB	$(^2/_3)^3 = ^8/_{27}$

Или въ общемъ видѣ при вѣроятностяхъ p и q и числѣ испытаній s :

появленіе s разъ A . . .	p^s
s—1 разъ A, одинъ разъ B	$sp^{s-1}q$
s—2 » A, два раза B	$\frac{s(s-1)}{1.2} p^{s-2} q^2$
. . . .	
m разъ A, n разъ B . .	$\frac{s(s-1)....2.1}{1.2...n.1.2...m} p^m q^n$
. . . .	
два раза A, s—2 раза B	$\frac{s(s-1)}{1.2} p^2 q^{s-2}$
одинъ разъ A, s—1 разъ B	spq^{s-1}
s разъ B	q^s

Въ приведенныхъ формулахъ дроби типа $p^m q^n$ обозначаютъ вѣроятность наступленія m разъ событія A и n разъ событія B въ одномъ опредѣленномъ порядкѣ (по теоремѣ умноженія); коэффиціенты же типа $\frac{s!}{m!n!}$ (знакъ ! послѣ s обозначаетъ рядъ множителелей отъ 1.2... и т. д. включительно до s) представляетъ (по теоремѣ сложенія) число исходовъ, въ которыхъ A повторяется m разъ и B — n разъ, независимо отъ порядка наступленія того и другого событія.

Изъ разсмотрѣнія коэффиціентовъ на численныхъ примѣрахъ видно, что коэффиціенты располагаются симметрично возрастая къ серединѣ; такъ для бинома

2-ой степени	1	2	1				
3-ой »	1	3	3	1			
4-ей »	1	4	6	4	1		
5-ой »	1	5	10	10	5	1	
6-ой »	1	6	15	20	15	6	1

. . . .

При этомъ получается таблица, въ которой коэффиціенты каждой слѣдующей строчки получаются изъ суммы соотвѣтствующаго и предшествующаго коэффиціентовъ выше лежащей строчки.

При *равенствѣ вѣроятностей* $p=q$ выраженія p^s, $p^{s-1}q, \ldots p^m q^n \ldots pq^{s-1}$, q^s даютъ также симметрическій рядъ, такъ что члены разложенія бинома вообще располагаются симметрично, возрастая къ серединѣ. Если же p неравно q, то рядъ возрастаетъ, а затѣмъ убываетъ, но при этомъ оказывается несимметричнымъ.

Теорема Бернулли. Основной теоремой теоріи вѣроятности является теорема Бернулли. Содержаніе ея сводятся къ слѣдующему.

Пусть двумъ взаимно исключающимъ другъ друга событіямъ A и B (напримѣръ, появленію бѣлаго или черпаго шара при выниманіи одного шара наугадъ изъ ящика содержащаго *a* бѣлыхъ и *в* черныхъ шаровъ) свойственны вѣроятности p и $q=1-p$; при s послѣдовательныхъ испытаніяхъ мыслимы всѣ возможныя сочетанія появленія событія A и B: мыслимо, что всѣ s испытаній дадутъ A, или $s-1$ испытаній дадутъ появленіе A и одно появленіе B; $s-2$ — появленіе A, 2 — появленіе B и т. д. Въ числѣ всѣхъ мыслимыхъ результатовъ будутъ и такіе, въ которыхъ отношеніе числа случаевъ наступленія A и числа случаевъ наступленія B пропорціонально ихъ вѣроятностямъ p и q, т. е. A повторится sp разъ и B sq разъ (или если sp и sq не цѣлыя числа, то ближайшее къ нимъ цѣлое число разъ). Можно доказать, что наиболѣе вѣроятнымъ изъ всѣхъ возможныхъ исходовъ оказывается именно это сочетаніе событій, и вѣроятность его (y_0) выражается величиной

$$y_0 = \frac{s!}{sp!\ sq!}\ p^{sp}\ q^{sq} = \frac{1}{\sqrt{2\pi spq}}\ ;$$

вѣроятность же всякаго иного результата, въ которомъ A повторяется $sp-i$ разъ и B $sq+i$ разъ (полагая, что уклоненіе i можетъ равняться любому числу отъ 1 до sp и отъ — 1 до —sq) выражается величиной

$$y_i = y_0 \times 2{\cdot}71828^{-i^2/2spq}$$

величина эта быстро убываетъ съ возростаніемъ i; т. е. чѣмъ больше сочетаніе числа случаевъ появленія A и B отличается отъ сочетанія наиболѣе вѣроятнаго, тѣмъ съ съ еще большей быстротой убываетъ вѣроятность такого сочетанія. Сумма вѣроятностей всѣхъ возможныхъ сочетаній должна равняться 1 или

$$y_0 \Sigma (2{\cdot}71828^{-i^2/2spq}) = 1 ,$$

принимая i послѣдовательно равнымъ отъ sp до —sq.

По мѣрѣ того, какъ число опытовъ s возрастаетъ, число всѣхъ возможныхъ сочетаній также возрастаетъ; а такъ сумма вѣроятностей всѣхъ сочетаній вообще остается всегда равной 1, то съ возрастаніемъ числа сочетаній на долю сочетаній крайнихъ, наиболѣе уклоняющихся отъ сочетанія наиболѣе вѣроятнаго, изъ этой 1 приходится все меньшая и меньшая часть; и напротивъ, на долю сочетанія наиболѣе вѣроятнаго и близкихъ къ нему изъ общей суммы приходится все бо́льшая часть. Благодаря этой особенности формулы, можно выбрать число случаевъ s столь большое, что вѣроятность появленія событій въ сочетаніяхъ наиболѣе вѣроятномъ (т. е. ps и qs) и близкихъ къ нему (т. е. отличающихся на сравнительно небольшую величину α, или $ps \pm \alpha$ и $qs \mp \alpha$) была какъ угодно близка къ 1; какъ равнымъ образомъ при достаточно большомъ s предѣлы уклоненія (или α) можно сдѣлать какъ угодно малыми, сравнительно съ ps и qs.

Такимъ образомъ теорема Бернулли можетъ быть выражена слѣдующимъ образомъ: при достаточно большомъ числѣ испытаній наибольшую вѣроятность, приближающуюся къ 1 (достовѣрности), представляетъ тотъ мыслимый исходъ испытаній, въ которомъ отношеніе числа случаевъ появленія событій отвѣчаетъ ихъ вѣроятностямъ или уклоняется отъ послѣднихъ на какъ угодно малую величину $\left(\alpha/_{s}\right)$. Выводъ этотъ примѣнимъ не только къ двумъ событіямъ, но и къ любому числу событій.

Законъ большихъ чиселъ Пуассона. Въ теоремѣ Бернулли предполагается, что вѣроятности наступленія событій остаются неизмѣнными въ теченіе всего ряда предполагаемыхъ испытаній. Пуассонъ устранилъ это ограничительное условіе и предположилъ, что вѣроятность событій можетъ измѣняться въ теченіе испытаній. Его постановка задачи сводится къ слѣдующему. Въ каждомъ изъ s испытаній должно произойти либо A, либо B; вѣроятность появленія событія A бываетъ либо p_1, либо p_2, либо p_3 . . . . . либо p_n (и соотвѣтственно событія B — q_1, q_2, q_3, ... q_n); какая именно изъ указанныхъ вѣроятностей будетъ дѣйствовать въ данномъ испытаніи, опредѣляется, въ свою очередь, вѣроятностью ея наступленія; такъ вѣроятность наступленія p_1 составляетъ g_1, вѣроятность наступленія p_2 составляетъ g_2 и т. д., причемъ, такъ какъ которое-либо $p_1, p_2 \ldots p_n$ во всякомъ случаѣ наступитъ, то

$$g_1 + g_2 \cdot \cdot + g_n = 1.$$

Перечисленныя условія могутъ быть наглядно пояснены слѣдующимъ воображаемымъ примѣромъ. Пусть будетъ N ящиковъ, изъ нихъ n_1 съ такимъ содержаніемъ бѣлыхъ и черныхъ шаровъ, которое сообщаетъ выходу бѣлаго вѣроятность p_1 и выходу чернаго—q_2; n_2 ящиковъ съ вѣроятностью выхода бѣлаго шара p_2 и чернаго q_2; n_3 съ вѣроятностями p_3 и q_3 и т. д.; пусть выборъ ящика производится передъ каждымъ испытаніемъ наугадъ, а затѣмъ изъ выбраннаго ящика наугадъ производится тиражъ шара; послѣ каждаго испытанія шаръ возвращается въ ящикъ, содержимое котораго перемѣшивается, дабы послѣдующій тиражъ не находился ни въ какой зависимости отъ исхода предшествующаго опыта, какъ равнымъ образомъ съ тою же цѣлью перемѣшиваются всѣ ящики. Опытъ повторяется s разъ. При этихъ условіяхъ вѣроятность выхода ящика съ отношеніемъ шаровъ$=p_1/_{q_1}$ составляетъ $\frac{n_1}{N}=g_1$; ящика съ отношеніемъ $p_2/_{q_2}$ $\frac{n_2}{N}=g_2$ и т. д., и

$$\frac{n_1}{N}+\frac{n_2}{N}+\frac{n_3}{N}+\ldots=1$$

или

$$g_1+g_2+g_3\ldots\ldots=1.$$

Въ подобномъ случаѣ общая вѣроятность появленія бѣлаго шара (p_0), согласно теоретическому исчисленію, составляетъ

$$p_0=g_1p_1+g_2p_2+\cdot\cdot+g_n p_n$$

и соотвѣтственно общая вѣроятность появленія чернаго шара (q_0) —

$$q_0=g_1q_1+g_2q_2+\ldots+g_n q_n.$$

При указанныхъ условіяхъ воображаемыхъ испытаній, если число испытанія s выбрано достаточно большое, наибольшую вѣроятность, близкую къ 1, представитъ (какъ и по теоремѣ Бернулли) тотъ исходъ, въ которомъ отношеніе числа случаевъ появленія A и B пропорціонально ихъ (среднимъ) вѣроятностямъ p_0 и q_0 или отличается отъ нихъ на какъ угодно малую величину. Отличіе отъ теоремы Бернулли состоитъ лишь въ томъ, что въ теоремѣ Бернулли мы имѣемъ дѣло съ элементарно—простой вѣроятностью, тогда какъ здѣсь съ средней (общей) вѣроятностью. Указанный выводъ въ формулировкѣ Пуассона носитъ въ теоріи вѣроятности названіе «закона большихъ чиселъ».

Такимъ образомъ по данной, заранѣе (a priori) извѣстной вѣроятности (будь она элементарной или средней) при достаточно большомъ числѣ испытаній можно съ большою вѣроятностью сказать, что при s опытахъ отношеніе числа (m) случаевъ появленія событія A къ числу (s) всѣхъ опытовъ равно вѣроятности событія A или весьма мало отъ нея отличается:

$$m = sp \text{ или } = sp \pm \alpha$$
$$m/s = p \text{ » } = p \pm \alpha/s;$$

Обратный выводъ, вѣроятность *a posteriori.*

отсюда можно сдѣлать и обратный выводъ: по данному исходу мыслимаго опыта, давшему m разъ событіе A на s испытаній, при достаточно большомъ s, съ большою вѣроятностью можно утверждать, что результатъ m/s либо равняется вѣроятности событія A, либо весьма мало отличается отъ этой вѣроятности. (Къ этому же результату можно прійти и на основаніи теоремы Байэса). Благодаря этому, даже не зная a priori вѣроятности событія, можно съ любымъ приближеніемъ опредѣлить ея величину по результату s опытовъ, если только s взять достаточно большимъ. Этотъ именно способъ апостеріорнаго опредѣленія вѣроятности на основаніи результата наблюденій даетъ возможность примѣнять построенія теоріи вѣроятности къ случаямъ дѣйствительныхъ явленій.

Примѣненіе къ дѣйствительнымъ явленіямъ.

Положенія теоріи вѣроятности, какъ было указано, основаны не на опытѣ—наблюденіи или испытаніи, а на извѣстныхъ апріорныхъ цѣлесообразно принятыхъ допущеніяхъ; если теорія вѣроятности ссылается на испытанія, то испытанія эти мыслимыя, воображаемыя и служатъ лишь для нагляднаго изображенія положеній теоріи, а не для ихъ опытнаго обоснованія. Для того, чтобы имѣть возможность приложить такого рода теоретическія положенія къ дѣйствительнымъ явленіямъ, необходимо, чтобы условія, при которыхъ наступаютъ эти явленія, отвѣчали допущеніямъ теоріи. Необходимо, чтобы въ основѣ достаточно большой совокупности случаевъ лежала комплексно-множественная причина или система такихъ причинъ, и чтобы условія явленій обезпечивали равновозможность и независимость дѣйствія всѣхъ моментовъ такой причины (или причинъ). Въ такомъ случаѣ мы вправѣ ожидать, что дѣйствительныя наблюденія дадутъ результаты, отвѣчающіе теоретическимъ выводамъ. Поэтому, если намъ извѣстенъ составъ моментовъ комплексно-множественной причины, мы ранѣе наблюденія можемъ предсказать наиболѣе вѣроятный результатъ достаточно большого числа наблюденій; разумѣется, дѣйствительная совокупность нѣкотораго большого числа наблюденій можетъ дать результатъ болѣе или менѣе уклоняющійся отъ наиболѣе вѣроятнаго,—и это также отвѣчаетъ теоріи, такъ какъ послѣдняя допускаетъ возможность всякаго сочетанія событій, въ томъ числѣ и такихъ сочетаній, которыя представляютъ крайнія уклоненія отъ наиболѣе вѣроятнаго сочетанія. Но каждое такое уклоненіе въ теоріи имѣетъ свою вѣроятность; поэтому если въ дѣйствительномъ наблюденіи мы получимъ совокупность съ сочетаніемъ событій, уклоняющимся отъ наиболѣе вѣроятнаго, мы вправѣ ожидать, что если мы наберемъ достаточное число другихъ совокупностей съ такимъ же числомъ наблюденій, то среди нихъ совокупности, дающія тѣ или иныя уклоненія отъ наиболѣе вѣроятнаго сочетанія, будутъ встрѣчаться такъ часто, какъ это отвѣчаетъ теоретической вѣроятности, сообразно съ величиной уклоненія. Если вѣроятность наступленія единичныхъ событій неизвѣстна намъ до наблюденія, мы можемъ судить о ея приближенной величинѣ съ опредѣленною вѣроятностью по исходу достаточно большого числа наблюденій.

Въ испытаніяхъ. Однако случаи, когда намъ было бы напередъ извѣстно, что условія наступленія событій отвѣчаютъ апріорнымъ допущеніямъ теоріи, возможны лишь при экспериментѣ. Такого рода экспериментами являются прежде всего испытанія, нарочно предпринимаемыя для опытнаго подтвержденія дедуктивныхъ расчетовъ теоріи; таковы опыты съ подбрасываніемъ монетъ, выниманіемъ наугадъ шаровъ и т. п. Сюда же могутъ быть отнесены искусственно устраиваемыя въ цѣляхъ забавы или съ коммерческимъ расчетомъ игры, какъ кости, рулетка, лотто.

Во всѣхъ подобныхъ случаяхъ мы имѣемъ такое нарочитое устройство эксперимента, которое должно осуществлять, насколько это технически возможно, теоретическія допущенія; мы имѣемъ, слѣдовательно, ту или иную заранѣе извѣстную систему комплексно-множественныхъ моментовъ съ возможнымъ обезпеченіемъ ихъ равновозможности и независимости. Такъ, при выниманіи наугадъ шара мы имѣемъ въ урнѣ опредѣленное количество шаровъ каждаго цвѣта; изъ нихъ долженъ быть вынутъ всякій разъ, допустимъ, только одинъ шаръ; равновозможность выхода всѣхъ шаровъ обезпечивается тѣмъ, что шары по своей формѣ, поверхности, вѣсу и другимъ качествамъ, отъ которыхъ могло бы зависѣть ихъ положеніе въ урнѣ, тождественны и различаются лишь по цвѣту, причемъ эта разница не должна быть ощутима для вынимающей руки; шары не имѣютъ никакого систематическаго расположенія въ урнѣ, и движенія вынимающей руки не могутъ имѣть какой бы то ни было систематической связи ни съ какимъ изъ свойствъ шара; наконецъ, независимость каждаго событія при послѣдовательныхъ опытахъ достигается тщательнымъ перемѣшиваніемъ шаровъ.

Опыты и наблюденія, производимые съ соблюденіемъ подобныхъ условій, показываютъ, что результаты испытаній весьма близко отвѣчаютъ теоретически ожидаемому результату.

Въ наблюденіяхъ.

Однако за предѣлами эксперимента, когда дѣло касается простого наблюденія наступленія событій въ естественныхъ условіяхъ дѣйствительности, мы находимся въ совершенно иномъ положеніи: намъ, во-первыхъ, никогда не извѣстна апріорная величина вѣроятности наступленія событій, и, во вторыхъ, мы вообще лишены возможности знать условія, въ которыхъ происходятъ явленія, съ такой точностью, чтобы судить, осуществляются ли, и насколько, въ этихъ условіяхъ допущенія теоріи. Единственнымъ источникомъ для сужденія какъ о соотвѣтствіи дѣйствительности теоретическимъ допущеніямъ, такъ и о величинѣ вѣроятности являются результаты наблюденія. При такихъ условіяхъ, для рѣшенія вопроса о возможности примѣненія положеній теоріи вѣроятности къ наблюдаемымъ явленіямъ, намъ остается прибѣгнуть къ гипотезѣ; если расчеты, основанные на принятой гипотезѣ, оправдываются наблюдаемыми соотношеніями явленій, гипотеза получаетъ оправданіе и приближается къ достовѣрности въ той мѣрѣ, въ какой факты совпадаютъ съ расчетами.

Такимъ образомъ, если мы имѣемъ совокупность достаточно большого числа случаевъ появленія тѣхъ или другихъ событій, и въ условіяхъ наступленія событій намъ неизвѣстны обстоятельства, которыя противорѣчили бы допущеніямъ теоріи вѣроятности, мы можемъ гипотетически принять, что условія совокупности отвѣчаютъ апріорнымъ допущеніямъ теоріи. Отсюда мы получаемъ возможность найти a posteriori на основаніи общаго результата всѣхъ наблюденій наиболѣе вѣроятное выраженіе вѣроятности, лежащей въ основѣ совокупности наблюденныхъ случаевъ. Исходя изъ полученной величины вѣроятности, мы можемъ провѣрить, насколько основанные на ней теоретическіе расчеты совпадаютъ съ дѣйствительными соотношеніями. Съ этою цѣлью мы можемъ прибѣгнуть къ различнымъ расчетамъ и въ томъ числѣ, напримѣръ, къ слѣдующему. Разбивъ всю совокупность случаевъ на достаточно большой рядъ частичныхъ совокупностей съ одинаковымъ, но все же немалымъ числомъ случаевъ (что возможно лишь при достаточно большомъ числѣ случаевъ въ общей совокупности), мы получимъ въ каждой частичной совокупности то или иное сочетаніе событій; если уклоненія этихъ сочетаній отъ того, которое должно быть наиболѣе вѣроятнымъ, согласно принятой вѣроятности, встрѣчаются сравнительно такъ часто, какъ это отвѣчаетъ теоретической вѣроятности каждаго уклоненія, сообразно съ его величиной, мы имѣемъ основаніе думать, что условія наблюдаемыхъ случаевъ построены сходно съ теоретическими допущеніями, и данная совокупность допускаетъ пріемы изученія, основанные на теоріи вѣроятности, а въ томъ числѣ и примѣненіе основного положенія теоріи—закона большихъ чиселъ.

Законъ большихъ чиселъ въ статистикѣ.

Сообразно съ указанными особенностями условій наблюденія, законъ большихъ чиселъ въ примѣненіи къ наблюденію получаетъ иное выраженіе, чѣмъ въ тѣхъ случаяхъ, когда возможность приложенія исчисленія вѣроятности извѣстна независимо отъ результатовъ наблюденія. Въ статистикѣ «законъ большихъ чиселъ» имѣетъ слѣдующее содержаніе:

если отношеніе числа случаевъ появленія событія къ числу всѣхъ случаевъ наблюденія, вычисленное на основаніи достаточно большой совокупности случаевъ, можетъ быть принято за апостеріорное выраженіе вѣроятности, то при достаточно большомъ числѣ новыхъ однородныхъ наблюденій съ наибольшей вѣроятностью слѣдуетъ ожидать такого отношенія числа случаевъ появленія событія къ числу всѣхъ наблюденій, которое отличается отъ принятой вѣроятности на какъ угодно малую величину.

Вѣроятность качественно-различныхъ событій.

Событія A и B, вызываемыя дѣйствіемъ комплексно-множественной причины, могутъ представлять собою явленія, различающіяся качественно или количественно. Въ случаѣ качественнаго различія, если явленіе A представляетъ собою появленіе какого-либо событія, то явленіе B имѣетъ значеніе непоявленія событія, или появленія какого-либо другого качественно-отличнаго явленія; напримѣръ, A можетъ означать случай вступленія въ бракъ, B—случай отсутствія этого событія, или А—рожденіе мальчика и В—рожденіе дѣвочки. Такъ какъ событія А и В обладаютъ опредѣленной вѣроятностью, то въ достаточно большомъ числѣ наблюденій обнаружатся всѣ тѣ правильныя соотношенія, какія свойственны вѣроятности, и съ тою характерной для вѣроятности особенностью, что эти соотношенія обнаруживаются лишь при большомъ числѣ случаевъ, тогда какъ каждый единичный случай подчиненъ своему особому стеченію обстоятельствъ. Такимъ образомъ получается рядъ качественно-отличныхъ событій, закономѣрность наступленія которыхъ обнаруживается лишь въ достаточно большой совокупности единичныхъ случаевъ; общимъ выраженіемъ этой закономѣрности является сводная величина—отношеніе числа случаевъ появленія одного или другого событія ко всему числу случаевъ—съ ея постоянствомъ, основаннымъ на законѣ большихъ чиселъ.

Вѣроятность количественно-различныхъ событій.

Событія А и В могутъ представлять какъ мы сказали, и количественныя явленія. Изъ всѣхъ возможныхъ явленій количественнаго характера остановимся на одномъ особомъ случаѣ и положимъ, что событія А и В имѣютъ значеніе элементарной ошибки.

Элементарная ошибка.

Элементарной ошибкой мы называемъ погрѣшность измѣренія или вообще наблюденія, обладающую слѣдующими свойствами: она случайна, т. е. уклоняетъ результатъ наблюденія отъ истинной величины какъ въ одну, такъ и въ другую сторону съ одинаковой или почти одинаковой вѣроятностью, чѣмъ отличается отъ ошибки систематической, уклоняющей результатъ наблюденія всегда въ одну и ту же сторону; она обладаетъ чрезвычайно малой величиной, сравнительно съ наблюдаемой или измѣряемой величиной, и зависитъ отъ неустранимыхъ несовершенствъ инструментовъ, пріемовъ наблюденія, органовъ чувствъ и отъ вліянія окружающихъ обстоятельствъ; этимъ случайная ошибка отличается отъ грубыхъ погрѣшностей, имѣющихъ сравнительно большую величину и зависящихъ отъ устранимыхъ недостатковъ наблюденія; наконецъ, элементарная ошибка—это ошибка, производимая въ наблюденіи каждой отдѣльной причиной неточности, а такъ какъ каждому наблюденію присущъ цѣлый рядъ такихъ источниковъ погрѣшностей и притомъ другъ отъ друга независимыхъ, то въ каждомъ единичномъ случаѣ наблюденія ошибки отдѣльныхъ источниковъ могутъ сочетаться другъ съ другомъ всѣми возможными способами.

Ряды случайныхъ уклоненій.

Для удобства изложенія пріймемъ, что абсолютная величина элементарной ошибка равняется 1; при равной вѣроятности ея значенія какъ положительнаго $+1$, такъ и отрицательнаго -1, т. е. при $p=1/2$ и $q=1/2$, и при допущеніи, скажемъ, 10 отдѣльныхъ источниковъ погрѣшностей мы получимъ слѣдующіе возможные результаты сочетаній элементарныхъ ошибокъ:

10, 8, 6, 4, 2, 0, —2, —4, —6, —8, —10

и число случаевъ каждаго сочетанія:

1, 10, 45, 120, 210, 252, 210, 120, 45, 10, 1.

Если допустимъ, что измѣряемая величина равна 100, то съ присоединеніемъ возможныхъ ошибокъ мы получимъ слѣдующій рядъ уклоняющихся другъ отъ друга результатовъ наблюденія съ соотвѣтствующей частотой каждаго результата:

возможные результаты	110,	108,	106,	104,	102,	100,	98,	96,	94,	92,	90
ихъ сравнительная частота	1,	10,	45,	120,	210,	252,	210,	120,	45,	10,	1.

Такимъ образомъ мы получаемъ схему совокупности, состоящей изъ отличныхъ другъ отъ друга количественно случаевъ; величина каждаго случая зависитъ отъ индивидуально присущаго ему стеченія обстоятельствъ, вся же совокупность представляетъ свою особую закономѣрность; выраженіемъ этой закономѣрности служитъ, прежде всего, сводная величина — въ данномъ случаѣ средняя ариѳметическая; въ ней погрѣшности отдѣльныхъ наблюденій взаимно уничтожаются и получается истинная величина наблюдаемаго случая; при этомъ такой результатъ долженъ обладать извѣстнымъ постоянствомъ, въ силу закона большихъ чиселъ, которому подчинены вѣроятности элементарной ошибки.

При равной вѣроятности уклоненія какъ въ положительную, такъ и отрицательную сторону, рядъ уклоненій и рядъ результатовъ измѣренія получается вполнѣ симметричный, и средняя ариѳметическая свободна отъ всякой погрѣшности; при неравной вѣроятности положительной и отрицательной ошибки, рядъ становится въ меньшей или большей степени асимметричнымъ, и средняя ариѳметическая снабжена средней погрѣшностью.

Ошибки и уклоненія. Элементарную ошибку мы можемъ разсматривать не только какъ результатъ погрѣшности наблюденія, но и какъ уклоненіе, въ дѣйствительности присущее самой измѣряемой величинѣ. Въ такомъ случаѣ количественный характеръ ряда остается неизмѣннымъ со всѣми свойственными ему сводными правильностями, измѣняется лишь логическое значеніе величинъ ряда. Въ этомъ случаѣ мы должны представить себѣ, что въ основѣ ряда лежитъ неизмѣнное дѣйствіе нѣкотораго постояннаго фактора, стремящагося въ каждомъ единичномъ случаѣ произвести соотвѣтствующій количественный результатъ; но одновременно въ совокупности случаевъ дѣйствуетъ нѣкоторое число дополнительныхъ, независимыхъ отъ основного и взаимно другъ отъ друга факторовъ, стремящихся произвести свойственный имъ количественный результатъ сравнительно малой величины положительнаго и отрицательнаго значенія съ равной или неравной вѣроятностью. При такихъ условіяхъ сводная величина средней будетъ стремиться обнаружить величину постояннаго фактора въ чистомъ видѣ при симметріи ряда, или съ той или иной примѣсью уклоненія при асимметріи.

Сводныя закономѣрности. Совокупности качественно или количественно различныхъ явленій мы можемъ разсматривать двояко. Во-первыхъ, мы можемъ разсматривать ихъ какъ самостоятельный объектъ изученія, безъ отношенія къ какимъ либо другимъ обстоятельствамъ, явленіямъ или моментамъ. Въ такомъ случаѣ цѣлью изученія является открытіе правильности наступленія различныхъ событій въ массѣ случаевъ, когда мы имѣемъ дѣло съ совокупностью качественно различныхъ случаевъ, или открытіе правильности появленія различныхъ величинъ, когда мы имѣемъ дѣло съ совокупностью количественно различныхъ случаевъ. Такъ напримѣръ, мы изучаемъ правильность отношенія числа рожденій мальчиковъ къ числу рожденій дѣвочекъ; или правильность распредѣленія варіацій количественнаго характера въ животномъ и растительномъ мірѣ и т. п. Въ подобныхъ случаяхъ совокупность явленій, обладающихъ сводными правильностями, служитъ самодавлѣющимъ предметомъ изученія, и цѣлью изученія являются тѣ сводныя закономѣрности, которыя свойственны именно самимъ явленіямъ, собираемымъ въ совокупности. Если каждая изъ изслѣдуемыхъ совокупностей относится къ опредѣленному мѣсту, времени, связана съ опредѣленнымъ родомъ явленій, связь эта либо, по возможности, устраняется, либо учитывается, какъ обстоятельство, вносящее ту или иную модификацію въ изучаемыя сводныя закономѣрности, или составляющее признакъ совокупности. Результатомъ подобнаго рода изученія являются сводныя закономѣрности наступленія или появленія качественныхъ или количественныхъ событій опредѣленнаго рода; закономѣрность появленія шаровъ разнаго цвѣта при выниманіи ихъ наугадъ, закономѣрность отношенія рожденій, закономѣрность появленія варіацій той или иной величины.

Сводные признаки. Во-вторыхъ, сводныя величины мы можемъ разсматривать въ связи съ опредѣленнымъ мѣстомъ и временемъ или родомъ предметовъ;

въ такомъ случаѣ сводныя величины являются признаками мѣста и времени, рода предметовъ или явленія, имѣющаго характеръ совокупности—онѣ становятся сводными *признаками;* если сводная величина состоитъ изъ качественно различныхъ событій, сводный признакъ получаетъ видъ относительной величины; если она состоитъ изъ количественно различныхъ явленій, сводный признакъ получаетъ видъ средней величины. Такимъ образомъ мы разсматриваемъ относительное число рожденій, смертныхъ случаевъ, браковъ, какъ характеристику того или другого мѣста, государства, или какъ характеристику мѣста, государства въ тѣ или иные промежутки времени; средній ростъ разсматриваемъ, какъ признакъ той или иной народности; среднюю продолжительность жизни, какъ характеристику разныхъ профессій и т. п.

Статическія сводныя величины.

Какъ въ первомъ случаѣ, когда совокупности изслѣдуются самостоятельно, такъ и во второмъ, когда они изслѣдуются въ связи или зависимости отъ другихъ явленій, мы имѣемъ дѣло съ сводными величинами въ ихъ статическомъ состояніи, и разсматриваемъ ихъ, какъ одно общее выраженіе нѣкотораго комплекса производящихъ ихъ условій. Правда, самыя условія могутъ мѣняться, такъ, напримѣръ, въ совокупности нѣкотораго количества тиражей шара можетъ измѣняться въ теченіе опыта вѣроятность появленія бѣлаго шара; но задачей изслѣдованія въ данномъ случаѣ является не установленіе зависимости сводныхъ величинъ отъ измѣненія тѣхъ или иныхъ факторовъ въ предѣлахъ данной совокупности, а установленіе общаго результата дѣйствія кякъ мѣняющихся, такъ и неизмѣнныхъ факторовъ. Всѣ факторы, какъ мѣняющіеся такъ и неизмѣнные, будучи заключены въ рамки данной совокупности, разсматриваются лишь какъ моменты образованія одной общей сводной величины, одной общей системы сводной правильности. Въ новой, другой совокупности дѣйствующіе факторы могутъ измѣниться и дать другой сводный результатъ, но и въ этой другой совокупности сводная величина точно также выражаетъ одинъ общій окончательный результатъ дѣйствія заключенныхъ въ совокупности факторовъ. Такимъ образомъ, напримѣръ, коэффиціентъ рождаемости представляетъ собою для совокупности, на основаніи которой онъ вычисленъ, или къ которой отнесенъ, одно общее выраженіе дѣйствія всѣхъ моментовъ въ предѣлахъ совокупности, какъ мѣняющихся, такъ и постоянныхъ, и не содержитъ въ себѣ выраженія способа измѣненія этой величины въ зависимости отъ какихъ-либо измѣненій производящихъ ее факторовъ.

Статическая зависимость. Мы можемъ сдѣлать предметовъ изученія и самый процессъ измѣненія сводной величины, въ зависимости отъ измѣненія какихъ-либо другихъ обстоятельствъ. Если это другое обстоятельство имѣетъ качественный характеръ, то вопросъ сводится къ установленію связи между сводной величиной и качественнымъ видовымъ признакомъ совокупности. Такъ напримѣръ, если вопросъ касается зависимости средней продолжительности жизни отъ профессіи, то общую совокупность умершихъ мы дѣлимъ на частныя совокупности по признаку профессіи и для каждой частной совокупности находимъ средній возрастъ умершихъ; различія этой величины для разныхъ совокупностей указываютъ на зависимость ея отъ даннаго признака профессіи. Здѣсь мы имѣемъ дѣло съ примѣненіемъ къ сводной величинѣ обычной индукціи, и пріемы статистическаго изслѣдованія необходимы лишь для полученія сводныхъ величинъ и для рѣшенія вопроса, насколько наблюдаемая разность двухъ сводныхъ величинъ выходитъ за допустимыя для среднихъ колебанія и указываетъ на наличность дѣйствительной разницы среднихъ.

Подобно тому, какъ мы дѣлимъ общую совокупность на частныя по качественному признаку, мы можемъ раздѣлить ее по количественному признаку и изслѣдовать измѣненія сводной величины въ зависимости отъ измѣненія количественнаго признака; такъ напримѣръ, мы можемъ изслѣдовать измѣненія средняго количества осадковъ въ зависимости отъ высоты станціи надъ уровнемъ океана.

Однако дѣло принимаетъ иной характеръ, когда количественный признакъ самъ обладаетъ своднымъ характеромъ, т. е., будучи различенъ въ индивидуальныхъ случаяхъ, обнаруживаетъ постоянство и правильность въ совокупности. Тогда возникаетъ вопросъ о зависимости между двумя признаками своднаго характера внутри одной и той же общей совокупности. Такого рода зависимость получаетъ вполнѣ статистическій характеръ: она незамѣтна въ единичныхъ случаяхъ и можетъ быть обнаружена лишь сопоставленіемъ единичныхъ случаевъ въ общей совокупности Представимъ себѣ, что между двумя явленіями, изъ которыхъ каждое можетъ принимать различныя значенія, существуетъ вполнѣ опредѣленная зависимость, выражаемая посредствомъ функціи

$$y=3+2x\ ;$$

тогда значеніямъ независимой перемѣнной x будутъ отвѣчать вполнѣ опредѣленныя значенія зависимой перемѣнной y, какъ то:

x	y
1	5
2	7
3	9
4	11

Если величина y подъ вліяніемъ случайныхъ обстоятельствъ подвергается колебаніямъ, напримѣръ, на единицу въ обѣ стороны (вмѣсто 5 — 4 и 6 и т. д.), то мы получимъ слѣдующіе случаи связи x и y:

x — y	x — y
1 — 4	3 — 8
1 — 6	3 — 10
2 — 6	4 — 10
2 — 8	4 — 12

Если же къ уклоненіямъ y присоединятся и уклоненія x (вмѣсто 1 — 0 и 2 и т. д.), то получимъ совокупность случаевъ связи, представленную въ слѣдующей таблицѣ:

При x=1	y=
0	4, 6
1	6, 8
2	4, 6, 8, 10,
3	6, 8, 10, 12.
4	8, 10.
5	10, 12.

Такого рода совокупность характеризуется той особенностью, что въ единичныхъ случаяхъ связь незамѣтна или отсутствуетъ, напримѣръ, для $x = 2$, y можетъ быть 4, 6, 8 или 10, и обратно, для $y = 6$, x можетъ имѣть значеніе 0, 1, 2 или 3; но зависимость обнаруживается какъ только мы сведемъ индивидуальные случаи въ частныя совокупности и послѣднія сопоставимъ въ одной общей совокупности, какъ это уже замѣтно въ представленной таблицѣ. Мы можемъ образовать частныя совокупности по одинаковымъ значеніямъ x и для каждой такой частной совокупности получить сводную величину y; или обратно — по одинаковымъ величинамъ y и для каждой совокупности y-а получить сводную величину x; въ первомъ и во второмъ случаѣ получимъ слѣдующіе ряды:

I		II	
x	y	x	y
0	5	1.0	4
1	7	1.5	6
2	7	2.5	8
3	9	3.5	10
4	9	4.0	12
5	11		

Обѣ таблицы могутъ быть сведены въ одну (комбинаціонную) таблицу; случаи соединенія x съ соотвѣтствующими y и обратно обозначены въ соотвѣтственныхъ клѣткахъ единицами, а среднія значенія x и y обозначены въ графахъ x_m и y_m

x \ y	4	6	8	10	12	y_m
0	1	1				5
1		1	1			7
2	1	1	1	1		7
3		1	1	1	1	9
4			1	1		9
5				1	1	11
x_m	1	1.5	2.5	3.5	4	

Въ тѣхъ случаяхъ, когда намъ даны подобнаго рода ряды, и не дана въ точномъ видѣ связывающая ихъ зависимость, задачей изслѣдованія является установленіе наиболѣе вѣроятной формы этой зависимости на основаніи эмпирическаго матеріала. Однимъ изъ способовъ рѣшенія такой задачи является опредѣленіе т. н. корреляціи, т. е. отысканіе наиболѣе вѣроятнаго выраженія связи между уклоненіями величинъ ряда.

Какъ зависимость сводной величины отъ качественныхъ признаковъ, такъ и сводная зависимость двухъ количественныхъ признаковъ представляютъ собою зависимости статическаго характера: въ первомъ случаѣ различія качественныхъ признаковъ не имѣютъ, сами по себѣ, динамическаго характера; во второмъ—количественныя различія признаковъ представляютъ собою не процессъ развитія какой-либо величины въ опредѣленномъ направленіи, а лишь случайныя уклоненія отъ сводной величины, имѣющей статическій характеръ съ точки зрѣнія данной совокупности.

Динамическая зависимость.

Отъ статической зависимости отличается зависимость динамическая; она представляетъ собою зависимость двухъ количественныхъ моментовъ, измѣняющихся въ опредѣленномъ направленіи, какъ напримѣръ, зависимость между урожаемъ и количествомъ удобренія. Если обѣ величины, измѣняясь въ опредѣленныхъ направленіяхъ, въ единичныхъ случаяхъ своего движенія подвергаются случайнымъ уклоненіямъ, зависимость получаетъ сводный характеръ и изслѣдованіе ея пріобрѣтаетъ статистическій характеръ. Задача изслѣдованія состоитъ здѣсь въ томъ, чтобы найти функцію, которая выражала бы связь эмпирическихъ перемѣнныхъ величинъ въ ея наиболѣе вѣроятномъ чистомъ видѣ, свободнымъ отъ случайныхъ уклоненій.

Характеръ эмпирическихъ сводныхъ цыфръ.

Пріемы статистическаго метода основаны на положеніяхъ теоріи вѣроятности и потому примѣненіе этого метода къ изслѣдованію дѣйствительныхъ явленій предполагаетъ, что въ изслѣдуемыхъ совокупностяхъ мы имѣемъ дѣло съ качественными или количественными событіями, появленіе которыхъ опредѣляется нѣкоторыми вѣроятностями. Какъ уже было упомянуто, доказать, что въ той или иной конкретной совокупности имѣются на лицо требуемыя теоріей условія независимости событій, равновозможности исходовъ, опредѣленнаго состава комплексно-дѣйствующихъ моментовъ, представляется вообще мало осуществимымъ; въ этомъ отношеніи мы принуждены довольствоваться гипотетическимъ допущеніемъ, что такія условія существуютъ, если этой гипотезѣ не противорѣчатъ какіе-либо извѣстные намъ факты, а результаты наблюденія ее подтверждаютъ. Но независимо отъ того, съ какой степенью увѣренности мы примѣняемъ къ той или иной совокупности пріемы, основанныя на положеніяхъ вѣроятности, съ достовѣрнымъ ли знаніемъ ихъ приложимости, или только гипотетически,—во всякомъ случаѣ логическое строеніе совокупности должно быть таково, чтобы формально она допускала примѣненіе понятій и категорій вѣроятности.

Нѣкоторыя совокупности, съ которыми имѣетъ дѣло статистика на практикѣ, дѣйствительно имѣютъ съ формальной стороны такое логическое строеніе. Такъ, если составляется совокупность индивидовъ, перешедшихъ 20-лѣтній возрастъ и устанавливается, что до момента наступленія 21 года x индивидовъ умерло, то отношеніе числа умершихъ къ начальной цыфрѣ живущихъ представляетъ собою выраженіе, съ формальной стороны имѣющее характеръ вѣроятности—вѣроятность для лица 20-лѣтняго возраста умереть въ теченіе ближайшаго года. При этомъ можетъ представляться сомнительнымъ, являются ли, напримѣръ, всѣ случаи равновозможными, будутъ ли и въ какой степени послѣдующіе случаи независимыми отъ предъидущихъ,—но это вопросы факта; они могутъ быть рѣшены положительно или отрицательно, и это уже доказываетъ, что съ чисто формальной стороны совокупность случаевъ можетъ допускать приложеніе понятій вѣроятности. Съ другой стороны, несомнѣнно, что въ основѣ такой совокупности лежитъ не одна простая элементарная вѣроятность, но для разныхъ группъ индивидовъ различная вѣроятность смерти, и что для нѣкоторыхъ индивидовъ вѣроятность эта мѣняется въ теченіе самого наблюденія. Однако и это обстоятельство не мѣшаетъ формальному примѣненію категорій вѣроятности, такъ какъ теорія также допускаетъ случаи совокупнаго дѣйствія разныхъ вѣроятностей и ихъ измѣненія. Въ такихъ случаяхъ сводная величина получаетъ лишь сложно-составной характеръ, и интересы теоріи и практики могутъ побуждать изслѣдователя искать способовъ разложенія общей совокупности на такія частныя, которыя давали бы болѣе простой видъ своднымъ величинамъ. То же самое можно сказать, напримѣръ, относительно изслѣдованія средняго роста индивидовъ какой-либо точно, по возможности, опредѣленной народности. Въ отличіе отъ предшествующаго случая, гдѣ мы имѣли дѣло съ качественными событіями, здѣсь вопросъ касается количественнаго явленія. Мы можемъ допустить, что съ признаками опредѣленной народности связанъ комплексъ причинъ, стремящихся привести ростъ индивида въ опредѣленный возрастъ его жизни къ извѣстной постоянной величинѣ, которая подвергается случайнымъ уклоненіямъ подъ вліяніемъ всевозможныхъ сочетаній большого числа независимыхъ факторовъ; быть можетъ, на самомъ дѣлѣ мы имѣемъ дѣло не съ одной постоянно дѣйствующей величиной, а нѣкоторымъ ихъ количествомъ, съ своей особой вѣроятностью для каждой. Все это дѣлаетъ сводную величину средняго роста выраженіемъ чрезвычайно сложнаго результата дѣйствія многихъ фактически неуловимыхъ причинъ, но съ точки зрѣнія формальной не препятствуетъ представленію, что индивидуальныя различія величинъ являются результатомъ сочетаній комплексно-множественной системы случайныхъ уклоненій.

Однако не всѣ случаи совокупностей и сводныхъ величинъ, къ которымъ на практикѣ прилагаются статистическіе пріемы, имѣютъ такой формальный характеръ. Такого рода цыфры, какъ коэффиціентъ рождаемости, смертности, браковъ; процентное отношеніе числа самоубійствъ и т. п., не отвѣчаютъ по своему составу понятію вѣроятности. Коэффиціентъ смертности, если онъ представляетъ собою отношеніе числа умершихъ въ промежутокъ одного года къ цыфрѣ населенія середины этого года, не представляетъ понятія вѣроятности по самому характеру своего строенія, такъ какъ числитель — число умершихъ, не умѣщается цѣликомъ въ той совокупности живущихъ, которая помѣщена въ знаменателѣ. Такого рода коэффиціенты представляютъ лишь болѣе или менѣе сложную функцію вѣроятности, и для полученія послѣдней въ чистомъ видѣ необходимо частично измѣнить составъ сравниваемыхъ совокупностей. Примѣненіе статистическихъ пріемовъ изслѣдованія къ такого рода цыфрамъ имѣетъ въ большей или меньшей степени условный, формальный характеръ.

Наконецъ, въ статистической практикѣ имѣются и такіе случаи среднихъ величинъ и отношеній, гдѣ эти пріемы вычисленія представляютъ собою просто ариѳметическій способъ сведенія ряда различныхъ величинъ или качественныхъ явленій къ одному удобному выраженію. Таковы, напримѣръ, отношеніе площади территоріи къ числу жителей, показательныя цѣны ряда товаровъ разнаго рода и т. п.

---

3. Среднія величины.

Варіанты, частоты. Средней величиной вообще называется величина, не выходящая за предѣлы ряда данныхъ величинъ. Величины ряда обозначаютъ количественный признакъ (мѣру, число) явленія, получающій въ единичныхъ случаяхъ различное значеніе, варіирующій; какъ, напримѣръ, число головъ скота или десятинъ земли крестьянскаго двора; поэтому каждую величину ряда мы можемъ назвать *варіантой*. Число разъ, которое варіанта повторяется въ рядѣ, называется ея *частотою* (вѣсомъ). Варіанты и соотвѣтствующія имъ частоты образуютъ два ряда: первый даютъ варіанты, расположенныя въ порядкѣ возрастанія или убыванія ихъ значенія, второй — частоты, соотвѣтствующія каждому значенію варіанты, напримѣръ:

v (Число лошадей въ хозяйствѣ двора).	f Число дворовъ. *
0	6 321
1	8 834
2	1 357
3	215
4 и болѣе	90

Варіанты могутъ представлять либо цѣлыя числа, отстоящія другъ отъ друга не менѣе, какъ на 1, или же значенія ихъ могутъ падать и въ любое мѣсто промежутка между двумя цѣлыми числами, образуя какъ бы непрерывный рядъ. Но даже въ тѣхъ случаяхъ, когда первоначальныя варіанты имѣютъ видъ цѣлыхъ чиселъ,

*) Балахнинскій у. Нижег. губ., Матеріалы и пр. 1896.

среднія величины могутъ лежать въ промежуткѣ; это даетъ основаніе принимать всякій вообще рядъ варіантъ за непрерывный.

Интервалы. Когда число отличныхъ другъ отъ друга значеній варіантъ велико, то по необходимости приходится дѣлить рядъ на интервалы: отъ — до; независимо отъ того, дѣленіе на интервалы вообще обнаруживаетъ характеръ ряда и облегчаетъ вычисленія. Интервалы могутъ быть равные и неравные. Примѣромъ равныхъ интерваловъ можетъ служить общепринятое въ статистикѣ населенія дѣленіе возраста на 5 или 10-лѣтнія группы; примѣромъ неравныхъ интерваловъ будетъ дѣленіе, скажемъ, отъ 0 до 6; отъ 6 до 7, отъ 7 до 12 и т. п. Правильнымъ и желательнымъ является дѣленіе на равные интервалы, во-первыхъ, потому что составить представленіе о томъ, въ какомъ интервалѣ вмѣщается больше и въ какомъ меньше варіантъ, возможно только тогда, когда самая мѣра интерваловъ не мѣняется; во-вторыхъ, потому что равные интервалы облегчаютъ вычисленія, въ томъ числѣ въ особенности интерполяцію, посредствомъ которой возможно опредѣлить частоту въ любыхъ предѣлахъ варіантъ, въ большинствѣ случаевъ съ достаточною точностью. Дѣленіе на неравные интервалы вызывается всегда какой-либо одной опредѣленной цѣлью; напримѣръ, при учетѣ возраста—необходимостью выдѣлить рабочій, полурабочій возрастъ, или школьный возрастъ и т. п. Если матеріалъ имѣетъ общее значеніе, то приведеніе его въ видъ, служащій только одной цѣли и малопригодный для всѣхъ другихъ, представляется крайне нежелательнымъ.

При дѣленіи ряда на интервалы существеннымъ вопросомъ является величина интервала. Отъ величины интервала въ значительной степени зависитъ общая картина распредѣленія частотъ и въ той или другой степени величина вычисляемыхъ результатовъ. Такъ, напримѣръ, число браковъ на 10 000 православнаго населенія въ 50 губерніяхъ Россіи (1903 Ц. С. К.) даетъ слѣдующее распредѣленіе при разныхъ интервалахъ:

v I	f	v II	f
до 56	6	до 56	1
56— 66	3	56— 71	6
66— 76	4	71— 86	9
76— 86	8	86—101	21
86— 96	14	101—116	13
96—106	17		
106—116	3		

(въ первомъ случаѣ при вычисленіи по интерваламъ средняя ариѳметическая=89·8, во второмъ 90·2, при точномъ вычисленіи 89·5).

Къ сожалѣнію, вопросъ мало разработанъ *, и на практикѣ выборъ величины интерваловъ зависитъ отъ произвола. Здѣсь можно сдѣлать лишь одно практическое указаніе. Цѣлью сведенія въ интервалы является, прежде всего, необходимость получить представленіе объ общемъ характерѣ распредѣленія варіантъ—ихъ сравнительной частотѣ; поэтому и размѣры интервала должны быть приноровлены къ тому, чтобы обнаружить эту правильность; если при этомъ среднеариѳметическая каждаго интервала мало уклоняется отъ варіанты центральной точки интервала, принятое распредѣленіе обезпечиваетъ достаточную точность вычисленій. Въ приведенномъ выше примѣрѣ числа браковъ первое распредѣленіе, по указаннымъ здѣсь соображеніямъ, заслуживаетъ предпочтенія.

Распредѣленіе частотъ. Въ очень большомъ числѣ случаевъ, раздѣливъ рядъ на равные интервалы, при достаточно большомъ числѣ наблюденій мы получаемъ очень

* Предлагаемые теоретическіе критеріи отчасти недостаточно обоснованы, главнымъ образомъ требуютъ практической провѣрки. Ср. А. Леонтовичъ, Пособіе, ч. 1, стр. 27—30.

характерное распредѣленіе частоты варіантъ, а именно, наибольшее число варіантъ скопляется въ среднихъ интервалахъ ряда и затѣмъ частоты правильно убываютъ въ обѣ стороны. Иногда распредѣленіе почти симетрично, но чаще всего обнаруживаетъ большую или меньшую асиметрію. Подобное распредѣленіе является весьма обычнымъ какъ въ области явленій естественныхъ (въ ботаникѣ, зоологіи, антропологіи, біологіи), такъ и въ явленіяхъ хозяйства и общественной жизни вообще; оно органически связано съ варіаціями естественныхъ признаковъ, функцій и сознательныхъ дѣйствій человѣка *.

Изъ числа возможныхъ среднихъ величинъ нѣкоторыя имѣютъ особое практическое и теоретическое значеніе. Къ числу ихъ прежде всего относится средняя ариѳметическая, издавна занявшая прочное положеніе въ статистикѣ; съ тѣхъ поръ, какъ вниманіе было обращено на свойственную многимъ рядамъ асиметрію, получили значеніе и двѣ другія среднія: средняя наибольшей частоты (нормальная, обычная или мода) и средняя центральная или медіана.

Средняя ариѳметическая.

Средняя ариѳметическая, которую будемъ обозначать Ме, получается отъ дѣленія суммы варіантъ на ихъ число; когда варіанты имѣютъ различныя частоты, нахожденіе Ме сводится къ умноженію каждой варіанты на ея частоту, суммированію произведеній и дѣленію результата на сумму частотъ. Въ тѣхъ случаяхъ, когда варіанты даны по интерваламъ съ общей частотой для всего интервала, мы можемъ замѣнить распредѣленныя по всему интервалу въ неизвѣстномъ порядкѣ варіанты одной центральной варіантой, составляющей середину интервала, и къ ней отнести частоту интервала; если интервалы не велики и варіанты заполняютъ интервалъ болѣе или менѣе равномѣрно, результатъ получается почти тождественный съ

* См. мою работу о „сводныхъ признакахъ", 1910, часть 3.

тѣмъ, какъ если мы каждую варіанту будемъ умножать на ея частоту; но и при довольно большихъ интервалахъ и даже неравномѣрномъ распредѣленіи варіантъ, разница въ результатахъ по большей части такова, что при вычисленіяхъ для практической цѣли значенія не имѣетъ. Поэтому даже въ тѣхъ случаяхъ, когда частоты каждой варіанты извѣстны, цѣлесообразно примѣнять сокращенный способъ вычисленія по интерваламъ. Такъ въ примѣрѣ I на стр. 52 средняя, вычисленная по интерваламъ, сравнительно даже большимъ въ 10 единицъ, даетъ разницу отъ точнаго вычисленія лишь на 0·3. Если бы въ каждомъ интервалѣ мы взяли среднеариѳметическую всѣхъ варіантъ этого интервала и къ ней отнесли общую частоту интервала, то результатъ долженъ былъ бы получиться тожественнымъ съ обычнымъ способомъ вычисленія средней; поэтому чѣмъ меньше отличается варіанта, отвѣчающая серединѣ интервала, отъ среднеариѳметической варіанты этого интервала, тѣмъ ошибка общей средней будетъ меньше.

346 дворовъ Арефинской в. Яр. губ. по количеству земли распредѣляется слѣд. образомъ:

Десятинъ.	Центр. варіанта.	Дворовъ.	
v	v_i	f	$v_i \cdot f$
0— 5	2·5	29	72·5
5—10	7·5	101	757·5
10—15	12·5	92	1150
15—20	17·5	68	1190
20—	22·5	56	1260
		346	4430·0

$$Me = \frac{4430}{346} = 12·8$$

Способъ моментовъ. Такъ какъ при вычисленіи средней и другихъ величинъ приходится умножать иногда большія цифры частотъ на нерѣдко дробныя величины варіантъ, то для упрощенія вычисленій удобно пользоваться системой т. н. моментовъ (Pearson). Пріймемъ центральную варіанту какого-нибудь интервала—удобнѣе всего одного изъ среднихъ—за исходную точку и обозначимъ ее 0; разстояніе между центральными варіантами двухъ сосѣднихъ интерваловъ пріймемъ равнымъ 1, и разстоянія вверхъ отъ исходной точки будемъ считать отрицательными, внизъ—положительными. Тогда величина центральной варіанты каждаго интервала замѣниться цѣлымъ и небольшимъ числомъ, указывающимъ разстояніе ея отъ исходной точки. Умножая каждое такое разстояніе на соотвѣтствующую частоту, складывая произведенія и дѣля результатъ на сумму частотъ, мы получимъ среднеариѳметическое разстояніе, которое носитъ названіе перваго момента и обозначается ν_1. На этомъ разстояніи отъ исходной точки должна лежать среднеариѳметическая варіанта. Такъ какъ разстояніе между центральными точками интерваловъ принято нами за 1, а на самомъ дѣлѣ оно равно величинѣ интервала или i, то достаточно первый моментъ умножить на величину интервала и прибавить (или отнять, смотря съ какой стороны нулевой точки приходится первый моментъ) къ величинѣ варіанты, принятой за исходную точку, и мы получимъ среднюю ариѳметическую. Ходъ вычисленій очень легко уясняется на примѣрѣ.

v	v_i	f	z	f×z
0—5	2·5	29	—2	—58
5—10	7·5	101	—1	—101
10—15	12·5	92	0	
15—20	17·5	68	1	68
20—25	22·5	56	2	112
		346		+180
				—159
				21

$$\nu_1 = {}^{21}/_{346} = 0{\cdot}06$$

$$Me = v_0 + \nu_1 \times i = 12{\cdot}5 + 0{\cdot}06 \times 5 = 12{\cdot}8.$$

Доказательство. $\dfrac{f_1\,z_1+f_2\,z_2+\ldots+f_n\,z_n}{f_1+f_2\ldots+f_n}=$

$$\frac{f_1\,(v_1-v_0)+f_2\,(v_2-v_0)+\ldots+f_n\,(f_n-v_0)}{f_1+f_2+\ldots\ldots+f_n}=$$

$$\frac{f_1\,v_1+f_2\,v_2+\ldots+f_n\,v_n\cancel{+}v_0(f_1+\ldots+f_n)}{f_1+f_2+\ldots+f_n}=\frac{S\,fv}{S\,f}-v_0\frac{S\,f}{S\,f}$$

$\nu_1=Me-v_0$ или $v_0+\nu_1=Me$.

Если каждое разстояніе z возвести въ квадратъ и произвести тѣ же дѣйствія, какъ при полученіи перваго момента, или, что то же, каждое $f\times z$ помножить еще разъ на z, суммировать и раздѣлить на сумму f, то мы получимъ второй моментъ или ν_2; если взять z^3, получимъ 3-ій моментъ и т. д. Моменты второй и т. д. степени понадобятся при дальнѣйшемъ изложеніи.

Среднеквадратическое уклоненіе. Разница между каждой варіантой и среднеариѳметической называется уклоненіемъ (ошибкой); сумма всѣхъ уклоненій равна 0; если взять уклоненія въ ихъ абсолютныхъ величинахъ, безъ знаковъ, то среднеариѳметическая всѣхъ уклоненій даетъ т. н. среднее уклоненіе (обозначается d); если же каждое уклоненіе возвести въ квадратъ, взять среднеариѳметическое полученныхъ квадратовъ и изъ результата извлечь квадратный корень, получимъ среднеквадратическое уклоненіе (обозначается σ). Если уклоненіе отъ средней обозначимъ черезъ x, то указанныя величины выразятся слѣдующимъ образомъ:

$$d=\frac{S\,fx}{S\,f}\qquad\sigma=\sqrt{\frac{S\,fx^2}{S\,f}}\,.$$

Вычисленіе среднеквадратическаго уклоненія чрезвычайно упрощается, если пользоваться способомъ моментовъ; а именно, среднеквадратическое уклоненіе, обозначаемое въ системѣ моментовъ $\sqrt{\mu_2}$, получается изъ:

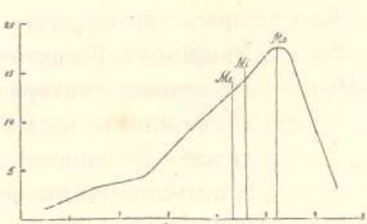
$$\sigma^2=\mu_2=\nu_2-\nu_1{}^2.$$

Пріемъ доказательства тотъ же, что идея нахожденія средней по способу моментовъ (56 стр.).

Среднеквадратическое уклоненіе является весьма характернымъ для средней ариѳметической; если поставить задачу: найти для даннаго ряда такую величину, для которой сумма квадратовъ уклоненій ея отъ данныхъ величинъ получаетъ наименьшее значеніе, то рѣшеніе задачи приводитъ къ среднеариѳметической.

Среднеквадратическое уклоненіе служитъ мѣрой точности ряда : чѣмъ оно меньше, тѣмъ рядъ въ общемъ тѣснѣе расположенъ около средней; той же цѣли можетъ служить и среднее уклоненіе, которое $=0{\cdot}797885\delta$.

Медіана, мода. При симетричномъ расположеніи частотъ среднеариѳметическая лежитъ въ центрѣ ряда и вмѣстѣ съ тѣмъ является варіантой съ наибольшей частотой. Если однако распредѣленіе частотъ несиметрично, то всѣ три варіанты расходятся: центральная варіанта, медіана, остается въ центральной точкѣ ряда, имѣя съ каждой стороны равное число варіантъ (равное число частотъ); мода располагается въ мѣстѣ наибольшаго скопленія варіантъ (мѣсто наибольшей частоты), а среднеариѳметическая отодвигается отъ медіаны въ противоположномъ направленіи сравнительно съ модой (на нѣсколько меньшее разстояніе). Если варіанты отложить на оси абсцисъ, а частоты изобразить въ видѣ ординатъ, то сравнительное расположеніе ихъ можно видѣть на прилагаемомъ чертежѣ.



Медіана. Медіана дѣлитъ рядъ на двѣ равныя части. Если каждая варіанта ряда, расположеннаго въ восходящемъ или нисходящемъ порядкѣ, имѣетъ особую величину, то, при нечетномъ числѣ варіантъ, медіаной будетъ серединная варіанта, при нечетномъ—среднеариѳметическая изъ двухъ сосѣднихъ варіантъ, лежащихъ по серединѣ ряда. Если рядъ раздѣленъ на интервалы, то значеніе медіаны можно опредѣлить приблизительно, при помощи вычисленія или графически. Начиная съ одного или другого края ряда, складываемъ частоты слѣдующихъ другъ за другамъ интерваловъ; когда дойдемъ до такой суммы, что съ прибавленіемъ частоты слѣдующаго интервала сумма перешагнула бы за половину всего числа ~~частотъ ряда,~~ мы останавливаемся на этой суммѣ (недостающей до $^1/_2$ всѣхъ частотъ): въ слѣдующемъ за ней интервалѣ должна находиться медіана. Чтобы найти ея положеніе отъ начала интервала, пріймемъ во вниманіе, какая часть интервала приходится на 1 частоты въ немъ, т. е. если въ интервалѣ, гдѣ должна находиться медіана, частота $= f_i$, то дѣлимъ i на f_i; частное множимъ на число единицъ, которое не хватаетъ ранѣе полученной суммѣ до $^1/_2$ всего числа частотъ; полученный результатъ представляетъ ту пропорціональную часть интервала, которую надо прибавить къ началу интервала, передъ которымъ остановилось суммированіе, чтобы получить медіану. Вычисленія показаны на примѣрѣ:

v	f	sf
— 56	1	1
56— 66	3	4
66— 76	4	8
76— 86	8	16
86— 96	14	
96—106	17	
106—116	3	
	50	

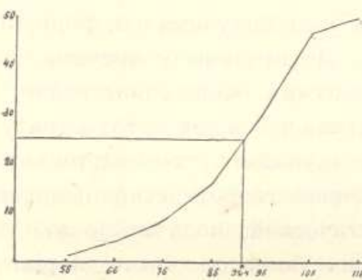
$\frac{1}{2} \times 50 = 25$; $25 - 16 = 9$, т. е. не хватаетъ до $^1/_2$ всего числа частотъ 9 варіантъ. Въ интервалѣ 86 - 96 на 10 единицъ интервала приходится 14 варіантъ; на 1 варіанту приходится $\frac{10}{14}$ единицъ интервала, и на 9 недостающихъ варіантъ:

$$\frac{10 \times 9}{14} = 6.4\ ;$$

слѣдовательно 25 варіантъ исчерпываются на варіантѣ

$$86+6{\cdot}4=92{\cdot}4=\mathrm{Mi}\ .$$

Этотъ же самый способъ вычисленія можетъ быть изображенъ графически. Нанесемъ интервалы на ось абсцисъ; послѣдовательныя же суммы изобразимъ въ конечной точкѣ каждого интервала въ видѣ ординатъ, верхніе концы которыхъ соединимъ кривой; взявъ на оси ординатъ точку, отвѣчающую $^1/_2$ всего числа частотъ, проводимъ линію паралельно оси абсцисъ; она встрѣтитъ кривую какъ разъ въ той точкѣ, которая отвѣчаетъ суммѣ половины всего числа частотъ; перпендикуляръ, опущенный изъ этой точки кривой на ось абсциссъ, покажетъ значеніе медіаны.



Медіана характеризуется тѣмъ свойствомъ, что сумма уклоненій ея отъ каждой величины ряда въ 0-ой степени равна 0 (т. к. всякое число въ 0-ой степени =1, и уклоненій въ одну и въ другую сторону равное число), сумма же уклоненій, взятыхъ въ ихъ абсолютной величинѣ, въ первой степени имѣетъ наименьшее значеніе.

Мода. Положеніе моды (Mo) опредѣляется мѣстомъ наибольшаго скопленія варіантъ (максимальной ординатой); поэтому положеніе ея выясняется лишь послѣ сведенія варіантъ въ интервалы; если интервалы взяты малые, то часто получаются двѣ точки скопленія варіантъ, раздѣленныя промежуткомъ, который исчезаетъ съ расширеніемъ интерваловъ, если послѣднее допускается характеромъ ряда. Теоретически положеніе моды, при неслишкомъ большой ассиметріи, опредѣляется слѣдующимъ образомъ: мода должна находиться приблизительно въ точкѣ, отстоящей отъ медіаны на тройное разстояніе медіаны отъ среднеариѳметической и въ сторону, противоположную среднеариѳметической, или

$$Mo = Me - 3\ (Me - Mi)\ ;$$

такъ, если въ предшествующемъ примѣрѣ $Me = 89{\cdot}8$ и $Mi = 92{\cdot}4$, то

$$Mo = 89{\cdot}8 - 3 \times (89{\cdot}8 - 92{\cdot}4) = 98{\cdot}6.$$

Болѣе точное опредѣленіе положенія моды (также какъ и медіаны) можетъ быть получено изъ формулы кривой ряда.

Нѣкоторыя замѣчанія о примѣненіи среднихъ.

До недавняго времени средняя ариѳметическая была единственной средней, примѣнявшейся въ статистикѣ, и въ настоящее время она занимаетъ господствующее положеніе. Вполнѣ опредѣленное теоретическое основаніе примѣненіе средней ариѳметической получаетъ въ тѣхъ случаяхъ, когда достаточно большое число варіантъ образуетъ вполнѣ симетрическій рядъ съ постепенно наростающими въ направленіи къ серединѣ частотами. Ряды такого строенія получаются при условіи, если къ нѣкоторой постоянной абсолютной величинѣ въ единичныхъ случаяхъ присоединяются всевозможныя сочетаніи двухъ (или болѣе) уклоненій или ошибокъ, обратныхъ по знаку, но одинаковыхъ по абсолютной величинѣ и съ равной вѣроятностью. Такъ, двѣ ошибки $+1$ и -1, съ вѣроятностью каждая равной $1/2$, сочетаясь по шести разъ въ каждомъ наблюденіи, даютъ слѣдующій рядъ уклоненій:

Сочетаніе уклоненій		Результатъ сочетанія	Его вѣроятность по формулѣ
$+1$	-1		$\left(\frac{1}{2}+\frac{1}{2}\right)^6$
6 разъ	0 разъ	6	$\left(\frac{1}{2}\right)^6 = {}^1/_{64}$
5 »	1 »	4	$6\times\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right) = {}^6/_{64}$
4 »	2 »	2	$15\times\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^2 = {}^{15}/_{64}$
3 »	3 »	0	$20\times\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^3 = {}^{20}/_{64}$
2 »	4 »	—2	$15\times\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^4 = {}^{15}/_{64}$
1 »	5 »	—4	$6\times\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^5 = {}^6/_{64}$
0 »	6 »	—6	$\left(\frac{1}{2}\right)^6 = {}^1/_{64}$

Или:

Величина уклоненія.	Число случаевъ.
6	1
4	6
2	15
0	20
—2	15
—4	6
—6	1
	64

Такъ какъ уклоненія располагаются симетрично, то въ средней ариѳметической они взаимно уничтожаются; и въ результатѣ получается постоянная величина, свободная отъ всякой ошибки. Поэтому всякій разъ, когда эмпирическій рядъ приближается къ подобному строенію, средне-ариѳметическая можетъ быть разсматриваема, какъ наиболѣе вѣроятное выраженіе истинной величины, лежащей въ основѣ ряда. Въ этомъ именно обстоятельствѣ лежитъ главное теоретическое обоснованіе примѣненія ариѳметической средней.

При ассиметріи. Если однако ошибки, равныя по абсолютной величинѣ, но съ обратными знаками имѣютъ разную вѣроятность, напримѣръ, ${}^1/_3$ и ${}^2/_3$, результаты сочетаній получаются тѣ же, но вѣроятность ихъ будетъ иная и рядъ получаетъ ассиметрію:

Сочетанія уклоненій		Ихъ результатъ	Ихъ вѣроятность по формулѣ
$+1$	-1		$(^1/_3+^2/_3)^6$ L3
6 разъ	0 разъ	6	$(^1/_3)^6 = {}^1/_{729}$
5 »	1 »	4	$6\times(^1/_3)^5(^2/_3) = {}^{12}/_{729}$
4 »	2 »	2	$15\times(^1/_3)^4(^2/_3)^2 = {}^{60}/_{729}$
3 »	3 »	0	$20\times(^1/_3)^3(^2/_3)^3 = {}^{160}/_{729}$
2 »	4 »	—2	$15\times(^1/_3)^2(^2/_3)^4 = {}^{240}/_{729}$
1 »	5 »	—4	$6\times(^1/_3)\,(^2/_3)^5 = {}^{192}/_{729}$
0 »	6 »	—6	$(^2/_3)^6 = {}^{64}/_{729}$

Или:

Уклоненія.	Число случаевъ
6	1
4	12
2	60
0	160
—2	240
—4	192
—6	64
	729

Въ этомъ случаѣ среднеариѳметическая не будетъ свободна отъ ошибки и представляетъ истинную величину, снабженную среднеариѳметической всѣхъ уклоненій, въ данномъ примѣрѣ равной —2. Въ данномъ случаѣ система уклоненій, благодаря неравной вѣроятности положительныхъ и отрицательныхъ уклоненій, систематически уклоняетъ единичные случаи отъ истинной величины, является такимъ образомъ также систематически дѣйствующей причиной, и средняя ариѳметическая является выраженіемъ истинной величины съ прибавкой наиболѣе вѣроятнаго результата систематически дѣйствующей системы уклоненій.

То же будетъ и въ томъ случаѣ, если при равныхъ вѣроятностяхъ уклоненій ($p = {}^1/_2$, $q = {}^1/_2$) мы возьмемъ разной величины уклоненія $+1$ и -2, или соединимъ различную вѣроятность съ разной величиной уклоненій, напримѣръ, уклоненіе $+1$ съ вѣроятностью $p = {}^1/_3$ и уклоненіе -2 съ вѣроятностью $q = {}^2/_3$; въ послѣднемъ случаѣ мы получимъ слѣдующій рядъ:

уклоненія:	6	3	0	—3	—6	—9	—12
частоты:	1	12	60	160	240	192	64

Въ немъ наиболѣе вѣроятное уклоненіе $= -6$, среднеариѳметическая же изъ всѣхъ уклоненіе тоже даетъ -6, такимъ образомъ среднеариѳметическая ряда даетъ выраженіе истинной величины съ наиболѣе вѣроятнымъ сочетаніемъ уклоненій.

Примѣчаніе. Величина наиболѣе вѣроятнаго уклоненія опредѣляется наиболѣе вѣроятнымъ сочетаніемъ уклоненій, которое, по теоремѣ Бернулли, отвѣчаетъ члену бинома

$$p^{sp}\, q^{sq} \;;$$

если же sp не цѣлое число, то число (μ) повтореній p представляетъ цѣлое число, опредѣляемое неравенствомъ:

$$sp - q < \mu < sp + p \;.$$

Въ данномъ случаѣ, при $s = 6$, $p = {}^1/_3$ и $q = {}^2/_3$

$$({}^1/_3)^2\ ({}^2/_3)^4 \;,$$

т. е. уклоненіе, вѣроятность котораго $= {}^1/_3$, повторяется 2 раза, уклоненіе съ вѣроятностью ${}^2/_3$ повторяется 4 раза, что даетъ

$$2 \times 1 + 4 \times (-2) = -6.$$

Положеніе же средней ариѳметической можетъ быть найдено по формулѣ

$$m = \frac{s}{2}\left(1 - \frac{p-q}{p+q}\right) + 1 \;,$$

гдѣ m обозначаетъ разстояніе (а не значеніе варіанты), на которомъ находится среднеариѳметическая варіанта, считая отъ начала ряда; такъ въ данномъ примѣрѣ

$$m = \frac{6}{2}\left(1 - \frac{{}^1/_3 - {}^2/_3}{{}^1/_3 + {}^2/_3}\right) + 1 = 5\,,$$

т. е. среднеариѳметическая представляетъ 5-ый членъ ряда отъ начала, или — 6.

Однако рядъ можетъ быть построенъ такъ, что наиболѣе вѣроятное уклоненіе не совпадаетъ съ среднеариѳметической всѣхъ уклоненій; такъ, если уклоненіямъ $+1$ и -1 отвѣчаютъ вѣроятности $p = {}^1/_8$ и $q = {}^7/_8$, то наиболѣе вѣроятное цѣлое число разъ повтореній $+1$ будетъ опредѣляться неравенствомъ:

$$ps - q < \mu < ps + p$$

или

$$\frac{6}{8} - \frac{7}{8} < \mu < \frac{6}{8} + \frac{7}{8}\,,$$

т. е.

$$\mu = 0$$

и наиболѣе вѣроятный членъ выразится въ видѣ $({}^1/_8)^0({}^7/_8)^6$; соотвѣтствующее этому члену наиболѣе вѣроятное уклоненіе составитъ

$$0 \times 1 + 6 \times (-1) = -6\,;$$

среднеариѳметическое уклоненіе по формулѣ

$$m = \frac{6}{2}\left(1 - \frac{\frac{1-7}{8}}{\frac{1+7}{8}}\right) + 1 = 6{}^1/_4$$

находится въ точкѣ, отстоящей отъ начала ряда на $6{}^1/_4$ единицъ, считая за 1 интервалъ между варіантами, т. е. за шестой варіантой (-4) займетъ еще ${}^1/_4$ интервала или среднеариѳметическая уклоненій составитъ — 4·5.

Ряды съ асиметріей. Въ тѣхъ случаяхъ, когда среднеариѳметическая не можетъ быть принята за свободное отъ случайныхъ ошибокъ выраженіе истинной величины и въ то же время не совпадаетъ съ наиболѣе вѣроятной варіантой ряда, значеніе ея становится сомнительнымъ. Это всегда бываетъ, когда рядъ обнаруживаетъ ясно выраженную асиметрію; когда преобладающая масса варіантъ располагается болѣе или менѣе симетрично по обѣ стороны точки наибольшей частоты, нисходящія же вѣтви кривой одна заканчивается сравнительно недалеко отъ этой точки, другая же тянется на большое разстояніе, почти совпадая съ осью абсциссъ. Сравнительное небольшое число крайнихъ варіантъ этой стороны притягиваетъ къ себѣ среднеариѳметическую, заставляя ее падать въ мѣстѣ сравнительно малого скопленія варіантъ; вслѣдствіе этого преобладающая масса варіантъ ряда уклоняется отъ среднеариѳметической въ одну и ту же сторону, систематически; тогда какъ, съ другой стороны, противоположныя уклоненія, хотя и немногочисленны, но зато достигаютъ нерѣдко неумѣстно большой величины. Благодаря этому среднеариѳметическая оказывается лишенной своего логическаго основанія и жизненного значенія.

Отбрасываніе крайнихъ членовъ. Чтобы вернуть среднеариѳметическую въ мѣсто большаго скопленія варіантъ, иногда отбрасываютъ крайнія варіанты. Если асиметрія произошла отъ смѣшенія разнородныхъ рядовъ, относящихся къ явленіямъ разной величины, какъ это нерѣдко случается (напримѣръ, урожай при разномъ количествѣ удобренія), то возникаетъ необходимость не отсѣченія крайнихъ варіантъ, а раздѣленія матеріала. Если же асиметрія не можетъ быть отнесена на счетъ разнородности матеріала, то выбрасываніе крайнихъ варіантъ также не можетъ быть обосновано; если же отбрасываніе предоставить субъективному усмотрѣнію, то рѣшая вопросъ, какія варіанты представляются неподходящими, мы въ бо́льшей или меньшей степени предрѣшаемъ вопросъ о томъ, какую среднюю мы считаемъ подходящей. Правда, предлагаются и объективные критеріи опредѣленія предѣльнаго значенія крайнихъ варіантъ или допустимаго предѣльнаго протяженія абсциссы варіантъ *. Но эти критеріи исходятъ изъ предположенія, что распредѣленіе должно быть нормально-симетричнымъ, и потому примѣнимы лишь къ тѣмъ случаямъ, когда уклоненіе отъ симетріи можно считать случайнымъ. Между тѣмъ въ дѣйствительности существуетъ множество рядовъ, къ которымъ гипотеза симетріи неприложима, и гдѣ асиметрія является органически связанной съ природой явленій. Такого рода рядами являются ряды данныхъ біологическаго и соціальнаго характера. Асиметрія является здѣсь, повидимому, выраженіемъ того явленія, что естественныя и общественныя условія жизни организма и человѣка допускаютъ варіаціи въ устройствѣ органовъ, функціяхъ и дѣйствіяхъ въ одну сторону бо́льшія, чѣмъ въ другую. Примѣромъ можетъ служить протяженіе возраста половой жизни, возможныя варіаціи заработной платы и т. п.

Значеніе моды. Поэтому въ тѣхъ случаяхъ, когда мы имѣемъ дѣло съ рядами органически асиметрическими, средняя ариѳметическая теряетъ свое значеніе первенствующей и единственной средней; она сохраняетъ весьма важное показательное, симптоматическое значеніе и часто является необходимой вспомогательной величиной. Болѣе реальное значеніе получаетъ, прежде всего, мода. Она является выраженіемъ наиболѣе вѣроятнаго результата совмѣстнаго дѣйствія систематическихъ факторовъ, служитъ наиболѣе распространеннымъ обычнымъ типомъ величины явленія. По отношенію къ ограниченной части варіантъ, симетрически расположенныхъ около моды, она является среднеариѳметической и такимъ

* См. C. B. Davenport, Statistical Methods, 1904, стр. 12 и 25.

образомъ, отсѣченіе части крайнихъ варіантъ и принятіе для остальной части среднеариѳметической является на самомъ дѣлѣ переходомъ отъ среднеариѳметической для всего ряда къ модѣ ряда.

Положеніе моды, хотя и опредѣляется мѣстомъ скопленія варіантъ, но находится въ зависимости отъ всѣхъ варіантъ ряда; эта зависимость обнаруживается между прочимъ въ томъ, что между положеніемъ моды и среднеариѳметической существуетъ всегда опредѣленное соотношеніе. Лишь при эмпирическомъ опредѣленіи значеніе отдѣльныхъ варіантъ не принимается во вниманіе и положеніе моды обусловливается исключительно мѣстомъ наибольшаго скопленія.

Но какъ бы точно не опредѣлялось положеніе моды, логически она все же является выраженіемъ только господствующаго, преобладающаго типа величинъ и даетъ непосредственно представленіе объ одномъ опредѣленномъ участкѣ всего протяженія варіантъ. Между тѣмъ часто является необходимость въ такой средней, которая по своему положенію въ рядѣ варіантъ и величинѣ находилась бы въ бо́льшей зависимости отъ всего числа варіантъ и служила бы представленіемъ всего ихъ протяженія, исполняя роль, аналогичную средней ариѳметической.

Значеніе медіаны.

Такой величиной является медіана. Подобно средней ариѳметической она непосредственно зависитъ отъ всѣхъ величинъ ряда; но въ отличіе отъ нея, зависитъ только отъ ихъ числа, но не зависитъ отъ ихъ значенія. Прибавленіе къ ряду варіанты съ сколь угодно большимъ значеніемъ передвигаетъ медіану лишь на одно значеніе внутри ряда, тогда какъ для среднеариѳметической прибавленіе очень большихъ варіантъ можетъ извратить ея значеніе. Поэтому для рядовъ очень разсѣянныхъ, съ большимъ протяженіемъ и неравномѣрными величинами варіантъ, медіана является болѣе подходящей средней, сравнительно съ средне-ариѳметической; такъ, напримѣръ, при опредѣленіи средняго дохода какогс-либо класса, она мало измѣняется отъ присоединенія доходовъ исключительной высоты, учитывая однако самый фактъ наличности большого дохода, но не величину послѣдняго, чѣмъ отличается съ одной стороны отъ моды, дающей представленіе лишь объ обычномъ доходѣ, и, съ другой, отъ среднеариѳметической, распредѣляющей доходъ милліардера—капиталиста или исключительно оплачиваемаго работника на весь классъ.

Медіана интересна еще и съ той специфической стороны, что способъ нахожденія ея требуетъ лишь счета числа случаевъ, расположенныхъ въ возрастающемъ или убывающемъ порядкѣ значеній варіантъ, но не требуетъ измѣренія самого значенія; поэтому она примѣнима даже въ томъ случаѣ, когда значенія варіантъ поддаются лишь количественному сравненію, хотя бы оно не могло быть точно измѣрено, напримѣръ, степень развитія учениковъ; въ этомъ отношеніи мода идетъ на встрѣчу стремленію привлекать къ статистической обработкѣ качественныя явленія, не поддающіяся по своему характеру обычнымъ способамъ измѣренія.

О вѣсѣ при вычисленіи средней ариѳметической. Когда въ результатѣ измѣренія или наблюденія какого либо явленія получается рядъ варіантъ, то число случаевъ, приходящееся на одну и ту же варіанту или на варіанты одного и того же интервала, имѣетъ существенное значеніе. Непринятіе во вниманіе частоты варіантъ при вычисленіи средней ариѳметической было бы равносильно исключенію части варіантъ. Принимая во вниманіе вѣсъ варіантъ, мы придаемъ этимъ бо́льшее значеніе варіантамъ, имѣющимъ бо́льшую частоту, что вполнѣ цѣлесообразно, когда отдѣльныя варіанты не имѣютъ иного значенія, кромѣ того, что онѣ представляютъ единичные случаи одного общаго наблюденія. Однако не слѣдуетъ дѣлать общаго правила изъ этого положенія и утверждать, что во всѣхъ случаяхъ безъ исключенія средняя ариѳметическая, при вычисленіи которой былъ принятъ вѣсъ, лучше вычисленной безъ вѣса. Примѣненіе или непримѣненіе вѣса зависитъ отъ логической цѣли, которую имѣетъ въ виду средняя, и часто примѣненіе вѣса можетъ придать средней совершенно несоотвѣтствующій смыслъ.

Представимъ себѣ, что при собираніи свѣдѣній объ урожаѣ, мы отмѣчаемъ величину участка, къ которому относится зарегистрированная величина урожая. Если регистрація обнимала собою всѣ участки въ предѣлахъ опредѣленной территоріи, то принимая площадь участка, къ которой относится каждое показаніе объ урожаѣ, за вѣсъ этого показанія, мы совершенно правильно вычислимъ среднюю величину урожая данной территоріи. Если однако изслѣдованіе носило выборочный характеръ, будетъ ли имѣть смыслъ принимать площадь каждаго участка за вѣсъ показанія?

Это имѣло бы смыслъ въ томъ случаѣ, если бы величина участка съ урожаемъ x репрезентировала собою всю ту площадь территоріи, на которой встрѣчается урожай x; но величина зарегистрированнаго участка не можетъ репрезентировать собою долю территоріи; если урожай самъ-5 отмѣченъ на участкѣ 10 дес., а самъ-3 на участкѣ 20 дес., то это не значитъ, что послѣдній урожай въ предѣлахъ всей территоріи также занимаетъ площадь въ 2 раза большую, чѣмъ урожай самъ-5. Равнымъ образомъ нѣтъ основанія думать, что показанія большихъ участковъ лучше репрезентируютъ урожай окрестнаго района земель и хозяйствъ, чѣмъ участки малые. Они могутъ лишь репрезентировать урожай качественно иначе, чѣмъ участки малые; но дѣлать изъ этого обстоятельства количественный вѣсъ показанія нѣтъ смысла. То же самое въ сущности можно сказать и о числѣ показаній, когда мы собираемъ ихъ по заранѣе намѣченнымъ отграниченнымъ районамъ. Если въ одномъ районѣ собрано 10 показаній, а въ другомъ 30, то нельзя эти цыфры принимать за вѣсъ района при вычисленіи общей средней. Соберемъ ли мы по данному району много показаній или мало, но всѣ они репрезентируютъ только одну цыфру— величину явленія въ районѣ; разумѣется, большее количество показаній репрезентируетъ эту цыфру вѣрнѣе, чѣмъ малое, и потому желательно, чтобы всѣ районы были представлены достаточнымъ числомъ показаній; но тройное количество показаній не должно дѣлать изъ одной порайонной цыфры тройную при вычисленіи средней для всѣхъ районовъ. При опредѣленіи цѣны какого-либо продукта на опредѣленномъ базарѣ мы вполнѣ цѣлесообразно стремимся установить количество продукта, проданнаго по разнымъ цѣнамъ, исходя изъ представленія, что болѣе вѣрной цѣной будетъ та, по которой продано бо́льшее количество товара. Если однако мы опредѣляемъ среднеариѳметическую цѣну товара по всѣмъ базарамъ губерніи, то принятіе или непринятіе во вниманіе вѣса сообщаетъ разный смыслъ средней; принимая во вниманіе количество проданнаго товара, мы даемъ преобладаніе рынку съ большимъ оборотомъ даннаго продукта и умаляемъ значеніе остальныхъ; мы изъ всѣхъ базаровъ губерніи дѣлаемъ одинъ общій рынокъ; это отвѣчаетъ дѣйствительному отношенію въ томъ случаѣ, если дѣйствительно цѣна крупнаго рынка регулируетъ цѣны всей территоріи, и цѣны отдѣльныхъ базаровъ можно разсматривать лишь какъ единичныя случайно уклоняющіяся проявленія той самой цѣны, которая въ большомъ оборотѣ проявляется на крупномъ рынкѣ; если однако отдѣльные рынки экономически самостоятельны въ той или иной степени, то и при выводѣ средней цѣны по всей губерніи, они должны фигурировать въ качествѣ самостоятельныхъ и равнозначныхъ единицъ. Правда, средняя вычисленная такимъ образомъ будетъ имѣть лишь показательное значеніе, но надо принять во вниманіе, что и невозможно найти типичное выраженіе одной цѣны, когда ея не существуетъ въ дѣйствительности; и мы не можемъ сдѣлать ее существующей, исходя изъ невѣрной гипотезы.

Или, напримѣръ, въ томъ случаѣ, когда ежегодно мы имѣемъ нѣкоторое число показаній о цѣнѣ продукта, и число этихъ показаній по годамъ различно; при выводѣ общей средней цѣны за рядъ лѣтъ ошибочно было бы принимать годичныя среднія съ вѣсомъ, соотвѣтствующимъ числу показаній, изъ которыхъ онѣ выведены; ибо среднія за извѣстные промежутки времени разсматриваются, какъ равноцѣнныя. Иначе было бы въ томъ случаѣ, если бы величина явленія съ теченіемъ времени не мѣнялась, т. е. была бы независима отъ времени; въ такомъ случаѣ единичныя показанія относились бы къ одному и тому же явленію, независимо отъ времени, и для вывода общей средней частныя среднія слѣдовало бы взять съ ихъ вѣсомъ.

Вообще принятіе или непринятіе во вниманіе вѣса зависитъ отъ двухъ моментовъ: во 1-хъ, отъ исчерпывающаго или выборочного характера данныхъ и, во 2-хъ, отъ зависимости или независимости явленія отъ пространства или времени. Съ точки зрѣнія перваго момента, если данныя имѣютъ исчерпывающій характеръ, вѣсъ слѣдуетъ принимать во вниманіе, какъ равнымъ образомъ и въ томъ случаѣ, когда выборочныя данныя имѣютъ безспорно репрезентативный характеръ; если же выборочныя данныя не репрезентируютъ сплошнаго матеріала, вѣсъ не долженъ быть принимаемъ во вниманіе. Съ точки зрѣнія второго момента вѣсъ слѣдуетъ принимать во вниманіе въ тѣхъ случаяхъ, когда всѣ единичныя данныя представляютъ собою случаи одного и того же явленія, распространяющагося на все пространство или весь періодъ времени безразлично; если же явленіе получаетъ различное выраженіе въ отдѣльныхъ частяхъ пространства или промежуткахъ времени, то эти частныя выраженія явленія при выводѣ общей средней являются равноцѣнными, независимо отъ числа показаній, на основаніи которыхъ онѣ выведены, и слѣдовательно въ общую среднюю приходятъ съ одинаковымъ вѣсомъ.

Въ конкретныхъ случаяхъ могутъ представляться различныя комбинаціи этихъ двухъ принциповъ.

Вѣсъ показательныхъ чиселъ. Много сомнѣній возбуждаетъ вопросъ, слѣдуетъ ли брать вѣсъ и какой именно при вычисленіи такъ называемыхъ показательныхъ чиселъ. И здѣсь рѣшеніе вопроса зависитъ отъ логическаго содержанія, которое связывается съ показательнымъ числамъ. Если мы беремъ среднюю изъ ряда цѣнъ различнаго рода товаровъ (хлѣбъ, мясо, перецъ, желѣзо, сукно, полотно, лѣсъ и т. п.), то рѣшеніе вопроса, слѣдуетъ ли принимать какой-либо вѣсъ для каждаго изъ товаровъ, зависитъ отъ цѣли, съ какою образуется сводная показательная цѣна. Цѣлью этой можетъ быть, во-первыхъ, опредѣленіе измѣненія цѣны денегъ. Разсматривая вопросъ абстрактно, независимо отъ измѣненій въ условіяхъ производства и сбыта отдѣльныхъ продуктовъ, мы находимъ, что измѣненія цѣны денегъ должны выражаться въ измѣненіи цѣнъ всѣхъ товаровъ безъ исключенія и притомъ совершенно одинаково, какъ въ тѣхъ товарахъ, которые продаются въ маломъ количествѣ, такъ и тѣхъ, которые имѣютъ обширный сбытъ; поэтому нѣтъ основанія, съ этой точки зрѣнія, принимать во вниманіе какія-либо различныя количества товаровъ, въ качествѣ ихъ вѣса, какъ обстоятельство безразличное; напротивъ, важно брать какъ можно большее число товаровъ, такъ какъ этимъ увеличиваются шансы открытія общей всѣмъ товарамъ причины измѣненія ихъ цѣны. Въ дѣйствительности однако измѣненія цѣнъ товаровъ, зависящія отъ денегъ, для отдѣльныхъ товаровъ скрадываются тѣми измѣненіями, которыя вызываются условіями производства и сбыта этихъ именно товаровъ, и если какой-либо изъ такихъ товаровъ взять съ вѣсомъ, то взятый вѣсъ только усиливаетъ значеніе этихъ постороннихъ вліяній. Такимъ образомъ принятіе во вниманіе вѣса единичныхъ товаровъ въ этомъ случаѣ нисколько не даетъ болѣе правильной средней величины, а, напротивъ, можетъ сдѣлать ее неправильной. Иначе обстоитъ дѣло въ томъ случаѣ, когда вопросъ идетъ о вздорожанiи жизни; здѣсь цѣлью образованiя сводной средней цѣны является обнаруженiе измѣненiй покупной силы единицы денегъ, производимыхъ не только измѣненiемъ цѣны самихъ денегъ, но и измѣненiемъ цѣны продуктовъ и другихъ средствъ удовлетворенiя потребностей; само собою разумѣется, что измѣненiе цѣнъ отдѣльныхъ предметовъ имѣютъ тѣмъ большее значенiе, чѣмъ въ большемъ количествѣ они прiобрѣтаются; поэтому по существа дѣла является логически правильнымъ принимать во вниманiе, въ качествѣ вѣса, то количество каждаго товара, какое входитъ въ бюджетъ потребленiя, или хотя бы условнымъ коэффицiентомъ опредѣлять относительное значенiе различныхъ категорiй предметовъ.

Типы кривыхъ. Какъ было указано, распредѣленiе варiантъ съ ихъ частотами можетъ быть симетрично или несиметрично. Симетричное распредѣленiе получается отъ разложенiя бинома $(p+q)^s$ при $p=q={}^1/_2$, т. е. при допущенiи равной вѣроятности какъ положительныхъ, такъ и отрицательныхъ уклоненiй одинаковой абсолютной величины. Если s сдѣлать какъ угодно большимъ, то получается предѣльный случай разложенiя—такъ наз. нормальная кривая Гаусса; она характеризуется тѣмъ, что, при полной симетрiи, среднеариѳметическая, мода и медiана совпадаютъ, и ординаты, изображающiе частоты, убывая въ обѣ стороны, приближаются къ оси абсциссъ на какъ угодно малое разстоянiе, не совпадая съ нею. Нормальная кривая на практикѣ примѣняется не только въ случаяхъ полной симетрiи, но и при нѣкоторой асиметрiи матерiала, давая практически удовлетворительные результаты для бо́льшей части кривой и служа критерiемъ для опредѣле́нiя случайнаго характера уклоненiй.

Нормальная кривая. Изслѣдованіе строенія кривой * даетъ слѣдующія опредѣленія характеризующихъ ея величинъ.

Если y обозначаетъ ординаты вообще, y_0 — ординату максимальной частоты, y_x — ординату варіанты, отстоящей на x отъ наиболѣе частой или среднеариѳметической; x — уклоненіе данной варіанты отъ средне-ариѳметической; $\delta^2 = \frac{S fx^2}{S f}$ среднеквадратическое уклоненіе, оно же равно $\mu_2 = \nu_2 - \nu_1^2$ (если принять обозначенія въ терминахъ моментовъ);
$e = 2 \cdot 71\ 82\ 81\ 8$ — основаніе неперовыхъ логариѳмовъ; $\pi = 3.\ 14\ 15\ 92\ 7$ — отношеніе окружности къ діаметру; N — сумма всѣхъ случаевъ $= S f$, то

$$y_0 = \frac{N}{\sqrt{2\pi\mu_2}}$$

$$y_x = y_0 e^{-x^2/2\mu_2}$$

Примѣръ. Duncker ** приводитъ слѣдующій примѣръ нормальнаго распредѣленія:

v	f	z	fz	fzz
11	1	—3	—3	9
12	2	—2	—4	8
13	189	—1	—189	189
14	1234	0		
15	454	1	454	454
16	20	2	40	80
s=	1900	= N	+298	740

$$\nu_1 = + \frac{298}{1900} = 0 \cdot 1568$$

* См. W. Palin Elderton, Frequency-Curves and Correlation, London.

** Die Methode der Variationsstatistik 1899, стр. 70—71, число иглъ на хвостовомъ плавникѣ Acerina cernua.

$$\nu_2 = +\frac{740}{1900} = 0\cdot3895$$

$$\mu_2 = 0\cdot3895 - 0\cdot1568^2 = 0\cdot3650$$

$$y_0 = \frac{1900}{\sqrt{0\cdot3650}\sqrt{2\pi}} = 1254\cdot9$$

$$y_x = y_0 \cdot e^{-x^2/0\cdot73}$$

Принимая во вниманіе, что ординаты y_0 отвѣчаетъ среднеариѳметическая варіанта

$$Me = V_0 + \nu_1 = 14\cdot1568\,,$$

уклоненія варіантъ ряда отъ Me, или x получаютъ значеніе

$$x_{11} = -3\cdot16$$
$$x_{12} = -2\cdot16$$
$$\cdots\cdots$$
$$x_{14} = -0\cdot16$$
$$x_{15} = 0\cdot84$$
$$x_{16} = 1\cdot84$$

Отсюда для варіанты, равной 14, теоретически вычисленная величина составитъ:

$$y_{14} = 1254\cdot9 \times 2\cdot71\;828^{-0\cdot16^2/2\times0\cdot3650}$$

или

$$\lg y_{14} = 3\cdot09\;83\;71 - 0\cdot0256 \times \frac{0\cdot43\;42\;945}{0\cdot73}$$

и

$$y = 1212.$$

Другія значенія y даны въ слѣдующей таблицѣ:

v	Эмпир. y	вычисл. y'
11	1	0·0
12	2	2·1
13	189	2000·4
14	1234	1212.4
15	454	473.3
16	20	11.9
	1900	1900·1

При вычисленіяхъ необходимо помнить, что вычисленіе моментовъ ведется при помощи условныхъ интерваловъ z=1; при переходѣ же къ дѣйствительнымъ величинамъ средней ариѳметической, среднеквадратическаго уклоненія и пр. условные интервалы слѣдуетъ переводить на дѣйствительные интервалы ряда (стр. 55). Въ данномъ случаѣ и тѣ и другіе равны и составляютъ 1.

Такимъ образомъ зная формулу кривой, мы можемъ для каждой варіанты (или для всякаго уклоненія варіанты отъ среднеариѳметической) вычислить теоретическую частоту и сравнить ее съ эмпирически данной; равнымъ образомъ мы можемъ взять сумму эмпирическихъ частотъ до даннаго уклоненія включительно и сравнить ее съ теоретической, или, наконецъ, сравнить эмпирическую сравнительную частоту уклоненій отъ среднеариѳметической до какого либо даннаго уклоненія съ теоретической вѣроятностью въ тѣхъ же предѣлахъ.

Таблица значеній функціи $F(u)$

Для облегченія такихъ сравненій можно пользоваться таблицой, въ которой вѣроятность уклоненій до извѣстныхъ предѣловъ заранѣе вычислена для нормальной кривой, площадь которой принята за 1, а уклоненія выражены въ доляхъ среднеквадратическаго уклоненія; вѣроятность всѣхъ уклоненій до какого-либо уклоненія (y) вычислена въ ней на основаніи формулы кривой

$$Fu = \frac{2}{\sqrt{\pi}} \int_0^u e^{-t^2} dt;$$

если показанное въ таблицѣ уклоненіе (u) умножить на величину $c = \sqrt{2\delta^2}$, которую называютъ модулемъ, то въ предѣлахъ такого уклоненія отъ средней въ обѣ стороны теоретически должно находиться число случаевъ, равное произведенію N (сумма частотъ эмпирическаго ряда) на соотвѣтствующую u вѣроятность Fu. Приводимъ нѣкоторыя цыфры этой таблицы

u	F(u)
0·18	0·201
0·23	0·255
0·37	0·399
0·48	0 503
0·59	0·596
0·81	0·748
0·91	0·802
1·16	0·899
1·38	0·949
1·82	0·990
2·50	0·999 6
3·00	0·9999 779

Изъ этой таблицы видно, что въ предѣлахъ уклоненій отъ средней на $\pm$ 1·8 с должно лежать 0·201 всего числа случаевъ, принятаго за 1, или на 1000 случаевъ 201, т. е. почти 20%; 50% всего числа уклоненій лежитъ въ предѣлахъ $\pm 0·48c$; 4 случая на 10 000 выходятъ за предѣлы $\pm 2·50c$ и всего 2 случая изъ 100 000 за предѣлы тройнаго уклоненія с.

Отсюда, между прочимъ, явствуетъ, что уклоненіе $\pm 2·5c$ или $3c$ можно считать критеріемъ крайнихъ уклоненій при распредѣленіи по нормальной кривой: бо́льшія уклоненія отъ средней могутъ встрѣтиться развѣ въ видѣ единичныхъ случаевъ при очень большомъ числѣ наблюденій; когда же число такихъ уклоненій сравнительно велико, ихъ нельзя считать случайными уклоненіями типа нормальной кривой.

Для сравненія эмпирическихъ частотъ съ теоретическими величинами можно принимать за мѣру уклоненія, вмѣсто $c = \delta\sqrt{2}$, величину δ, т. е среднеквадратическое уклоненіе. Тогда таблица теоретическихъ вѣроятности получитъ слѣдующій видъ:

u	F(u)	u	F(u)	u	F(u)	u	F(u)
0·1	0·08	0·8	0·58	1·5	0·87	2·2	0·97
0·2	0·16	0·9	0·63	1·6	0·89	2·3	0·98
0 3	0·24	1·0	0·68	1·7	0·91	2·5	0·99
0·4	0·31	1 1	0·73	1·8	0·93	3 0	0·99 7
0·5	0·38	1·2	0·77	1·9	0·94	3·5	0·9999 4
0·6	0·45	1·3	0·81	2·0	0·95		
0·7	0·52	1·4	0·84	2·1	0·96		

Среднеквадратическое уклоненіе. Величина среднеквадратическаго уклоненія, какъ видно изъ его выраженія $\delta = \sqrt{\frac{S fx^2}{S f}}$, уменьшается съ увеличеніемъ числа случаевъ наблюденій и обратно; она обратно пропорціональна корню квадратному изъ числа наблюденій (N). Такъ какъ среднеариѳметическая составлена изъ N наблюденій, то точность ея должна быть въ $\sqrt{N}$ больше, чѣмъ средняя мѣра точности единичнаго наблюденія и среднеквадратическое уклоненіе средней или δ_m составляетъ

средней ариѳметической

$$\delta_m = \frac{\delta}{\sqrt{N}}.$$

Если для одной величины среднеквадратическое уклоненіе составляетъ δ_1 и для другой δ_2, то для производной величины, представляющей сумму или разность двухъ данныхъ, среднеквадратическое уклоненіе составитъ

суммы или разности

$$\sqrt{\delta_1^2 + \delta_2^2} \ ;$$

если же данныя величины вычислены на основаніи разнаго числа наблюденій n_1 и n_2 то среднеквадратическое уклоненіе ихъ выразится въ видѣ

при разномъ вѣсѣ

$$\sqrt{\delta_1^2/n_1 + \delta_2^2/n_2} \ .$$

Если среднеквадратическое уклоненіе для данной величины а равно δ, то для величины въ n разъ бо́льшей или n разъ меньшей оно равно

$$n\delta \text{ и } \delta/_n .$$

Кривыя болѣе сложныхъ типовъ. Для асиметрическаго распредѣленія необходимы болѣе сложныя формулы, которыя могутъ быть построены различно *; наиболѣе общее рѣшеніе вопроса далъ Pearson. Онъ показалъ, что нормальная кривая получается, какъ частный случай изъ функціи

$$\frac{dy}{y} = \frac{xdx}{b_1 + b_2x + b_3x},$$

если положить $b_2 = b_3 = o$; при иныхъ допущеніяхъ, интегрированіе этого уравненія даетъ другіе результаты; такимъ образомъ можно получить нѣсколько типовъ кривыхъ. Въ дѣйствительности сравнительно чаще другихъ встрѣчаются ряды, подходящіе подъ I и IV типъ кривыхъ Пирсона. Для примѣра приведемъ способъ вычисленія кривой I типа.

Для вычисленія кривыхъ Пирсона вообще, въ частности и даннаго типа, необходимо найти предварительно по способу моментовъ слѣдующія величины.

1. Первые четыре момента отъ любой исходной точки въ серединѣ ряда, $\nu_1, \nu_2, \nu_3, \nu_4$;

2. Полученные моменты надо отнести къ точкѣ средней ариѳметической; зависимость же между моментами отъ любой точки и моментами (μ) отъ среднеариѳметической дается слѣдующими формулами:

$$\mu_2 = \nu_2 - \nu_1^2$$
$$\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3$$
$$\mu_4 = \nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 3\nu_1^4 .$$

3. Необходимо образовать вспомогательныя величины

$$\beta_1 = {\mu_3^2}/_{\mu_2^3} \qquad \beta_2 = {\mu_4}/_{\mu_2^2}$$

* Ср. A. L. Bowley, Elements of statistics, 1907, стр. 329—334

и на основаніи ихъ

$$F = 2\beta_2 - 3\beta_1 - 6$$

4. Эти три вспомогательныя величины служатъ критеріемъ для опредѣленія, какому типу отвѣчаетъ распредѣленіе эмпирическаго ряда, а именно:

Критерій типа кривой.

a) Когда F отрицательное, кривая относится либо къ
I типу, если $\beta_1 > 0$, либо
II типу, если $\beta_1 = 0$ и $\beta_2 < 3$.

b) Когда $F = 0$, либо къ
III типу, при $\beta_1 > 0$ и $\beta_2 > 3$ либо къ
нормальной кривой, при $\beta_1 = 0$, $\beta_2 = 3$.

Когда F положительное къ
IV типу, при $\beta_1 > 0$ и $\beta_2 > 3$.

Формулы кривыхъ этихъ типовъ приведено ниже.

Формула кривой I типа:

Типъ I.

$$y = y_0 \left(1 + \frac{x}{a_1}\right)^{m_1} \left(1 - \frac{x}{a_2}\right)^{m_2}.$$

Для вычисленія a_1, a_2, m_1 m_2 нужны слѣдующія вспомогательныя величины:

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

мѣра асимметріи

$$\alpha = \frac{1}{2}\sqrt{\beta_1}\,\frac{s+2}{s-2}\cdot\frac{\mu_3}{+\sqrt{\mu_3^2}};$$

если D разстояніе между Me и Mo то

$$\alpha = {}^{D}/_{\sqrt{\mu_2}} \text{ и } D = \alpha\sqrt{\mu_2}.$$

На основаніи этихъ величинъ получаемъ

протяженіе абсциссы a

$$a = \frac{\sqrt{\mu_2}}{2}\sqrt{\left\{\beta_1(s+2)^2 + 16(s+1)\right\}}$$

и части ея a_1 и a_2 въ одну и другую сторону отъ y_0:

$$a_1 = \frac{1}{2}(a - Ds) \text{ и } a_2 = a - a_1 ;$$

$$m_1 = \frac{a_1}{a}(s-2) , \; m_1 + m_2 = s - 2 .$$

Исходная для вычисленій ордината y_0 — *мода*.

$$y_0 = \frac{N}{a} \frac{(m_1+m_2+1)\sqrt{m_1+m_2}}{\sqrt{2\pi m_1 m_2}} e^{1/_{12}\left(\frac{1}{m_1+m_2} - \frac{1}{m_1} - \frac{1}{m_2}\right)}$$

Примѣръ. Число Мюллеровыхъ желѣзъ у 2000 экземпляровъ свиньи

v	f	z	fz	fzz	fzzz	fz^4
0	15	—4	—60	240	—960	3840
1	209	—3	—627	1881	—5643	16 929
2	365	—2	—730	1460	—2920	5840
3	482	—1	—482	482	—482	482
4	414	0				
5	277	1	277	277	277	277
6	134	2	268	536	1072	2144
7	72	3	216	648	1944	5832
8	22	4	88	352	1408	5632
9	8	5	40	200	1000	5000
10	2	6	12	72	432	2592
	2000		—998	6148	—3872	48 568

$$\nu_1 = -\frac{998}{2000} = -0{\cdot}499$$

$$\nu_2 = \frac{6148}{2000} = 3{\cdot}074$$

$$\nu_3 = -\frac{3872}{2000} = -1{\cdot}936$$

$$\nu_4 = \frac{48568}{2000} = 24{\cdot}284$$

$Me = 4 - 0{\cdot}499 = 3{\cdot}501$

$\mu_2 = 3{\cdot}074 - (-0{\cdot}499)^2 = 2{\cdot}824\,999 \qquad \sqrt{\mu_2} = 1{\cdot}680\,774$

$\mu_3 = 1{\cdot}936 - 3(-0{\cdot}499 \times 3{\cdot}074) + 2(-0{\cdot}499)^3 = 2{\cdot}417\,278$

$\mu_4 = 24{\cdot}284 - 4(-0{\cdot}499 \times -1{\cdot}936) + 6(0{\cdot}249\,001 \times 3{\cdot}074) - 3(-0{\cdot}499)^4 = 24{\cdot}826\,297.$

$$\beta_1 = \frac{(2{\cdot}417\,278)^2}{(2{\cdot}824\,999)^3} = 0{\cdot}259\,178$$

$$\beta_2 = \frac{24{\cdot}826\,297}{(2{\cdot}824\,999)^2} = 3{\cdot}110\,823$$

$F = 6{\cdot}222 - 0{\cdot}778 - 6 = -0{\cdot}556$

Критеріи типа:

F отрицательное, $\beta_1 > 0$ — *типъ I.*

$$s = \frac{6(3{\cdot}110\,82 - 0{\cdot}259\,18 - 1)}{0{\cdot}55589} = 19{\cdot}9857$$

$$\alpha = {}^1/_2 \sqrt{0{\cdot}259\,178} \; \frac{21{\cdot}9857}{17{\cdot}9857} = 0{\cdot}31115$$

$D = 1{\cdot}680\,774 \times 0{\cdot}3111 = 0{\cdot}5230$.

$Ds = 0{\cdot}5230 \times 19{\cdot}9857 = 10{\cdot}4519$

$a = 0{\cdot}840\,387 \sqrt{\left\{16 \times 20{\cdot}9857 + 0{\cdot}25918 \times (21{\cdot}9857)^2\right\}} = 18{\cdot}0448$

$$a_1 = \frac{18{\cdot}0448 - 10{\cdot}4519}{2} = 3{\cdot}7965$$

$a_2 = 18{\cdot}0448 - 3{\cdot}7965 = 14{\cdot}2483$

$$m_1 = \frac{3{\cdot}7965 \times 17{\cdot}9857}{18{\cdot}0448} = 3{\cdot}78401$$

$$m_2 = \frac{14{\cdot}2483 \times 17{\cdot}9857}{18{\cdot}0448} = 14{\cdot}2006$$

$y_0 =$

$$\frac{2000}{18{\cdot}0448} \; \frac{18.9846 \sqrt{17{\cdot}9846}}{\sqrt{2\pi \times 3{\cdot}784 \times 14{\cdot}2006}} \, 2{\cdot}171828^{0{\cdot}0833\,(0{\cdot}0556 - 0{\cdot}2643 - 0{\cdot}0704)} = 475{\cdot}24$$

Такъ какъ среднеариѳметическая $Me = 3{\cdot}501$ и мода лежитъ отъ нея на разстояніи $D = 0{\cdot}523$, то $Mo = 3{\cdot}501 - 0{\cdot}523 = 2{\cdot}978$.

Формула кривой:

$$y = 475{\cdot}24 \left(1 + \frac{x}{3{\cdot}7965}\right)^{3{\cdot}78401} \left(1 - \frac{x}{14{\cdot}2483}\right)^{14{\cdot}2006}$$

Вычислимъ для примѣра ординаты частотъ двухъ ближайшихъ къ модѣ варiантъ. Такъ какъ мода $= 2{\cdot}978$, то ближайшими варiантами будутъ, съ одной стороны 2 и, съ другой 3; въ первомъ случаѣ

$$x = 2 - 2{\cdot}978 = -0{\cdot}978,$$

во второмъ

$$x = 3 - 2{\cdot}978 = 0{\cdot}022;$$

необходимо принять во вниманiе, что за исходную точку кривой этого типа принимается мода, отъ которой и исчисляются разстоянiя x, и что эти разстоянiя измѣряются въ условныхъ единицахъ, т. е. разстоянiе между двумя сосѣдними данными варiантами принимается равнымъ 1. Для $x = -0{\cdot}978$:

$$\lg\left(-\frac{0{\cdot}978}{3{\cdot}7965}\right) = \bar{1}{\cdot}99034 - 0{\cdot}57938 = \bar{1}{\cdot}41096 = \lg(-0{\cdot}2576)$$

$$1 - 0{\cdot}2576 = 0{\cdot}7424$$

$$\lg\,(0{\cdot}7426)^{3{\cdot}78401} = 3\,78401 \times \bar{1}{\cdot}87064 = -0{\cdot}12936 \times 3{\cdot}78401 =$$
$$= -0{\cdot}4894995 = \bar{1}{\cdot}51050$$

$$\lg\left(1 + \frac{-0{\cdot}978}{3{\cdot}7965}\right)^{3{\cdot}78401} = \bar{1}{\cdot}51050$$

$$\lg\left(-\frac{0{\cdot}978}{14{\cdot}2483}\right) = \bar{1}{\cdot}99034 - 1{\cdot}15379 = \bar{2}{\cdot}83655 =$$
$$\lg\,(-0{\cdot}068635)$$

$$1 - (-0{\cdot}068635) = 1{\cdot}068635$$

$$\lg\,(1{\cdot}068635)^{14{\cdot}2006} = 14{\cdot}2006 \times 0{\cdot}02883 = 0{\cdot}40940$$

$$\lg\left(1 - \frac{-0{\cdot}978}{14{\cdot}2483}\right)^{14{\cdot}2006} = 0{\cdot}40940$$

и $\lg\,475{\cdot}24 = 2{\cdot}67692$

Такимъ образомъ

$$\begin{array}{r} 2{\cdot}67692 \\ +\ \bar{1}{\cdot}51050 \\ 0{\cdot}40940 \\ \hline \lg y = 2{\cdot}59682 \end{array}$$

откуда y, ордината варіанты 2, равняется 395.2.

Точно также для $x = 0{\cdot}022$:

$$\frac{0{\cdot}022}{3{\cdot}7965} = 0{\cdot}0058$$

$$\lg\,(1 + 0{\cdot}0058)^{3{\cdot}78401} = 3{\cdot}78401 \times 0.00251 = \underline{0{\cdot}00950};$$

$$\frac{0{\cdot}022}{14{\cdot}2483} = 0{\cdot}00154$$

$$\lg\,(1 - 0{\cdot}00154) = \lg 0{\cdot}99845 = \bar{1}{\cdot}99932; - 0{\cdot}00068 \times 14{\cdot}2006 =$$

$$- 0{\cdot}0095564 = \underline{\bar{1}{\cdot}99034}$$

$$\begin{array}{r} 2{\cdot}67692 \\ + \; 0{\cdot}00950 \\ \bar{1}{\cdot}99034 \\ \hline \end{array}$$

$$\text{и } \lg y = 2{\cdot}67676 = \lg 475$$

или y, ордината варіанты 3, равняется 475.

При опредѣленіи значеній моды, средней арифметической, среднеквадратич. уклоненія въ дѣйствительныхъ варіантахъ слѣдуетъ помнить, что дѣйствительные интервалы могутъ быть отличны отъ единицы, и правильный результатъ получается путемъ умноженія соотвѣтствующихъ величинъ на интервалъ.

Формулы кривыхъ другихъ типовъ имѣютъ слѣдующій видъ *

$$\text{II} \quad y = y_0\,(1 - x^2/a^2)^m$$

$$\text{III} \quad y = y_0\,(1 + x/a)^p\, e^{x/d}$$

$$\text{IV} \quad y = y_0\,(\cos\theta)^{2m}\, e^{-\nu\theta}, \text{ при } \tan\theta = x/a,$$

$$x = a\tan\theta$$

$$\theta = \tan^{-1}\frac{x}{a}$$

---

* Подробнѣе см. А. Леонтовичъ, Элементарное пособіе къ примѣненію методовъ Gauss'а и Pearson'а, ч. II, стр. 14—34; w. Palin Elderton, Frequency-Curves and Correlation.

Типъ IV. Кривая представляется асиметричной съ неограниченной въ обѣ стороны абсциссой. Особенностью ея является то, что исходная ордината (y_0) начинается не въ точкѣ наиболѣе частой варіанты и не въ точкѣ средней ариѳметической, а на разстояніи mD отъ средней ариѳметической точки, причемъ D разстояніе между максимальной ординатой и ординатой средней ариѳметической точки — равняется

$$D = \frac{\mu_3}{2\mu_2} \cdot \frac{s-2}{s+2}$$

и

$$m = \frac{1}{2}(s+2)$$

или:

$$mD = \frac{\mu_3}{\mu_2} \cdot \frac{s-2}{4} .$$

Далѣе:

$$s = \frac{6(\beta_2 - \beta_1 - 1)}{2\beta_2 - 3\beta_1 - 6}$$

(въ знаменателѣ обратные знаки сравнительно съ типомъ I).

$$a = \frac{\sqrt{\mu_2}}{4} \sqrt{16(s-1) - \beta_1 (s-2)^2}$$

$$\nu = \frac{s(-2)\sqrt{\beta_1}}{\sqrt{16(s-1) - \beta_1 (s-2)^2}}$$

причемъ ν имѣетъ знакъ обратный знаку μ_3.

$$y_0 = \frac{N}{a} \sqrt{\frac{s}{2\pi}} \frac{e^{\frac{(\cos\varphi)}{3s} - \frac{1}{12s} - \nu\varphi}}{(\cos\varphi)^{s+1}}$$

гдѣ

$$\tan\varphi = \frac{\nu}{s} .$$

Въ выраженіяхъ $\theta\nu$, $\varphi\nu$ величина угла въ градусахъ переводится на отношеніе дуги къ радіусу = 1, т. е.

$$\theta = \frac{\theta^\circ \pi}{180} .$$

При маломъ ν ($\nu < 2$) формула y_0 должна быть замѣнена болѣе сложнымъ выраженіемъ *.

Примѣромъ кривой IV типа могутъ служить данныя измѣренія роста дѣвочекъ 8 лѣтъ въ штатахъ С. Луи — приведено у Pearson'a.

Ростъ въ см.	Число.	Ростъ въ см.	Число.
141 и 142	1	119 и 120	342
139 » 140	0	117 » 118	321
137 » 138	1	115 » 116	297
135 » 136	5	113 » 114	222
133 » 124	10	111 » 112	137
131 » 132	21	109 » 110	84
129 » 130	28	107 » 108	42
127 » 128	79	105 » 106	27
125 » 126	138	103 » 104	8
123 » 124	183	101 » 102	2
121 » 122	243	99 » 100	1

Принимая 2 сант. за единицу разстоянія по оси абсциссъ, получимъ слѣдующія значенія моментовъ:

Среднiй ростъ (по ν_1) 118·27 см.

$\mu_2 = 7{\cdot}70739$

$\mu_3 = -2{\cdot}38064$ $\quad \beta_1 = 0{\cdot}0123784$

$\mu_4 = 192{\cdot}17419$ $\quad \beta_2 = 3{\cdot}235045$

$F = 2\beta_2 - 3\beta_1 - 6 > 0$, типъ IV.

Значенія постоянныхъ представятся слѣдующимъ образомъ:

$$D = 0{\cdot}135606$$
$$s = 30{\cdot}8023$$
$$\nu = 4{\cdot}56967$$
$$a = 14{\cdot}9917$$
$$m = 16\ 4011$$

* Способъ вычисленія котораго см. Elderton (l. c. стр. 69).

Откуда

$$y = y_0 \cos\theta^{32\cdot8023} e^{-4\cdot56967\,\theta}$$

и

$$\tan\theta = \frac{x}{14\cdot9917}\;.$$

Опредѣленіе y_0:

$$y_0 = \frac{N}{a}\sqrt{\frac{s}{2\pi}}\;\frac{e^{\frac{(\cos\varphi)^2}{3\,s} - \frac{1}{12\,s} - \nu\varphi}}{(\cos\varphi)^{s+1}}$$

гдѣ $\tan\varphi = \frac{\nu}{s}$.

Обозначивъ показателя при e чрезъ z и логариѳмируя выраженіе, получимъ:

$$\log y = \log N - \log a + \frac{1}{2}\log s + \log\frac{1}{\sqrt{2\pi}} +$$

$$+ z\log e - (s+1)\log\cos\varphi.$$

$$\log N = \log\ 2192 = 3\cdot34084$$

$$\log a = \lg\ 14\cdot9917 = 1\cdot17585$$

$$\log s = \log 30\cdot8023 = 1\cdot48858$$

$$1/2 \log s = \ldots\ldots = 0\cdot74429$$

$$\log\frac{1}{\sqrt{2\pi}} = \ldots\ldots = \bar{1}\cdot60091$$

$$\log e = \log 2\cdot71828 = 0\cdot43429$$

$$\log\nu = \log 4\cdot56967 = 0\cdot65987$$

$$\tan\varphi = \frac{4\cdot56967}{30\cdot8023}$$

$$\log\tan\varphi = \log 4\cdot56967 - \log 30\cdot8023 =$$

$$= 0\cdot65987 - 1\cdot48858 = \bar{1}\cdot17129.$$

По $\log\tan\varphi = \bar{1}\cdot17129$ находимъ по таблицамъ логариѳмовъ

$$\varphi = 8^0 26\cdot3' \quad \text{и}$$

$$\log\cos\varphi = \bar{1}\cdot99527$$

$$\log(\cos\varphi)^2 \quad = 2\log\cos\varphi = \overline{1}\cdot99054$$
$$\log(3\,s) \quad = 1\cdot96570$$
$$\log\frac{(\cos\varphi)^2}{3\,s} = \overline{2}\cdot02484$$
$$(\cos\varphi)^2 \div 3s = 0\cdot0105885$$
$$\frac{1}{12s} = 0\cdot00270544$$
$$\varphi = 8^0 26\cdot3' = \frac{506\cdot3^0}{60}$$
$$\operatorname{arc}\varphi = \frac{506\cdot3\,\pi}{60\times180}.$$
$$\log\varphi = \log\ 506\cdot3 + \log\pi - \log\ 60 - \log\ 180$$
$$= \overline{1}\cdot16814$$
$$\log(\varphi\nu) = \overline{1}\cdot82801$$
$$\varphi\nu \quad = 0\cdot6729855$$
$$z \quad = \frac{(\cos\varphi)^2}{3s} - \frac{1}{12s} - \nu\varphi = -0\cdot6651$$
$$\log e^{-0\cdot6651} = -0\cdot6651\,\log e = -0\cdot6651\times0\cdot43429$$
$$\log(0\cdot6651\times0\cdot43429) = \overline{1}\cdot46067 = \log 0\cdot28885$$
$$-0\cdot6651\ \log e = -0\cdot28885 = \overline{1}\cdot71114$$
$$(s+1)\ \log\cos\varphi = 31.8023\times\overline{1}\cdot99527 =$$
$$31\cdot8023\times-0\cdot00473 = -0\cdot15042 = \overline{1}\cdot84958$$

Сводя результаты, получимъ:

$$\begin{aligned}
&\log N && = 3\cdot34084\\
&\operatorname{colog} a && = \overline{2}\cdot82414\\
&{}^1/_2 \log s && = 0\cdot74429\\
&\log\frac{1}{\sqrt{2\pi}} && = \overline{1}\cdot60091\\
&z \log e && = \overline{1}\cdot71114\\
&\operatorname{colog}(\cos\varphi)^{s+1} && = 0\cdot15042\\
\hline
&\log y_0 && = 2\cdot37174\\
&y_0 && = 235\cdot363
\end{aligned}$$

Зная величину исходной ординаты, остальныя получаемъ по формулѣ

$$y = 235{\cdot}363 \cos\theta^{32{\cdot}8023} e^{-4{\cdot}56967\,\theta}$$

Зная, что

$$x = a \tan\theta$$

мы можемъ либо придавать опредѣленныя значенія x и вычислять $\tan\theta$, либо принимая послѣдовательно $\theta = 1^0$, $2^0 \ldots$ или $\theta = 10^0, 20^0, \ldots$ находить соотвѣтствующія значенія x. Ходъ вычисленій виденъ изъ слѣдующихъ примѣровъ.

При $x = 0$ (исходная ордината), $a \tan\theta = 0$, т. е. $\tan\theta = 0$, слѣдовательно $\cos\theta = 1$ и $\theta = 0$ и въ формулѣ кривой

$$y - y_0 = 235{\cdot}363$$

Дальнѣйшія вычисленія могутъ быть расположены въ слѣдующемъ порядкѣ:

θ	0	1^0	..	..
x	0	.	.	.
$-\theta\nu \log e$		.	.	.
$(s+2)\log\cos\theta$		.	.	.
$\log y_0$	2·37174	.	.	.
$\log y_x =$	Σ	.	.	.

При $\theta = 1^0$, $x = 14{\cdot}99 \tan 1^0$

$$\log x = \log 14{\cdot}99 + \log \tan 1^0 =$$
$$= 1{\cdot}17580 + \bar{2}{\cdot}24192 = \bar{1}{\cdot}41772$$

$$x = 0{\cdot}26.$$

Такъ какъ за 1 x-овъ принято 2 сантм., то разстояніе $x = 0{\cdot}26$ составляетъ 52 сантм.

$$\operatorname{arc}\theta = \frac{1^0 \cdot \pi}{180}, \quad \log\theta = \bar{2}{\cdot}24188;$$
$$\log\nu = \log 4{\cdot}57 \quad = 0{\cdot}65992,$$
$$\log\theta\nu = \bar{2}{\cdot}90180; \quad \theta\nu \quad = 0\,07976.$$

$$\log e^{-\nu\theta} = 0{\cdot}07976 \times 0{\cdot}43429$$

$$\log\ 0{\cdot}43429 = \bar{1}{\cdot}63778$$

$$\log\ 0{\cdot}07976 = \bar{2}{\cdot}90180$$

$$\bar{2}{\cdot}53958 = \log 0{\cdot}03464$$

$$-\nu\theta \log e = -0{\cdot}03464 = \bar{1}{\cdot}96536$$

$$-0{\cdot}07976\ \log e = \bar{1}{\cdot}95536$$

$$\log \cos\ 1^0 = \bar{1}{\cdot}99993$$

$$s+2 = 32{\cdot}8$$

$$\log \cos \theta^{s+2} = 32{\cdot}8 \times \bar{1}{\cdot}99993 = -0{\cdot}00007 \times 32{\cdot}8 =$$

$$= 0{\cdot}002296 = \bar{1}{\cdot}99770$$

$$32{\cdot}8\ \lg \cos \theta = \bar{1}{\cdot}99770.$$

Сложеніемъ найденныхъ логариѳмовъ получаемъ:

$$-\theta\nu \lg e \ldots\ldots \bar{1}{\cdot}96536$$

$$(s+2) \log \cos \theta \ldots \bar{1}{\cdot}99770$$

$$\log y_0 \ldots\ldots 2{\cdot}37174$$

$$\log y = 2{\cdot}33480$$

$$y = 216{\cdot}2$$

Такъ какъ среднiй ростъ равенъ 118·27 сантим., исходная ордината находится на разстояніи $Dm = 0{\cdot}1356 \times 16{\cdot}4 = 2{\cdot}22$ и единица x-овъ равна 2 сантиметрамъ, то исходная ордината соотвѣтствуетъ росту $118{\cdot}27 - 2 \times 2{\cdot}22 = 113{\cdot}83$ (варіанты вправо убываютъ, см. чертежъ на стр. 92). Въ положительную сторону (вправо) отъ исходной точки, на разстояніи $x = 0{\cdot}26 = 0{\cdot}52$ сантим., лежитъ точка, отвѣчающая росту $113{\cdot}83 - 0{\cdot}52 = 113{\cdot}31$; этому росту соотвѣтствуетъ найденная ордината 216·2 случая. Продолжая вычисленія далѣе, мы можемъ найти рядъ послѣдовательныхъ ординатъ и по нимъ построить кривую, а равнымъ образомъ опредѣлить число случаевъ, отвѣчающихъ требуемому промежутку.

Придавая θ значенія, возрастающія на равные промежутки, мы получаемъ неравно отстоящія значенія x. Вычисленіе по значеніямъ x представляется въ слѣдующемъ видѣ. Положимъ.

$$x = -2{\cdot}0884,$$

каковое условіе даетъ максимальную ординату, такъ какъ разстояніе ея влѣво отъ исходной ординаты равняется

$$Dm = 0{\cdot}135606 \times 16{\cdot}4011 = 2{\cdot}0884.$$

Въ такомъ случаѣ

$$\tan \theta = -\frac{2{\cdot}0884}{14{\cdot}9917};$$

отрицательное значеніе тангенса показываетъ, что уголъ взятъ въ послѣдней четверти круга, тангенсъ его по абсолютной величинѣ равенъ тангенсу той-же величины угла въ первой четверти; при отрицательномъ θ произведеніе $\theta\nu$ на $\log e$ слѣдуетъ брать съ обратнымъ знакомъ. Поэтому оставляя въ сторонѣ знакъ минусъ, получимъ

$$\log \tan \theta = \bar{1}{\cdot}143982$$

$$\theta = 7^{0}55' \frac{782}{925}$$

$$\log \cos \theta = \bar{1}{\cdot}995839$$

$$32{\cdot}8023 \log \cos \theta = 32{\cdot}8023 \times \bar{1}{\cdot}995839 =$$
$$= 32{\cdot}8023 \times -0{\cdot}004161$$

$$\begin{array}{ll} \log 32{\cdot}8023 \;\ldots & 1{\cdot}515904 \\ \log 0{\cdot}004161 \;\ldots & \bar{3}{\cdot}619198 \\ \hline & \bar{1}{\cdot}135102 = \end{array}$$
$$= \log 0{\cdot}136490$$

$$32{\cdot}8023 \log \cos \theta = -0{\cdot}136490 = \bar{1}{\cdot}863510$$

$$\theta = 7^{0}55' \frac{782}{925} = \frac{(55\times925+782+7\times925\times60)\pi}{925\times60\times180} =$$
$$= \frac{44{\cdot}0157\,\pi}{999}$$

$$\begin{array}{ll} \log \theta & = \bar{1}{\cdot}141192 \\ \log \nu & = 0{\cdot}659885 \\ \hline \log \theta\nu & = \bar{1}{\cdot}801077 \end{array}$$

$$\log(\log e) = \bar{1}\cdot637784$$
$$\log \theta_v = \bar{1}\cdot801077$$
$$\bar{1}\cdot438861 = \log 0\cdot274701$$

Принимая во вниманіе отрицательный знакъ θ, вмѣсто— $\theta_v \log e$ возьмемъ

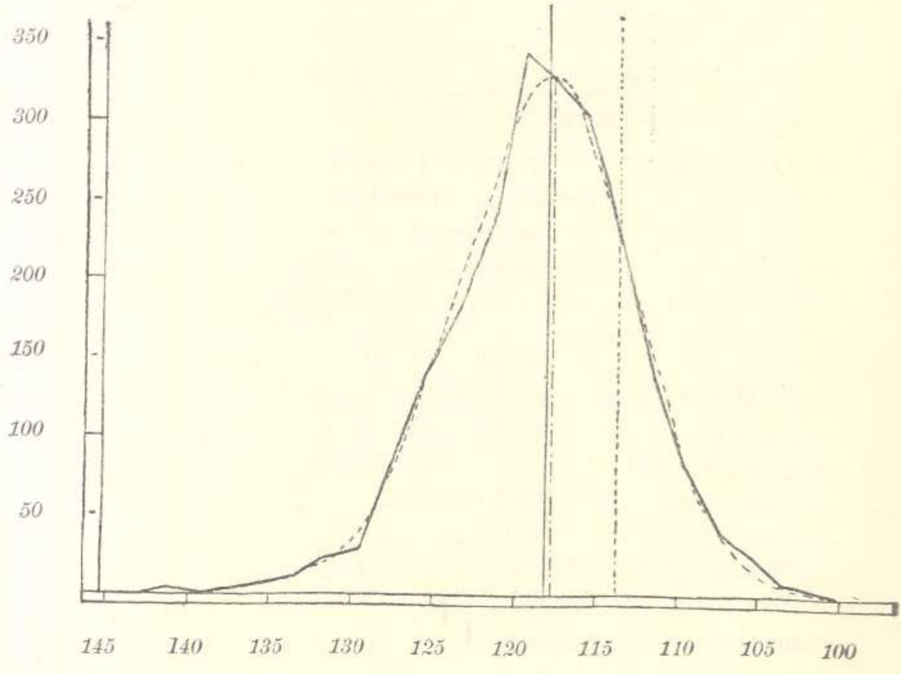
$$\theta_v \log e = 0\cdot274701$$

Суммируя, получимъ:

$$\begin{array}{ll} (s+2)\log\cos\theta \ldots\ldots & \bar{1}\cdot863510 \\ \theta_v \log e \ldots\ldots\ldots & 0\cdot274701 \\ \log 235\cdot363 \ldots\ldots & 2\cdot371738 \\ \hline \log y \ldots\ldots\ldots : & 2\cdot509949 \end{array}$$

$$y = 323\cdot556.$$

Эмпирическія данныя и теоретическая кривая воспроизведены на прилагаемомъ чертежѣ (Pearson).



4. Относительныя числа.

Если имѣется N наблюденій или случаевъ опредѣленнаго рода, и изъ нихъ въ n случаяхъ появляется нѣкоторое событіе или признакъ, то отношеніе $^n/_N$ называется относительнымъ числомъ; оно показываетъ сравнительную частоту (частость) появленія даннаго событія или признака и въ этомъ смыслѣ сходно съ эмпирическимъ апостеріорнымъ выраженіемъ вѣроятности событія; на этомъ сходствѣ основывается методологическое значеніе относительныхъ чиселъ.

Общеотносительное и частноотносительное.

Относительное число можетъ быть получено либо на основаніи всей совокупности случаевъ непосредственно, или же выведено изъ относительныхъ, полученныхъ изъ частныхъ или частичныхъ совокупностей; равнымъ образомъ и обратно, общая совокупность случаевъ, на основаніи которой выведено относительное число, можетъ быть разбита на части, изъ которыхъ каждая даетъ, въ свою очередь, относительное число. Въ этомъ смыслѣ мы различаемъ общеотносительное число и частноотносительныя числа. Число случаевъ въ отдѣльныхъ частныхъ совокупностяхъ можетъ быть одинаково или различно, и сообразно съ этимъ отдѣльныя частноотносительныя числа могутъ привходить для образованія общеотносительнаго съ одинаковымъ или разнымъ вѣсомъ. Пусть будетъ z совокупностей съ s случаями въ каждой и m_1, $m_2 \ldots m_z$ случаями появленія событія; въ такомъ случаѣ получимъ рядъ частноотносительныхъ

$$\frac{m_1}{s}, \frac{m_2}{s} \ldots \frac{m_z}{s}$$

и на основаніи общей совокупности $s+s+\ldots=zs$ случаевъ и $m_1+m_2\ldots+m_z$ случаевъ появленія событія, получимъ общеотносительную:

$$\frac{m_1+m_2+\ldots+m_z}{zs}=\frac{1}{z}\frac{m_1}{s}+\frac{1}{z}\frac{m_2}{s}+\ldots+\frac{1}{z}\frac{m_z}{s}$$

При неравномъ числѣ случаевъ въ частныхъ совокупностяхъ, частноотносительныя получатъ видъ

$$\frac{m_1}{s_1},\ \frac{m_2}{s_2}\ \ldots\ldots\ \frac{m_z}{s_z}$$

и общеотносительная:

$$\frac{m_1+m_2+\ldots+m_z}{s_1+s_2+\ldots+s_z}=\frac{s_1}{N}\frac{m_1}{s_1}+\frac{s_2}{N}\frac{m_2}{s_2}+\ldots+\frac{s_z}{N}\frac{m_z}{s_z}$$

если $N=s_1+s_2+\ldots+s_z$.

Если выраженіе $\frac{m_i}{s_i}$ представляетъ сравнительную частоту появленія событія, или вѣроятность его въ совокупности s_i, то аналогичное выраженіе $\frac{s_i}{N}$ представляетъ относительную частоту или вѣроятность появленія случаевъ совокупности s_i среди всѣхъ случаевъ общей совокупности N.

Общеотносительная можетъ быть разсматриваема также, какъ средняя ариѳметическая (съ вѣсомъ) изъ частноотносительныхъ и соотвѣтствуетъ понятію средней вѣроятности. Если частноотносительныя $^{m_1}/_{s_1}$, $^{m_2}/_{s_2}\ldots$ обозначимъ p_1, $p_2\ldots$, ихъ вѣсъ $^{s_1}/_N$, $^{s_2}/_N\ldots g_1$, $g_2\ldots$, то общеотносительная p_0 получитъ слѣдующее выраженіе:

$$p_0=g_1\,p_1+g_2\,p_2+\ldots+g_z\,p_z\,.$$

При одинаковомъ внѣшнемъ выраженіи отношеніе общеотносительной къ частноотносительнымъ можетъ быть различно по существу.

Относительная элементарнаго состава. 1. Первый наиболѣе простой случай имѣетъ мѣсто тогда, когда въ основѣ каждаго случая общей совокупности, а слѣдовательно и всякаго случая каждой частной совокупности лежитъ одна и та же неменяющаяся вѣроятность p; тогда частноотносительная каждой отдѣльной совокупности (p_1, p_2 . . .) будетъ служить эмпирическимъ приближеннымъ выраженіемъ этой постоянной вѣроятности, съ одинаковой точностью, если число случаевъ каждой совокупности одинаково, или разной—при различномъ ихъ вѣсѣ; общеотносительная же выражаетъ ту же вѣроятность съ наибольшею точностью. Такого рода общеотносительныя мы можемъ назвать общеотносительными элементарнаго состава. Такъ какъ общеотносительная является также среднеариѳметической изъ частноотносительныхъ, то среднюю мѣру уклоненія членовъ ряда частноотносительныхъ мы можемъ измѣрить обычнымъ способомъ посредствомъ среднеквадратическаго уклоненія. Такъ, взявъ болѣе простой случай группъ равнаго вѣса, имѣемъ рядъ частноотносительныхъ

$$\frac{m_1}{s}, \frac{m_2}{s}, \ldots \frac{m_z}{s}$$

и среднюю изъ нихъ или p_0:

$$p_0 = \frac{1}{z} \cdot \frac{m_1}{s} + \frac{1}{z}\frac{m_1}{s} + \ldots + \frac{1}{z} \cdot \frac{m_z}{s} = \frac{p_1 + p_2 \cdot \cdot + p_z}{z};$$

уклоненія составятъ:

$$p_1 - p_0 = e_1$$
$$p_2 - p_0 = e_2$$
$$\cdot \cdot \cdot \cdot \cdot \cdot \cdot$$
$$p_z - p_0 = e_z$$

и среднеквадратическое уклоненіе:

$$\sqrt{\frac{s\, e^2}{z}} = \delta .$$

Среднеквадратическое уклоненіе σ. Но то же же уклоненіе можетъ быть выражено иначе. Пусть мы имѣемъ s случаевъ испытанія и въ каждомъ случаѣ вѣроятность появленія событія p; по теоремѣ Бернулли, наиболѣе вѣроятное число случаевъ появленія событія должно быть

sp

поэтому представимъ, что при достаточно большомъ числѣ испытаній событіе наступило въ m случаяхъ, причемъ $m = ps$ и не наступило въ $n = s - m$ случаяхъ; когда событіе наступило, то изъ вѣроятнаго оно для даннаго случая становится достовѣрнымъ, или иначе вѣроятность его получаетъ значеніе достовѣрности, т. е. 1; въ случаѣ же ненаступленія, она получаетъ значеніе 0; тогда мы получаемъ рядъ, состоящій изъ m единицъ и n нулей, слѣдующихъ другъ за другомъ въ случайномъ порядкѣ

$$1, 0, 0, 1, 1, 1, 1, 0, 1 \ldots\ldots ;$$

среднеариѳметическая такого ряда составитъ

$$\frac{m \times 1}{s} = \frac{m}{s},$$

уклоненія же каждаго члена отъ средней дадутъ:

$$1 - \frac{m}{s} = 1 - p \text{ въ } m \text{ случаяхъ}$$

и

$$0 - \frac{m}{s} = \quad - p \text{ въ } n \text{ случаяхъ};$$

откуда среднеквадратическое уклоненіе членовъ ряда составитъ

$$\delta = \sqrt{\frac{m(1-p)^2 + n(-p)^2}{m+n}} ;$$

принимая во вниманіе, что

$$\frac{m}{m+n} = p \quad \text{и} \quad 1 - p = q ,$$

получимъ

$$\delta = \sqrt{pq^2 + qp^2} = \sqrt{pq} \quad ;$$

среднеквадратическое же уклоненіе самой средней $\frac{m}{s} = p$, составленной изъ s случаевъ, будетъ въ $\sqrt{s}$ разъ меньше (см. стр. 78); обозначая его въ этомъ новомъ видѣ черезъ σ, имѣемъ

$$\sigma = \sqrt{\frac{pq}{s}}$$

Если истинная величина p неизвѣстна и дана лишь въ эмпирическихъ приближеніяхъ p_1, p_2 и т. д , то наиболѣе вѣроятнымъ выраженіемъ ея служитъ общеотносительная p_0 , и среднеквадратическое уклоненіе отдѣльныхъ членовъ ряда выразится въ видѣ

$$\sigma_0 = \sqrt{\frac{p_0\, q_0}{s}} \; ;$$

такъ какъ то же самое уклоненіе ранѣе было выражено въ видѣ

$$\delta = \sqrt{\frac{s\, e^2}{z}} \, ,$$

находимъ, что при элементарномъ составѣ вѣроятности среднеквадратическое уклоненіе, вычисленное на основаніи эмпирическихъ уклоненій частныхъ вѣроятностей отъ общей и на основаніи расчетовъ теоріи вѣроятности, совпадаютъ.

Нормальное разсѣяніе. Такъ какъ при элементарномъ составѣ вѣроятности, уклоненія числа случаевъ наступленія событія въ отдѣльныхъ совокупностяхъ отъ наиболѣе вѣроятнаго результата (sp_0) подчинены закону случайныхъ ошибокъ нормальной кривой, то при такомъ распредѣленіи уклоненій мы говоримъ о нормальномъ разсѣяніи ряда частноотносительныхъ, и признакомъ такого разсѣянія служитъ равенство средней мѣры уклоненій, вычисленной двуми различными способами:

$$\delta = \sqrt{\frac{s\, e^2}{z}} = \sqrt{\frac{p_0 q_0}{s}} = \sigma_0 \; ;$$

или

$$\frac{\delta}{\sigma} = \sqrt{\frac{S\frac{e^2}{z}}{\frac{p_0\, q_0}{s}}} = Q = 1.$$

Коэффиціентъ расхожденія. Величина Q, которая при нормальномъ разсѣяніи равна 1 или стремится къ этому предѣлу, носитъ названіе коэффиціента расхожденія и служитъ показателемъ характера разсѣянія — нормальнаго или отличнаго отъ нормальнаго.

Нормальное разсѣяніе, кромѣ случая элементарной вѣроятности, встрѣчается также и въ случаѣ сложной вѣроятности при слѣдующихъ условіяхъ.

Относительныя случайно-перемѣннаго состава. 2. Пусть частноотносительныя p_1, p_2 . . . p_z представляютъ собою не случайныя уклоненія одной и той же вѣроятности, а эмпирическія выраженія различныхъ вѣроятностей, иначе говоря, пусть отдѣльныя частныя совокупности характеризуются различной вѣроятностью появленія событія; и, съ другой стороны, каждая совокупность имѣетъ свою особую сравнительную частоту появленія или особую вѣроятность ея появленія: g_1 g_2 . . . g_z. Такимъ образомъ вѣроятность наступленія событія въ единичномъ случаѣ наблюденія опредѣляется, прежде всего, вѣроятностью появленія въ наблюденіи той или иной совокупности; когда же совокупность опредѣлилась, то вѣроятностью появленія событія, свойственной данной совокупности. Такого рода условія отвѣчаютъ т. н. теоремѣ закона большихъ чиселъ Пуассона, по которой вѣроятность появленія событія опредѣляется средней вѣроятностью; эмпирическимъ приближеннымъ выраженіемъ послѣдней, является общеотносительная:

$$p_0 = g_1\, p_1 + g_2\, p_2 + \; . \; . \; . \; + g_z\, p_z\,.$$

Такъ, если сравнительную частоту (вѣсъ) одной совокупности $^2/_5$ и другой $^3/_5$ примемъ за вѣроятности появленія совокупностей, и сравнительную частоту появленія событія въ первой $^3/_4$ и во второй $^1/_3$ примемъ за вѣроятность появленія событія, то, по закону большихъ чиселъ Пуассона, при достаточно большомъ числѣ наблюденій наиболѣе вѣроятный результатъ долженъ соотвѣтствовать средней вѣроятности, и общеотносительная

$$p_0 = \frac{2}{5} \times \frac{3}{4} + \frac{3}{5} \times \frac{1}{3} = \frac{5}{10}$$

является эмпирическимъ выраженіемъ этой средней вѣроятности.

Общеотносительныя, составленныя подобнымъ образомъ, можно назвать общеотносительными случайно-перемѣннаго состава (въ отличіе отъ общеотносительныхъ постояннаго состава, о которыхъ ниже).

Нормальное разсѣяніе. Всякій случай средней вѣроятности такого случайно-перемѣннаго состава можно выразить посредствомъ элементарной вѣроятности, дающей ту же вѣроятность появленія событія, что и средняя вѣроятность. Такъ въ приведенномъ выше численномъ примѣрѣ предположимъ, что сдѣлано 100 случаевъ наблюденій; наиболѣе вѣроятное распредѣленіе этихъ 100 случаевъ между случаями совокупности перваго и второго рода будетъ, согласно теоремѣ Бернулли, пропорціонально вѣроятностямъ появленія этихъ совокупностей, или $^2/_5$ и $^3/_5$; или случаевъ первой совокупности будетъ $100 \times {}^2/_5 = 40$ и второй $100 \times {}^3/_5 = 60$; въ 40 случаяхъ, выпавшихъ на долю первой совокупности, наиболѣе вѣроятное число появленій событія будетъ пропорціонально его вѣроятности, т. е. $40 \times {}^3/_4 = 30$ случаевъ и непоявленія событія $40 - 30 = 10$ случаевъ; въ 60 же случаяхъ второго рода наиболѣе вѣроятное число случаевъ появленія событія составитъ $60 \times {}^1/_3 = 20$ и непоявленія $60 - 20 = 40$ случаевъ; въ общемъ на 100 случаевъ наблюденій наиболѣе вѣроятный результатъ дастъ $30 + 20$ случаевъ появленія событія. Но тотъ же самый результатъ мы могли бы получить иначе. Представимъ себѣ, что въ одной урнѣ мы имѣемъ 100 шаровъ, изъ которыхъ 40 имѣютъ обозначеніе № 1 и 60—№ 2; тогда вѣроятность вынуть шаръ съ обозначеніемъ № 1 составляетъ $\frac{40}{100}$ и съ обозначеніемъ № 2 — $\frac{60}{100}$; пусть изъ 40 шаровъ съ № 1 30 будутъ бѣлаго цвѣта и 10 чернаго; и изъ 60 шаровъ съ № 2—20 бѣлаго и 40 чернаго цвѣта; тогда вѣроятность, что если будетъ вынутъ одинъ изъ шаровъ № 1, онъ окажется бѣлымъ, составитъ, по теоремѣ умноженія $\frac{40}{100} \cdot \frac{30}{40}$; и вѣроятность, что, если будетъ вынутъ шаръ № 2, онъ окажется бѣлымъ, составитъ $\frac{60}{100} \cdot \frac{20}{60}$; вѣроятность же, что тѣмъ или другимъ путемъ вынется бѣлый шаръ, по теоремѣ сложенія, составитъ

$$p = \frac{40}{100} \cdot \frac{30}{40} + \frac{60}{100} \cdot \frac{20}{60} = \frac{2}{5} \cdot \frac{3}{4} + \frac{3}{5} \cdot \frac{1}{3} = \frac{5}{10},$$

т. е. результатъ будетъ выражаться совершенно также, какъ и въ предыдущемъ случаѣ. Однако въ приведенномъ примѣрѣ мы можемъ опредѣлить вѣроятность появленія бѣлаго шара непосредственно, принимая во вниманіе, что на 100 всѣхъ шаровъ бѣлыхъ приходится 50, причемъ очевидно, что здѣсь мы имѣемъ дѣло съ простой элементарной вѣроятностью. Такъ какъ вѣроятность появленія бѣлаго шара объективно одна и та же, будемъ ли мы выражать ее въ элементарно-простомъ видѣ, или въ сложной формѣ, тождественной съ выраженіемъ вѣроятности случайно-перемѣннаго состава, то ясно, что дѣйствіе такой сложной вѣроятности мы всегда можемъ свести къ дѣйствію вѣроятности элементарной. Отсюда слѣдуетъ, что и среднеквадратическое уклоненіе въ случаѣ сложной вѣроятности случайно перемѣннаго состава имѣетъ такой же видъ, какъ и въ случаѣ элементарной вѣроятности, т. е.

$$\sigma_0 = \sqrt{\frac{p_0 q_0}{s}}.$$

Поэтому при сравненіи σ_0 съ δ, вычисленнымъ на основаніи конкретно данныхъ уклоненій, мы, какъ и въ случаѣ элементарной вѣроятности получаемъ

$$\sigma_0 = \delta$$

или коэффиціентъ расхожденія Q и въ этомъ случаѣ будетъ стремиться къ 1; иначе говоря, мы имѣемъ здѣсь второй случай нормальнаго разсѣянія.

Относительныя постояннаго состава: 3. Общеотносительная можетъ быть составлена изъ частноотносительныхъ такимъ образомъ, что каждая частная совокупность, дающая различную вѣроятность событія, привходитъ для образованія общей совокупности опредѣленное, чѣмъ-либо нормированное число разъ; въ такомъ случаѣ вѣса частныхъ совокупностей выражаютъ не вѣроятность ихъ появленія, а постоянные коэффиціенты, и появленіе той или другой частной совокупности въ наблюденіи опредѣляется не случаемъ, а опредѣленной нормой. Общеотносительныя подобнаго рода называются общеотносительными постояннаго состава. Такого рода общеотносительная можетъ образоваться, напримѣръ, въ томъ случаѣ, если при вычисленіи частоты какого либо явленія общая совокупность наблюдаемыхъ случаевъ составляется изъ опредѣленнаго всякій разъ одного и того же числа случаевъ одной совокупности и опредѣленнаго и неизмѣннаго числа случаевъ другой совокупности.

Поднормальное разсѣяніе. По т. н. теоремѣ Пуассона, и здѣсь наиболѣе вѣроятный результатъ наблюденій долженъ соотвѣтствовать средней вѣроятности. Если первая совокупность появляется n_1 разъ и вторая n_2 разъ, и вѣроятность появленія событія въ первой совокупности составляетъ p_1 и во второй p_2, то наиболѣе вѣроятный результатъ испытаній долженъ соотвѣтствовать средней вѣроятности

$$p_0 = \frac{n_1}{N} p_1 + \frac{n_2}{N} p_2$$

гдѣ $$N = n_1 + n_2 \text{ и } \frac{n_1}{N} = g_1, \ \frac{n_2}{N} = g_2.$$

Среднеквадратическое уклоненіе для N испытаній въ этомъ случаѣ составитъ, согласно теоремѣ Пуассона

$$\sigma_c = \sqrt{\frac{n_1}{N}\frac{p_1q_1}{N} + \frac{n_2}{N}\frac{p_2q_2}{N}}$$

(или въ иномъ видѣ:

$$\sigma_c = \sqrt{g_1^2\frac{p_1q_1}{n_1} + g_2^2\frac{p_2q_2}{n_2}}) .$$

Можно доказать, что эта величина меньше

$$\sigma_0 = \sqrt{\frac{p_0q_0}{N}}$$

и отношеніе

$$\frac{\sigma_c}{\sigma_0} = Q \text{ д. б.} < 1 ,$$

Поэтому если сравнить среднеквадратическое уклоненіе δ, вычисленное на основаніи дѣйствительныхъ уклоненій, съ величиною σ_0, вычисленной въ предположеніи элементарнаго характера вѣроятности p_0, мы получимъ

$$\delta < \sigma_0 , \text{ и } \frac{\delta}{\sigma_0} < 1 \text{ или } Q < 1 ,$$

т. е. мы будемъ имѣть случай поднормальнаго разсѣянія; если же составить уклоненіе σ_c, принимая во вниманіе постоянный составъ средней вѣроятности, мы должны получить равенство

$$\delta = \sigma_c .$$

Такимъ образомъ нормированіе числа случаевъ дѣйствія каждой частной вѣроятности содерживаетъ колебанія въ болѣе узкихъ предѣлахъ, сравнительно съ тѣмъ случаемъ, когда появленіе частныхъ совокупностей опредѣляется случаемъ и свойственной имъ вѣроятностью.

Сверхнормальное разсѣяніе.

3. Представимъ себѣ, что частноотносительныя

$$\frac{m_1}{s}, \frac{m_2}{s}, \dots \frac{m_z}{s}$$

являются эмпирическимъ выраженіемъ вѣроятностей p_1 $p_2 \ldots p_z$, которыя, въ свою очередь, являются случайными уклоненіями нѣкоторой основной вѣроятности p; въ такомъ случаѣ общеотносительная

$$p_0 = \frac{m_1 + m_2 + \ldots + m_z}{sz}$$

будемъ служить эмпирическимъ выраженіемъ вѣроятности p.

Уклоненія частноотносительныхъ отъ общеотносительной, напримѣръ

$$\frac{m_i}{s} - p_0$$

будутъ слагаться, во-первыхъ, изъ уклоненія p_i отъ p_0, и во-вторыхъ, изъ уклоненій $\frac{m_i}{s}$ отъ p_i, т. е. изъ случайныхъ уклоненій основной вѣроятности и случайныхъ уклоненій частныхъ вѣроятностей. Можно доказать, что общій итогъ тѣхъ и другихъ уклоненій долженъ выразиться въ видѣ

$$\sqrt{\frac{pq}{s} + \frac{S(p_i - p)^2}{z}} \quad ;$$

если же вмѣсто истинной неизвѣстной величины вѣроятности p подставить ея эмпирическое значеніе p_0, то уклоненіе получитъ видъ

$$\sigma_c = \sqrt{\frac{p_0 q_0}{s} + \frac{S(p_i - p_0)^2}{z}} \quad .$$

Это уклоненіе очевидно больше $\sigma_0 = \sqrt{\frac{p_0 q_0}{s}}$; поэтому если вычислить среднеквадратическое уклоненіе на основаніи дѣйствительной разницы между частноотносительными и общеотносительной, т. е. δ, то оно будетъ больше уклоненія σ_0, вычисленною въ предположеніи элементарнаго состава p_0, и мы получаемъ случай сверхнормальнаго разсѣянія. Внѣшнимъ выраженіемъ послѣдняго является значеніе коэффиціента расхожденія

$$Q > 1.$$

Физикальный и комбинаторный способъ. Способъ опредѣленія среднеквадратическаго уклоненія на основаніи дѣйствительныхъ уклоненій частноотносительныхъ отъ общеотносительныхъ (или δ) Lexis назвалъ способомъ физикальнымъ; способъ же нахожденія уклоненія по данной общеотносительной (или σ) — комбинаторнымъ. Сравненіе результатовъ, полученныхъ тѣмъ и другимъ способомъ даетъ, въ видѣ коэффиціента расхожденія Q, критерій для сужденія о характерѣ разсѣянія ряда по отношенію къ общеотносительной величинѣ ряда, какъ наиболѣе вѣроятному эмпирическому выраженію вѣроятности, лежащей въ основаніи ряда; характеръ же разсѣянія служитъ до извѣстной степени показателемъ природы вѣроятности.

Среднекв. уклоненіе для Q. Коэффиціентъ расхожденія Q, подобно всякой сводной величинѣ, допускаетъ случайныя колебанія; поэтому возникаетъ вопросъ, до какого предѣла уклоненіе Q отъ 1 должны быть признаваемы случайными и съ какого предѣла они служатъ показателемъ сверхнормальнаго или поднормальнаго разсѣянія? Вопросъ сводится къ тому, чтобы для Q найти мѣру случайныхъ уклоненій отъ его наиболѣе вѣроятнаго значенія 1, въ предположеніи случая нормальнаго разсѣянія. Представимъ себѣ, что мы имѣемъ z относительныхъ величинъ

$$p_1, p_2, \ldots . p_z,$$

представляющихъ нормальное разсѣяніе по отношенію къ общеотносительной p_0, вслѣдствіе чего $Q = 1$ или уклоняется отъ 1 на случайную погрѣшность; пусть будутъ и другіе такіе же ряды съ нормальнымъ разсѣяніемъ по отношенію къ той же общеотносительной p_0; если такихъ рядовъ n, то получится n значеній Q, случайно уклоняющихся отъ своей истинной величины, равной 1. Если эти уклоненія случайны и подчиняются, напримѣръ, закону случайныхъ ошибокъ, то мѣрою ихъ можетъ служить среднеквадратическое уклоненіе или

$$\sqrt{\left\{\frac{S(1-Q_i)^2}{n}\right\}};$$

дѣйствительныя уклоненія, которыя превышали бы эту величину въ 2·5 или 3 раза являлись бы неслучайными и свидѣтельствовали бы о томъ, что соотвѣтствующее значеніе Q не имѣетъ своимъ истиннымъ выраженіемъ 1, или что Q больше или меньше 1, т. е. что разсѣяніе сверхнормально или поднормально. Проф. Bortkiewicz показалъ, что среднеквадратическое уклоненіе для Q можетъ быть выражено въ видѣ

$$\frac{1}{\sqrt{2z}}$$

Такъ, если рядъ, дающій дисперсію состоитъ изъ 32 членовъ, то среднеквадратическое уклоненіе для Q составитъ

$$\frac{1}{\sqrt{2\times 32}}=\frac{1}{8}=0{\cdot}125\ ,$$

и величины Q при нормальной дисперсіи можетъ колебаться въ предѣлахъ

$$1\pm 0{\cdot}375\ .$$

Законъ малыхъ чиселъ. Проф. В. Борткевичъ обратилъ вниманіе на то, что при сверхнормальномъ разстояніи величина коэффиціента расхожденія зависитъ, при прочихъ неизмѣнныхъ условіяхъ, отъ числа случаевъ s, изъ которыхъ состоитъ каждая частная совокупность; а именно: при большомъ числѣ случаевъ s коэффиціентъ Q отличается отъ единицы больше, чѣмъ при маломъ s. Необходимость этого вывода явствуетъ изъ слѣдующаго.

При сверхнормальномъ разсѣяніи среднеквадратическое уклоненіе выражается въ видѣ

$$\sigma_c^2 = \frac{p_0 q_0}{s} + \frac{S(p_i - p_0)^2}{z};$$

среднеквадратическое же уклоненіе въ предположеніи элементарнаго состава p_0 выражается въ видѣ

$$\sigma^2 = \frac{p_0 q_0}{s};$$

откуда величина коэффиціента расхожденія составитъ:

$$Q^2 = \frac{\sigma_c^2}{\sigma^2} = 1 + \frac{S(p_i - p_0)^2}{z} \cdot \frac{s}{p_0 q_0}.$$

Изъ этой формулы ясно, что Q тѣмъ больше отличается отъ 1, чѣмъ больше величина

$$\frac{S(p_i - p_0)^2}{z} \cdot \frac{s}{p_0 q_0},$$

послѣдняя же, при прочихъ равныхъ условіяхъ, увеличивается и уменьшается съ увеличеніемъ и уменьшеніемъ s. По этому, если при данномъ s коэффиціентъ расхожденія, будучи > 1, показываетъ сверхнормальное разсѣяніе, послѣднее можетъ быть понижено уменьшеніемъ s, т. е. сокращеніемъ числа случаевъ, изъ которыхъ составляются частныя совокупности для вычисленія частноотносительныхъ величинъ. Такъ, если каждая частноотносительная состоитъ изъ $s = 10\,000$ случаевъ наблюденія и $Q = 2$, то

$$Q^2 = 1 + \frac{S(p_i - p_0)^2}{z} \cdot \frac{10\,000}{p_0 q_0} = 4 = 1 + 3;$$

если вмѣсто 10 000 взять совокупность въ 100 случаевъ наблюденія, т. е. уменьшить 10 000 въ 100 разъ, то съ уменьшеніемъ въ 100 разъ выраженія

$$\frac{S(p_i - p_0)^2}{z} \frac{10\,000}{p_0 q_0} \cdot \frac{1}{100},$$

послѣдній членъ (3) правой части равенства долженъ быть уменьшенъ татже во сто разъ и

$$Q^2 = 1 + 0{\cdot}03$$

или

$$Q = 1{\cdot}015.$$

Отсюда получается парадоксальный на первый взглядъ выводъ, что частноотносительныя, выведенныя на основаніи группъ съ малымъ числомъ наблюденій обнаруживаютъ болѣе близкое къ нормальному распредѣленіе около общеотносительной, чѣмъ частноотносительныя для того же самаго явленія, но лишь выведенныя изъ группъ съ большимъ числомъ наблюденій. Эту особенность «малыхъ группъ» называютъ «закономъ малыхъ чиселъ». Причина указанной особенности станетъ понятной, если мы обратимъ вниманіе на слѣдующее. Среднеквадратическое уклоненіе при сверхнормальномъ разсѣяніи состоитъ изъ двухъ частей; первая имѣетъ видъ уклоненія при нормальномъ типѣ разсѣянія, вычисленнаго по комбинаторному способу:

$$\frac{p_0\, q_0}{s};$$

вторая

$$\frac{\Sigma (p_i - p_0)^2}{z}$$

представляетъ дополнительное уклоненіе физикальнаго типа, соотвѣтствующее уклоненіямъ частныхъ вѣроятностей, лежащихъ въ основѣ каждой серіи наблюденій, отъ общей основной вѣроятности, выраженіемъ которой служитъ средняя изъ частныхъ вѣроятностей. Степень приближенія Q къ 1 зависитъ отъ отношенія уклоненія физикальнаго къ комбинаторному; число случаевъ s въ каждой серіи оказываетъ вліяніе на комбинаторное уклоненіе и остается безъ вліянія на уклоненіе физикальнаго типа; если это число мало, нормальныя уклоненія комбинаторнаго типа достигаютъ сравнительно большой величины и покрываютъ собою дополнительныя уклоненія частныхъ вѣроятностей каждой серіи отъ общей вѣроятности, и общее уклоненіе приближается къ нормальному. Когда же число случаевъ въ каждой серіи возрастаетъ, нормальныя уклоненія комбинаторнаго типа уменьшаются и выступаютъ наружу уклоненія физикальныя.

Отсюда мы должны сдѣлать слѣдующій выводъ, важный въ практическомъ отношеніи: если совокупности съ малымъ числомъ наблюденій обнаруживаютъ разсѣяніе, близкое къ нормальному, это обстоятельство не даетъ права заключать, что мы имѣемъ дѣло съ вѣроятностью элементарнаго (теорема Бернулли) или случайно-перемѣннаго состава (законъ большихъ чиселъ Пуассона); такой выводъ будетъ правильнымъ только въ томъ случаѣ, если, и при образованіи совокупностей изъ бо́льшаго числа наблюденій, разсѣяніе сохраняетъ нормальный характеръ; если же при этомъ условіи разсѣяніе становится сверхнормальнымъ, то нормальный характеръ его ранѣе обуславливался не простымъ или случайно-перемѣннымъ составомъ вѣроятности, а лишь особенностью малыхъ совокупностей.

Отъ этого случая слѣдуетъ отличать другой случай устойчивости рядовъ, близкой къ нормальному разсѣянію, который проф. Борткевичъ также называетъ «закономъ малыхъ чиселъ». Здѣсь устойчивость обуславливается не малымъ числомъ случаевъ наблюденія, а малымъ числомъ случаевъ появленія событія, его рѣдкостью, или иначе говоря, малой величиной вѣроятности появленія событія. Такого рода устойчивость объясняется тѣмъ, что рѣдко наступающія событія раздѣлены большими промежутками времени или пространства и вслѣдствіе этого въ значительной степени независимы другъ отъ друга; случаи этого рода не подходятъ подъ схему сверхнормальнаго разсѣянія, гдѣ предполагается, что цѣлая серія наблюденій находится сплошь подъ дѣйствіемъ одинаковыхъ причинъ, сообщающихъ появленію событія для всей серіи одну и ту же вѣроятность, которая мѣняется лишь отъ серіи къ серіи; данный же случай рѣдкихъ событій, благодаря разобщенности ихъ, скорѣе подходитъ подъ ту схему, когда причина, сообщающая событію опредѣленную вѣроятность появленія, опредѣляется случайно и особо для каждаго испытанія, т. е. подъ схему закона большихъ чиселъ Пуассона, почему и уклоненіе должно приближаться здѣсь къ нормальному. Аналитически это условіе выражается въ характерѣ формулы среднеквадратическаго уклоненія

$$\frac{p_0 q_0}{s} + \frac{k-1}{s} \frac{S(p_i - p_0)^2}{z},$$

s по прежнему обозначаетъ число испытаній, изъ котораго образуется каждая частная совокупность, и, среди числа s испытаній, к обозначаетъ случаи, связанные дѣйствіемъ одной и той же причины; при $k = 1$, уклоненіе превращается въ нормальное

$$\frac{p_0 q_0}{s}$$

и задача отвѣчаетъ условіямъ вѣроятности случайно-перемѣннаго состава; при $k = s$, формула превращается въ формулу сверхнормальнаго уклоненія (отличаясь отъ ранѣе приведеннаго вида лишь болѣе точнымъ выраженіемъ— присоединеніемъ коэффиціента $\frac{s-1}{s}$, который при достаточно большомъ s близокъ къ 1).

Для малыхъ чиселъ появленія событій среднеквадратическое уклоненіе комбинаторнаго типа можетъ быть выражено (при p очень маломъ, q почти равномъ 1 и достаточно большомъ n) въ видѣ

$$\sqrt{m},$$

гдѣ $m = np$. Поэтому коэффиціентъ расхожденія получаетъ здѣсь видъ

$$Q = \sqrt{\frac{\Sigma(m_i - m)^2}{z}} : \sqrt{m}$$

Примѣръ. Въ шести русскихъ книгахъ * буква «о» встрѣчается на 100 буквъ текста слѣдующее число разъ:

На 100 буквъ — «о» встрѣчается разъ.	Число сотенъ съ такой частотой «о».
3	1
4	3
6	4
7	6
8	14
9	3
10	13
11	9
12	6
13	5
14	7
15	3
16	1
17	1
18	1
19	2
21	1

На основаніи всего матеріала общая вѣроятность буквы «о» составитъ

$$p = 0{\cdot}105\,;$$

(точнѣе $p = 0{\cdot}104\ 875$)

для отдѣльныхъ источниковъ вѣроятность (на основаніи 1000 буквъ въ каждомъ случаѣ) составитъ:

$p_1 = 0{\cdot}093$ $p_3 = 0{\cdot}086$ $p_5 = 0{\cdot}133$ $p_7 = 0{\cdot}101$

$p_2 = 0{\cdot}103$ $p_4 = 0{\cdot}121$ $p_6 = 0{\cdot}118$ $p_8 = 0{\cdot}084$

Изъ указанныхъ 80 сотенъ сдѣланъ 50 разъ тиражъ наугадъ по одной сотнѣ и въ результатѣ получились слѣдующіе цыфры: (v — число буквъ «о» въ сотнѣ буквъ, f — число сотенъ съ соотвѣтствующей величиной v, x — уклоненіе каждаго v отъ средней ариѳметической)

* Пушкинъ (Скупой рыцарь), Гоголь (Ревизоръ), Гончаровъ (въ Университетѣ), Тургеневъ (Воспоминаніе), Короленко (Тѣни), Проблеммы Идеализма, Очерки реалистическаго мировозрѣнія, Законы Гражд. (т. X, ч. 1, изд. 1887 г.); счеть съ начала текста.

v	f	fv	x	fxx
4	3	12	—6·8	138·72
5	—			
6	2	12	—4·8	46·08
7	2	14	—3·8	28·88
8	6	48	—2·8	47·04
9	1	9	—1·8	3·24
10	8	80	—0·8	5·12
11	10	110	0·2	0·40
12	5	60	1·2	7·20
13	4	52	2·2	19·36
14	5	70	3·2	51·20
15	—			
16	1	16	5·2	27·04
17	—			
18	—			
19	3	57	8·2	201·72
Σ	50	540		576·00

$$Me = \frac{540}{50} = 10·8$$

$$\delta^2 = \frac{576}{50} = 11·52 \qquad \delta = 3·4$$

Среднеквадратическое уклоненіе по способу комбинаторному, въ предположенiи элементарной вѣроятности, составитъ:

$$\sigma^2 = 100 \times 0·104\ 875 \times 0·895\ 125 = 9·387\ 623$$

$$\text{или} \quad \sigma = 3·06 \text{ и } Q = \frac{3·4}{3·1} = 1·1$$

Такимъ образомъ разсѣяніе оказывается не выходящимъ за предѣлы нормальнаго. Составимъ однако группы изъ 200 буквъ, для чего по жребію будемъ опредѣлять всякій разъ автора и для даннаго автора по жребію же одну и другую сотню. Опытъ далъ слѣдующія вѣроятности появленія «о» въ каждой группѣ:

v	f
0·045	1
0·060	1
0·065	1
0·070	1
0·075	1
0·080	1
0·090	2
0·100	1
0·105	5
0·110	2
0·115	1
0·120	1
0·125	1
0·130	1
0·150	1
0·160	2
0·165	1

$\Sigma f = 25$, $\Sigma fv = 266$, $p = 0{\cdot}1064$

Далѣе:

$$\Sigma fx^2 = 0{\cdot}023577 \text{ и } \delta = \sqrt{0{\cdot}000943} = 0{\cdot}0307$$

При $n = 200$

$$\sigma^2 = \frac{0{\cdot}093\,876}{200} = 0{\cdot}00\,04\,69\,38 \text{ и } \sigma = 0{\cdot}0217 \; ;$$

откуда

$$Q = \frac{0{\cdot}0307}{0{\cdot}0217} = 1{\cdot}4 \; ,$$

что указываетъ на возможность сверхнормальнаго разсѣянія. Если допустить, что относительная каждой группы является снабженной двумя источниками уклоненій, а именно: уклоненіемъ частноотносительной (вѣроятность появленія (о) у того или другого автора) отъ общеотносительной и уклоненіемъ эмпирической относительной отъ соотвѣтствующей ей наиболѣе вѣроятной частноотносительной, то среднеквадратическое уклоненіе по комбинаторному способу выразится въ видѣ:

$$\sigma_c{}^2 = \frac{pq}{n} + \frac{\Sigma (p_i - p)^2}{z} \; ,$$

что для даннаго случая составитъ

$$0{\cdot}000\,469\,38 + \frac{0{\cdot}002\,174\,88}{8}$$

откуда

$$\sigma_c = 0{\cdot}0272$$

и

$$Q = 1{\cdot}12$$

Если вмѣсто двухъ сотенъ, каждую группу составить такимъ же образомъ изъ четырехъ сотенъ, то опытъ даетъ слѣдующіе результаты (v — вѣроятность появленія «о» въ каждой группѣ, f — число группъ, x уклоненіе отъ средней):

v	f	x	fxx
0·0675	1	—0408	00166464
·0775	1	—0308	00094864
·0800	1	—0283	... 80089
·0850	3	—0233	.. 162867
·0875	2	—0208	... 86528
·0900	1	—0183	... 33489
·0925	1	—0158	... 24964
·0950	1	—0133	... 17689
·1025	2	—0058	... 06728
·1050	1	—0033	... 01089
·1125	2	0042	... 03528
·1200	1	0117	... 13689
·1250	1	0067	... 27889
·1275	1	0192	... 36864
·1325	1	0242	... 58564
·1400	2	0317	.. 200978
·1450	1	0367	.. 134689
·1500	1	0417	.. 173889
·1600	1	0517	.. 267289
Me = 0·1083	25		S = ·0159215

$$\delta^2 = \frac{0{\cdot}0159215}{25} = 0{\cdot}00063586$$

$$\delta = 0{\cdot}0252$$

$$\sigma^2 = \frac{0{\cdot}093876}{400} = 0{\cdot}00023469$$

$$\sigma = 0\ 0153$$

откуда

$$Q = \frac{0{\cdot}0252}{0{\cdot}0153} = 1{\cdot}6.$$

Если образовать σ въ предположеніи составнаго характера вѣроятности, мы получимъ:

$$\sigma^2 = 0{\cdot}00023469 + 0{\cdot}00027186 = 0\ 00050655$$

$$\sigma = 0{\cdot}0225$$

и

$$Q = \frac{0\ 0252}{0{\cdot}0225} = 1{\cdot}1.$$

Изъ сопоставленія трехъ приведенныхъ случаевъ, мы видимъ, что для группъ изъ одной, двухъ, четырехъ сотенъ Q соотвѣтственно ровно

$$1{\cdot}1;\ 1{\cdot}4;\ 1{\cdot}6,$$

т. е. чѣмъ совокупность состоитъ изъ большаго числа случаевъ, тѣмъ больше проявляется сверхнормальность разсѣянія. Разсматривая среднеквадратическое уклоненіе σ_c, вычисленное въ предположеніи составнаго характера вѣроятности, мы находимъ, что для группъ изъ одной сотни оно состоитъ для σ_c^2 изъ двухъ слагаемыхъ

$$\frac{pq}{100} = 0{\cdot}0009388 \quad \text{и} \quad \frac{\Sigma\,(p_i - p)}{8} = 0{\cdot}0002719\ ,$$

причемъ второе составляетъ 0·3 перваго, такъ что σ_c^2 больше σ^2 (т. е. уклоненія, вычисленного въ предположеніи элементарнаго состава вѣроятности) въ отношеніи 1·3 къ 1 или

$$\frac{\sigma_c}{\sigma} = 1{\cdot}1\ ,$$

иначе говоря, первый членъ составнаго уклоненія почти покрываетъ собою вторую часть его. Для группъ, составленныхъ изъ четырехъ сотенъ, вторая часть въ выраженіи σ_c^2 остается безъ измѣненія, первый же членъ уменьшается въ 4 раза, вслѣдствіе чего вторая часть уже нѣсколько превышаетъ первую, и квадратъ σ_c превышаетъ квадратъ σ въ два слишкомъ раза.

Образуемъ теперь группы по пяти сотенъ въ каждой, но способомъ, отличнымъ отъ предшествующаго; а именно, вмѣсто того, чтобы, опредѣливъ жребіемъ источникъ, брать по жребію же изъ этого источника всѣ пять сотенъ, будемъ брать каждую сотню всякій разъ по жребію изъ всѣхъ 80 сотенъ и изъ пяти послѣдовательно вынутыхъ сотенъ образуемъ одну общую группу. Въ такомъ случаѣ среднеквадратическое уклоненіе, вмѣсто прежняго вида

$$\sigma_c^2 = \frac{pq}{n} + \frac{n-1}{n} \cdot \frac{\Sigma\,(p_i - p)^2}{z}\ ,$$

получимъ выраженіе

$$\frac{pq}{n} + \frac{k-1}{n} \cdot \frac{\Sigma (p_i - p)^2}{z},$$

гдѣ $n = 500$ и $k = 100$.

Опытъ даетъ слѣдующіе цыфры:

	v	x	x^2
	0·102	0·014	0·000196
	0·108	—0·006	36
	0·094	—0·014	196
	0·128	0·020	400
	0·118	0·010	100
	0·136	0·028	784
	0·088	—0·020	400
	0·086	—0·022	484
	0·098	—0·010	0·000100
Me =	0·108		0·002696

$$\delta^2 = \frac{0{\cdot}002696}{10} = 0{\cdot}0002696$$

$$\delta = 0{\cdot}0164.$$

Среднеквадратическое σ для элементарнаго состава вѣроятности составитъ

$$\sigma^2 = \frac{0{\cdot}00093876}{5} = 0{\cdot}00018775$$

$$\sigma = 0{\cdot}0137.$$

Если однако вычислить σ въ предположеніи сложнаго состава и принимая во вниманіе условія опыта, мы получимъ:

$$\sigma_c^2 = 0{\cdot}00018775 + \frac{1}{5} \times 0{\cdot}00027186 = 0{\cdot}00024212,$$

и $\sigma_c^2 = 0{\cdot}0156$,

т. е. величину болѣе близкую къ δ. При такомъ образованіи группъ разсѣяніе почти совпадаетъ съ нормальнымъ.

5. Корреляція.

Сводная зависимость. Зависимость между двумя явленіями можетъ быть изучаема либо общими пріемами изслѣдованія либо статистически. Первый способъ примѣнимъ въ тѣхъ случаяхъ, когда зависимость проявляется тождественно въ каждомъ случаѣ наблюденія, или когда изъ наблюденія путемъ эксперимента могутъ быть устранены всѣ обстоятельства, видоизмѣняющія и нарушающія однообразное проявленіе зависимости. При отсутствіи указанныхъ условій, зависимость двухъ явленій скрывается или извращается въ отдѣльныхъ случаяхъ привходящими обстоятельствами и можетъ быть подмѣчена лишь сведеніемъ многихъ случаевъ въ одну общую совокупность. При количественномъ характерѣ явленій, такого рода способъ изслѣдованія составляетъ одну изъ задачъ статистическаго метода.

Сопоставленіе рядовъ. Для того, чтобы найти, какого рода соотношеніе существуетъ между количественными значеніями двухъ явленій, мы, прежде всего, располагаемъ случаи наблюденія въ рядъ по возрастающей или убывающей величинѣ одного явленія и противъ каждаго члена ряда ставимъ соотвѣтствующую ему величину второго явленія. Если въ первомъ рядѣ какая-либо величина встрѣчается нѣсколько разъ и ей отвѣчаютъ разныя величины второго ряда, то возьмемъ среднюю ариѳметическую изъ послѣднихъ. Если между рядами существуетъ зависимость, она обнаруживается въ томъ, что, съ возрастаніемъ перваго ряда, величины второго ряда такъ же возрастаютъ (прямая зависимость) или убываютъ (обратная зависимость); при отсутствіи зависимости величины второго ряда чередуются безъ всякой правильности. Если зависимость выражена строго, она обнаруживается наглядно;

но при болѣе или менѣе неправильномъ чередованіи величинъ второго ряда, сужденіе о характерѣ соотношенія рядовъ становится затруднительнымъ. Наглядное сравненіе рядовъ облегчается, если вмѣсто того, чтобы сравнивать величины въ ихъ непосредственно данномъ видѣ, за основаніе для сравненія взять въ каждомъ рядѣ одну какую-либо величину—начальную, конечную, среднюю или сумму членовъ, и принявъ ее за 100, 1000 и т. п., остальныя выразить пропорціонально принятому основанію; ряды становятся нагляднѣе, но опредѣленіе общаго соотношенія при неправильномъ чередованіи величинъ все же мало облегчается.

Дѣленіе рядовъ на части. Эта цѣль въ извѣстной степени достигается дѣленіемъ рядовъ на двѣ, три, четыре... части, опредѣленіемъ средней ариѳметической для каждой части и сопоставленіемъ этихъ величинъ въ обоихъ рядахъ; если между рядами существуетъ зависимость—прямая или обратная, то она выразится въ томъ, что съ возрастаніемъ среднихъ перваго ряда возрастаютъ или убываютъ среднія второго ряда или то же съ убываніемъ среднихъ перваго ряда; при отсутствіи зависимости, среднія будутъ чередоваться безъ опредѣленнаго порядка. Пріемъ этотъ непримѣнимъ въ томъ случаѣ, когда число строкъ ряда невелико; но и тогда, когда онъ примѣнимъ, онъ не даетъ *мѣры* соотношенія рядовъ. Такую цѣлесообразно выбранную мѣру соотношенія рядовъ даетъ пріемъ изслѣдованія, называемый способомъ корреляціи.

Нахожденіе коэффиціента корреляціи. За основаніе для сравненія при опредѣленіи корреляціи принимается средне-ариѳметическая одного и другого ряда. Если мы возьмемъ уклоненія каждаго члена ряда отъ его средней, мы получаемъ возможность сравнивать уклоненія обоихъ рядовъ по ихъ знаку и величинѣ; однако сравненіе уклоненій непосредственно по данной ихъ величинѣ невозможно, такъ какъ явленія одного и другого ряда (а слѣдовательно, и ихъ уклоненія) могутъ быть выражены въ разныхъ мѣрахъ и наименованіяхъ; поэтому уклоненія каждаго ряда необходимо предварительно выразить въ какой-либо однородной мѣрѣ; цѣлесообразно для этой цѣли вмѣсто абсолютной величины уклоненія взять отношеніе его къ среднеквадратическому уклоненію: тогда уклоненія, находящіяся въ одинаковомъ отношеніи къ среднеквадратическому уклоненію своего ряда, получатъ одинаковое численное значеніе.

Представимъ себѣ, что величины перваго ряда расположены въ возрастающемъ порядкѣ; тогда уклоненія этого ряда сперва будутъ имѣть отрицательный знакъ и будутъ убывать по абсолютной величинѣ по направленію къ точкѣ средней ариѳметической; перейдя же за эту точку, станутъ положительными и будутъ возрастать. Если между первымъ рядомъ и вторымъ существуетъ зависимость, то съ возрастаніемъ членовъ перваго ряда будутъ возрастать и члены второго ряда, при прямой зависимости, или обратно— при обратной. Возьмемъ первый случай; тогда отрицательнымъ уклоненіямъ перваго ряда будутъ отвѣчать отрицательныя уклоненія второго и положительнымъ—положительныя; такое совпаденіе знаковъ будетъ полнымъ, если среднеариѳметическія лежатъ въ одномъ и въ другомъ рядѣ въ одномъ и томъ же мѣстѣ ряда, и въ обоихъ рядахъ всѣ величины, меньше средней, будутъ лежать въ одной половинѣ ряда, а бо́льшія въ другой. При обратной зависимости и при тѣхъ же вышеупомянутыхъ условіяхъ положительнымъ уклоненіямъ перваго ряда будутъ отвѣчать отрицательныя второго и отрицательнымъ—положительныя. При такомъ полномъ совпаденіи знаковъ—прямомъ или обратномъ — если перемножить уклоненія попарно, то при прямой зависимости всѣ произведенія будутъ положительными, при обратной, всѣ будутъ отрицательными. Если же среднеариѳметическія лежатъ въ разныхъ точкахъ ряда, или во второмъ ряду среди членовъ меньшихъ, чѣмъ среднеариѳметическая, оказываются члены больше среднеариѳметической и обратно, то нѣкоторымъ положительнымъ уклоненіямъ перваго ряда при прямой зависимости будутъ отвѣчать отрицательныя уклоненія второго ряда и обратно; вслѣдствіе чего нѣкоторыя произведенія попарныхъ уклоненій будутъ имѣть отрицательный знакъ при прямой зависимости рядовъ и положительный при обратной.

Если же между рядами не существуетъ никакой зависимости, то положительнымъ уклоненіямъ перваго ряда будутъ отвѣчать то положительныя, то отрицательныя уклоненія второго ряда и то же относительно отрицательныхъ уклоненій перваго ряда; въ такомъ случаѣ и попарныя произведенія будутъ то положительными, то отрицательными. Если сложить всѣ произведенія и сумму раздѣлить на число слагаемыхъ, то полученная среднеариѳметическая всѣхъ произведеній, при прочихъ равныхъ условіяхъ, будетъ имѣть наибольшую величину при полномъ совпаденіи и положительный знакъ при прямой зависимости, отрицательный—при обратной; и меньшую величину—при неполномъ совпаденіи; при отсутствіи же всякой зависимости между рядами, она будетъ стремиться къ 0.

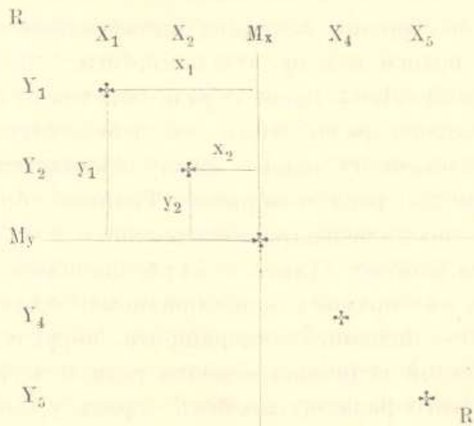
Съ другой стороны, величина произведеній, и слѣдовательно и средней изъ произведеній, будетъ больше, если бо́льшимъ уклоненіямъ одного ряда отвѣчаютъ бо́льшія уклоненія второго ряда, чѣмъ въ томъ случаѣ, когда бо́льшимъ уклоненіямъ одного ряда отвѣчаютъ меньшія уклоненія другого ряда и обратно. Такимъ образомъ по величинѣ и знаку среднеариѳметической изъ произведеній уклоненій мы можемъ судить о мѣрѣ зависимости рядовъ и характерѣ ея—прямомъ или обратномъ. Указанная величина носитъ названіе коэффиціента корреляціи. Если уклоненія каждой величины перваго ряда отъ среднеариѳметической этого ряда обозначимъ черезъ y_1, y_2 . . . y_n и второго ряда черезъ x_1 x_2 . . . x_n; среднеквадратическія уклоненія черезъ σ_x и σ_y и число членовъ ряда черезъ N, то коэффиціентъ корреляціи (r) получитъ слѣдующее выраженіе:

$$\frac{\frac{x_1}{\sigma_x}\frac{y_1}{\sigma_y}+\frac{x_2}{\sigma_x}\frac{y_2}{\sigma_y}+\ldots+\frac{x_n}{\sigma_x}\frac{y_n}{\sigma_y}}{N}=\frac{S(xy)}{N\sigma_x\sigma_y}=r\,.$$

Коэффиціентъ корреляціи, какъ мы увидимъ, обладаетъ тѣмъ свойствомъ, что величина его не выходитъ за предѣлы отъ $+1$ до -1; значеніе $+1$ отвѣчаетъ случаю полной прямой зависимости рядовъ, значеніе -1 — полной обратной; значеніе 0 указываетъ на отсутствіе всякой зависимости.

Графическое изображеніе. Представимъ себѣ случай идеальной зависимости двухъ рядовъ Y X:

Y X изобразимъ значеніе Y и X графически, какъ по-
1 1 казано ниже; линія M_y отвѣчаетъ среднеариѳмети-
2 2 ческой точкѣ ряда Y, линія M_x — такой же точкѣ
3 3 ряда X; обозначимъ крестами точки пересѣченія
4 4 линій Y_1 и X_1; Y_2 и X_2; и т. д.; всѣ эти точки
5 5 пересѣченія лежатъ на одной прямой RR.



x_1 — представляетъ уклоненіе X_1 отъ M_x, y_1 уклоненіе Y_1 отъ M_y и т. д. Изъ чертежа видно, что при полной зависимости рядовъ

$$x_1/y_1 = x_2/y_2 = \ldots = x_i/y_i = \ldots$$

и $x_1^2/y_1^2 = x_2^2/y_2^2 = \ldots \; x_i^2/y_i^2 = \dfrac{x_1^2+x_2^2+\ldots x_i^2+\ldots}{y_1^2+y_2^2+\ldots+y_i^2}$

или $$x_i/y_i = \sigma_x/\sigma_y$$

Такимъ образомъ при полной зависимости отношенія каждой пары уклоненій равны постоянной величинѣ, и потому всѣ уклоненія лежатъ на одной прямой. Въ тѣхъ случаяхъ когда зависимость между рядами неполная, точки пересѣченія каждой пары уклоненій уклоняются отъ прямой то въ одну, то въ другую сторону, и отношенія каждой пары уклоненій не даютъ одной и той же постоянной величины, но вмѣсто того мы получаемъ рядъ равенствъ:

Коэффиціентъ корреляціи и линія корреляціи.

$$x_1/y_1 = b_1$$
$$x_2/y_2 = b_2$$
$$\ldots\ldots\ldots$$
$$x_i/y_i = b_i$$
$$\ldots\ldots\ldots$$

Задача опредѣленія корреляціи состоитъ въ томъ, чтобы вмѣсто разныхъ b_1, b_2 . . . найти одно наиболѣе подходящее b_0, или, говоря иначе, чтобы провести такую прямую RR, которая возможно ближе проходила бы ко всѣмъ точкамъ пересѣченія въ ихъ общей совокупности, и такимъ образомъ замѣняла собой ломаную линію, соединяющую точки пересѣченія въ дѣйствительности. Задачу эту можно выполнить графически на глазъ, но это неудобно потому, что прямая RR можетъ получать разныя направленія, въ зависимости отъ субъективнаго усмотрѣнія изслѣдователя. Поэтому цѣлесообразно поставить какое либо объективное условіе, которое опредѣляло бы направленіе линіи RR. Подходящимъ условіемъ будетъ то, чтобы сумма квадратовъ уклоненій дѣйствительныхъ точекъ пересѣченія отъ замѣняющихъ ихъ точекъ прямой RR была минимальной; условіе это сводится къ тому, чтобы выраженіе

$$(b_1 - b_0)^2 + (b_2 - b_0) + \ldots + (b_i - b_0)^2 + \ldots$$

имѣло наименьшую величину, или

$$(-x_1 + b_0\, y_1)^2 + (-x_2 + b_0\, y_2)^2 + \ldots = \min.$$

Рѣшеніе этой задачи (см. слѣдующую главу) даетъ для b_0 значеніе

$$b_0 = \frac{S(xy)}{S(y^2)} = \frac{S(xy)}{N\,\sigma_y^2} = \frac{S(xy)}{N\,\sigma_y^2} \cdot \frac{\sigma_x}{\sigma_x} = \frac{S(xy)}{N\,\sigma_x\,\sigma_y} \cdot \frac{\sigma_x}{\sigma_y} = r\,\frac{\sigma_x}{\sigma_y},$$

и поэтому

$$x_i = r\,\frac{\sigma_x}{\sigma_y}\,y_i\,.$$

Коэффиціентъ корреляціи имѣетъ такимъ образомъ то значеніе, что замѣняетъ дѣйствительныя уклоненія теоретическими съ такимъ расчетомъ, чтобы сумма квадратовъ разницъ между дѣйствительными и теоретическими значеніями уклоненій была минимальной, и ломаную линію корреляціи замѣняетъ прямой. При $x_i / y_i = \frac{\sigma_x}{\sigma_y}$, т. е. при условіи идеальной зависимости рядовъ, абсолютная величина r, какъ видно изъ формулы

$$x_i = r\,\frac{\sigma_x}{\sigma_y}\,y_i$$

должна равняться 1.

Примѣръ. Для поясненія возьмемъ ряды Y и X, представляющіе полную зависимость, но выраженные въ различныхъ абсолютныхъ величинахъ; дѣленіе каждаго уклоненія на среднеквадратическое обнаруживаетъ ихъ одинаковую относительную величину въ обоихъ рядахъ;

совершенная зависимость рядовъ выражается въ приблизительномъ равенствѣ $r = 1$:

Y	X	y	x	y/σ_y	x/σ_x	$\frac{y\ x}{\sigma_y\ \sigma_x}$
1	2	—2	—4	—1·42	—1·40	1·988
2	4	—1	—2	—0·71	-0·70	0·497
3	6	0	0	0·	0·	0·
4	8	1	2	0·71	0·70	0·497
5	10	2	4	1·42	1·40	1·988
						4·970

$$M_Y = 3 \quad \sigma_y = \sqrt{10/5} = 1·41$$

$$M_X = 6 \quad \sigma_x = \sqrt{40/5} = 2·83$$

$$r = \frac{4·97}{5} = 0·994$$

Въ слѣдующемъ примѣрѣ приведенъ случай менѣе совершенной корреляціи. Пусть даны ряды:

Y	X	y	x	y/σ_y	x/σ_x	$\frac{x\ y}{\sigma_x\ \sigma_y}$
1	1	—2	—2·8	—1·42	—1·54	2·1868
2	3	—1	—0·8	—0·71	—1·44	0·3124
3	3·5	0	—0·3	0·	—1·16	0·
4	5·5	1	1·7	0·71	0·94	0 6638
5	6	2	2·2	1·42	1·21	1·7182
						4·8812

$$M_y = 3 \quad \sigma_y = 1·41$$

$$M_x = 3·8 \quad \sigma_x = \sqrt{3·26} = 1·80$$

$$r = \frac{48·8}{5} - 0·97624$$

Если примемъ за основаніе рядъ Y и по даннымъ уклоненіямъ вычислимъ значенія x, соотвѣтствующія коэффиціенту корреляціи, то получимъ:

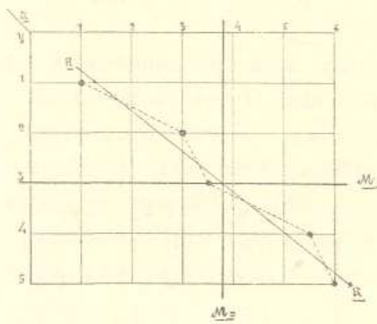
$$x_i = 0·97624 \frac{1·8}{1\ 41} y_i = 1\ 25\ y_i\ ;$$

выбравъ двѣ точки, мы опредѣлимъ положеніе кривой; такъ

$$\text{при } y_2 = -2,\ x_2 = -2{\cdot}5 \text{ и}$$
$$\text{при } y_3 = -2,\ x_3 = +2{\cdot}5$$

такъ какъ уклоненія считаются отъ средней, то величина $X_1 = 3{\cdot}8 - 2{\cdot}5 = 1{\cdot}3$ и $X_3 = 3{\cdot}8 + 2{\cdot}5 = 6{\cdot}3$. Проведя черезъ эти точки прямую, получимъ линію корреляціи RR.

Уголъ, который образуетъ линіи RR съ осями координатъ, указываетъ на величину корреляціи: при полной корреляціи эта линія идетъ по діогонали, при отсутствіи корреляціи становится перпендикулярной къ оси абсциссъ; при прямой зависимости она проходитъ въ направленіи, какъ показано на рисункѣ; при обратной зависимости она располагается въ обратномъ направленіи, по другой діогонали.



Корреляція односторонняя. При вычисленіи коэффиціента корреляціи, мы располагаемъ величины одного ряда по равнымъ интерваламъ, изъ соотвѣтствующихъ величинъ другого ряда беремъ среднеариѳметическія; если же этотъ второй рядъ расположить по равнымъ интерваламъ, а для перваго ряда взять изъ соотвѣтствующихъ величинъ среднеариѳметическія, то отъ перераспредѣленія отдѣльныхъ случаевъ между группами коэффиціентъ корреляціи мѣняется. Такимъ образомъ одни и тѣ же данныя, въ зависимости отъ способа сопоставленія рядовъ, даютъ два коэффиціента корреляціи; для того, чтобы получить все же одну величину, можно взять среднеариѳметическую изъ обоихъ коэффиціентовъ.

Двусторонняя съ весомъ. Но та же цѣль достигается лучше и правильнѣе, если принять во вниманіе не только соотвѣтствующія другъ другу величины, но и число случаевъ каждаго сочетанія. Для того, чтобы учесть всѣ имѣющіеся случаи различныхъ сочетаній двухъ количественныхъ явленій, удобно составить комбинаціонную таблицу, въ которой сверху нужное число графъ обозначаютъ интервалы ряда X и строки слѣва интервалы ряда Y, какъ показано въ изображенной таблицѣ; въ соотвѣтствующихъ клѣткахъ таблицы показаны частоты сочетанія каждаго интервала Y съ каждымъ интерваломъ X; интервалъ Y 4 встрѣчается въ 4 случаяхъ въ сочетаніи съ интерваломъ X 2, въ 10 случаяхъ съ интерваломъ 3, въ 5 случаяхъ съ интерваломъ 4 и т. п.; точно такъ же интервалъ X 1 встрѣчается въ 5 случаяхъ съ $Y = 1$, въ 4 случаяхъ съ $Y = 2$ и въ 1 случаѣ съ $Y = 3$. Съ правой стороны въ графѣ n_y проставлены частоты каждаго интервала Y, получающіяся суммированіемъ соотвѣтствующей строки; внизу въ графѣ n_x — частоты интерваловъ ряда X; въ послѣдней графѣ съ правой стороны (x_m) выведены среднеариѳметическія изъ X, соотвѣтствующія каждой строкѣ (интервалу) y; внизу то же сдѣлано для y_m по отношенію къ интерваламъ X (графамъ).

Сопоставляя интервалы Y съ соотвѣтствующими ему среднеариѳметическими изъ X (или x_m) получаемъ ряды I и, сопоставляя интервалы X съ y_m, получаемъ ряды II;

X / Y	1	2	3	4	5	n_y	X_m
1	5	2	1	1	1	10	2·1
2	4	8	3	2	1	18	2·3
3	1	11	20	7	1	40	2·9
4		4	10	5	1	20	3·2
5			1	10	1	12	4·0
n_x	10	25	35	25	5	100	
Y_m	1·6	2·7	3·2	3·8	3·0		

I		II	
Y	X	Y	X
1	2·1	1·6	1
2	2·3	2·7	2
3	2·9	3·2	3
4	3·2	3·8	4
5	4·0	3·0	5

Для каждой пары рядовъ корреляція можетъ быть опредѣлена особо.

Для I:	для II:
$Y_m = 3 \quad \sigma_y = \sqrt{10/5}$	$Y_m = 2{\cdot}86 \quad \sigma_y = \sqrt{\frac{2{,}632}{5}}$
$X_m = 2{\cdot}9 \quad \sigma_x = \sqrt{2{\cdot}3/5}$	$X_m = 3 \quad \sigma_x = \sqrt{10/5}$
$S\,(xy) = 4{\cdot}7$	$S\,(xy)\;3{\cdot}9$
$r = \frac{4{\cdot}7}{\sqrt{23}} = 0{\cdot}98$	$r = \frac{3{\cdot}9}{\sqrt{26{\cdot}32}} = 0{\cdot}76$

Средняя изъ обоихъ r даетъ

$$r_0 = 0{\cdot}87.$$

Этотъ способъ опредѣленія коэффиціента корреляціи неточенъ въ томъ отношеніи, что не принимаетъ во вниманіе частоту разныхъ сочетаній; принимая во вниманіе послѣднюю, мы не только избѣгаемъ указанной неточности, но и получаемъ только одно значеніе для r.

Среднеариѳметическая для X (съ вѣсомъ) равна 2·9; для Y — 3·06; уклоненія отдѣльныхъ варіантъ X и Y даютъ слѣдующую таблицу:

Y \ X	−1·9	−0·9	0·1	1·1	2·1
−2·06	5	2	1	1	1
−1·06	4	8	3	2	1
−0·06	1	11	20	7	1
0·94		4	10	5	1
1·94			1	10	1

Для полученія S (xy) необходимо сдѣлать слѣдующій рядъ вычисленій:

$$-1{\cdot}9\times-2{\cdot}06\times5+(-0{\cdot}9\times-2{\cdot}06\times2)+(0{\cdot}1\times-2{\cdot}06\times1)+$$
$$\text{и т. д.}+(1{\cdot}94\times0{\cdot}1\times1)+(1{\cdot}94\times1{\cdot}1\times10)+(1{\cdot}94\times2{\cdot}1),$$

или для сокращенія дѣйствій:

$$-2{\cdot}06\times(-1{\cdot}9\times5-0{\cdot}9\times2+0{\cdot}1+1{\cdot}1+2{\cdot}1)=-2{\cdot}06\times-8$$
$$-1{\cdot}06(-1{\cdot}9\times4-0{\cdot}9\times8+0{\cdot}1\times3+1{\cdot}1\times2+2{\cdot}1)=$$
$$-1{\cdot}06\times-10{\cdot}2$$
$$-0{\cdot}06(-1{\cdot}9-0{\cdot}9\times11+0{\cdot}1\times20+1{\cdot}1\times7+2{\cdot}1)=-0{\cdot}06\times0$$
$$0{\cdot}94(-0{\cdot}9\times4+0{\cdot}1\times10+1.1\times5+2{\cdot}1)=0{\cdot}94\times5$$
$$1{\cdot}94(0{\cdot}1+1{\cdot}1\times10+2{\cdot}1)=1{\cdot}94\times13{\cdot}2$$

$$\text{Сумма } S(fxy)=57{\cdot}6$$

Далѣе,

$$\sigma_x=0{\cdot}77;\ \sigma_y=1{\cdot}14 \text{ и } N=100$$

Откуда

$$r=\frac{57{\cdot}6}{100\times0{\cdot}77\times1{\cdot}14}=0{\cdot}656$$

Упрощеніе вычисленій посредствомъ моментовъ.

Такъ какъ вычисленіе произведеній уклоненій на частоты обременительно, въ особенности, когда уклоненія выражаются съ дробями и требуется точность, то чрезвычайно удобно пользоваться вычисленіемъ посредствомъ способа моментовъ. Для этого какой либо интервалъ X въ серединѣ ряда принимается за 0, а другіе соотвѣтственно обозначаются черезъ — 1, — 2, — 3... съ одной стороны и 1, 2, 3... съ другой; то же дѣлаетсв для Y. Рядъ этихъ цифръ обозначаетъ уклоненія интерваловъ X отъ принятой исходной точки 0 въ условныхъ разстояніяхъ, равныхъ 1, и то же для Y; каждому сочетанію этихъ уклоненій отвѣчаетъ частота, стоящая въ соотвѣтствующей клѣткѣ; одна графа частотъ отвѣчаетъ уклоненіямъ X, равнымъ 0, и одна строка частотъ отвѣчаетъ уклоненіямъ Y, равнымъ 0; этими взаимно перпендикулярными графой и строкой таблица дѣлится на четыре квадранта, изъ которыхъ два будутъ давать произведенія отрицательныхъ на отрицательныя уклоненія и положительныхъ на положительныя, два другихъ будутъ давать отрицательныя произведенія:

Y \ X	—2	—1	0	1	2
—2	5	2	1	1	1
—1	4	8	3	2	1
0	1	11	20	7	1
1		4	10	5	1
2			1	10	1

Найти произведеніе каждой пары уклоненій, умножить его на соотвѣтствующую частоту и суммировать результаты не представляетъ затрудненій:

I квадрантъ:

$$4\times5+2\times2+2\times4+1\times8 \qquad = \quad 40$$

IV квадрантъ:

$$1\times5+2\times1+2\times10+4\times1 \qquad = \quad 31$$

II квадрантъ:

$$-2\times1-4\times1-1\times2-2\times1 = -10$$

III квадрантъ:

$$-1\times4 \qquad = -4$$

$$\text{Сумма} = S\,(fxy) \qquad = \quad 57$$

и

$$\frac{S\,(fxy)}{N} = 0{\cdot}57$$

Чтобы перевести результатъ отъ условной точки 0 къ точкѣ среднеариѳметической для X и Y, необходимо изъ полученнаго результата вычесть произведеніе первыхъ моментовъ для X и Y; въ данномъ случаѣ

$$\nu_{x,\,1} = -\frac{10}{100} \text{ и } \nu_{x,\,2} = \frac{60}{100}$$

$$\nu_{y,\,1} = \frac{6}{100} \text{ и } \nu_{y,\,2} = \frac{126}{100}$$

и

$$\nu_{x,\,1} \times \nu_{y,\,1} = -\,0{\cdot}06,$$

отсюда отнесенное къ среднеариѳметическимъ точкамъ

$$\frac{S(fxy)}{N} = 0{\cdot}57 + 0{\cdot}06 = 0{\cdot}576.$$

Среднеквадратическія уклоненія по способу моментовъ даютъ

$$\text{для } X:\ \mu_2 = 0{\cdot}6 - (-0{\cdot}1)^2 = 0{\cdot}59 = \sigma_x{}^2$$
$$\text{и } \sigma_x = 0{\cdot}77$$
$$\text{для } Y:\ \mu_2 = 1{\cdot}26 - 0{\cdot}06^2 = 1{\cdot}2564$$
$$\text{и } \sigma_y = 1{\cdot}14.$$

Отсюда

$$r = \frac{0{\cdot}576}{0{\cdot}77 \times 1{\cdot}14} = 0{\cdot}656$$

По величинѣ r можно найти теоретическія значенія уклоненій x по даннымъ уклоненіямъ y и обратно; а именно

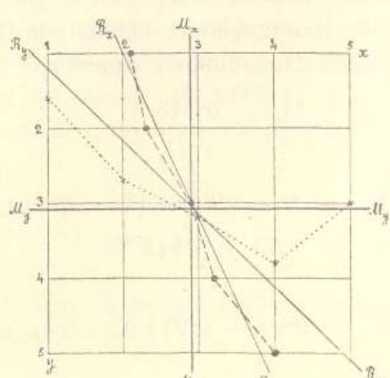
$$x = 0{\cdot}656\ \frac{0{\cdot}77}{1{\cdot}14}\ \ y = 0{\cdot}443\ y$$

и

$$y = 0{\cdot}656\ \frac{1{\cdot}14}{0{\cdot}77}\ \ x = 0{\cdot}971\ x.$$

Опредѣливъ два произвольно взятыхъ уклоненія и по нимъ отвѣчающія имъ значенія X и Y, получимъ возможность опредѣлить направленіе корреляціонныхъ линій; такъ

$$\begin{array}{llll} \text{для } y = -2{\cdot}06, & x = -0{\cdot}91 & \text{и } X = 2{\cdot}0 \\ \text{и } \ \ y = 1{\cdot}94, & x = 0{\cdot}86, & X = 3{\cdot}8; \\ \text{для } x = -1{\cdot}9, & y = -1{\cdot}85, & Y = 1{\cdot}2 \\ \text{и } \ \ x = 2{\cdot}1, & y = 2{\cdot}04 & Y = 5{\cdot}1. \end{array}$$



Переходъ къ зависимости между перемѣнными.

Указанный способъ опредѣленія коэффиціента корреляціи принимаетъ, что зависимость между уклоненіями x и y выражается уравненіемъ первой степени. Поэтому, когда мы переходимъ отъ уклоненій къ абсолютнымъ значеніямъ X и Y, послѣднія связываются функціей также первой степени; такъ, если

$$x_i = r \frac{\sigma_x}{\sigma_y} y_i \ ,$$

то подставивъ вмѣсто x_i его величину $X_i - M_x$ и, вмѣсто y_i, $Y_i - M_y$, получимъ

$$X_i - M_x = r \frac{\sigma_x}{\sigma_y} (Y_i - M_y) \ ;$$

и, замѣняя $r \frac{\sigma_x}{\sigma_y} = b$:

$$X_i = M_x - b\, M_y + b\, Y_i \ ;$$

такъ какъ среднеариѳметическія M_x и M_y опредѣляются на основаніи эмпирическаго матеріала, то принимая выраженіе

$$M_x - b\, M_y = a \ ,$$

получимъ

$$X_i = a + b\, Y_i \ .$$

Это выраженіе можетъ быть получено и непосредственно по способу наименьшихъ квадратовъ (см. сл. главу).

Для приведеннаго численнаго примѣра

$$x_i = 0{\cdot}443 \ y_i$$

или

$$X_i - 2{\cdot}9 = 0{\cdot}443 \ (Y_i - 3{\cdot}06)$$

$$X_i = 1{\cdot}55 + 0{\cdot}443 \ Y_i$$

и

$$Y_i = 0{\cdot}24 + 0{\cdot}971 \ X_i \ .$$

Однако прямолинейная зависимость между двумя перемѣнными не всегда отвѣчаетъ съ достаточною степенью приближенія эмпирическому матеріалу; нерѣдко, чтобы приблизить вычисляемыя значенія перемѣнныхъ къ эмпирическимъ, приходится прибѣгать къ выраженію ихъ зависимости въ формулахъ второй, третьей степеней; сообразно съ этимъ возможно и при вычисленіи коэффиціента корреляціи прибѣгать къ болѣе сложнымъ допущеніямъ; въ этихъ случаяхъ мы имѣемъ дѣло съ т. н. кривой корреляціей *.

Корреляція качественныхъ варіантъ. Въ заключеніе укажемъ упрощенный способъ вычисленія корреляціи, который заслуживаетъ вниманія еще и потому, что онъ приложимъ съ извѣстными допущеніями къ случаямъ качественной разницы явленій. Предположимъ, что варіанты ряда X раздѣлены всего на два интервала и то же варіанты Y; допуская, что оба интервала X равны между собою и также интервалы Y, возьмемъ исходную точку какъ разъ на границѣ интерваловъ и приравняемъ ее 0; тогда уклоненіе перваго интервала будетъ равно —0·5 и второго +0·5; въ остальномъ поступимъ, какъ и въ случаѣ вычисленія корреляціи съ частотами, т. е. вычислимъ S (fxy) сперва отъ условно принятой исходной точки, затѣмъ переведемъ результатъ къ точкѣ среднеариѳметической, вычтя произведеніе первыхъ моментовъ, и полученное раздѣлимъ на произведеніе среднеквадратическихъ уклоненій. При этомъ однако возможны значительныя сокращенія, такъ что окончательный результатъ можетъ быть полученъ гораздо проще. А именно, обозначивъ частоты буквами a, b, c, d, получимъ слѣдующую таблицу:

---

* См. А. Леонтовичъ, Элементарное пособіе къ примѣненію методовъ Gauss'a etc. ч. II, 1911, стр. 100—116.

y \ x	$-\frac{1}{2}$	$+\frac{1}{2}$	n_x
$-\frac{1}{2}$	a	b	E
$+\frac{1}{2}$	c	d	F
n_y	K	L	N

гдѣ $E = a + b$; $F = c + d$, $K = a + c$;

$L = b + d$ и $N = a + b + c + d$.

Въ такомъ случаѣ первый моментъ для X будетъ:

$$\nu_{1 \cdot x} = \frac{-\frac{1}{2} K + \frac{1}{2} L}{N} = \frac{-K + L}{2N} \text{ и}$$

$$\nu_{1 \cdot y} = \frac{-E + F}{2N} ;$$

такимъ же образомъ второй моментъ

$$\nu_{2 \cdot x} = \frac{\frac{1}{4} K + \frac{1}{4} L}{N} = \frac{1}{4} \text{ и}$$

$$\nu_{2 \cdot y} = \frac{1}{4} .$$

Поэтому

$$\mu_2 = \sigma_x^2 = \frac{1}{4} - \left(\frac{-K + L}{2N} \right)^2 \text{ и}$$

$$\mu_2 = \sigma_y^2 = \frac{1}{4} - \left(\frac{-E + F}{2N} \right)^2$$

Далѣе, сумма S (fxy) отъ исходной точки

$$\frac{S\,(fxy)}{N} = \left(\frac{1}{4} a + \frac{1}{4} d - \frac{1}{4} b - \frac{1}{4} c \right) \div N ,$$

съ поправкой же на среднеквадратическую

$$\frac{a + d - b - c}{4N} - \frac{(-K + L)^2}{4N^2} \times \frac{(-E + F)^2}{4N^2} =$$

$$\frac{N\,(a + d - b - c) - (-K + L)^2\,(-E + F)^2}{4N^2} ;$$

принимая во вниманіе значенія N, K, L, E, F, раскрывъ скобки и сдѣлавъ приведеніе, получимъ:

$$\frac{S(fxy)}{N} = \frac{ad - bc}{N^2}$$

и

$$r = \frac{S(fxy)}{N \sigma_x \sigma_y} = \frac{ad - bc}{\sqrt{E \times F \times K \times L}}$$

Изъ изложеннаго видно, что для вычисленія коэффиціента r необходимо лишь, чтобы рядъ X и Y былъ раздѣленъ на два равныхъ интервала каждый; но численное значеніе варіантъ, отвѣчающихъ каждому интервалу не входитъ въ вычисленія; поэтому, если даже интервалы не имѣютъ вовсе количественнаго значенія и различаются лишь качественно, вычисленіе коэффиціента корреляціи возможно, если только мы допустимъ, что качественныя различія ряда X и ряда Y находятся другъ къ другу въ такомъ же соотношеніи, какъ равные количественные интервалы. Для поясненія можетъ служить слѣдующій примѣръ.

При регистраціи домовъ въ городахъ обычно дѣлается отмѣтка о состояніи прочности; возникаетъ вопросъ, въ какомъ соотношеніи находится состояніе прочности жилого дома къ прочности холодныхъ надворныхъ построекъ. Раздѣливъ каждый родъ строеній на двѣ части—прочныя и непрочныя и допуская, что это дѣленіе отвѣчаетъ двумъ равнымъ интерваламъ варіантъ, находимъ: *

* М. Гуревичъ. Методъ оцѣнки городскихъ недвижимыхъ имуществъ, 1913, стр. 29.

Холодныя постройки.

Жилые дома.	Прочныя	Ветхія	Итого
Прочные	197	5	202
Ветхіе	3	129	132
Итого	200	134	334

$$r = \frac{197 \times 129 - 5 \times 3}{\sqrt{202 \times 132 \times 200 \times 134}};$$

или

$$\begin{aligned} \lg \text{(числителя)} &= 4{\cdot}404796 \\ \lg \text{(знаменателя)} &= 4{\cdot}427025 \\ \lg r &= \bar{1}{\cdot}977771 \end{aligned}$$

и

$$r = 0{\cdot}95.$$

Смыслъ корреляціи. Примѣненіе пріема корреляціи имѣетъ въ виду, главнымъ образомъ, слѣдующій логическій составъ изслѣдуемаго явленія. Въ совокупности случаевъ наблюдаются два количественныхъ признака. Каждый изъ нихъ имѣетъ опредѣленную истинную величину, вѣроятнымъ значеніемъ которой служитъ средняя величина этого признака для всей совокупности случаевъ; отъ этого вѣроятнаго значенія истинной величины въ отдѣльныхъ случаяхъ признакъ даетъ случайныя уклоненія, причемъ уклоненія эти случайны какъ для одного, такъ и для другого признака. Если совокупность раздѣлить на нѣсколько частей наугадъ и каждую часть расположить въ строку по равнымъ интерваламъ, съ указаніемъ частотъ каждаго интервала, то мы получимъ нѣсколько меньшихъ совокупностей, репрезентирующихъ каждая первоначальную общую совокупность; если каждая частичная совокупность состоитъ изъ достаточнаго числа случаевъ, распредѣленіе частотъ въ ней будетъ приближаться къ распредѣленію ихъ въ общей совокупности, и среднеариѳметическая каждой частичной совокупности будетъ лишь случайно уклоняться отъ общей среднеариѳметической то въ одну, то въ другую сторону; поэтому, если положеніе среднеариѳметической въ каждой строкѣ изобразить точкой на линіяхъ, паралелльныхъ оси абсциссъ и отстоящихъ другъ отъ друга на равныхъ разстояніяхъ, всѣ эти точки расположатся приблизительно на одной прямой, перпендикулярной къ оси абсциссъ и соотвѣтствующей точкѣ общей среднеариѳметической. Такое распредѣленіе свойственно какъ одному признаку (X), такъ и другому (Y), если для каждаго въ отдѣльности дѣленіе общей совокупности на части дѣлается случайно, наугадъ. Но если совокупность дѣлится на части не случайно, а по интерваламъ одного изъ признаковъ, напримѣръ, Y, тогда результатъ можетъ быть различный; если признакъ X не находится ни въ какой связи и зависимости съ признакомъ Y, то дѣленіе сохраняетъ для признака X свой случайный характеръ, распредѣленіе частотъ по строкамъ сохраняетъ свой репрезентативный характеръ, и среднеариѳметическія отдѣльныхъ строкъ, по прежнему, располагаются перпендикулярно къ оси абсциссъ по линіи общей среднеариѳметической; это послѣднее обстоятельство и служитъ указаніемъ на отсутствіе всякой корреляціи. Иначе будетъ въ томъ случаѣ, если признакъ Y не безразличенъ для X; тогда распредѣленіе частотъ по отдѣльнымъ строкамъ, отвѣчающимъ послѣдовательнымъ интерваламъ Y, перестанетъ быть однообразно—случайнымъ, и частоты каждой строки, хотя и съ случайными уклоненіями, будутъ перемѣщаться въ направленіи, опредѣляемомъ характеромъ зависимости даннаго признака отъ признака классифицирующаго; отсюда и среднеариѳметическія послѣдовательныхъ строкъ будутъ смѣщаться въ томъ же направленіи въ разной степени, давая линію корреляціи, уходящую отъ общей среднеариѳметической.

Статическая зависимость.

Такимъ образомъ въ подобныхъ случаяхъ мы имѣемъ дѣло съ совокупностью въ состояніи статическомъ: ей присущи два признака, изъ которой каждый представляетъ опредѣленную типическую устойчивую сводную величину и случайныя уклоненія отъ нея въ отдѣльныхъ наблюденіяхъ; но хотя уклоненія каждаго признака совершенно случайны, совпаденія уклоненій одного и другого признака могутъ быть неслучайны и въ такомъ случаѣ опредѣляются зависимостью признаковъ, которая, въ свою очередь, дѣйствуетъ какъ обстоятельство сводного характера, допускающее большія или меньшія уклоненія въ отдѣльныхъ случаяхъ. При такомъ строеніи совокупности отдѣльные случаи ея являются лишь случайными уклоненіями отъ нѣкоторой устойчивой постоянной величины признаковъ, находящей себѣ выраженіе въ двухъ среднихъ величинахъ.

Отъ этого случая устойчивой совокупности мы отличаемъ случаи динамическаго развитія явленій. Здѣсь количественное значеніе двухъ (или большаго числа) явленій въ каждомъ отдѣльномъ случаѣ представляетъ собою не случайное уклоненіе отъ нѣкоторой одной постоянной общей величины для всей совокупности, а приближенное неточное выраженіе разныхъ величинъ; самая величина явленія мѣняется отъ группы случаевъ къ группѣ и въ каждой группѣ проявляется съ нѣкоторымъ уклоненіемъ отъ своего истиннаго значенія подъ вліяніемъ привходящихъ и случайныхъ обстоятельствъ. Поэтому вся совокупность состоитъ здѣсь не изъ случайныхъ уклоненій отъ двухъ однѣхъ и тѣхъ же постоянныхъ величинъ, а изъ перемѣнныхъ величинъ, получающихъ различное значеніе въ рядѣ послѣдовательныхъ по мѣсту или времени случаевъ. Если одну изъ такихъ величинъ мы пріймемъ за независимую перемѣнную и расположимъ ея значенія въ какомъ-либо правильномъ порядкѣ, то они образуютъ рядъ динамическій, представляющій собою послѣдовательное систематическое измѣненіе величины явленія; съ другой стороны, величина другого явленія, выражаемаго зависимой перемѣнной, опредѣляется независимой перемѣнной и поэтому тоже не представляетъ случайныхъ уклоненій отъ какой-либо одной и той же величины, но образуетъ рядъ систематически измѣняющійся. Постояннымъ остается для всей совокупности лишь отношеніе, связывающее обѣ перемѣнныя, ихъ функціональная зависимость; случайными являются уклоненія отъ этой сводной зависимости въ отдѣльныхъ эмпирическихъ случаяхъ.

Такого рода функціональная зависимость двухъ перемѣнныхъ величинъ можетъ быть въ извѣстныхъ случаяхъ уложена и въ рамки корреляціоннаго отношенія; однако изслѣдованіе ея посредствомъ корреляціи становится неподходящимъ, когда зависимость получаетъ болѣе сложный видъ и ясно выраженный динамическій характеръ. Опредѣленіе зависимости въ этихъ случаяхъ требуетъ выраженія ея въ формѣ подходяще подысканной функціи, простой или сложной.

6. Функціональная зависимость рядовъ.

Мы допускаемъ, что между двумя или болѣе перемѣнными существуетъ опредѣленная зависимость, которая выражается той или иной формулой; въ отдѣльныхъ случаяхъ наблюденія эта зависимость затемняется случайными уклоненіями или ошибками наблюденія и можетъ быть обнаружена лишь на основаніи достаточнаго числа отдѣльныхъ случаевъ. Если связанные зависимостью факторы имѣютъ преобладающее значеніе, сравнительно съ другими привходящими условіями, зависимость этихъ факторовъ проявляется болѣе точно въ каждомъ отдѣльномъ наблюдаемомъ случаѣ, и установленіе ея на эмпирическомъ матеріалѣ требуетъ меньшаго числа наблюденій. Такъ, напримѣръ, если цѣна машины раціонально опредѣляется размѣромъ какого-либо ея признака, то характеръ связи между размѣромъ признака и величиной цѣны можетъ быть обнаруженъ путемъ сравненія немногихъ прейсъ-курантовъ, такъ какъ въ каждомъ изъ нихъ наиболѣе существенное значеніе при установленіи цѣны имѣетъ именно эта зависимость. Если же, напротивъ, связанные опредѣленной зависимостью факторы въ каждомъ отдѣльномъ случаѣ дѣйствуютъ совмѣстно съ бо́льшимъ или меньшимъ количествомъ другихъ независимыхъ обстоятельствъ, зависимость менѣе ясна или даже незамѣтна въ единичныхъ случаяхъ, и для обнаруженія ея требуется исключеніе постороннихъ случайныхъ обстоятельствъ путемъ сведенія болѣе или менѣе значительнаго числа случаевъ въ одну общую совокупность. Задача изслѣдованія сводится въ данномъ случаѣ, какъ и въ другихъ случаяхъ приложенія статистическаго метода, къ тому, чтобы для всей совокупности найти такое сводное выраженіе зависимости, которое, возможно лучше отвѣчая эмпирическому матеріалу, являлось бы наиболѣе вѣроятнымъ выраженіемъ истинной формы зависимости.

Функція 1-й степени. Наиболѣе простой случай зависимости выражается графически прямой линіей и аналитически простѣйшимъ уравненіемъ первой степени. Если значенія независимой перемѣнной (x) мы изобразимъ по опредѣленному масштабу въ видѣ точекъ на абсциссѣ, а значеніе зависимой перемѣнной пропорціональной длины перпендикулярами изъ этихъ точекъ, то при прямолинейной зависимости вершины перпендикуляровъ будутъ лежать на одной прямой. Чтобы выразить зависимость аналистически, достаточно взять разстояніе между двумя любыми точками на абсциссѣ и разницу соотвѣтствующихъ этимъ точкамъ ординатъ; дѣленіемъ послѣдней на первую мы находимъ, какая величина измѣненія (увеличенія или уменьшенія) зависимой перемѣнной приходится на единицу измѣненія независимой; если прямая проходитъ не въ нулевой точкѣ абсциссы, а начинается на извѣстной высотѣ отъ нея, или ниже нулевой точки, то кромѣ пропорціонально измѣняющейся части зависимая перемѣнная содержитъ еще постоянную часть, положительную или отрицательную, и общее выраженіе зависимости будетъ имѣть видъ

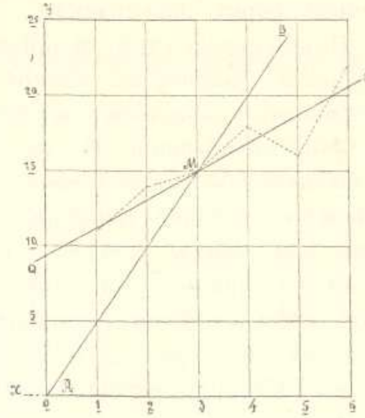
$$y = a + bx$$

На приложенномъ чертежѣ для прямой AB при $x = o$, $y = o$ и при $x = 3$, $y = 15$, т. е. увеличенію x на 3 единицы отвѣчаетъ увеличеніе y на 15 единицъ или

$$y = 5x \;;$$

для прямой QQ при $x = o$, $y = 9.6$ при $x = 3$, $y = 15$ т. е. на единицу x приходится $\frac{15 - 9{\cdot}6}{3} = 1{\cdot}8$, и y начинается съ $9{\cdot}6$ или

$$y = 9{\cdot}6 + 1{\cdot}8\,x$$



Нахожденіе постоянныхъ. Если однако точки, изображающія значенія зависимой перемѣнной лежатъ, не на одной прямой, но образуютъ ломаную линію, то задача состоитъ въ томъ, чтобы ломаную линію замѣнить прямой, возможно ближе проходящей къ точкамъ ломаной, съ условіемъ, чтобы сумма эмпирически данныхъ ординатъ совпадала съ суммой замѣняющихъ ихъ ординатъ прямой, иначе говоря, чтобы площадь, ограниченная крайними ординатами, осью абсциссъ и эмпирически данной ломаной совпадала съ площадью, ограниченной тѣми же линіями и вычисленной прямой. Пусть даны значенія зависимой перемѣнной, показанныя на чертежѣ:

X	Y
1	11
2	14
3	15
4	18
5	16
6	22

Зависимость должна выражаться формулой

$$y = a + bx\,;$$

для отысканія значеній a и b мы можемъ исходить изъ слѣдующихъ простыхъ соображеній: сумма всѣхъ y (Sy) заключаетъ въ себѣ 6 разъ a и bSx; если весь рядъ раздѣлимъ на двѣ равныя по числу членовъ части, то въ суммѣ значеній y каждой части будетъ заключаться:

въ первой половинѣ: $40 = 3a + 6b$

и во второй $56 = 3a + 15b;$

вычтя меньшую сумму изъ большей, получимъ, что разница на 9b даетъ 16 единицъ разницы въ y-ахъ, или

$$b = 1{\cdot}78$$

и

$$a = 9{\cdot}77;$$

уравненіе же кривой выражается формулой

$$y = 9{\cdot}77 + 1{\cdot}78\,x.$$

Въ общемъ видѣ a и b получаютъ слѣдующія значенія

$$a = \frac{Sy}{n} - \frac{Sx}{n}\,b$$

$$b = \frac{S_2y - S_1y}{S_2x - S_1x} = \frac{ym_{,2} - ym_{,1}}{xm_{,2} - xm_{,1}},$$

т. е. b выражаетъ отношеніе разницы между среднеариѳметическими значеніями y въ каждой половинѣ къ такой же разницѣ между значеніями x въ каждой половинѣ.

Способъ наименьшихъ квадратовъ. Другой способъ опредѣленія значеній a и b состоитъ въ слѣдующемъ. Пусть величина y опредѣляется функціей

$$y = a + bx$$

и уклоненія вычисленныхъ значеній y' отъ эмпирическихъ составляютъ

$$\begin{aligned} -y_1 + a + bx_1 &= z_1 \\ -y_2 + a + bx_2 &= z_2 \\ \cdots\cdots & \\ -y_n + a + bx_n &= z_n\,. \end{aligned}$$

Для того, чтобы опредѣлить величины a и b, поставимъ слѣдующія условія: суммы вычисленныхъ и эмпирическихъ значеній должны совпадать, т. с. сумма уклоненій Sz должна равняться 0, и сумма квадратовъ уклоненій, т. е. $S\,z^2$ должна быть наименьшей. Для рѣшенія задачи возьмемъ $S\,z^2$:

$$(-y_1+a+bx_1)^2+(-y_2+a+bx_2)^2+\ldots = y_1^2+a^2+b^2x_1^2 - 2\,ay_1 - 2by_1x_1 + 2abx_1 + y_2^2 + a^2 + b^2x_2^2 - 2ay_2 - 2by_2x_2 + 2abx_2 + \ldots$$

Для того, чтобы найти величины a и b, которыя обращаютъ это выраженіе къ минимумъ, необходимо продифференцировать его по a и затѣмъ по b и результаты приравнять 0. Дифференцируя выраженіе типа

$$y_1^2 + a^2 + b^2 x_1^2 - 2ay_1 - 2by_1x_1 + 2abx_1$$

по a получимъ

$$2a - 2y_1 + 2bx_1 ;$$

суммируя для всѣхъ y и x, приравнявъ результатъ 0 и сокративъ на 2, получимъ:

$$na - S(y) + b\,S(x) = 0 \;.\;.\;.\;.\; (I).$$

Дифференцируя по b, получимъ

$$2x_1^2 - 2y_1x_1 + 2ax_1 \;,$$

и послѣ надлежащихъ дѣйствій:

$$S(x^2) - S(xy) + a\,S(x) = 0 \;.\;.\;.\; (II).$$

Эти два уравненія даютъ возможность найти значенія двухъ неизвѣстныхъ, а именно:

$$a = \frac{S(y)}{n} - \frac{S(x)}{n}\,b$$

$$b = \frac{S(xy) - \dfrac{S(y)}{n}\,S(x)}{S(x^2) - \dfrac{S(x)}{n}\,S(x)}$$

или обозначая среднеариѳметическія

$$\frac{S(y)}{n} = y_m, \quad \frac{S(x)}{n} = x_m$$

получимъ:

$$a = y_m - x_m\,b$$

$$b = \frac{S(xy) - y_m\,S(x)}{S(x^2) - x_m\,S(x)} \,.$$

Послѣднее выраженіе можетъ быть представлено также въ видѣ

$$b = \frac{S((y - y_m)x)}{S((x - x_m)x)} .$$

Примѣръ. На численномъ примѣрѣ вычисленія имѣютъ слѣдующій видъ:

y	x	yx	x^2
11	1	11	1
14	2	28	4
15	3	45	9
18	4	72	16
16	5	80	25
22	6	132	36
SS: 96	21	368	91

$$\text{I:} \; -96 + 6a + 21b = 0$$
$$\text{II:} \; -368 + 21a + 91b = 0$$
$$a = \frac{96}{6} - \frac{21}{6} b = 16 - \frac{7}{2} b$$
$$-368 + 336 - 73^1/_2 \, b + 91 b = 0$$
$$b = 1{\cdot}8286$$
$$a = 16 - 6{\cdot}4 = 9{\cdot}6$$

и $y = 9{\cdot}6 + 1{\cdot}83\,x$.

Прямая, отвѣчающая этому выраженію, показана на чертежѣ (QQ). Такъ какъ здѣсь для постоянныхъ a и b подыскиваются величины, сообщающія наименьшее значеніе суммѣ квадратовъ уклоненій, то такой способъ называется способомъ наименьшихъ квадратовъ.

Когда при разныхъ значеніяхъ X соотвѣтствующія значенія y встрѣчаются не по одному разу или неравное число разъ, а съ различной частотою (f), и по условіямъ задачи мы считаемъ правильнымъ принять во вниманіе вѣсъ, то приведенныя въ предшествующемъ изложеніи формулы получаютъ слѣдующій видъ:

$$-Sfy + aSf + bSfx = 0 \quad . \; . \; . \quad \text{I}$$
$$-Sfxy + aSfx + bSfx^2 = 0 \quad . \; . \; . \quad \text{II}$$

Ходъ вычисленій показанъ на слѣдующемъ примѣрѣ: x — урожай-самъ ржи, y — урожай-самъ овса, f — число зарегистрированныхъ случаевъ.

f	x	y	fy	fx	fxy	fxx
325	4	2·9	942·5	1300	3802·5	5200
478	7	3·6	1720·8	3346	12045·6	23422
167	10	3·8	634·6	1670	6346·0	16700
33	13	4·1	135 3	429	1758·9	5577
1003			3433·3	6745	23953·0	50899
Sf			Sfy	Sfx	Sfxy	Sfx^2

$$y = 1{\cdot}62 + 0{\cdot}27\,x.$$

Та же зависимость, выведенная безъ принятія во вниманіе вѣса, выражается въ видѣ

$$y = 2{\cdot}52 + 0{\cdot}13\,x.$$

Когда перемѣнную y необходимо поставить въ зависимость отъ двухъ независимыхъ перемѣнныхъ x и z, по формулѣ

$$y = a + bx + cz\,,$$

постоянныя опредѣляются по способу наименьшихъ квадратовъ изъ слѣдующихъ уравненій:

$$- Sy + na + bSx + cSz = 0 \quad . \; . \; . \; \mathrm{I}$$
$$- Sxy + aSx + bSx^2 + cSxz = 0 \quad . \; . \; . \; \mathrm{II}$$
$$- Szy + aSz + bSxz + cSz^2 = 0 \quad . \; . \; . \; \mathrm{III}$$

Переходъ къ кривымъ.

Нерѣдко однако эмпирическія цыфры располагаются такимъ образомъ, что выравниваніе ихъ при помощи прямой явно не отвѣчаетъ природѣ явленія; внѣшнимъ признакомъ такого несоотвѣтствія служитъ систематическое направленіе уклоненій; такъ, если въ началѣ и концѣ ряда всѣ уклоненія эмпирически данныхъ величинъ отъ вычисленныхъ ихъ значеній имѣютъ отрицательный знакъ, то очевидно, что расположенію эмпирическихъ цыфръ болѣе отвѣчаетъ выпуклая кривая. Въ подобныхъ случаяхъ возникаетъ необходимость перехода къ болѣе сложнымъ формуламъ. Изъ простѣйшей формулы прямой линіи могутъ получаться формулы кривыхъ посредствомъ повышенія или пониженія степени независимой перемѣнной x; на прилагаемомъ чертежѣ показана прямая формула которой

$$y = 50 + 3\,x$$

для значеній $x = 0,\ 1,\ 2,\ 3,\ 4,\ 5$ и рядъ кривыхъ, которыя получаются, если вмѣсто x^1 брать послѣдовательно

$$x^{3/2} = \sqrt{x^3}\,,\ x^2,\ x^3$$

$$x^{1/2} = \sqrt{x}\,,\ x^{1/3}\,.$$

Каждая кривая даетъ соотвѣтственныя значенія для у, которыя приведены ниже; для того, чтобы при разныхъ показателяхъ для х, не измѣнялось положеніе начальной точки у, равной 50, какъ равнымъ образомъ не измѣнялась и сумма всѣхъ у, равная 345, необходимо соотвѣтствующимъ образомъ измѣнять значеніе b; измѣненіе опредѣляется тѣмъ условіемъ, что сумма всѣхъ приращеній у-а, сверхъ начальной точки, равной 50, должна оставаться неизмѣнной. Въ первоначальномъ рядѣ (I) сумма приращеній составляетъ

$$b\,(x_1+x_2+x_3+x_4+x_5) = 3\,(1+2+3+4+5) = 45;$$

если вмѣсто х взять x^2, то и здѣсь сумма приращеній должна дать ту же величину, т. е.

$$b_1\,(x_1{}^2+x_2{}^2+x_3{}^2+x_4{}^2+x_5{}^2) = b\,(1+4+9+16+25) = 45$$

или

$$b_1 = 0{\cdot}82\;;$$

для $x^{3/2} = \sqrt{x^3}$:

$$b_1\,(1+2{\cdot}828+5{\cdot}196+8+11.180) = 45$$

$$b_1 = 1{\cdot}59.$$

Если не измѣнять коэффиціента b, то необходимо измѣнить величину a; не измѣняя значеній a и b, будемъ получать различную сумму у, т. е. различныя площади для разныхъ кривыхъ.

Формулы кривыхъ получаютъ слѣдующій видъ:

I $\quad y = 50 + 3x$

II $\quad y = 50 + 1{\cdot}59x^{3/2}$

III $\quad y = 50 + 0{\cdot}82x^2$

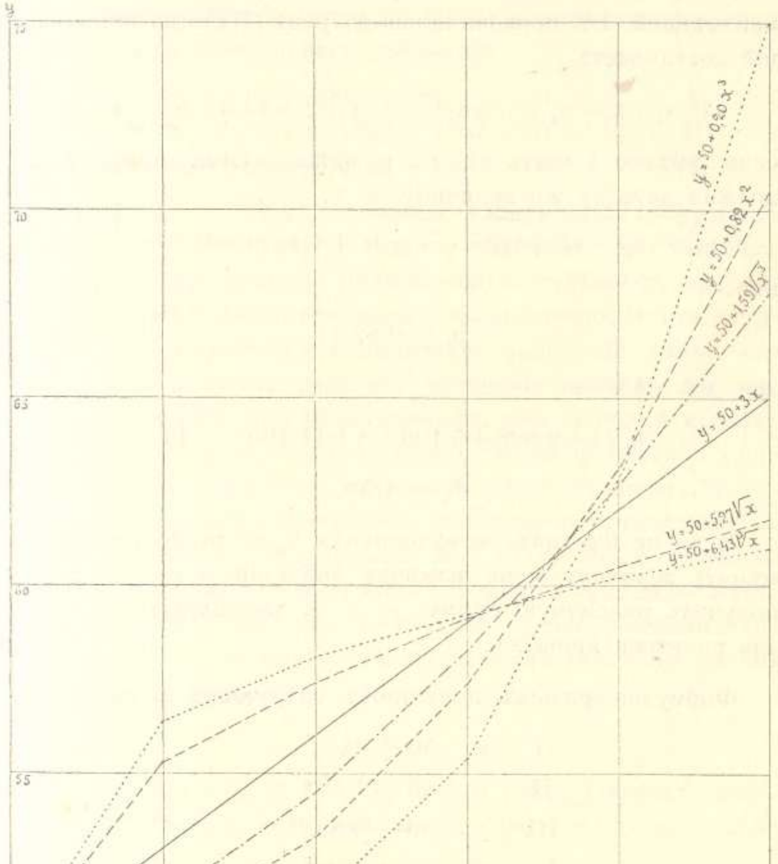
IV $\quad y = 50 + 0{\cdot}2x^3$

V $\quad y = 50 + 5{\cdot}27x^{1/2}$

VI $\quad y = 50 + 6{\cdot}43x^{1/3}$

Численныя значенія у даютъ слѣдующіе ряды

	I	II	III	IV	V	VI
x	у при $(x^{3/2})$		(x^2)	(x^3)	$(x^{1/2})$	$(x^{1/3})$
0	50	50	50	50	50	50
1	53	51·6	50·8	50·2	55·3	56·4
2	56	54·5	53·3	51·6	57·4	58·1
3	59	58·3	57·4	55·4	59·1	59·3
4	62	62·7	63·1	62·8	60·5	60·2
5	65	67·8	70·5	75·0	61·8	61·0



Выборъ вида кривой. Выборъ подходящей кривой можетъ опредѣляться прежде всего знаніемъ природы явленія, независимо отъ подлежащаго изслѣдованію статистическаго матеріала. Такъ, напримѣръ, мы знаемъ, что производительность почвы, начиная съ извѣстнаго предѣла, увеличивается медленнѣе затратъ на производство и, достигши извѣстнаго размѣра, остается почти неизмѣнной (до какого либо существеннаго усовершенствованія); такого рода зависимость можетъ быть выражена логариѳмической кривой

$$y = a + \lg x\ ,$$

гдѣ, съ увеличеніемъ x въ геометрической прогрессіи, y возрастаетъ въ ариѳметической. Или при опредѣленіи пониженія цѣнности строеній отъ времени мы можемъ исходить изъ апріорнаго предположенія, что процессъ обветшанія и изнашиванія аналогиченъ учету по сложнымъ процентамъ и можетъ быть выраженъ формулой

$$W_0 \left(1 - \frac{p}{100}\right)^t$$

гдѣ независимая перемѣнная t означаетъ время, W_0 первоначальную цѣнность и p процентъ скидки. При вычисленіи населенія между моментами переписи мы исходимъ изъ гипотезы прироста въ геометрической прогрессіи, и послѣдній выражается формулой

$$y = \sqrt[t]{\frac{N_t}{N_1}} - 1.$$

гдѣ N_t и N_1 означаютъ число населенія въ моменты переписей, раздѣлены числомъ t лѣтъ.

Въ тѣхъ случаяхъ, когда такого рода выборъ формы функціональной зависимости невозможенъ, за отсутствіемъ предварительнаго знакомства съ явленіемъ, выборъ приходится дѣлать на основаніи ознакомленія съ характеромъ эмпирическаго статистическаго матеріала; для этой цѣли можно изобразить величины при помощи системы координатъ, и расположеніе точекъ даетъ указаніе на характеръ подходящей кривой.

Кривыя статическаго типа.

Кривыя можно раздѣлить прежде всего на кривыя статическихъ явленій и кривыя динамическихъ явленій; къ первымъ относятся нормальная кривая Гаусса и кривыя Пирсона; характерной особенностью ихъ является то, что эти кривыя изображаютъ явленія, величина которыхъ случайно колеблется около нѣкотораго средняго устойчиваго значенія, съ равной или различной вѣроятностью въ обѣ стороны, и притомъ число случаевъ каждаго уклоненія находится въ зависимости отъ величины этого уклоненія; обычно оно тѣмъ больше, чѣмъ уклоненіе меньше. Поэтому кривыя этого рода представляются сначала возрастающими, а затѣмъ убывающими и весьма часто представляютъ площадь, замкнутую (или почти замкнутую) осью абсциссъ. Выборъ типа кривой рѣшается здѣсь приблизительно внѣшнимъ видомъ эмпирическихъ данныхъ—симетрическимъ или асиметрическимъ—и точно посредствомъ вычисленія критерія кривой. Формула нормальной кривой

$$y = \frac{N}{\sqrt{2\pi\mu}} e^{-\frac{x^2}{2\mu_2}};$$

несиметрическія кривыя приблизительно выражаются формулой бинома

$$(p+q)^c$$

въ которомъ величины p, q, и c опредѣляются на основаніи эмпирическаго матеріала по слѣдующимъ соотношеніямъ:

$\frac{p-q}{p+q} = \delta = {}^{\mu_3}/_{\mu_2}$ $c = \frac{4\mu_2}{1-\delta^2}$; точка среднеариѳметической на абсциссѣ $M = \frac{c}{2}(1-\delta)+1$, среднеквадратическое уклоненіе составляетъ $\frac{cpq}{(p+q)^2} = \frac{c}{4}(1-\delta^2)$.

Болѣе точно несимметрическія кривыя выражаются формулами Пирсона.

Кривыя динамическаго типа. Отъ статистическихъ кривыхъ отличны кривыя динамическихъ явленій; эти кривыя служатъ для изображенія тѣхъ случаевъ, когда постояннымъ остается отношеніе между двумя (или болѣе) явленіями, самыя же величины явленій не имѣютъ постояннаго типа и могутъ неопредѣленно возрастать или убывать, начинаясь и кончаясь въ любыхъ точкахъ по отношенію къ абсциссѣ. Съ точки зрѣнія статистическаго использованія эти кривыя, въ свою очередь, важно раздѣлить на кривыя, не имѣющія максимальнаго или минимальнаго значенія и кривыя, получающія такое значеніе.

Типическимъ примѣромъ линій перваго рода является прямая линія, идущая въ одномъ направленіи на неопредѣленное разстояніе и ни при какомъ значеніи x

$$y = a + bx$$

не дающая ни максимума ни минимума. Такимъ же характеромъ обладаетъ каивая

$$y = a + bx^3$$

напримѣръ

$$y = 10 + 2x^3,$$

которая при $x = 0$ получаетъ значеніе $y = 10$, прн $x > 0$ даетъ возрастающія величины, уходящія вверхъ отъ абсциссы, при $x = -1$ даетъ $y = 8$ и при $x = -2$ даетъ $y = -6$ и далѣе идетъ внизъ; въ точкѣ

$$-10 = 2x^3 \text{ или } x = -1{\cdot}71$$

она пересѣкаетъ абсциссу, давая $y = 0$.

Напротивъ, кривая

$$y = 10 + 2x^2$$

принимаетъ минимальное значеніе при $x = 0$; дѣйствительно,

при $x = 0$ $y = 10$
при $x = +1$ $y = 12$
при $x = -1$ $y = 12$ и т. д.

Минимумъ и максимумъ. Для опредѣленія, имѣетъ ли кривая максимумъ или минимумъ и при какомъ именно значеніи независимаго перемѣннаго, мы можемъ исходить изъ слѣдующаго разсужденія. Если съ увеличеніемъ x на какую либо величину y тоже увеличивается, то отношеніе приращенія y или dy къ приращенію x или dx даетъ положительную величину т. е.

$$\frac{dy}{dx} > 0 \; ;$$

хотя бы эта величина была и незначительна, но все же y возрастаетъ; при убываніи y съ возрастаніемъ x, это отношеніе должно быть меньше нуля; слѣдовательно въ поворотномъ пунктѣ, гдѣ возрастаніе переходитъ въ убываніе или наоборотъ это отношеніе должно равняться 0 или

$$\frac{dy}{dx} = 0 \; ;$$

при этомъ однако надо принять во вниманіе, что поворотный пунктъ представляетъ собою не протяженіе кривой, а точку, и потому приращеніе x-а, которое соотвѣтствовало бы не нѣкоторому протяженію кривой, а одной ея точкѣ, должно быть менѣе всякой данной величины и должно имѣть своимъ предѣломъ 0.

Сообразно съ этимъ дадимъ независимому перемѣнному x приращеніе h; такъ, если функція имѣетъ видъ

$$y = a + bx^2, \quad . \; . \; . \; . \; . \; . \; . \; . \quad (I)$$

замѣнивъ x черезъ x+h получимъ

$$a + b\,(x+h)^2 \; ; \quad . \; . \; . \; . \; . \; . \; . \quad (II)$$

чтобы узнать приращеніе функціи y, найдемъ разницу между II и I выраженіемъ:

$$dy = a + b\,(x+h)^2 - a - bx^2 = b\,[\,(x+h)^2 - x^2] = b\,(h^2 + 2xh);$$

возьмемъ отношеніе приращенія dy къ приращенію dx или h:

$$\frac{dy}{dx} = \frac{b(h^2 + 2xh)}{h} = bh + 2bx \ ;$$

принимая во вниманіе, что h по условію должно быть какъ угодно мало, для предѣльнаго его значеніе ($= 0$) получимъ

$$\frac{dy}{dx} = 2bx \ .$$

Такого рода отношеніе называется первой производной функціи y вычисленной по (приращенію) x и обозначается черезъ y' или $f'(x)$; если повторить ту же операцію съ первой производной, получимъ вторую производную y'' или $f''(x)$. Въ простѣйшемъ случаѣ этотъ пріемъ дифференцированія сводится къ слѣдующему правилу: члены, не заключающіе x, исчезаютъ, степени x переходятъ въ коэффиціентъ и показатель при x понижается на 1.

Теперь остается первую производную приравнять 0 или

$$\frac{dy}{dx} = y' = 2bx = 0 \ ;$$

величина x, удовлетворяющая этому равенству, 0; такимъ образомъ искомая точка кривой получится при $x = 0$; для того, чтобы опредѣлить, будетъ ли эта точка максимумъ или минимумъ, надо узнать, даетъ ли кривая далѣе пониженіе (тогда точка будетъ максимумъ) или повышеніе (тогда она будетъ минимумъ); при пониженіи отношеніе слѣдующаго dy къ dx должно быть

$$\frac{dy}{dx} < 0$$

при повышеніи обратно

$$\frac{dy}{dx} > 0 \ ;$$

поэтому надо взять вторую производную, и если она даетъ величину отрицательную (полагая x равнымъ значенію, полученному изъ первой производной), то найденная точка будетъ максимальнымъ значеніемъ функціи, если положительную — минимальнымъ; въ данномъ примѣрѣ

$$y'' = 2b ,$$

т. е. при положительномъ b больше 0, слѣдовательно мы имѣемъ дѣло съ минимумомъ, какъ это и видно изъ численнаго примѣра.

Если вторая производная, послѣ подстановки въ нее для x величины, найденной изъ первой производной, также обращается въ 0, то рѣшаетъ вопросъ ближайшая производная четнаго порядка, не обращающаяся въ 0; если же такой необращающейся въ 0 будетъ производная нечетнаго порядка, функція не имѣетъ максимума и минимума.

Для функціи

$$y = (x - a)^4 + b$$

$f'(x) = 4(x - a)^3 = 0$ и $x = a$

$f''(x) = 12(x - a)^2$ также обращается въ 0 при $x = a$; тоже $f'''(x) = 24(x - a)$; и наконецъ $f''''(x) = 24$, слѣдовательно y обращается въ минимумъ при $x = a$ и тогда

$$y = b.$$

Если формула кривой присущими ей свойствами осмысленно отвѣчаетъ характеру изображаемаго явленія, это служитъ однимъ изъ указаній на удачный ея выборъ. Но весьма часто представляется невозможнымъ установить такое соотношеніе характера формулы съ характеромъ явленія, и вниманіе обращается лишь на совпаденіе кривой съ эмпирическими данными хотя бы въ той части кривой, которая отвѣчаетъ протяженію эмпирическаго матеріала. Въ выборѣ подходящей кривой въ этихъ случаяхъ рѣшающее значеніе имѣетъ то соображеніе, чтобы на протяженіи эмпирической абсциссы точки кривой лежали возможно ближе къ точкамъ эмпирической ломаной линіи, и чтобы вычисленіе постоянныхъ величинъ представляло возможно меньше затрудненій. Въ этомъ отношеніи въ преобладающемъ числѣ случаевъ вполнѣ удовлетворительные результаты даетъ даже простѣйшая формула прямой. Если она не укладывается хорошо на эмпирическій матеріалъ, то во многихъ случаяхъ бываетъ достаточно, вмѣсто того, чтобы переходить къ формуламъ болѣе сложнымъ, раздѣлить эмпирическій матеріалъ на подходящее число отдѣльныхъ участковъ и для каждаго участка вычислить свою особую прямую; неудобство этого способа даетъ себя чувствовать лишь на границѣ смежныхъ участковъ, гдѣ приходится переходить отъ одной формулы къ другой, отчего нерѣдко получаются рѣзкіе скачки.

Нахожденіе постоянныхъ. Когда прямая не удовлетворяетъ условіямъ эмпирическаго матеріала, и нѣтъ основанія прибѣгать къ формуламъ какого-либо особаго вида, пользуются общей формулой

$$y = a + bx + cx^2 + dx^3 + \ldots$$

беря въ ней столько членовъ, сколько требуется для того, чтобы получить возможно лучшее совпаденіе. Такъ какъ нахожденіе коэффиціентовъ для высшихъ степеней x сильно усложняетъ вычисленія и сравнительно мало измѣняетъ результатъ, то безъ крайней необходимости нѣтъ основанія брать лишнее число членовъ.

По способу наименьшихъ квадратовъ. Для нахожденія постоянныхъ мы прежде всего можемъ воспользоваться способомъ наименьшихъ квадратовъ. Если принята формула

$$y' = a + bx + cx^2 \ ,$$

то вычисленныя на основаніи ея значенія y' будутъ давать уклоненія отъ эмпирическихъ y, а именно:

$$-y_1 + a + bx_1 + cx_1^2 = z_1 \ .$$

$$-y_2 + a + bx_2 + cx_2^2 = z_2 \text{ и т. д.}$$

Пусть Sz^2 должна имѣть минимальное значеніе, или

$$\underset{i=1}{\overset{i=n}{S}} \left(-y + a + bx_i + cx_i^2 \right) \text{ д. б. } \min \text{ ;}$$

для опредѣленія, какое значеніе a удовлетворяетъ этому условію, найдемъ первую производную указаннаго выраженія по a и приравняемъ ее 0.

Квадратъ каждаго уклоненія имѣетъ видъ:

$$y^2 + a^2 + b^2x^2 + c^2x^4 - 2ya - 2ybx - 2ycx^2 + 2abx + \\ + 2acx^2 + 2bcx^3 \text{ ;}$$

первая производная по a даетъ:

$$2a - 2y + 2bx + 2cx^2 \text{ ,}$$

суммированіе для всѣхъ x и сокращеніе на 2 даетъ:

$$- Sy + na + bSx + cSx^2 = 0 \quad . \quad . \quad . \quad . \quad \text{I.}$$

Первая производная по b для каждаго уклоненія выразится:

$$2bx^2 - 2yx + 2ax + 2cx^3$$

и послѣ суммированія:

$$- Syx + aSx + bSx^2 + cSx^3 = 0 \quad . \quad . \quad . \quad \text{II.}$$

Наконецъ, дифференцируя по c, получимъ третье уравненіе:

$$- Syx^2 + aSx^2 + bSx^3 + cSx^4 = 0 \quad . \quad . \quad . \quad \text{III.}$$

Три уравненія I, II, III даютъ возможность опредѣлить значенія a, b, c. Для нахожденія ихъ необходимо слѣдовательно сдѣлать слѣдующія вычисленія на основаніи эмпирическаго матеріала для каждаго даннаго значенія x и y:

$$y, \ x, \ x^2, \ x^3, \ x^4, \ yx, \ yx^2$$

и затѣмъ суммировать каждую графу для всѣхъ значеній.

По способу моментовъ.

Другой способъ нахожденія подходящихъ значеній постоянныхъ въ формулѣ

$$y = a + bx + cx^2 + \ . \ . \ .$$

способъ моментовъ. Этотъ способъ основанъ на допущеніи, что нулевой, первый, второй и т. д. моменты для вычисленныхъ значеній y должны быть равны тѣмъ же моментамъ для эмпирическаго ряда. При такомъ допущеніи мы можемъ образовать нужное число уравненій для отысканія любого числа постоянныхъ. Ограничиваясь тремя членами, мы имѣемъ для y выраженіе въ видѣ

$$y = a + bx + cx^2;$$

пусть протяженіе абсциссы равняется 2 l, и возьмемъ исходной точкой ея середину; въ одну сторону отъ исходной точки x-ы будутъ имѣть отрицательное значеніе и крайній $x = -l$; въ другую сторону — положительное, и крайній $x = +l$; въ такомъ случаѣ для нѣкотораго $x = -i$ соотвѣтствующій y будетъ имѣть видъ

$$a - bi + ci^2$$

и для $x = +i$:

$$a + bi + ci^2;$$

n-ый моментъ, если n четное даетъ:

$$ai^n - bi^{n+1} + ci^{n+2} \text{ (для } x = -i)$$

и

$$ai^n + bi^{n+1} + ci^{n+2} \text{ (для } x = +i)$$

При суммированіи строкъ для всѣхъ значеній x отъ $-l$ до $+l$, члены, содержащіе b, сократятся, и n-ый моментъ получитъ видъ:

$$Syx^n = a\,2Sx^n + c2Sx^{n+2};$$

для нечетнаго момента n+1 будемъ имѣть

$$\text{для } x = -i: \quad -ai^{n+1} + bi^{n+2} - ci^{n+3}$$

$$\text{и для } x = -i: \quad ai^{n+1} + bi^{n+2} + ci^{n+3};$$

при суммированіи исчезнутъ члены, содержащіе a и c и (n+1)-ый моментъ получитъ видъ

$$Syx^{n+1} = b2Sx^{n+2} .$$

Суммы типа

$$1^2+2^2+3^2+\ldots+l^2 = S_2$$
$$1^3+2^3+3^3+\ldots+l^3 = S_3$$
$$1^4+2^4+3^4+\ldots+l^4 = S_4$$
$$\ldots\ldots\ldots\ldots$$

обладаютъ тѣмъ свойствомъ, что при дѣленіи соотвѣтственно на l^3, l^4, l^5 . . . даютъ въ предѣлѣ, при $l = \infty$,

$$\frac{S_1}{l^2} = \frac{1}{2}\,;\ \frac{S_2}{l^3} = \frac{1}{3}\,,\ \frac{S_3}{l^4} - \frac{1}{4} \text{ и т. д.}$$

и вообще

$$\frac{S_n}{l^{n+1}} = \frac{1}{n+1}\,.$$

Поэтому $Sx^2 = 1^2+2^2+3^2+\ldots+l^2$ можетъ быть представлена въ видѣ

$$\frac{l^3\,Sx^2}{l^3} = \frac{l^3}{3}\,,$$

равнымъ образомъ

$$S\,x^n = 1^n + 2^n + \ldots + l^n$$

можетъ быть представлена въ видѣ

$$\frac{l^{n+1}\,S\,x^n}{l^{n+1}} = \frac{l^{n+1}}{n+1}\,.$$

Пользуясь этимъ свойствомъ и полагая, что интервалы какъ угодно малы, можно перейти къ предѣльнымъ значеніямъ указанныхъ выше выраженій и получить въ общемъ видѣ слѣдующій рядъ уравненій:

$$\frac{S\,y}{2l} = a + \frac{cl^2}{3} + \frac{el^4}{5} + \dots$$

$$\frac{1}{2l} \times \frac{S\,yx}{l} = \frac{bl}{3} + \frac{dl^3}{5} + \frac{fl^5}{7}$$

$$\frac{1}{2l} \times \frac{S\,yx^2}{l^2} = \frac{a}{3} + \frac{cl^2}{5} + \frac{el^4}{7}$$

$$\frac{1}{2l} \times \frac{S\,yx^3}{l^3} = \frac{bl}{5} + \frac{dl^3}{7} + \frac{fl^7}{9}$$

$$\frac{1}{2l} \times \frac{S\,yx^4}{l^4} = \frac{a}{5} + \frac{cl^2}{7} + \frac{el^9}{9}$$

Если надо опредѣлить постоянныя a, b, c, то полагая $d = e = f = \ldots = 0$, получимъ три уравненія, рѣшеніе которыхъ даетъ:

$$a = \frac{3}{4}\left(\frac{3}{2l} \cdot S\,y - \frac{5}{2l} \cdot \frac{S\,yx^2}{l^2}\right)$$

$$b = \frac{3}{l} \cdot \frac{1}{2l} \cdot \frac{S\,yx}{l}$$

$$c = \frac{15}{4l^2}\left(-\frac{1}{2l} \cdot S\,y + \frac{3}{2l} \cdot \frac{S\,yx^2}{l^2}\right)$$

Для двухъ постоянныхъ формулы

$$y = a + bx$$

получимъ

$$a = \frac{S\,y}{2l}$$

$$b = \frac{3}{l \times 2l} \cdot \frac{S\,xy}{l}.$$

Примѣры.

Прилагая этотъ способъ къ примѣру на стр. 140 получимъ слѣдующій рядъ вычисленій

y	x	yx
11	—2·5	—27·5
14	—1·5	—21·0
15	—0·5	— 7·5
18	0·5	9·0
16	1·5	24·0
22	2·5	55·0
96		+88
		—56
		32

$$S\,y = 96$$

$$S\,xy = 32$$

$$2l = 6;\ l = 3$$

$$a = \frac{96}{6} = 16$$

$$b = \frac{3 \times 32}{3 \times 6 \times 3} = 1{\cdot}78$$

уравненіе $y = 16 + 1{\cdot}78\,x$

для соотвѣтствующихъ x даетъ слѣдующій рядъ

эмпир.	вычисл.
11	11·6
14	13.3
15	15·1
18	16 9
16	18·7
22	20·4
96	96 0

Вычислять значеніе y въ данномъ примѣрѣ при помощи болѣе сложной формулы

$$y = a + bx + cx^2$$

нѣтъ расчета, такъ какъ величина $c = -0{\cdot}14$ оказываетъ ничтожное вліяніе на y. Для сравненія пріемовъ вычисленія по способу наименьшихъ квадратовъ и по способу моментовъ возьмемъ слѣдующій примѣръ:

По способу наименьшихъ квадратовъ:

y	x	yx	yx^2	x^2	x^3	x^4
3	1	3	3	1	1	1
7	2	14	28	4	8	16
9	3	27	81	9	27	81
10	4	40	160	16	64	256
11	5	55	275	25	125	625
40	15	139	547	55	225	979
Sy	Sx	Syx	Syx^2	Sx^2	Sx^3	Sx^4

$$\text{I} : -40 + 5a + 15b + 55c = 0$$
$$\text{II} \quad -139 + 15a + 55b + 225c = 0$$
$$\text{III} \quad -547 + 55a + 225b + 979c = 0$$

$$(\text{I}) \qquad a = 8 - 3b - 11c$$
$$(\text{II}) \; -139 + 15(8 - 3b - 11c) + 55b + 225c = 0$$
$$\text{или } -19 + 10b + 60c = 0$$
$$b = 1{\cdot}9 - 6c$$
$$a = 2{\cdot}3 - 7c$$
$$(\text{III}) \; -547 + 55(2{\cdot}3 + 7c) + 225(1{\cdot}9 - 6c) + 979c = 0$$
$$\text{или} \qquad 7 + 14c = 0$$
$$c = -\tfrac{1}{2}$$
$$b = \quad 4{\cdot}9$$
$$a = -1{\cdot}2$$
$$y = -1{\cdot}2 + 4{\cdot}9x - \tfrac{1}{2}x^2 .$$

Отсюда вычисленныя (y') значенія y:

y	y'
3	3·2
7	6·6
9	9·0
10	10·4
11	10·8
	40·0

Тоже по способу моментовъ:

y	x	yx	yx^2
3	—2	—6	12
7	—1	—7	7
9	0		
10	1	10	10
11	2	22	44
40		+32	73
		—13	
		19	

$S\,y = 40$

$S\,yx = 19$

$S\,yx^2 = 73$

$2l = 5$

$l = 5/2$

$$a = \frac{3}{4}\left(\frac{3}{5} \times 40 - \frac{5}{5} \times \frac{73 \times 4}{25}\right)$$
$$= 18 - 8{\cdot}76 = 9{\cdot}24$$

$$b = \frac{3 \times 2 \times 19 \times 2}{5 \times 5 \times 5} = 1{\cdot}824$$

$$c = \frac{15 \times 4}{4 \times 25}\left(-\frac{40}{5} + \frac{3 \times 73 \times 4}{5 \times 25}\right)$$
$$= \frac{60}{100}(-8 + 7{\cdot}008) = -0{\cdot}5952$$

и

$$y = 9{\cdot}24 + 1{\cdot}824x - 0{\cdot}5952x^2\ ,$$

откуда вычисленныя (y') значенія y будутъ:

y	y'
3	3·2
7	6·8
9	9·2
10	10·5
11	10·5
40	40·2

Для облегченія вычисленій по способу наименьшихъ квадратовъ Парето предлагаетъ слѣдующій пріемъ:

раздѣливъ рядъ на двѣ равныя части, обозначимъ разстояніе (x) каждаго члена ряда отъ середины, какъ и при способѣ моментовъ, черезъ $-1, -2 \ldots, +1, 2, \ldots$ (при нечетномъ ихъ числѣ) или $-1{\cdot}5, -0{\cdot}5 \ldots, +0{\cdot}5, 1{\cdot}5 \ldots$ (при четномъ числѣ членовъ); затѣмъ составимъ функціи отъ x слѣдующаго вида:

$$\psi_I = x$$

$$\psi_{II} = x^2 - \frac{n^2 - 1}{12}$$

$$\psi_{III} = x^3 - \frac{3n^2 - 7}{20}\,x$$

(гдѣ n — число членовъ ряда), а также суммы квадратовъ:

$$S\,\psi_{II}^{\,2}\ ,\ S\,\psi_{III}^{\,2}.$$

Въ такомъ случаѣ формулу

$$y = a + bx + cx^2 + \ldots$$

можно замѣнить посредствомъ

$$y = a + b\psi_I + c\psi_{II} + d\psi_{III} + \ldots,$$

гдѣ

$$a = \frac{S\,y}{n}$$

$$b = \frac{S\,xy}{S\,x^2} = \frac{S\,y\psi^2}{S\,\psi_I}$$

$$c = \frac{S\,y\psi_{II}}{S\,\psi_{II}^{\,2}}$$

$$d = \frac{S\,y\psi_{III}}{S\,\psi_{III}^{\,2}}$$

Такъ, для приведеннаго выше примѣра:

y	$x=\psi_I$	ψ_{II}	ψ_{III}	$y\psi_I$	$y\psi_{II}$	$y\psi_{III}$
3	-2	2	$-1{\cdot}2$	-6	6	$-3{\cdot}6$
7	-1	-1	$2{\cdot}4$	-7	-7	$16{\cdot}8$
9	0	-2	0		-18	
10	1	-1	$-2{\cdot}4$	10	-10	$-24{\cdot}0$
11	2	2	$+1{\cdot}2$	22	22	$13{\cdot}2$
40				$+32$	-35	$+30{\cdot}0$
$S\psi_I^2 = 10$		$Sy = 40$		-13	$+28$	$-27{\cdot}6$
$S\psi_{II}^2 = 14$		$n = 5$		19	-7	$+2{\cdot}4$
$S\psi_{III}^2 = 14{\cdot}4$				$Sy\psi_I$	$Sy\psi_{II}$	$Sy\psi_{III}$

$$a = \frac{40}{5} = 8$$

$$b = \frac{19}{10} = 1{\cdot}9$$

$$c = -\frac{7}{14} = -\frac{1}{2}$$

и

$$y = 8 + 1{\cdot}9\,\psi_I - \frac{1}{2}\psi_{II}$$

что даетъ слѣдующій рядъ:

y	y'
3	$3{\cdot}2$
7	$6{\cdot}6$
9	$9{\cdot}0$
10	$10{\cdot}4$
11	$10{\cdot}8$
40	$40{\cdot}0$

7. Мѣра совпаденія.

Вычисленныя по формуламъ значенія обычно не совпадаютъ съ эмпирически данными величинами и часто возникаетъ необходимость оцѣнить степень совпаденія или несовпаденія теоретическаго ряда съ эмпирическимъ. Обыкновенно для этой цѣли довольствуются общимъ впечатлѣніемъ отъ нагляднаго сопоставленія обоихъ рядовъ, и если общее направленіе ихъ сходно, въ отдѣльныхъ же членахъ большихъ уклоненій нѣтъ, то говорятъ о достаточномъ совпаденіи. Въ виду субъективности представленія о «большомъ» или «маломъ» уклоненіи, предпочтительнымъ является выразить мѣру совпаденія количественно Для этой цѣли можно, напримѣръ, взять разницу, между каждой парой членовъ теоретическаго и эмпирическаго ряда, выразить ее въ % къ среднеариѳметическому изъ соотвѣтствующей пары членовъ и затѣмъ образовать среднеариѳметическую изъ всѣхъ разницъ. Подобнымъ же образомъ можно сравнивать площади, вычисляя несовпадающія площади посредствомъ планиметра и сравнивая ихъ въ % съ совпадающими площадями.

Такимъ образомъ мѣра совпаденія или несовпаденія получаетъ объективное количественное выраженіе, въ видѣ опредѣленнаго процента или отношенія. Но при этомъ все же субъективнымъ остается то, какой процентъ расхожденія или уклоненія считать существеннымъ или несущественнымъ. Весьма часто этотъ вопросъ рѣшается практическою цѣлью, для которой дѣлаются вычисленія. Пирсонъ предложилъ способъ замѣны субъективной оцѣнки мѣры расхожденія вычисленіемъ вѣроятности расхожденія рядовъ. Онъ исходитъ изъ допущенія, что эмпирически данный рядъ величинъ представляетъ собою приближенное выраженіе нѣкотораго ряда истинныхъ значеній этихъ величинъ, снабженное случайными уклоненіями. Если бы вмѣсто даннаго ряда эмпирическихъ величинъ взять другой эмпирическій рядъ величинъ того же рода, онъ точно также случайно уклонялся бы отъ истинныхъ значеній, а слѣдовательно и отъ перваго ряда или вообще всякаго другого эмпирическаго ряда тѣхъ же величинъ. Если уклоненія случайны и подчиняются какому-либо закону, напримѣръ такому, какъ уклоненія нормальной кривой Гаусса, то можно опредѣлить вѣроятность уклоненія любого размѣра, а равнымъ образомъ можно опредѣлить размѣръ уклоненія, которое съ опредѣленной вѣроятностъю, напримѣръ, очень большой, можно считать случайнымъ. Отсюда ясно, что если считать теоретическій рядъ также приближеннымъ выраженіемъ ряда истинныхъ величинъ, то можно опредѣлить, съ какою вѣроятностью слѣдуетъ считать разницу между эмпирическимъ и теоретическимъ рядомъ случайной.

Мѣру уклоненія Пирсонъ выражаетъ посредствомъ суммы квадратовъ уклоненій каждаго эмпирическаго значенія отъ теоретическаго, дѣленныхъ на соотвѣтствующее теоретическое значеніе; если члены эмпирическаго ряда обозначить черезъ

$$m_1', \ m_2' \ . \ . \ . \ m'_{n+1}$$

и теоретическаго:

$$m_1, \ m_2 \ . \ . \ . \ m_{n+1},$$

и разницы

$$m_i' - m_i = e_i,$$

то мѣра уклоненія f выразится

$$f^2 = S\left(\frac{e^2}{m}\right).$$

По этой мѣрѣ расхожденія требуется опредѣлить вѣроятность расхожденія такой величины, если допустить, что оба ряда представляютъ собою лишь случайно уклоняющіяся выраженія одного и того же истиннаго ряда. Для величины этой вѣроятности (P) Пирсонъ нашелъ слѣдующее выраженіе

$$P = e^{-1/_2 \chi^2} \left(1 + \frac{\chi^2}{2} + \frac{\chi^4}{2 \cdot 4} + \frac{\chi^6}{2 \cdot 4 \cdot 6} + \ldots + \frac{\chi^{n+2}}{2 \cdot 4 \ldots (n-2)}\right)$$

гдѣ $n+1$ число членовъ ряда нечетное (или можетъ быть сдѣлано нечетнымъ прибавленіемъ члена съ значеніемъ 0). Таблицы значеній P вычислены Elderton'омъ (Biometrica, I) и воспроизведены у Леонтовича (ч. 1).

Примѣръ. При метаніи 12 игральныхъ костей поверхность съ 5 или 6 очками была получена на 0, 1, 2 . . . 12 костяхъ въ m' числѣ случаевъ; теоретическая вѣроятность появленія поверхностей составляетъ m; разница m' — m = d

v'	m'	m	d	d²	$\frac{d^2}{m}$
0	185	203	18	324	1·59 606
1	1 149	1 217	68	4 624	3·79 951
2	3 265	3 345	80	6 400	1·91 330
3	5 475	5 576	101	10 201	1·82 945
4	6 114	6 273	159	25 281	4·03 013
5	5 194	5 018	176	30 976	6·17 298
6	3 067	2 927	140	19 600	6·69 628
7	1 331	1 254	77	5 929	4·72 807
8	403	392	11	121	0 30 903
9	105	87	18	324	3·72 414
10	14	13	1	1	0 07 346
11	4	1	3	9	9 00 000
12	0	0	0	0	00 000
				Σ	48·87 241

$$\chi^2 = 48{\cdot}87\ 241$$
$$\chi = 6{\cdot}623\ 625$$
$$P = 0{\cdot}000\ 016.$$

Результатъ выражаетъ, что при случайномъ распредѣленіи подобное данному уклоненіе (или большое) допускаютъ только 16 случаевъ изъ 1 000 000, или на 62 499 случаевъ всевозможныхъ уклоненій, если они обязаны своимъ происхожденіемъ случаю, только 1 разъ мы могли бы получить уклоненіе, не меньшее, чѣмъ данное; во всѣхъ же остальныхъ случаяхъ уклоненіе должно было бы быть меньшее. Такимъ образомъ наше субъективное впечатлѣніе о случайности или неслучайности уклоненія одного ряда отъ другого получаетъ опредѣленное численіе выраженіе, не допускающее разумнаго сомнѣнія въ томъ, что эмпирическій рядъ не подходитъ подъ теоретическую формулу.

75-00