Глава V
СТАТИСТИЧЕСКИЕ ВЫВОДЫ: ОЦЕНИВАНИЕ И ПРОВЕРКА ГИПОТЕЗ

1. Генеральная и выборочная совокупность. Оценка ошибки выборки

Объектом социологических исследований обычно являются различные социальные общности. Если изучаются все индивиды данной совокупности, то говорят о сплошном исследовании, если же только часть, то о выборочном. Как правило, социологические исследования носят выборочный характер.

Это связано прежде всего с тем, что экономические и временные ограничения не позволяют провести сплошное исследование (затраты на проведение Всесоюзной переписи, например, составляют десятки миллионов рублей и требуют более 700 тысяч интервьюеров[1])

Но даже в тех случаях, когда сплошное исследование практически осуществимо, зачастую рентабельней проводить выборочное. Его экономичность позволяет увеличить затраты на совершенствование инструмента исследования и компенсировать тем самым падение надежности за счет того, что исследование не сплошное, а выборочное — в итоге исследователь имеет возможность получить более полную и надежную информацию.

Основания, которые позволяют нам по изучению части судить о целом, связаны_с некоторыми вероятностными законами[2]. Если, например, вытаскивать из урны[3], в которой находятся хорошо перемешанные белые и черные камешки (50 белых и 50 черных), 20 камешков, то вероятность того, что нам попадутся все черные камешки очень мала

[167]

вероятность того, что первый камешек будет черным – 50/100, что второй — 49/99, так как в урне осталось всего 99 камешков, из них 49 черных; тогда вероятность, что первые два — черные, равна 50/100 * 49/99. Продолжая эти рассуждения, получаем, что вероятность того, что все 20 камешков черные, равна 50/100 * 49/99 * .... * 31/81= 0,00000009. Вероятность того, что одних камешков будет намного больше, чем других, тоже мала. Наиболее часто будет встречаться такая ситуация, при которой число черных камешков приблизительно равно числу белых. Аналогично, если в городе половина населения имеет одни ценностные ориентации, а половина - другие, то маловероятно в выборочном исследовании (при случайном отборе[4]) получить, что лиц с одними ценностными ориентациями намного больше, чем с другими.

Однако осуществить случайный отбор очень трудно. Даже в случае с камешками требуется обеспечить хорошее перемешивание, одинаковые размеры камешков и гарантию, что тот, кто вытаскивает, не видит цвета камешка (примером идеально организованного случайного отбора являются тиражи спортлото). Эксперименты, проведенные с отбором камешков одного цвета, лежащих на столе, показали, что испытуемые непроизвольно выбирают более крупные камешки - они, видимо, чаще попадаются под руку[5]. Несравнимо трудней обеспечить случайный отбор в социологических исследованиях. Широко известен пример неудачного прогноза результатов выборов президента в США в 1936 г.: журнал «Литэрари Дайджест» по телефонным книгам отобрал свыше двух миллионов адресатов, получив тем самым, казалось бы, случайную выборку. По адресам были разосланы открытки с просьбой ответить — Рузвельту или Ландону отдаст свой голос респондент. По результатам опроса журнал предсказал победу с большим перевесом Ландона. Интересно, что социологи Дж. Гэлап и Эл. Роупер правильно предсказали победу Рузвельта,

---

[1] Всесоюзная перепись населения — всенародное дело. М., 1978, с. 41.

[2] Примеры с урной широко используются в теории вероятностей и восходят, видимо, к принятой в Древней Греции процедуре голосования.

[3] Примеры с урной широко используются в теории вероятностей и восходят, видимо, к принятой в Древней Греции процедуре голосования.

[4] Случайным отбором называют такой, при котором все элементы исследуемой совокупности имеют равную вероятность попасть в выборку.

[5] *Пейте Фрэнк.* Выборочный метод в переписях и обследованиях. М., 1965, с. 34.

основываясь на анализе в 500 раз меньшего массива – четырех тысяч анкет. Ошибка в прогнозе «Литэрари Дайджест» объясняется тем, что выборка по телефонным книгам не была случайной, она не обеспечивала

[168]

равную вероятность попасть в выборку для всех лиц, имеющих избирательное право, так как в 1936 г. телефоны были преимущественно у обеспеченных слоев населения, предпочитавших Ландона.

Свойство выборки отражать характеристики изучаемой совокупности называется *репрезентативностью.* Иногда вместо выборки говорят *выборочная* совокупность, а изучаемую совокупность называют *генеральной.* Можно сказать, что генеральная совокупность – это та, на которую исследователь намерен распространять выводы, сделанные при изучении выборки. Различие характеристик выборочной и генеральной совокупности называют ошибкой репрезентативности. Можно выделить два вида таких ошибок — систематические и случайные.

Систематические ошибки — это ошибки такого типа, как допущенные журналом «Литэрари Дайджест», т.е. некоторое постоянное смещение, которое не уменьшается при увеличении числа опрошенных (если бы журнал опросил не два, а четыре миллиона обладателей телефонов, это не спасло бы его от ошибки).

Случайные ошибки — это те, которые при повторных измерениях изменяются по вероятностным законам. В частности, если мы определяем некоторую характеристику выборки, например, среднее арифметическое, то извлекая все новые и новые выборки того же размера будем получать, что эта характеристика отклоняется то в одну, то в другую сторону от истинного значения (т.е. от значения в генсовокупности) приблизительно с одинаковой частотой и при увеличении числа выборок средняя арифметическая ошибка приближается к нулю. Систематическую ошибку можно устранить, изменяя процедуру формирования выборки; случайная ошибка будет присутствовать всегда, при любом выборочном опросе. Тем не менее систематическая ошибка значительно опасней, так как по выборке ее невозможно оценить. Случайная же ошибка подчиняется определенным законам и поддается оценке. Вообще репрезентативность выборки характеризуется двумя взаимосвязанными параметрами – уровнем ошибки и вероятностью. Говорить о какой-либо выборке, что она репрезентативна, не совсем точно, так как любая выборка имеет определенный уровень репрезентативности (хотя этот уровень может нас совершенно не устраивать). Более точно говорят, что ошибка репрезентативности данной выборки с вероятностью P не превышает Δ (вероятность P называют доверительной).

[169]

Для случайной выборки существуют методы, позволяющие оценить эту ошибку (мы рассмотрим их при изложении способов проверки гипотез). При планировании социологического исследования обычно решают иную задачу — задаются некоторым устраивающим исследователя уровнем точности результата, т.е. допустимой ошибкой и доверительной вероятностью, и определяют для этих параметров необходимый объем выборки. В частности, объем выборки для определения доли некоторого признака X в генсовокупности определяется формулой[6]:

$$n = \frac{1}{\dfrac{\Delta^2}{t^2 \nu(1-\nu)} + \dfrac{1}{N}},$$

где N - объем генеральной совокупности, n — объем выборки, t - коэффициент, соответствующий доверительной вероятности (см. табл. И Приложения 3; если $n > 60$, то при

[6] *Кокрен У.* Методы выборочного исследования. М., 1976, с. 89.

P=0,954 t=2, а при P=0,997 t=3 и т.д.), ν - доля признака X в генсовокупности, Δ — величина допустимой ошибки (в долях).

Если, например, исследователь хочет получить с вероятностью 0,95 (t = 2) данные о доле признака X в генсовокупности (пусть N = 10000) с ошибкой, не превышающей 5% (Δ = 0,05), и ему известно, что искомая доля составляет приблизительно *20%* (ν = 0,20), то по формуле (V,1,1) получим, что требуется опросить 256 человек *(n* = 256).

Неудобство пользования формулой заключается в том, что она требует хотя бы приближенной информации о доле признака в генеральной совокупности, т.е. как раз о том, что исследователю требуется определить. Чтобы избавиться от этого неудобства, заметим, что при ν=0,5 произведение $\nu(1 - \nu)$ максимально, следовательно, n тоже максимально. Поэтому если в (V, 1,1) вместо ν подставить 0,5, мы получим формулу, которой можно пользоваться при любых значениях доли признака в генеральной совокупности (объем выборки при этом будет получаться с некоторым запасом). Положив также значение доверительной вероятности равным 0,954, т.е. t = 2, получим

$$n = \frac{1}{\Delta^2 + \frac{1}{N}}, \text{ (V,1,1')}$$

[170]

Воспользовавшись этой формулой, определим, как объем выборки зависит от объема генеральной совокупности и от величины допустимой в исследовании ошибки.

Из таблицы видно, что для обеспечения заданной репрезентативности при исследовании города с населением 100 тыс. жителей надо опросить 398 человек, а при исследовании всей страны практически столько же — 400 чел. Для обеспечения одного и того же уровня репрезентативности (5 %) требуется опросить такие доли генсовокупности:

Таблица 37

Зависимость объема выборки от объема генсовокупности при допустимой ошибке 5% доверительная вероятность — 0,954)

Объем генсовокупности	500	1000	2000	3000	4000	5000	10000	100000	Бесконечная
Объем выборки	222	286	333	350	360	370	385	398	400

для N = 500-222 человека, т.е. приблизительно 44% генсовокупности, для N = 5000 - 7,4%, а для N, равном 4 миллионам (например, население Ленинграда) — сотую долю процента. Поэтому изредка встречающиеся в публикациях характеристики выборки типа «было опрошено 15% генсовокупности» или «опрашивался каждый двадцатый школьник» ничего не говорят о репрезентативности выборки. Вообще из таблицы видно, что начиная с некоторого момента увеличение объема генеральной совокупности не оказывает существенного влияния на увеличение объема выборки, поэтому при больших генеральных совокупностях (скажем, при N > 5000) величиной $1/N$ в формуле (V,1,1') можно пренебречь. Тогда формула (V,1,1') примет вид:

$n = \frac{1}{\Delta^2}$, откуда $\Delta = \sqrt{\frac{1}{n}}$. Рис. 22, показывающий связь между объемом и ошибкой выборки, может использоваться для принятия решения о требуемом объеме выборки.

При планировании объема выборки следует иметь в виду следующее. Приведенные выше формулы позволяют получить заданную точность при анализе выборки в целом, т.е. если мы не будем расчленять ее на части. Если, например, требуется определить долю лиц, состоящих в браке, для крупного города, то опросив 400 случайным образом

[171]

отобранных человек мы определим искомую долю с ошибкой, не превышающей 5% (с вероятностью 0,954). Но если мы хотим определить эту долю не для всего массива в целом, а для женщин и для мужчин, то нам необходимо, чтобы в выборке было 400 мужчин и 400 женщин, т.е. 800 человек. Чем больше будет дробиться массив при анализе информации, тем больший объем выборки потребуется. Программа

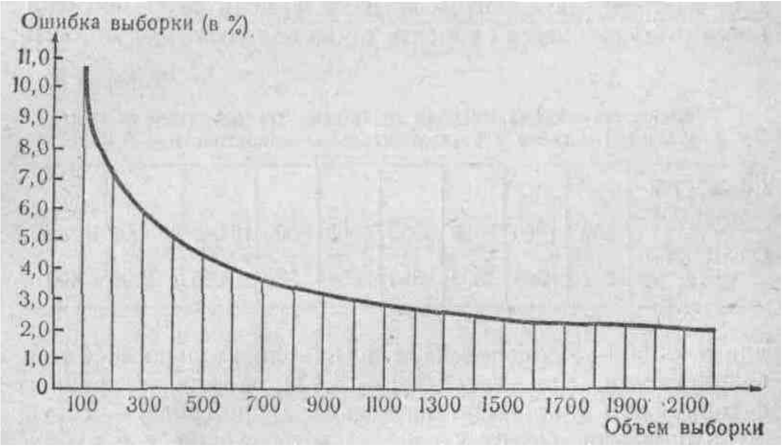

Рис. 22. Зависимость между объемом и ошибкой выборки для $P = 0{,}954$, $p=q=1/2$ и бесконечно большой генеральной совокупности

исследования, план обработки и анализа информации существенно влияют на объем выборки.

Но основная сложность при планировании выборки заключается, пожалуй, в том, что во многих случаях, особенно при крупномасштабных исследованиях, случайную выборку сформировать очень сложно. Так, в проведенном нами репрезентативном исследовании работающего населения г. Киева[7] ($P = 0{,}954$, Δ не более 3—5%), данные которого уже использовались в примерах, была применена следующая процедура, моделирующая случайный отбор. Из избирательных списков для каждого участка города (всего их более семисот[8]) отбирались с определенным шагом адреса избира-

[172]

телей: например, адрес 1-й, 100-й, 200-й и т.д. (если шаг — 100).

При этом могла быть допущена систематическая ошибка такого рода. Фамилии в списках расположены в алфавитном порядке; начиная с 1-го номера, мы почти всегда включаем в выборочный список фамилию на букву А, т.е. в списке доля лиц с фамилией на букву А будет выше, чем в генсовокупности. Если у лиц какой-либо национальности (например, армянской) фамилии чаще начинаются на А, чем у лиц другой национальности, то их доля в выборке выше, чем в генсовокупности.

Поэтому на всякий случай мы «стохастизировали»[9] выбор первой фамилии в списке, сделав его случайным, равномерно распределенным внутри шага выборки (можно, например, начинать с номера равного целой части числа $\frac{k}{7}+1$, где k—номер избирательного

[7] Исследование проводилось отделом конкретных социологических исследований Института философии АН УССР (руководитель исследования В. Ф. Черноволенко, проект выборки разработан В. И. Паниотто).

[8] Списки всех идбираюльных участков ранона хранятся в райисполкомах 12-и районов Киева, что очень упрощает риботу.

[9] От слова стохастический, т.е. случайный.

участка, изменяющийся от 1 до 700: тогда в 7 избирательных участках фамилии будут отбираться с 1-го номера, в 7 — со 2-го и т.д. до сотого номера.

Поскольку нас интересовало все работающее население, а не только лица, старше 18 лет и занесенные в списки для голосования, опрашивалось не обязательно лицо, указанное в избирательном списке. Список использовался лишь как выборка адресов. При посещении семьи, проживающей по данному адресу, интервьюер переписывал всех проживающих в определенном порядке и по специальной процедуре[10], стохастизирующей выбор, определял, кого надо опросить[11]. Можно показать, что полученная выборка является случайной, для оценки ошибки выборки применима формула (V,1,1).

Но что делать в случае, когда такой список невозможно составить, например, при построении выборки, репрезентативной для Советского Союза? В таких случаях прибегают к многоступенчатому отбору: сначала (первая ступень) из множества всех областей страны отбирают случайным образом области (область в данном случае — это единица).

[173]

отбора на первом шаге). На второй ступени из выбранных областей отбирают районы. На третьей из каждого района — населенные пункты. На четвертой из населенных пунктов — лиц, подлежащих опросу (на первых трех ступенях выбирались единицы отбора, на последней — единицы исследования). При таком построении выборки формула (IV, 1,1) непригодна для оценки ошибки, требуются формулы, позволяющие оценить ошибку, возникающую на каждой ступени[12].

Чтобы оптимизировать процедуру построения выборки, исследователи на первых ступенях отбора используют специальные процедуры. Например, при построении всесоюзной выборки для исследования читателей газеты «Правда»[13] на первой ступени в качестве единиц отбора (или, как их иногда называют, гнезд) использовались области либо республики, если республика не имела областного деления. Всего было выделено 130 гнезд: 71 область РСФСР, 25 областей Украины, 6 областей Белоруссии, 17 областей Казахстана и 11 союзных республик. Прежде чем отбирать гнезда, они были сгруппированы таким образом, чтобы в одну группу попадали территориальные единицы, близкие по урбанизированности, степени развития сельской субкультуры, уровню развития инфраструктуры и промышленного развития области (эти четыре группы переменных описывались 105-ю показателями).

Разбиение на группы (называемое также стратификацией, или районированием) производилось на ЭВМ с помощью так называемого лингвистического метода (это один из методов таксономии, см. главу VI). В результате 130 гнезд были разбиты на 12 страт. Из каждой страты отбиралось некоторое число гнезд (пропорционально численности в данной страте), всего их было отобрано 25. Эта процедура эффективнее случайного отбора 25 из 130 гнезд, так как гарантирует, что в выборку попадут гнезда разных типов, представители всех страт. N u n здесь невелики, поэтому ошибка случайного отбора была бы слишком большой (если бы требовалось отобрать 250 из 1300 гнезд, то можно было бы использовать случайную выборку).

На следующем этапе каждое из гнезд (в данном случае областей) выступало в качестве генеральной совокупности. В каждой области выделялись свои гнезда — города област-

[10] Использовалась некоторая модификация процедуры Киша. *(Кокрен У.* Методы выборочного исследования. М., 1976, с. 384, 385; *Петренко Е. С., Ярошенко Т. М.* Социально-демографические показатели в социологических исследованиях. М., 1979, с. 96.)

[11] Для реализации случайных отборов иногда используются специальные таблицы случайных чисел (см. табл. М Приложения 3)

[12] *Кокрен У.* Методы выборочного исследования. М., 1976, гл. 10.

[13] Территориальная выборка в социологических исследованиях. М., 1980, гл. 3.

[174]

ного и республиканского подчинения, районы. Выделенные гнезда группировались в страты, из каждой страты отбиралось определенное число гнезд и т.д. Всего выборка состояла из 6 ступеней (3-я ступень — районы города и села, 4-я — жилищные организации, т.е. ЖЭКи, ЖКК, общежития и т.п.); 5-я — семьи; 6-я — респонденты в семье).

Другим примером многоступенчатой выборки является трехступенчатый отбор респондентов, осуществленный при исследовании городов Татарии[14].

Вообще всякое выборочное исследование может быть охарактеризовано следующими параметрами: числом ступеней, способом выделения гнезд, способом их группировки (стратификации) и способом отбора гнезд на каждой ступени. Очевидно, что во многих случаях способ организации выборки весьма далек от одноступенчатого случайного или даже многоступенчатого случайного отбора. В этих случаях вообще не существует формул для оценки ошибки выборки. Если для некоторых из изучаемых признаков есть контрольные цифры по генсовокупности, полученные из государственной или ведомственной статистики, то можно оценить ошибку выборки по этим признакам. Сопоставляя величину ошибки с той, которая была бы, если бы выборка строилась как одноступенчатая случайная (т.е. с ошибкой, рассчитанной по формуле (V,1,1)), можно оценить отклонения построенной выборки от случайной и внести коррективы в результаты, полученные с помощью формул, основанных на предположении о случайности выборки.

Весь последующий материал этой главы дан в предположении, что из генеральной совокупности извлекается одноступенчатая выборка.

2. Выборочное распределение

Статистический вывод — это некоторое утверждение об изучаемой генеральной совокупности на основании изучения выборки (т.е. рассуждение от частного к общему, индукция). Математическая статистика рассматривает, разумеется, не любые утверждения о генеральной совокупности, а лишь касающиеся числовых характеристик, рассмотренных выше (средние, меры вариации, коэффициенты корреляции,

[175]

и т.п.). Числовые характеристики, описывающие генеральную совокупность, называются *параметрами.* Те же самые характеристики, но рассчитанные для выборки, называются *статистиками.* Мы будем обозначать параметры через $M^Г$, $(\sigma^Г)^2$, $r^Г$ и т.д. (генеральное среднее, дисперсия, коэффициент корреляции), а статистики через M^B, $(\sigma^B)^2$, r^B и т.д. (выборочное среднее, дисперсия, коэффициент корреляции)[15]. Таким образом, статистический вывод — это утверждение о параметрах генеральной совокупности на основании изучения статистики. Такие утверждения носят вероятностный характер и подразделяются на три вида: статистическое оценивание точечное, статистическое оценивание интервальное и проверка гипотез.

Статистическое оценивание заключается в том, что исследователь по выборке ищет показатель, наиболее близкий к оцениваемому параметру, или интервал, в границах которого с большой вероятностью лежит этот параметр. Другой разновидностью статистического вывода является проверка гипотез: исследователь заранее формулирует некоторое утверждение о параметрах генеральной совокупности (гипотезу), затем оценивает степень соответствия результатов, полученных в выборочном исследовании, сформулированной гипотезе и принимает решение об истинности или ложности гипотезы. Методологию

[14] *Рукавишников В. О., Елисеева И. И.* Проблемы проектирования социологического исследования крупных территориальных объектов.- В кн.: Проектирование и организация выборочного социологического исследования, М., 1977.

[15] В литературе можно встретить также обозначение параметров греческими, а статистик—латинскими буквами. Например, μ, δ^2, π и M^2, s^2, p для среднего, дисперсии и доли признака соответственно.

проверки статистических гипотез мы рассмотрим подробнее, что позволит уточнить различие и сходство видов статистического вывода. Сейчас для нас важно, что статистическое оценивание и проверка гипотез основываются на идее так называемого *выборочного распределения* (обращаем внимание читателя на важность этого понятия для понимания сущности статистического вывода).

Рассмотрение выборочного распределения начнем с примера. Предположим, что известно распределение оценок 1000 абитуриентов некоторого вуза на экзамене по математике. Пусть 400 человек получили 2, 200—3, 300—4 и 100—5, полигон распределения приведен на рис. 23. Легко подсчитать, что средний балл равен $M^Г$ = 3,1, а дисперсия $(\sigma^Г)^2$=0,ll. Насколько вероятно получить в выборке значение, существенно отличающееся от генерального среднего? Определим, например, вероятность того, что для вы-

[176]

борки из 5 человек выборочое среднее M^B будет отличаться от генерального не менее, чем на 0,5, т.е. $|M^Г - M^B| \geq$ 0,5. С этой целью станем формировать выборки по 5 человек многократно и вычислять для каждой из них средний балл. Тогда отношение таких выборок, для которых $| M^Г - M^B | \geq 0{,}5$, к общему числу извлеченных выборок даст частоту, близкую к искомой вероятности. При увеличении числа выборок эта частота неограниченно приближается к вероятности того, что $| M^Г - M^B | \geq$ 0,5. Построим полигон рас-

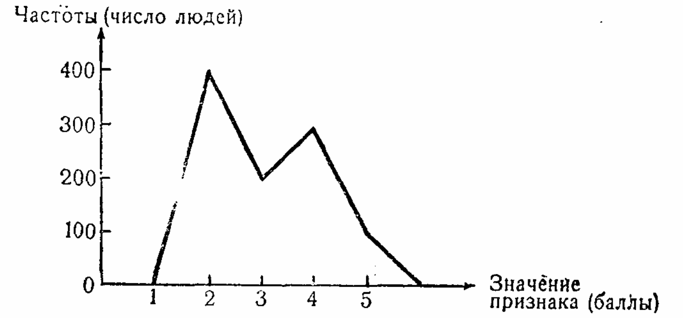

Рис. 23. Пример полигона распределения (оценки по математике)

пределения выборочных средних (по оси абсцисс откладываются значения M^B, по оси ординат — частоты): при увеличении числа выборок до бесконечности получим полигон[16], называемый выборочным распределением статистики M^B (рис. 24). Искомая вероятность равна отношению площади под кривой выборочного распределения, заштрихованный на рис. 24 к площади всей кривой (напоминаем, что площадь под кривой между двумя точками а и & на оси абсцисс равна сумме частот для значений признака, лежащих между *a* и *b*. Зная выборочное распределение, можно решать и другой вопрос — зафиксировать не интервалы изменения статистики, а вероятность, и искать, в каких пределах с заданной вероятностью лежит статистика. Например, можно определить, в каком интервале с вероятностью 0,95 лежит выборочное среднее M^B. Для этого на графике выборочного распределения влево и вправо от генерального среднего $M^Г$ откладывается такой отрезок Δ, чтобы между

[177]

$M^Г$ - Δ и $M^Г$ + Δ было заключено 95% площади (на рис. 25 площадь под кривой, лежащей в указанных пределах, заштрихована). Таким образом, зная выборочное распределение, можно определить Δ так, что с вероятностью 0,95 выполняется неравенство: $|M^Г — M^B|<\Delta$.

Предположим теперь, что мы не знаем генерального среднего $M^Г$, но знаем выборочное

[16] Полигон при этом превратится в плавную кривую.

распределение статистики $M^В$

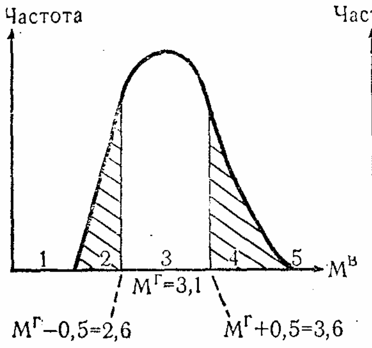

Рис. 24. Выборочное распределение $M^В$ для распределения оценок, изображенных на рис. 23

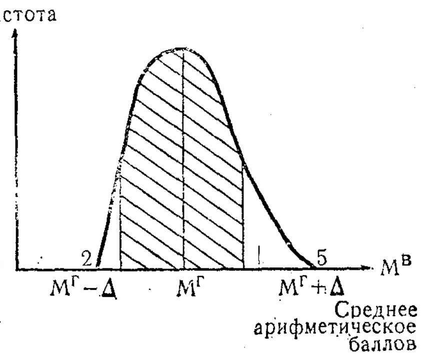

Рис. 25. Выборочное распределение $M^В$, доверительный интервал, включающий 95% площади (заштрихована)

и по нему нашли Δ таким образом, что $|M^Г — M^В|<\Delta$ с вероятностью 0,95. Если мы получим, что в некоторой выборке средний балл абитуриентов равен, например, 3,4, то с вероятностью 0,95 мы можем утверждать, что неизвестное нам генеральное среднее $M^Г$ отличается от найденного значения не больше, чем на Δ, т.е. с вероятностью 0,95 $|M^Г — 3,4|<\Delta$ или $3,4 — \Delta < M^Г < 3,4 + \Delta$. Таким образом, мы получили интервальную оценку для неизвестного параметра $M^Г$.

Но какой смысл в такой оценке, если для того, чтобы ее получить, надо экспериментально определить выборочное распределение, что практически неосуществимо? Оказывается, однако, что во многих случаях связь между параметром и выборочным распределением статистики носит такой характер, что выборочное распределение статистики можно построить теоретически (более того, для этого часто не требуется никакой информации о форме генерального распределения).

[178]

Рассмотрим этот вопрос более подробно. Пусть из бесконечно большой совокупности с параметрами $M^Г$ (среднее арифметическое) и $(\sigma^Г)^2$ (дисперсия) извлекаются случайные выборки объема n. Можно доказать, что при достаточно больших n выборочное распределение средних $M^В$ будет описываться законом, близким к нормальному, причем среднее арифметическое всех выборочных средних будет равно $M^Г$, а дисперсия выборочных средних будет равна $\frac{(\sigma^Г)^2}{n}$. Это утверждение называется центральной предельной теоремой (слово «центральная» характеризует ее роль в теории статистического вывода). Для предыдущего примера это означает, что выборочное распределение на рис. 25 приближается к нормальному (если бы мы извлекали выборки объема 100, а не 5, то сходство с нормальным распределением было бы значительно большим), причем среднее выборочного распределения равно 3,1, а дисперсия равна 0,11/5 (где 0,11 —дисперсия генерального распределения). Удивительным является тот факт, что распределение выборочных средних близко к нормальному независимо от формы распределения генеральной совокупности. Как уже говорилось (§ 3, гл. I), для нормального распределения существуют таблицы, показывающие, какая доля площади лежит под кривой между любыми двумя точками $M — z\sigma$ и $M + z\sigma$ оси абсцисс (см. табл. А Приложения 3). Известно, например, что в пределах $\sigma^В$ от среднего арифметического (т.е. от $M—\sigma^В$ до $M+\sigma^В$) лежит 68,2% площади кривой; в пределах $2\sigma^В$ — 95,4%;

в пределах $3\sigma^B$—99,7%. Другими словами, вероятность того, что отобрав из генеральной совокупности *n* единиц и рассчитав M^B, мы получим, что $|\mathrm{M}^{\mathrm{B}} - \mathrm{M}^{\Gamma}| < \frac{2\sigma^\Gamma}{\sqrt{n}}$ с вероятностью 0,682; вероятность того, что $|\mathrm{M}^{\mathrm{B}} - \mathrm{M}^{\Gamma}| < \frac{\sigma^\Gamma}{\sqrt{n}}$ равна 0,954; вероятность того, что $|\mathrm{M}^{\mathrm{B}} - \mathrm{M}^{\Gamma}| < \frac{3\sigma^\Gamma}{\sqrt{n}}$ равна 0,997. Это и дает возможность по выборке оценивать генеральную совокупность.

Предположим, что в нашем примере мы знаем генеральную дисперсию $(\sigma^\Gamma)^2 = 0{,}11$, или $\sigma^\Gamma = 0{,}33$, но не знаем генерального среднего[17]. Пусть для выборки, состоящей из 5 единиц, оказалось, что $M^B = 3{,}3$. Принятая в социологии степень надежности высказываемых утверждений *P* обычно

[179]

равна[18] 0,954 или 0,997. Для $P = 0{,}954$, в силу приведенных выше неравенств: $\left|3{,}3 - \mathrm{M}^{\Gamma}\right| < 2 \cdot \frac{0{,}33}{\sqrt{5}} = 0{,}25$, т.е. $3{,}05 < M^\Gamma < 3{,}55$. Интервал, в который попадает значение оцениваемого параметра, называется доверительным, а вероятность того, что доверительный интервал содержит этот параметр, называется доверительной вероятностью. В нашем примере интервал от 3,05 до 3,55 является доверительным для среднего арифметического генеральной совокупности, построенным с доверительной вероятностью 0,954. Можно также сказать, что (3,05; 3,55) — это 95,4%-и доверительный интервал для M^Γ в окрестности точки 3,3.

Может быть, полезной для понимания окажется следующая аналогия. В урне для голосования лежат 20 белых и черных камешков. Предположим, что из теоретических соображений известно, что 19 из них одного цвета, а 1 —другого. Мы вытаскиваем 1 камешек и оказывается, что он белый.Тогда с вероятностью $\frac{19}{20} = 0{,}95$ можно утверждать, что в урне 19 белых и 1 черный камешек. При этом вероятность, что мы ошиблись и в урне 19 черных и 1 белый — тот единственный, который мы вытащили — равна $\frac{1}{20} = 0{,}05$

Упражнение 87. В проведенном нами выборочном исследовании 3500 жителей г. Киева средняя зарплата в выборке равна 150,0 руб. ($M^B = 150{,}0$). Предположим, что исследования прошлых лет показали устойчивость дисперсии генеральной совокупности и мы полагаем, что она нам известна (пусть $(\sigma^\Gamma)^2 == 3700$). Найти 95%-й доверительный интервал для средней зарплаты. Ответ: 148 руб.; 152 руб.

Итак, доверительный интервал для неизвестного среднего генеральной совокупности M^Γ при известной дисперсии $(\sigma^\Gamma)^2$ строится следующим образом:

1. Из генеральной совокупности извлекается выборка достаточно большого объема *n* (100 и более) и рассчитывается среднее $\bar{x}$.

2. Выбирается некоторая доверительная вероятность и по специальной таблице (табл. А Приложения 3) находится коэффициент *z*, соответствующий этой вероятности.

---

[17] Обычно в социологических исследованиях нам неизвестны ни средняя, ни дисперсия генсовокупности. В данном случае мы приняли это допущение, чтобы упростить ситуацию. Ниже будет рассмотрен случаи оценивания, когда неизвестны никакие параметры генсовокупности.

[18] Вероятности эти выбраны еще и из тех соображений, что при вероятности 0,954 в неравенстве $|M^B - M^\Gamma| < \sigma^\Gamma/\sqrt{n}$, z = 2; при вероятности 0,997: z = 3; часто используются также вероятности 0,95 и 0,99 — при этом z равно 1,96 и 2,58 соответственно.

[180]

3. Тогда неизвестный параметр $M^Г$ с вероятностью p лежит в пределах:

$$М^в - z\frac{\sigma^Г}{\sqrt{n}} < M^Г < М^в + z\frac{\sigma^Г}{\sqrt{n}}$$

3. Точечное и интервальное оценивание

Предположим, что по выборке нужно найти не интервал, в котором находится параметр, а одно число, которое ближе всего к параметру (его мы хотим использовать для дальнейших вычислений вместо неизвестного параметра). Казалось бы, естественно предположить, что в качестве наилучшего приближения для $M^Г$ следует выбрать M^B для $(\sigma^Г)^2$ величину $(\sigma^B)^2$ и т.д. Однако это не совсем так. Но прежде, чем рассмотреть этот вопрос подробнее, определим, что мы понимаем под наилучшей оценкой (под оценкой понимается любое число, рассчитанное по выборке и характеризующее параметр).

Принято различать следующие свойства оценок. *Несмещенность* — свойство, состоящее в том, что среднее выборочного распределения оценки равно величине параметра. Это нужно понимать следующим образом: если для оценки некоторого параметра α из генеральной совокупности мы извлечем k выборок объема n, для каждой выборки рассчитаем оценку параметра a_i, и найдем среднее арифметическое этих оценок $\bar{a} = \frac{1}{k}\sum_{i=1}^{k} a_i$,

то оно будет близко к параметру α, причем при увеличении k среднее оценок $\bar{a}$ будет стремиться к α . Например, оказывается, что среднее арифметическое выборки $\bar{x} = M^B$ является несмещенной оценкой $M^Г$. Это следует из сформулированной ранее центральной предельной теоремы.

Выборочная дисперсия $(s^B)^2$ оказывается смещенной оценкой[19] параметра $(\sigma^Г)^2$, сумма $\frac{1}{k}\sum_{i=1}^{k}(s_I^B)^2$ при увеличении k стремится к числу, несколько меньшему, чем $(\sigma^Г)^2$, а именно к $\frac{n-1}{n}(\sigma^Г)^2$. Если, например, из генеральной совокупности извлекаются выборки объема 2, то оценка $(s^B)^2$ стремится к величине вдвое меньшей, чем $(\sigma^Г)^2$, если выборка

[181]

объема 3, то $(s^B)^2$ стремится к $\frac{3-1}{3}(\sigma^Г)^2 = \frac{2}{3}(\sigma^Г)^2$ и т.д.

При увеличении n это различие существенно уменьшается. Свойство оценки при увеличении объема выборки приближаться к значению оцениваемого параметра называется *состоятельностью* оценки. Таким образом, $(\sigma^B)^2$ является смещенной, но состоятельной оценкой $(\sigma^B)^2$. Отметим, что статистика $\frac{n}{n-1}(\sigma^B)^2$ дает несмещенную оценку для $(\sigma^Г)^2$.

Обозначим ее через s^2.

Третьим свойством точечных оценок является *эффективность,* мерой которой является дисперсия выборочного распределения статистики. Чем ниже дисперсия, т.е. чем меньше отличаются оценки, полученные в разных выборках, тем выше эффективность. Эти три свойства характеризует качество оценки. Можно показать, что медиана, так же как и среднее арифметическое выборки M^B, является несмещенной оценкой среднего арифметического генсовокупности $M^Г$. Медиана является также и состоятельной оценкой

[19] *Кендалл М, Дж., Стьюарт А.* Статистические выводы и связи. М., 1973, с. 18.

$M^Г$. Можно, однако, показать, что дисперсия выборочного распределения медианы приблизительно в полтора раза больше, чем дисперсия выборочного распределения, среднего арифметического, т.е. среднее арифметическое является более эффективной оценкой, чем медиана. Если зафиксировать интервал таким образом, чтобы из 100 выборок в 95-м среднее лежало внутри выделенного интервала, то окажется, что лишь в 61 выборке в этих же пределах лежит и медиана. Оказывается, что выборочное среднее арифметическое обладает наибольшей эффективностью из всех несмещенных и состоятельных оценок $M^Г$ и является, следовательно, наилучшей оценкой для $M^Г$. Аналогично происходит выбор точечных оценок для других параметров генеральной совокупности.

Вернемся теперь к более важному, как нам представляется, для социологии методу интервального оценивания. Подчеркнем еще раз, что основой интервального оценивания, а также методов проверки гипотез, изложенных в следующих параграфах, является выборочное распределение статистики. Если у читателя нет четкого представления о различии генерального распределения признака и выборочного распределения статистики, то рекомендуем ему внимательно прочесть еще раз страницы, где вводится выборочное распределение.

[182]

Упражнение 88. А. Сформулировать определения: 1) генеральное распределение признака; 2) выборочное распределение признака; 3) выборочное распределение статистики.

Б. Пусть из некоторой бесконечной генеральной совокупности извлечено 1000 выборок по одному элементу каждая, измерено значение признака *X* для каждого элемента и построено распределение. Какое из названных в п. *А* распределений мы получим?

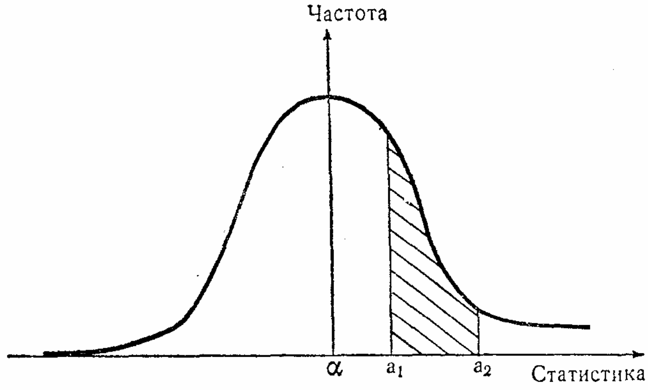

Рис. 26. Распределение статистики a для выборок объема n из генеральной совокупности с параметром α

В. Если в предыдущем случае в каждой из выборок извлекать не по одному, а по два элемента (получим X_1 и X_2) и затем построить распределение. Какое из названных распределений мы получим?

Г. В предыдущем случае в каждой выборке будем вычислять $Y == 1/2\ (X_1+ X_2)$ и построим распределение *Y*.

Какое из названных распределений мы получим?

Ответ: Б и В — выборочное распределение признака, Г —распределение статистики.

В предыдущем параграфе мы рассмотрели, как строится интервальная оценка генерального среднего. Аналогично получаются интервальные оценки и других параметров. В общем виде схему интервального оценивания можно представить следующим образом. Пусть дана генеральная совокупность с некоторым неизвестным параметром и, значение которого требуется оценить (это может быть доля признака, среднее, мера вариации, коэффициент корреляции, разность средних, разность коэффициентов корреляции и т.п.).

[183]

Для этого выбирается оценка параметра α — некоторая статистика (как правило, наилучшая, т.е. несмещенная, состоятельная и максимально эффективная оценка). Нужно найти выборочное распределение этой статистики. Для этого исследователь не извлекает выборки объема n из генеральной совокупности, не вычисляет статистику a_i, (i — номер выборки) и не строит распределение значений a_i, а определяет, каким будет распределение, теоретически. Пусть такое распределение найдено (рис. 26). Важно отметить, что статистическая теория дает также возможность определить по выборочному распределению, каково значение неизвестного параметра. Например, как нам уже известно, если построено выборочное распределение среднего арифметического, то среднее выборочного распределения (т.е. среднее средних арифметических каждой выборки) равно параметру. Это свойство выполняется во всех случаях, когда в качестве статистики выбирается несмещенная оценка параметра генеральной совокупности (по определению несмещенности).

Далее полученное распределение табулируется: для каждого значения Δ, взятого с некоторым шагом, определяют долю площади, которая лежит под кривой выборочного распределения[20] от $\alpha - \Delta$ до $\alpha + \Delta$. Эта доля, т.е. отношение этой площади ко всей площади, лежащей под кривой, равна вероятности того, что в выборке значение статистики a_i, будет больше $\alpha - \Delta$ и меньше $\alpha + \Delta$. Например, если выборочное распределение нормально, то по таблице А Приложения 3 можно определить, что от -0,1 до 0,1 лежит 0,00798% площади кривой, от -1,2 до 1,2 — 0,76986% и т.д. (в таблице приведены значения для случая, когда $\alpha = 0$, $\sigma = 1$; если же $\alpha \neq 0$, $\sigma \neq 1$, то в пределах от $\alpha - 0{,}1\sigma$ до $\alpha + 0{,}1\sigma$; лежит 0,00798% площади кривой, от $\alpha - 0{,}2\sigma$ до $\alpha + 0{,}2\sigma$; лежит 0,76986% площади кривой и т.д.).

По таблице можно также определить долю площади, лежащей между любыми двумя точками a_1 и a_2 (т.е. вероятность того, что значение статистики a_i в выборке будет лежать в интервале от a_1 до a_2). Для этого сначала определяют доли площади, лежащие от 0 до a_1 и от 0 до a_2 (площадь, лежащая между 0 и a, это половина площади, лежащей между — a и a). Тогда площадь между a_1 и a_2, определится

[184]

как разность площади, лежащей от 0 до a_2, и площади, лежащей от 0 до a_1.

Упражнение 89. Пусть выборочное распределение переменной x описывается нормальным распределением со средним 0 и дисперсией 1 (т.е. таблицей А Приложения 3). Определите, какова вероятность, что в выборке мы получим значение, лежащее между а) -2 и 2; б) -2,58 и 2,58; в) -1,4 и 1,4; г) 1 и 1,4; д) -1 и 1,4.

Вся описанная в п. 1 работа выполняется не социологом, а статистиком, для большинства стандартных случаев она уже проделана и наиболее часто встречающиеся выборочные распределения протабулированы (это нормальное распределение, распределение χ^2 (хи-квадрат), F — распределение Фишера и t — распределение Стьюдента, см. таблицы А, Б, И и Л Приложения 3). Поэтому социолог должен лишь определить, каким выборочным распределением описывается его статистика, т.е. какую из таблиц он должен выбрать (изложению этого и посвящены последующие параграфы этой главы). Таким образом социологу нужно:

1. Определить, какой статистикой следует пользоваться и найти соответствующую таблицу.

2. Задавшись некоторой доверительной вероятностью (например, 0,95), по выбранной таблице для заданной вероятности определить такое число Δ, чтобы в пределах от $\alpha - \Delta$ до $\alpha + \Delta$ лежало 95% площади кривой. Это означает, что с вероятностью 0,95 любое выборочное значение a лежит в этих пределах, т.е.

$$|\alpha - a| < \Delta$$

[20] Иногда табулируется площадь, лежащая между 0 и Δ или между Δ и ∞ - это зависит от выборочного распределения и от конкретной таблицы.

или: $a - \Delta < \alpha < a + \Delta$ (V,3,1)

3. Из генсовокупности извлекается случайная выборка и вычисляется значение статистики а. Из неравенства (V,3,1) тогда следует, что $(a - \Delta,\ a + \Delta)$ и есть искомый 95%-и доверительный интервал.

В последующих параграфах при изложении методов проверки гипотез будут рассмотрены конкретные случаи отыскания доверительных интервалов для процентов, средних, коэффициентов корреляции и т.п.

4. Проверка статистических гипотез

Предположим, что исследователь провел на некотором крупном предприятии сплошной опрос и оказалось, что средний балл удовлетворенности трудом равен α (пусть, например, $\alpha = 0,43$), а дисперсия равна $(\sigma^{Г})^2$ (пусть $(\sigma^{Г})^2 = 1,26$,

[185]

т.е. $\sigma^{Г} = 1,12$). По рекомендациям, разработанным социологом, на предприятии были проведены мероприятия, направленные на повышение удовлетворенности работников трудом. Через год для проверки эффективности мероприятий исследователь провел выборочное исследование объема n (пусть $n = 100$) и получил, что средний балл удовлетворенности в выборке равен b, причем b несколько выше, чем α. (например, $b=0,68$), а дисперсия не изменилась ($(\sigma^{В})^2 = 1,26$). Возникает вопрос, произошли ли какие-нибудь изменения

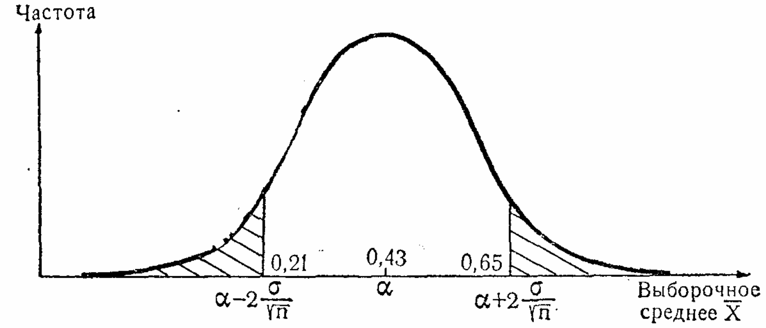

Рис. 27. Выборочное распределение: площадь, соответствующая вероятности совершить ошибку I рода (заштрихована)

на предприятии или тот факт, что $\beta > \alpha$, является случайным[21], и проведя сплошной опрос, мы никаких изменений не обнаружили бы (т.е. новое значение удовлетворенности ($\beta = \alpha$)? Другими словами, можно высказать две гипотезы о неизвестном параметре генеральной совокупности (обозначим их через H_0 и H_1):

1. H_0: $\beta = \alpha$. (так называемая, нулевая гипотеза)

2. H_1: $\beta \neq \alpha$ (альтернативная)

Какая из гипотез более обоснована? Для принятия решения в данном случае поступают следующим образом. Предположим, что гипотеза H_0 верна (т.е. ($\beta = 0,43$), а дисперсия $(\sigma^{Г})^2$ равна 1,26, как и ранее. Тогда выборочное распределение описывается нормальной кривой со средним $\beta = \alpha$ (т.е. 0,43) и дисперсией $(\sigma^{В})^2 == \frac{(\sigma^{Г})^2}{n}$ (рис. 27). Как известно, 95,4% площади нормального распределения лежит в пре-

[186]

[21] Введенные обозначения связаны с тем, что β — неизвестный параметр генеральной совокупности, а b — статистика.

делах двух среднеквадратических отклонений от среднего, а так как среднее квадратическое отклонение выборочного распределения в данном случае равно $\frac{\sigma^{\Gamma}}{\sqrt{n}}(\frac{1{,}12}{\sqrt{100}}=0{,}11)$, то, следовательно, 95,4% выборочных средних лежит в пределах от $\alpha-\frac{2\sigma}{\sqrt{n}}$ до $\alpha+\frac{2\sigma}{\sqrt{n}}$ (т.е. от 0,43 -0,22 = 0,21 до 0,43 + 0,22 = 0,65) и b = 0,68, лежит вне интервала (0,21; 0,65). Если H_o верна, то вероятность получить значение «b» вне этого интервала равна 0,046 (4,6%). Поэтому в данном случае разумно отклонить гипотезу H_o и принять гипотезу H_1 (риск, что мы совершили ошибку и что гипотеза H_0 верна составит лишь 4,6%). Такая ошибка — отклонить гипотезу H_0, когда она верна — называется *ошибкой I рода.*

Вероятность совершить ошибку I рода называется уровнем значимости (эту вероятность выражают и в процентах). Поэтому эквивалентным вышеприведенному является утверждение «гипотеза H_0 равенстве средней удовлетворенности работников предприятия до и после проведенных мероприятий отклоняется на уровне значимости 0,046». Чаще всего в социологических исследованиях задают 5- или 1%-ный уровень значимости. В нашем случае доверительный интервал для проверки на 1%-ном уровне значимости равен $\alpha-2{,}58\frac{\sigma^{\Gamma}}{\sqrt{n}};\alpha+2{,}58\frac{\sigma^{\Gamma}}{\sqrt{n}}$, т.е. (0,14; 0,72). Если полученное значение b было бы меньше 0,14 или больше 0,72, то мы с большей уверенностью могли бы утверждать, что произошли изменения, и отклонить гипотезу H_0 на 1 %-ном уровне.

Принятая при проверках гипотез терминология включает также понятия «критическая область» и «критическая точка». *Критическая область* — это те значения статистики, при которых отвергается гипотеза H_0 (в каком-то смысле это понятие дополнительное к понятию доверительного интервала). Так, для 5%-го уровня значимости критической областью при проверке гипотезы H_0 в приведенном примере являются значения, лежащие вне интервала (0,21; 0,65), т.е. значения, которые меньше 0,21 и больше 0,65 (площадь под кривой выборочного распределения, лежащая в этих пределах, на рис. 27 заштрихована). Точки, отделяющие критическую область от области принятия гипотезы (т.е. 0,21 и 0,65), называются *критическими.*

[187]

В этих терминах процесс проверки гипотез описывается следующим образом. Исследователь формулирует гипотезы H_0 (нулевую) и H_1 (альтернативную) и определяет, каким будет выборочное распределение статистики, служащей для оценки параметра, о котором сформулирована гипотеза H_0, если предположить, что она верна.

Следующий шаг — выбор уровня значимости и определение критической области для этого выборочного распреде ления

Таблица 38

Ошибки при проверке статистических гипотез

Исследователь принял решение	В действительности	
	H_0 верна (т.е. H_1 неверна)	H_1 верна (т.е. Но неверна)
Отклонить H_0 (т.е. принять H_1)	Ошибка I рода Вероятность (т.е. уровень значимости) q	Правильное решение Вероятность (т.е. мощность) $1 - p$
Отклонить H_0 (т.е. принять H_1)	Правильное решение Вероятность $1 - q$	Ошибка II рода Вероятность p

при данном уровне значимости. И последний шаг — проведение выборочного исследования и определение выборочного значения статистики: если полученное значение попадает в критическую область,— гипотеза *Но* отвергается и принимается альтернативная гипотеза H_1, в противном случае, говорят, что выборочное значение статистики попало в область принятия гипотезы и гипотеза принимается.

Важно отметить, что в принятой схеме проверки нуль-гипотеза и альтернативная неравноправны. Хотя в учебниках по статистике[22] нуль-гипотезу иногда определяют просто как любую проверяемую гипотезу, поменять ее местами с альтернативной гипотезой не всегда возможно. Если бы в приведенном нами примере в качестве нуль-гипотезы была бы принята гипотеза о неравенстве средних ($\beta \neq \alpha$), то мы столкнулись бы с серьезными трудностями.

Действительно, для проверки такой гипотезы надо пред-

[188]

положить, что она верна и определять затем выборочное распределение статистики *b* для проверки этого утверждения. Но таких распределений может быть не одно, как в случае, если ($\beta = \alpha$, а бесконечно много (ведь существует бесконечно много β не равных α)! Поэтому определить критическую область для проверки гипотезы H_1 в данном случае очень сложно. Ошибка отклонить альтернативную гипотезу H_1 при условии, что она верна, называется *ошибкой II рода,* а вероятность не допустить эту ошибку называется мощностью. Соотношение введенных понятий хорошо видно из таблицы 38.

Мы не будем рассматривать методы расчета ошибок второго рода из-за их сложности и потому, что они практически не используются социологами (по крайней мере, нам неизвестна ни одна отечественная публикация, в которой рассчитывалась бы ошибка II рода). Однако анализ различий этих видов ошибок позволяет лучше понять методы статистической проверки гипотез. Почему оценка вероятности ошибки I рода получила гораздо более широкое применение в социологических исследованиях (и, пожалуй, в исследовательской работе вообще), чем оценка вероятности ошибки II рода?

Целью исследователя является поиск различных закономерностей, связей изучаемых явлений, различий между изучаемыми группами респондентов и т.п. Используя математическую статистику, он стремится показать, что между изучаемыми переменными есть связь, что рассчитанный коэффициенты корреляции значимо отличаются от нуля, что между двумя процентами есть различия и т.д., т.е. исследователь стремится показать, что обнаруженные им различия в выборочных данных не обусловлены игрой случая, что в действительности (т.е. в генеральной совокупности) они существуют.

Путь, предлагаемый математической статистикой — это доказательство от противного: предположим, что никаких различий нет, т.е. проценты равны, средние арифметические не различаются между собой, коэффициент корреляции равен нулю и т.п. (именно такого рода гипотезы, а не любые произвольные формулируют в качестве нуль-гипотезы).

Если даже в действительности различий нет, то из-за случайных обстоятельств в выборке их можно все же получить — статистика позволяет теоретически оценить, насколько большими могут быть эти случайные различия, например, показать, что с вероятностью 0,99 при отсутствии различий в действительности, различия в выборке за счет

[189]

случайностей не могут превышать некоторого числа *k* (критическая точка).

---

[22] *Гласс Дж, Стэнли Дж.* Статистические методы в педагогике и психологии. М., 1976, с. 254; Статистические методы анализа информации в социологических исследованиях. М., 1979, с. 81.

Если оказывается, что в эмпирическом исследовании получено различие большее *k,* то это противоречит гипотезе об отсутствии различий и, следовательно, гипотеза H_1 принимается с определенной вероятностью (скажем, 0,99). При этом, естественно, есть риск, что исследователь ошибся и в действительности различий нет (вероятность этого 0,01). Это и есть ошибка I рода.

Допустить такую ошибку — все равно, что заявить, об открытии, которое оказывается фикцией. В приведенном выше примере, если бы социолог допустил ошибку I рода, то это значило бы, что он утверждал, что предложенные им мероприятия эффективны, хотя в действительности это не так. Что же касается ошибки II рода, т.е. принятие нулевой гипотезы — хотя в действительности есть различия — то она выглядит значительно менее неприятной.

К каким последствиям она приведет? К тому, что исследователь будет продолжать работу, увеличивать объем выборки, точность инструмента исследования и если различия есть, то он их в конце концов обнаружит. Принятие нулевой гипотезы означает, скорее, не строгое отсутствие различий, а то, что либо различий нет, либо они невелики. Здесь, как и в методе доказательства от противного, если мы получили противоречие сделанному допущению, то допущение неверно, но если мы путем некоторого рассуждения не получили противоречия, то это еще не значит, что допущение верно – другой ход рассуждений может привести к противоречию.

Такая ситуация типична для проверки научных теорий вообще — отсутствие фактов, противоречащих теории, еще не означает, что она верна — научный поиск может в конце концов привести к открытию нового факта, который противоречит теории, что потребует разработки новой теории, включающей и этот факт. Таким образом, принятие нулевой гипотезы означает не подтверждение ее, а неопровержение, поэтому ошибка II рода не так опасна, как ошибка I рода. Если ошибка I рода приводит к получению ложных фактов и вносит шум, помехи в научную информацию, то ошибка II рода несколько отдаляет получение нового факта и увеличивает затраты, но стимулирует исследователя к поиску новых, более совершенных и точных методов.

Вероятности ошибок I и II рода связаны обратной зависимостью; уменьшение вероятности ошибки I рода (умень-

[190]

шение уровня значимости) приводит к увеличению вероятности ошибки II рода. К счастью, с увеличением объема выборки падают вероятности ошибок I и II рода.

Таким образом, не вызывает сомнения важность оценки вероятности ошибок I рода. К сожалению, в социологической литературе часто встречаются работы, авторы которых довольно легкомысленно склонны трактовать даже незначительные различия в полученных ими эмпирических данных, не оценивая вероятности ошибок I рода. Нам представляется обязательным расчет уровней значимости — это дисциплинирует исследователя и позволяет как сквозь сито просеять эмпирические данные, оставить лишь наиболее надежные факты (напомним, что даже 5%-ный уровень значимости в среднем в одном случае из 20 дает ложный факт).

Перейдем теперь к рассмотрению конкретных, наиболее распространенных случаев статистического вывода, в частности, проверки гипотез и построения доверительных интервалов. Изложение ведется в такой последовательности:

1) формулирование нулевой и альтернативной гипотез;

2) определение выборочного распределения для проверки нулевой гипотезы[23]; 3) определение критических значений для проверки гипотез; 4) построение доверительных интервалов; 5) пример; 6) упражнение.

[23] В зависимости от вида распределения используется то или. иное обозначение для статистики: z — при нормальном распределении, t и *F* — при распределении Стьюдента и Фишера соответственно.

5. Значимость различий долей (процентов)

Эта задача, по-видимому, чаще всего встречается в социологических исследованиях: имеются две генеральные совокупности, из одной извлечена выборка объема n_1, из другой — независимая выборка объема n_2. Оказалось, что доля некоторого признака в одной выборке ν^{B}_{1}, а в другой ν^{B}_{2}. Возникает вопрос, не обусловлено ли различие ν^{B}_{1} и ν^{B}_{2} случайными факторами, т.е. различаются ли доли этого признака в генеральных совокупностях? В реальных исследованиях чаще встречается ситуация, когда извлекается одна выборка, которая затем разбивается на группы (например, по полу, по характеру труда и т.п.), и ставится задача определить, различаются ли выделенные группы по доле изучаемого признака (например, различаются ли

[191]

мужчины и женщины по доле рационализаторов). В этом случае можно считать каждую из групп выборкой из своей генеральной совокупности (например, при городском выборочном опросе мужчины, попавшие в выборку, представляют генеральную совокупность «мужское население города», а женщины — соответственно «женское население города»).

Итак, предположим, что даны две бесконечные генеральные совокупности (будем считать, что генеральная совокупность бесконечна, точнее, что мы имеем право пользоваться формулами, выведенными для бесконечной генеральной совокупности, если объем выборки *n* составляет менее 5% от объема генеральной совокупности *N,* но *N* при этом не менее 1000). Из первой генеральной совокупности извлечена выборка объема n_1, из второй — объема n_2, доля признака *X* в первой выборке ν^{B}_{1}, во второй ν^{B}_{2}, неизвестные доли признака *X* в генеральной совокупности составляют $\nu^{Г}_{1}$ и $\nu^{Г}_{2}$ соответственно.

1. Формулируем гипотезы:

1) H_0: $\nu_1^{Г} = \nu_2^{Г}$

2) H_1: $\nu_1^{Г} \neq \nu_2^{Г}$

2. Пусть $n_1 \geq 50, n_2 \geq 50, n_1\nu_1^{B} > 5, n_2\nu_2^{B} > 5, \quad n_1(1-\nu_1^{B}) > 5, \quad n_2(1-\nu_2^{B}) > 5.$ Тогда, если верна гипотеза H_0, приведенная ниже функция от ($\nu_1^{B} - \nu_2^{B}$) имеет нормальное распределение с нулевым средним и единичным среднеквадратичным отклонением:

$$z = \frac{\left|\nu_1^{B} - \nu_2^{B}\right|}{\sigma_z}, \text{(V,5,1)}$$

где $\sigma_z = \sqrt{\nu^{B}(1-\nu^{B}) \cdot (\frac{1}{n_1} + \frac{1}{n_2})},$

$$\nu^{B} = \frac{\nu_1^{B} n_1 + \nu_2^{B} n_2}{n_1 + n_2}$$

3. Задавшись некоторым уровнем значимости, по таблице А Приложения 3 определяем критические точки. Например, при 5%-м уровне значимости - z = 1,96, т.е. критические точки - 1,96 и +1,96, а на 1%-ном - 2,58 и +2,58. Область, лежащая между критическими точками, является областью принятия гипотезы, а вне этих точек — критической областью. Если, например, подставив полученные в эксперименте данные в формулу (V,5,1), мы по-

[192]

лучили, что -1,95< z <1,96, то гипотеза *Ho* принимается, а если $z \leq 1{,}96$ или $z \geq 1{,}96$, то гипотеза отвергается на 5%-м уровне значимости.

4. Доверительные интервалы для доли признака в каждой из выборок можно найти даже в случае конечной генеральной совокупности. Если, например, первая из рассмотренных генеральных совокупностей состоит из N_1 единиц, то доверительный интервал задается формулой (V,5,2)

$$\left(\nu_1^B-\frac{1}{2n_1}\right)-z\sqrt{\frac{\nu_1^B(1-\nu_1^B)}{n_1}\cdot\frac{N_1-n_1}{N_1-1}}<\nu_1^{Г}<\left(\nu_1^B-\frac{1}{2n_1}\right)+z\sqrt{\frac{\nu_1^B(1-\nu_1^B)}{n_1}\cdot\frac{N_1-n_1}{N_1-1}}$$

где z — коэффициент, который определяется по таблице А Приложения 3 (например, при 95%-ной доверительной вероятности, т.е. при 5%-ом уровне значимости, z = 1,96). Аналогично определяются доверительные интервалы для доли признака в любой генеральной совокупности[24].

Что же касается доверительного интервала разности долей, то он определяется формулой

$$\left(\nu_1^b-\nu_2^b\right)-z\sigma_z^b<\left(\nu_1^{Г}-\nu_2^{Г}\right)<\left(\nu_1^b-\nu_2^b\right)+z\sigma_z^b \text{(V,5,3)}$$

5. Рассмотрим *пример № 29.* Исследование общественного мнения об Олимпиаде-80 в Москве[25] показало, что по мере приближения к началу Олимпийских игр интерес москвичей к ним увеличивался: в двух исследованиях, проводившихся с интервалом в полгода, доля ответивших, что вопросы, связанные с подготовкой и проведением Олимпиады, их не интересуют, уменьшилась с 9% до *4°o.* Проверим на 5%-ом уровне значимости, что увеличение интереса к Олимпиаде действительно имело место. В указанной работе данные о выборке приводятся по второму исследованию *(*n_2= 919), о первом сказано лишь, что оно проводилось в двух районах Москвы. Предположим, что первая выборка включала 300 опрошенных *(*n_1= 300), ν_1^B = 0,09, ν_2^B = 0,04. Ясно, что n_1 = 300 > 50, n_2= 919 > 50, $n_1\ \nu_1^B$ = 27 > 5,

[193]

$n_2\nu_2^B=36{,}8>5, n_1(1-\nu_1^B)=300\cdot 0{,}91=273>4$ и, наконец, $n_2\left(1-\nu_1^B\right)=919\cdot 0{,}96=882>5.$

Проводим вычисления по формуле (V,5,1):

$$\nu^B=\frac{0{,}09\cdot 300+0{,}04\cdot 919}{300+919}=0{,}052$$

$$\sigma_z^B=\sqrt{0{,}052\cdot 0{,}948\left(\frac{1}{300}+\frac{1}{919}\right)}=0{,}0148$$

$$z=\frac{0{,}09-0{,}04}{0{,}0148}=3{,}38$$

Как видим, полученное значение больше, чем 1,96, поэтому нулевая гипотеза, заключающаяся в том, что за полгода никаких изменений в отношении москвичей к Олимпиаде не произошло, отвергается на 5%-ом уровне. Поскольку полученное значение больше, чем 2,58, гипотеза отвергается и на 1 %-ом уровне, наличие изменений можно считать доказанным.

Найдем теперь доверительные интервалы. Пусть доверительная вероятность 0,95 *(*z = 1,96). Тогда нижняя граница доверительного интервала доли лиц, не проявляющих интереса

[24] Приближенно 95-й и 99%-и доверительные интервалы можно определить, не проводя каких-либо расчетов, по таблице 3 Приложения 3.

[25] *Коробейников В. С., Воинова В. Д., Токаровский Г. Д.* Общественное мнение об Олимпиаде в Москве.— Социологические исследования, 1980, .№ 2, с. 153.

к Олимпиаде в первом опросе, равна (заметим, что поскольку генеральную совокупность — жители Москвы – можно считать бесконечной, выражение $\frac{N-n}{n-1}$ в формуле (V,5,2) равно1):

$$\left(0{,}09-\frac{1}{2\cdot 300}\right)-1{,}96\sqrt{\frac{0{,}09\cdot 0{,}91}{300}\cdot 1}=0{,}056$$

Верхняя граница соответственно равна

$$\left(0{,}09-\frac{1}{2\cdot 300}\right)+1{,}96\sqrt{\frac{0{,}09\cdot 0{,}91}{300}\cdot 1}=0{,}124$$

Таким образом, с вероятностью 0,95 доля лиц, не проявляющих интереса к Олимпиаде при первом опросе, лежит в пределах от 0,056 до 0,124 (или от 5,6% до 12,4%). Аналогично рассчитываем, что во втором опросе соответствующая доля лежит в пределах от 2,7% до 5,3%. Доверительные пределы для разности долей ищем по формуле (V,5,3)

(0,09 - 0,04) - 1,96 • 0,0148 < v_1^B — v_2^B < (0,09 — 0,04) + 4-1,96 • 0,0148.

[194]

Таким образом, с вероятностью 0,95 величина, на которую снизилась доля лиц, не проявляющих интерес к Олимпиаде, лежит между 2,1% и 7,9%.

6. *Упражнение 90.* В проведенном нами почтовом опросе работающего населения г. Киева были получены следующие данные (табл. 39).

Таблица 39

Семейное положение мужчин и женщин г. Киева (занятое население)

Пол	Число опрошенных	Семейное положение			Всего
		женат (замужем)	Сейчас не женат (не замужем), но ранее был (а)	Не женат (не замужем) и не был (а)	
Мужчины	1150	85,7%	4,3%	9,9%	100%
Женщины	1365	71,0 %	17,2%	11,8%	100%

Проверить на 5%-ном уровне значимости, отличаются ли доли женатых мужчин и замужних женщин и построить 95%-ные доверительные интервалы для долей женатых мужчин, замужних женщин и для разности этих долей.

6. Значимость различий средних арифметических

Даны две независимые выборки объема n_1 и n_2 из бесконечных генеральных совокупностей с неизвестными дисперсиями и неизвестными средними $M_1^Г$, $M_2^Г$. По каждой из выборок рассчитаны средние M_1^B и M_2^B и оценки[26] дисперсий s_1^2и s_2^2

1. Формулируем гипотезы

$H_0 = M_1^Г = M_2^Г$

$H_1 = M_1^Г \neq M_2^Г$

[195]

2. Рассчитываем показатели

[26] Несмещенной оценкой генеральной дисперсии является не выборочная дисперсия, а величина $s^2=\frac{n}{n-1}(\sigma^B)^2$, где $(\sigma^B)^2$ – выборочная дисперсия. Поэтому в формулах для проверки гипотез обычно используется s (впрочем, при n>100 различие s и σ^B несущественно).

$$t = \frac{\left|M_1^B - M_2^B\right|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (V,6,1)$$

$$f = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 + 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 + 1}} \quad (V,6,2)$$

Если верна гипотеза H_0, то показатель t, рассчитанный по формуле (V,6,1), имеет так называемое t-распределение Стьюдента с f степенями свободы.

3. По таблице И Приложения 3 определяем критические точки: задаемся некоторым уровнем значимости, например 5%-ным, и находим соответствующий столбец; берем ближайшее к полученному по формуле (V,6,2) значению f целое число и находим соответствующую строку — на пересечении найденной строки и столбца стоит критическое значение $t_{кр.}$ Например, для критерия значимости 5% и $f = 2$, $t_{кр} = 4{,}30$. Если полученное по формуле (V,6,1) значение t больше $t_{кр}$ гипотеза отвергается, если же $t < t_{кр}$, то гипотеза H_0 принимается.

4. Если дана выборка объема n из бесконечной генеральной совокупности, то доверительные границы с доверительной вероятностью (1—q) определяются по формуле

$$M^B - t_{n-1}\frac{s}{\sqrt{n}} < M^Г < M^B + t_{n-1}\frac{s}{\sqrt{n}}$$

где $M^Г$ — среднее арифметическое генеральной совокупности, M^B — среднее арифметическое выборки, q — уровень значимости, (1—q) — доверительная вероятность, s — несмещенная оценка дисперсии, которая рассчитывается по выборке, t_{n-1} — коэффициент, определяемый из таблицы распределения Стьюдента для столбца, соответствующего q и строки, соответствующей (n — 1) степени свободы.

5. *Пример 30.* Рассмотрим данные социально-демографического исследования молодоженов[27], подавших заявле-

[196]

ние о вступлении в брак в Киевский Дворец бракосочетаний в 1970 г.:

Число супружеских пар	Зарплата (или стипендия) жениха (руб.)	Средняя зарплата невесты (руб.)	Несмещенная оценка среднеквадратического отклонения зарплаты невесты (руб.)
132	до 50	62,8	3,4
144	50—100	81,6	2,7
461	100—150	84,9	3,3

Проверим, связана ли зарплата жениха с зарплатой невесты. Для этого определим сначала значимо ли различаются на 1 %-ном уровне зарплаты невест для двух групп женихов:

с зарплатой до 50 и от 50 до 100 руб.

Итак, $n_1 = 132$, $M^B_1 = 62{,}8$, $s^B_1 = 3{,}4$,

[27] *ЧуйкоЛ. В.* Браки и разводы. М., 1975, с. 88 (рассчитано по табл. 27).

$n_2 = 144$, $M^B{}_2 = 81{,}6$, $s^B{}_2 = 2{,}7$,

Тогда $\frac{\left(s_1^B\right)^2}{n_1} = 0{,}0875, \frac{\left(s_2^B\right)^2}{n_2} = 0{,}0506, t = \frac{18{,}8}{\sqrt{0{,}138}} = 50{,}6, \nu \approx 252$

Как видим, полученное значение t намного превышает требуемое для 1%-го уровня (2,58), т.е. связь есть.

Упражнение 91. По приведенным в примере данным проверить на 1%-ном уровне значимость различий средней зарплаты невест для двух групп женихов: с зарплатой 50- 100 руб. и 100—150 руб.

Ответ: t = 12,1; ν = 290, различие значимо.

7. Значимость различий дисперсии

Из двух бесконечных нормально распределенных генеральных совокупностей (предположение о нормальности распределений здесь существенно, если исследователь сомневается в его верности, следует использовать другие методы[28]) извлечены независимые выборки объема n_1 и n_2. Требуется определить, равны ли дисперсии генеральных совокупностей.

1. $$\begin{aligned} H_0 &: \left(\sigma_1^{\Gamma}\right)^2 = \left(\sigma_2^{\Gamma}\right)^2 \\ H_1 &: \left(\sigma_1^{\Gamma}\right)^2 \neq \left(\sigma_2^{\Gamma}\right)^2 \end{aligned}$$

[197]

2. Обозначим через $\left(\sigma_1^B\right)^2$ большую из выборочных дисперсий. Для проверки гипотезы H_0 против H_1 рассчитывается отношение выборочных дисперсий или оценок (дело в том, что $\frac{\left(\sigma_1^B\right)^2}{\left(\sigma_2^B\right)^2} = \frac{\left(s_1^2\right)}{\left(s_2^2\right)}$):

$$F = \frac{s_1^2}{s_2^2} \quad (V,7,1)$$

Jтношение дисперсий F имеет так называемое F-распределение с (n_1 — 1) и (n_2 — 1) степенями свободы (если верна гипотеза H_0).

3. Критические точки для заданного исследования уровня значимости определяются так: верхняя критическая точка F_B – по специальным таблицам (см. Приложение 3, табл. Л), а нижняя F_H из соотношения:

$$F = \frac{1}{F_B} \quad (V,7,2)$$

Гипотеза принимается, если рассчитанное по формуле (V,7,1) значение F лежит между F_H и F_B, *т.е.* $F_H < F < F_B$. Если же $F < F_H$ или $F > F_B$, то гипотеза отвергается на заданном исследователем уровне значимости.

4. Доверительный интервал с использованием найденных выше критических точек определяется по формуле:

$$\frac{s_1^2}{s_2^2} \cdot F_H < \frac{\left(\sigma_1^{\Gamma}\right)^2}{\left(\sigma_2^{\Gamma}\right)^2} < \frac{s_1^2}{s_2^2} \cdot F_B$$

[28] *Закс Л.* Статистическое оценивание. М., 1976, с. 242, 243. *Гласс Дж., Стэнли Дж.* Статистические методы в педагогике и психологии. М., 1976.

Пример 81. Пусть n_1 = 31, s_1^2 = 16, n_2 = 21, s_2^2 = 25. Отношение большей оценки к меньшей равно: $\frac{25}{16}$ = 1,56. F_B по таблице приблизительно равно 2,0; F_H = 0,5. Гипотеза H_0 принимается. $1{,}56 \cdot 0{,}5 < \frac{\left(\sigma_1^\Gamma\right)^2}{\left(\sigma_2^\Gamma\right)^2} < 1{,}56 \cdot 2$,т.е. $0{,}78\left(\sigma_1^\Gamma\right)^2 < \left(\sigma_2^\Gamma\right)^2 < 3{,}12\left(\sigma_1^\Gamma\right)^2$ - это доверительный интервал для большей дисперсии.

Упражнение 92.

Пусть n_1 = 60, s_1^2 = 10, n_2, == 140, s_2^2 == 5. Определить значимость различий на уровне 5- и 95% -и доверительный интервал для отношения генеральных дисперсий. Ответ: F = 2, F_B = 1,43, F_H= 0,70, д*оверительный* интервал

(1,40; 2,86).

[198]

8. Значимость коэффициентов корреляции *r*, ρ, τ и коэффициентов, основанных на χ^2

А. Коэффициент r

1. Требуется проверить значимость r, т.е. может ли при данном значении выборочного коэффициента r^B быть равным нулю коэффициент корреляции r^Γ для генеральной совокупности.

1. H_0: $r^\Gamma=0$

2. H_1: $r^\Gamma \neq 0$

2. Статистика *t,* рассчитываемая по формуле (V,8,1) имеет *t*-распределение с *(n*—2) степенями свободы:

$$t = \frac{r^B\sqrt{n-2}}{\sqrt{1-(r^B)^2}} \quad (V,8,1)$$

3. Критические точки[29] определяются по таблице И Приложения 3 для заданного уровня значимости *q,* при $|t| \geq t_{кр}$ гипотеза $H_0 : r^\Gamma = 0$ отклоняется на уровне значимости *q*.

4. Для построения доверительных интервалов выборочное значение r^B подвергается так называемому преобразованию Фишера:

$$z = \frac{1}{2}\ln\frac{1+r}{1-r} \quad (V8.2)$$

Дело в том, что r^B имеет нормальное распределение лишь в случае, когда r^Γ = 0, а вот z^B — при любых значениях z^Γ. Поэтому, рассчитав z^B по полученному в выборке значению r^B, строим доверительный интервал для r^Γ (*z* определяется по таблице нормального распределения А Приложения 3, например для 95% доверительного интервала *z* =1,96):

$$z^B - z\frac{1}{\sqrt{n-3}} < z^\Gamma < z^B + z\frac{1}{\sqrt{n-3}}$$

Получив нижнее и верхнее значения z, рассчитываем значения $r_{нижн}$ И $r_{верхн}$ по формуле

$$r = \frac{e^{2z}-1}{e^{2z}+1} \quad (V.8.3)$$

Преобразование (V,8,3) можно также осуществить по табл. К Приложения 3.

---

[29] Значимость *r* можно оценить также непосредственно по таблице Е Приложения 3 (без проведения каких-либо расчетов).

[199]

5. Рассмотрим следующий *пример* 32. В исследовании инженеров ленинградских проектноконструкторских организаций были получены данные, характеризующие связь удовлетворенностей профессией и работой[30]. Коэффициент корреляции, рассчитанный по оценкам 89 руководителей групп, между удовлетворенностью работой и удовлетворенностью профессией равен 0,23. Проверим значимость коэффициента на 1%-ном уровне и построим соответствующий (т.е. 99%-й) доверительный интервал.

Итак, n = 89, r^B = 0,23. По формуле (V,8,1) получаем:

$$t = \frac{0{,}23\sqrt{89-2}}{\sqrt{1-(0{,}23)^2}} \approx 2{,}2$$

По табл. И Приложения 3 находим, что для n — 2 = 87 степеней свободы 1%-я критическая точка равна приблизительно 2,64 (в таблице не приведены критические значения для 87 степеней свободы, но приведены для 60, равные 2,66 и для 120 степеней свободы — 2,62, т.е. искомое значение критической точки лежит приблизительно посредине между 2,62 и 2,66). Таким образом, хотя в цитированной книге, откуда взят этот пример, указано, что коэффициент значим на уровне 1%, он значим лишь на 5%-ном уровне (как видно из таблицы, критическое значение для уровня 5% $t_{кр}$ =1,99). Такой же результат дает использование таблицы Е Приложения 3.

Построим теперь 99%-и доверительный интервал. По формуле (V,8,2) получаем:

$$z^B = \frac{1}{2}\ln\frac{1+0{,}23}{1-0{,}23} \approx 0{,}234$$

По таблице нормального распределения находим, что для 1%-го уровня значимости z=2,58. Тогда нижняя граница для $z^Г$ равна $0{,}234 - 2{,}58\,\frac{1}{\sqrt{89-3}} = -0{,}04$ Верхняя граница равна $0{,}234 + 2{,}58\frac{1}{\sqrt{89-3}} = 0{,}513$.

Таким образом, —0,045 < $z^Г$ < 0,513. Но это доверительные границы для z. Теперь необходимо опять вернуться к коэффициентам корреляции. По формуле (V,8,3) получаем для

[200]

нижней границы:

$$\frac{e^{-2\cdot 0{,}045}-1}{e^{-2\cdot 0{,}045}+1} = -0{,}04$$

Для верхней границы аналогичные расчеты дают 0,47.

Вместо преобразований по формулам (V,8,2) и (V,8,3) приблизительно тот же результат можно получить по таблице К Приложения 3. Например, для преобразования 0,23 находим строку 0,2 и столбец 3 — на их пересечении стоит z, соответствующий r = 0,23, а именно 0,2342. А чтобы преобразовать z = 0,513, находим внутри таблицы наиболее близкое к нему число (это 0,5101). Оно стоит в строке 0,4 и столбце 7, следовательно, r = 0,47. Итак, 99% доверительный интервал равен:

— 0,04 < $r^Г$ < 0,47,

По таблице Ж Приложения 3 можно найти, не проводя расчетов, приближенно 95%-й интервал: 0,03 < $r^Г$ < 0,42.

[30] Социально-психологический портрет инженера. М., 1977, табл. 11 на с. 149.

6. *Упражнение 93.* В той же работе[31] указывается, что коэффициент корреляции, полученный для 52 главных инженеров проекта и равный 0,51, значим на уровне 1%. Проверить, так ли это, и построить доверительный интервал. Ответ: коэффициент значим *(t* = 4,19); $0,19 < r^{Г} < 0,73$.

Б. Коэффициент ранговой корреляции ρ

1. H_0: $\rho^{Г}=0$

2. H_1: $\rho^{Г} \neq 0$

2. Если H_0 верна и $n > 10$ (при $n \leq 10$ значимость ρ определяется другим способом по специальной таблице[32], то значение *t,* рассчитанное по формуле (V,8,4), имеет *t* — распределение Стьюдента с *(n* — 2) степенями свободы:

$$t = \frac{\rho^{B}}{\sqrt{\left[1-(\rho^{B})^{2}\right]\frac{1}{n-2}}} \quad (V,8,4)$$

3. Критические точки определяются по табл. И Приложения 3. Значимость ρ можно определить также непосредст-

[201]

венно по значению ρ (без расчета *t)* по таблице В Приложения 3.

4. Доверительные интервалы для коэффициента ρ не рассчитываются, так как, оказывается, получить выборочное распределение для ρ^{B} в случае, когда $\rho^{Г} \neq 0$, очень сложно[33].

Пример 33. В примере 19 был рассчитан коэффициент связи между положительными ответами на вопросы «интересная работа» и «образование соответствует работе» для 14 групп рабочих. Оказалось, что ρ^{B}=0,345. Определим на 5%-ном уровне значимость этого коэффициента.

$$t = \frac{0,345}{\sqrt{\left[1-\left(0,345\right)^{2}\right]\cdot\frac{1}{12}}} \approx 1.27$$

Поскольку это меньше критического значения 2,23 (для 10 степеней свободы, так как в таблице не приведены значения для 12), коэффициент незначим, хотя в цитированной работе он интерпретируется как значимый.

Упражнение 94. Определить, будет ли значим коэффициент, если он рассчитан для 42 групп. Ответ: Да *(t* = 2,32).

В. Коэффициент ранговой корреляции τ

Вопрос о существенности коэффициента т мы рассматривали ранее (§ 6, гл. II), там же показано, каким образом определять значимость τ при $n \leq 10$ (поскольку для проверки существенности используется S, мы сочли целесообразным рассмотреть этот вопрос сразу после введения S). Пусть $n > 10$.

1. H_0: $\tau^{Г}=0$

2. H_1: $\tau^{Г} \neq 0$

2. Если H_0 верна, величина *z* имеет нормальное распределение:

---

[31] Социально-психологический портрет инженера. М., 1977, с. 149.

[32] *Кендэл М.* Ранговые корреляции. М., 1975, с. 69, 188—191.

[33] *Кендэл М.* Ранговые корреляции. М., 1975, с. 102. См. также *Кендалл М. Дж., Стьюарт А.* Статистические выводы и связи. М., 1973, с. 637—644.

$S^* = S - 1$, если S>0

$z = \dfrac{S^*}{\sigma}$, где S*=0, если S=0; (VI,8,5)

$S^* = S + 1$, если S<0

$$\sigma = \sqrt{\frac{1}{18}n(n-1)(2n+5)},$$

[202]

если нет объединенных раногов и

$$\sigma^2 = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_r t_r(t_r-1)(2t_r+5) - \sum_s u_s(u_s-1)(2u_s+5)\right] +$$

$$+\frac{1}{9n(n-1)(n-2)}\cdot\left[\sum_r t_r(t_r-1)(t_r-2)\right]\cdot$$

$$\cdot\left[\sum_s u_s(u_s-1)(u_s-2)\right] + \frac{1}{2n(n-1)}\cdot\left[\sum_r t_r(t_r-1)\right]\cdot\left[\sum_s u_s(u_s-1)\right]$$

если есть объединенные ранги; σ, разумеется, равна корню квадратному из приведенного выражения; t_r и u_s — число объединенных рангов в r-м объединении по X и s-м объединении по Y соответственно.

3. Критические точки определяются по таблице нормального распределения, H_0 отвергается при | z| > $z_{кр}$. При отсутствии объединенных рангов значимость τ можно определить по таблице Д Приложения 3 (без расчета г).

4. Доверительные интервалы для τ не определяются из тех же соображений, что и доверительные интервалы для ρ.

5. *Пример 34.* Пусть n = 20, S = 52 (τ = 0,27). Определим значимость τ на уровне 5%. Так как S положительно, S^* равно 51.

$$z = \frac{51}{\sqrt{\frac{1}{18}\cdot 20\cdot 19\cdot 45}} \approx 1{,}65$$

Так как для 5%-ного уровня критическое значение равно 1,96 (табл А Приложения 3), гипотеза H_0 принимается, коэффициент незначим.

Упражнение 95. S = 33 (τ = 0,36), n = 14. Найти, значим ли коэффициент на уровне 5%. Ответ: незначим, z=1,75.

Г. Коэффициенты, основанные на χ^2

Как уже указывалось, существенность их проверяется с помощью χ^2. если значим χ^2, то значим и рассчитанный с его помощью коэффициент. Поэтому при расчетах, которые для таблиц $k \times l$, как правило, производятся на ЭВМ, желательно выпечатывать не только значение коэффициента,

[203]

но и значение χ^2. Если это не сделано, то χ^2 можно, разумеется, легко найти, преобразовав формулы для расчета коэффициентов. Например, преобразовав формулу для расчета коэффициента Чупрова T, получим:

$$\chi^2 = T^2\sqrt{n(k-1)(l-1)}$$

1. Нулевая гипотеза H_0 состоит в том, что $N_{ij} = N(x_i)N(y_j)\dfrac{1}{N}$ для всех *i* и *j*. Гипотеза H_1 заключается в том, что найдется хотя бы одна пара *i* и *j* такая, что $N_{ij} \neq N(x_i)N(y_j)\dfrac{1}{N}$. Критическая точка определяется для заданного исследователем уровня значимости *q* и для $(k—1)\cdot(l—1)$ степеней свободы по таблице Б Приложения 3.

Доверительные интервалы для χ^2 не вычисляются. *Пример 35.* Значение χ^2, рассчитанное для таблицы 21 (гл. II, §2), равно 92,2. Поскольку $(k—1)\cdot(l—1)=10$, χ_0^2 для 1%-го уровня значимости равно 23,21. Следовательно, χ^2 и все рассчитанные на его основе коэффициенты значимы.

Упражнение 96. Для таблицы 5×8 было получено значение, равное 45,4. Значимо ли это значение на уровне 5%? Ответ: да.

9. Значимость различий r_1 и r_2

Из двух бесконечных генеральных совокупностей извлечены выборки объема n_1 и n_2 и для некоторых признаков *X* и *Y* в каждой из выборок рассчитаны выборочные коэффициенты корреляции r_1^B и r_2^B.

$1. H_0 : r_1^Г = r_2^Г$

$2. H_1 : r_1^Г \neq r_2^Г$

2. Для проверки гипотезы H_0 применяется z-преобразование Фишера — см. формулу (V,8,2). Вычисляем r:

$$z = \frac{z_1^B - z_2^B}{\sqrt{\dfrac{1}{n_1 - 3} + \dfrac{1}{n_2 - 3}}} \quad \text{(V,9,1)}$$

Можно доказать, что *z* имеет нормальное распределение с нулевым средним и единичной дисперсией, если верна гипотеза H_0.

[204]

3. Критические точки по заданному уровню значимости *q* определяются по таблице А Приложения 3.

4. Найденные в п. 3 критические точки могут быть использованы для построения доверительных интервалов с доверительной вероятностью 1 — *q:*

$$\left(z_1^B - z_2^B\right) - z\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}} < \left(z_1^Г - z_2^Г\right) < \left(z_1^B - z_2^B\right) + z\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$$

После того, как найдены критические точки для z, они преобразуются в критические точки для *r* по формуле (V,8,3) или с помощью таблицы К Приложения 3.

5. *Пример 36.* В исследовании влияния престижа профессий среди школьников на привлекательность профессии[34] была выдвинута гипотеза, что престиж оказывает большее влияние на привлекательность профессии для школьников из семей интеллигенции, чем для школьников из семей рабочих. Для проверки этого предположения было проведено репрезентативное для г. Киева исследование школьников 10-х классов, в ходе которого для 43 профессий были получены оценки престижа и привлекательности их для школьников. Связь престижа и привлекательности характеризовалась коэффициентом корреляции этих

---

[34] *Черноволенко В. Ф., Оссовский В. Л., Паниотто В. И.* Престиж профессий и проблемы социально-профессиональной ориентации молодежи. Киев, 1979, с. 146, 147.

оценок. Оказалось, что этот коэффициент для школьников из семей интеллигентов равен 0,94, а из семей рабочих 0,82. Дают ли эти данные основание заключить, что гипотеза подтвердилась? Проверим значимость различий: $n_1=n_2=43$, $r^B{}_1 = 0{,}94$, $r^B{}_2 = 0{,}82$. Чтобы воспользоваться формулой (V,9,1), проводим сначала z-преобразование Фишера по формуле (V,8,2) или по таблице К Приложения 3: $z_1 = 1{,}738$, $z_2 == 1{,}157$. Тогда получим:

$$z = \frac{1{,}738 - 1{,}157}{\sqrt{\frac{1}{43-3} + \frac{1}{43-3}}} \approx 2{,}60.$$

Поскольку полученное значение выше, чем 2,58, можно утверждать, что различия значимы на уровне 1%, и гипотеза

[205]

исследователей, следовательно, получила эмпирическое подтверждение.

Для этого же уровня значимости (т.е. для 99% доверительной вероятности) найдем доверительный интервал:

$$(1{,}738 - 1{,}157) - 2{,}58\sqrt{\frac{2}{40}} < z_1^{\Gamma} - z_2^{\Gamma} < (1{,}738 - 1{,}157) + 2{,}58\sqrt{\frac{2}{40}}$$

или $0{,}004 < z_1^{\Gamma} - z_2^{\Gamma} < 1{,}158$

Переводя z в r по формуле (V,8,3) или по таблице К Приложения 3, получаем: $0{,}004 < z_1^{\Gamma} - z_2^{\Gamma} < 0{,}820$

6. *Упражнение 97.* В примере и упражнении § 8 А приведены результаты исследования[35], показавшего, что коэффициент корреляции между удовлетворенностями работой и профессией 89 руководителей групп равен 0,23, а этот же коэффициент, рассчитанный по оценкам 52 главных инженеров проекта, равен 0,51. Достаточны ли эти различия, чтобы утверждать, что более высокое должностное положение позволяет полнее реализовать профессиональные ожидания, значимо ли различие полученных коэффициентов корреляции? Проверить на 5%-м уровне значимости и построить для разности коэффициентов доверительный интервал.

Ответ: различие незначимо ($z = 1{,}83$), $-0{,}022 < \left| r_1^{\Gamma} - r_2^{\Gamma} \right| < 0{,}592$. Отметим связь гипотез с доверительными интервалами: если H_0 принимается, то доверительный интервал содержит 0; если H_0 отвергается — то не содержит. Как видим, в данном случае доверительный интервал содержит 0.

[206]

[35] Социально-психологический портрет инженера. М., 1977, табл. 11, с. 149