# 4.4. КЛАССИФИКАЦИЯ И СИСТЕМАТИЗАЦИЯ СТАТИСТИЧЕСКИХ ВЫБОРОЧНЫХ МЕТОДОВ В АУДИТЕ

Логиненков А.В., магистр экономики, аудитор

*ЗАО «БДО», Санкт-Петербург (филиал)*

*Перейти на Главное МЕНЮ*
*Вернуться к СОДЕРЖАНИЮ*

Проведен обзор статистических выборочных методов, применяемых в аудите. Произведена классификация и систематизация способов оценки результатов выборочного исследования в рамках аудита с увязкой к общим теоретическим положениям математической статистики.

## ВЕДЕНИЕ

Любая аудиторская компания вынуждена применять статистические методы при проверке более или менее крупного клиента. Даже средняя по размеру организация имеет несколько десятков тысяч операций за год, отраженных в бухгалтерском учете, что полностью исключает возможность сплошной проверки такой организации для аудиторской компании. Это неэффективно ни с точки зрения трудозатрат аудитора, ни с точки зрения финансовых затрат самого клиента. Поэтому единственный выход из данной ситуации это использование выборочных совокупностей.

Выборочный метод – один из классических методов статистического анализа, который относительно недавно получил новую область практического использования. Как показывает опыт крупнейших аудиторских фирм, выборка эффективно применяется в различных видах финансового контроля. В аудите нашли применение как стандартные, так и специальные приемы извлечения выборок и последующая их статистическая оценка.

Статья содержит описание, классификацию и систематизацию применяемых в аудите выборочных методов в увязке с рядом общетеоретических положений статистики.

## 1. ЦЕЛИ И МЕТОДЫ ПРОВЕДЕНИЯ ВЫБОРОЧНОГО АНАЛИЗА

Теория выборочных обследований и методики их применения разработаны весьма полно, в основном применительно к контролю качества продукции: оценивание доли брака или, наоборот, доля продукции высшего качества и т.п. Также широкое распространение выборочные методы получили в различных медико-биологических и социологических исследованиях. Проще говоря, там, где обследуемая совокупность слишком велика для того, чтобы полностью охватить ее, или там, где контроль связан с разрушением обследуемого объекта (в некоторых видах технического контроля). Вместе с тем для практических или научных целей необходимы некоторые параметры, характеризующие совокупность в целом. Необходимо обратить внимание, что во всех указанных областях объекты проверки, как правило, измеряются в натуральном выражении, а не по стоимости. Аудитор же в основном имеет дело со стоимостными характеристиками. Соответственно в аудите в зависимости от специфики изучаемых объектов и целей проверки применяются как стандартные выборочные методы, так и свои, специальные методы выборочного контроля (например, монетарный метод). До характеристики конкретных целей и методов выборочного обследования необходимо ввести несколько специальных терминов, относящихся к теории выборочного метода, с учетом того, что эти методы применяются в аудите.

Единицы наблюдения в аудите – письменные свидетельства выполненных сделок и операций, договоры, учетные регистры, счета, платежные поручения, накладные и прочие первичные документы, в которых содержаться сведения о проверяемых аудитором фактах и событиях. Далее для краткости все это – документы.

- Исследуемая (проверяемая, тестируемая) совокупность – все множество единиц наблюдения, относительно которого необходимо, основываясь на выборке, получить обобщение в виде выборочной оценки. В теории статистики такое множество обычно называют генеральная совокупность.
- Выборочная совокупность или выборка – отобранные для тестирования единицы наблюдения (документы).
- Отбор – процедура извлечения единиц наблюдения (документов) из тестируемой совокупности.
- Оценка (выборочная оценка) – размер некоторого параметра (средней суммы, доли и т.п.), полученный по данным выборки. Выборочная оценка параметра для тестируемой совокупности несколько отличается от неизвестного истинного его значения.
- Процедура оценивания (оценивание) – процесс вычисления оценки по выборочным данным.
- Объем выборки – количество единиц наблюдения, попавших в выборку.
- Доверительная вероятность (уровень надежности) – вероятность, достаточная по мнению аудитора для суждения о достоверности выборочной оценки параметра.
- Риск выборки – вероятность того, что вывод аудитора, сделанный на основании отобранной совокупности, отличается от вывода, который мог быть сделан, если к генеральной совокупности в целом были бы применены идентичные процедуры аудита.
- Точность оценки (или, наоборот, погрешность оценки) – близость оценки к истинному значению параметра. Оценка может быть весьма точной, но иметь низкую надежность и наоборот.

Пояснение остальных терминов будет дано по ходу рассмотрения соответствующих проблем.

Четкое определение единиц наблюдения и границ совокупности – важный шаг в применении выборочного метода. Обследуемая совокупность должна охватывать все единицы наблюдения, которые связаны с целью проверки и в одинаковой степени доступны для аудитора.

Когда аудитор считает, что аудиторский риск зависит от, например, географического размещения единиц наблюдения, отрасли кредитования и т.п., тестируемая совокупность может быть разбита на соответствующие подмножества (слои, группы, страты). Каждое такое подмножество рассматривается как самостоятельный объект тестирования.

По целям и условиям проведения выборочные методы можно разделить на репрезентативные (статистические) и нерепрезентативные (содержательные). Оба вида применяются в аудите. Статистическая оценка возможна только при репрезентативном методе.

Целью репрезентативного выборочного исследования является получение (оценивание) характеристик генеральной совокупности по выборочным данным. Например, оценивание по выборке общей суммы на счетах, оценивание суммарной стоимостной ошибки и т.п. Иначе говоря, полученная в ходе выборочного тестирования информация может быть распространена без больших погрешностей на всю проверяемую совокупность. В таких случаях говорят о репрезентативной (представительной) выборке. Далеко не всегда выборка является репрезентативной. Выборка будет репрезентативной только тогда, когда она получена специальными методами при соблюдении определенных условий, основным из которых является равная вероятность для каждой единицы наблюдения совокупности попасть в выборку.

В аудите применяют и нерепрезентативный отбор. Он также заключается в отборе некоторой части единиц наблюдения из всей обследуемой совокупности. Однако цели здесь иные и отбор производится по другим принципам, чем при репрезентативном отборе. Нерепрезентативный отбор обычно производится на основе субъективных суждений аудитора. Основанием для отбора могут служить:

- интуиция или чутье аудитора в отношении «подозрительности» документов или их источников, опыт прошлых проверок, дополнительная информация со стороны и т.п.;
- предположение аудитора о том, что отобранные единицы наблюдения наиболее характерны для проверяемой совокупности;
- ориентированность аудитора на проверке наиболее крупных по стоимости документов (метод основного массива).

Не исключены случаи, когда отбор производится произвольно, без соблюдения требования равной вероятности отбора для каждой единицы совокупности.

Сформировав тем или иным путем выборочную совокупность, аудитор при нерепрезентативном отборе использует полученные данные для непосредственного суждения о соответствующих процессах или характеристиках только в рамках выборки. Он не имеет основания для их распространения на исходную генеральную совокупность.

Часто это выглядит следующим образом. Аудитор из всей совокупности счетов отбирает, например, счета, поступившие только от фирмы «А». Обнаруживает, что определенный процент документов в выборке оказались с ошибками. Полученный результат относится только к выборочной совокупности (в данном случае к счетам фирмы «А»). Как видим, речь не идет об оценивании данных генеральной совокупности. Такой отбор называют направленным. Он решает определенные задачи, поставленные аудитором, однако как уже было сказано, он ограничен по своим возможностям, так как здесь нет оценки параметров обследуемой совокупности в целом.

В связи со сказанным уместно отметить некоторую асимметричность получаемых результатов при нерепрезентативном отборе. Действительно, если выборка обнаруживает отрицательные факты, то для аудитора это уже данные для заключения. Однако если направленный отбор, да еще при небольшом объеме выборки, не обнаружил отрицательные факты (например, ошибки или приписки в тестируемых документах), то такой результат мало что дает для аудитора.

В практике отечественного аудита в основном применяют направленный отбор, который, как было показано выше, является нерепрезентативным. В какой-то мере это можно объяснить тем, что такой отбор не имеет статистико-теоретического обоснования и поэтому очень прост при реализации. Не менее важным является и то, что применяемые здесь понятия не выходят за рамки бухгалтерского учета. В связи с этим он не вызывает вопросов у бухгалтера и им можно пользоваться в любой ситуации. Однако за простоту, как известно, надо платить. Платой является получение крайне ограниченной информации и, главное, невозможность ее распространения на совокупность. Иное дело, репрезентативная выборка, которая дает аудитору обобщенные данные, относящиеся к тестируемой совокупности. Вместе с тем для осуществления подобного рода выборок аудитору надо быть знакомым с рядом статистических понятий и методов анализа, которые никак не связаны с бухгалтерией. Такие понятия, как «доверительная вероятность», «верхний предел точности» и т.п., трудно увязать с бухучетом и соответствующим менталитетом. Последнее обстоятельство может вызвать естественное внутреннее сопротивление аудитора.

В связи со всем сказанным, в том числе и в введении, далее в этой работе будут рассматриваться исключительно репрезентативные выборки и соответственно статистические методы оценки.

Итак, выборочный метод в аудите кратко сводится:

- к извлечению из тестируемой или генеральной совокупности некоторой ее части (выборки);
- проверке по заданной программе единиц наблюдения (документов), попавших в выборку;
- обобщению результатов проверки и распространению выборочных результатов на совокупность.

Специфика исследуемых совокупностей и конкретных целей аудита определяет необходимость применения различных методов извлечения выборок и их последующего анализа. В теории и практике аудита обычно выделяют следующие виды выборочных исследований, различающихся по целям и соответственно по способам формирования выборок и методикам анализа:

- атрибутивные выборки при проверках на соответствие или оценивание риска выборки через отклонение [8];
- оценивание средней и суммы стоимостных показателей совокупности;
- оценивание размера средней абсолютной ошибки в генеральной совокупности или оценивание риска выборки через абсолютную ошибку;
- оценивание размера средней относительной ошибки в генеральной совокупности или оценивание риска выборки через относительную ошибку.

Для проверок на соответствие применяют атрибутивные выборки. Этот метод необходим в ситуациях, когда аудитора интересует частота наступления некоторых случайных событий. Речь идет о фиксации в выборке наличия (или отсутствия) некоторого свойства, атрибута. Примерами могут служить:

- проверка наличия на документах разрешительной подписи;
- полноты комплекта документов по сделке;
- выявление ошибок в оформлении и т.д.

Таким образом, речь идет о случаях, когда имеется два исхода – есть событие или его нет (да, нет).

В свою очередь при проверках по существу внимание аудитора нацелено на содержательную сторону документа – его стоимостные характеристики. Здесь применяются следующие виды выборочного контроля.

### *Оценивание средней и суммы стоимостных показателей исследуемой совокупности*

Необходимость в оценивании средней в основном возникает в тех случаях, когда общий размер исследуемой совокупности в денежном измерении неизвестен, поскольку стоимостные характеристики единиц наблюдения не регистрировались или они есть, но имеются сомнения в истинности размера этой суммы, представленной аудитору. Аудитору требуется оценка этой суммы с заданной надежностью и приемлемой погрешностью.

Для решения данных задач применяют классический выборочный метод и некоторые специальные методы, разработанные в аудите для проверки стоимостных показателей.

### *Оценивание абсолютных и относительных ошибок в стоимостных показателях документов*

Необходимые оценки получают на основе выборочных данных с помощью специальных методов. Также возможно оценивание риска выборки на основе этих показателей.

В последнее время в зарубежных публикациях по аудиту упоминаются и оценивание неизвестных параметров совокупности на основе регрессий [7, 5] и байесовских методов [18]. Однако сомнительно, что в российской практике в ближайшее время данные методы привлекут внимание отечественных аудиторов. В настоящей работе они не рассматриваются.

В связи с применением в аудите методов математической статистики возникла одна проблема методологического порядка. Дело в том, что аудитор обычно «вырастает» из бухгалтерского персонала, который, как известно, не работает с вероятностями и случайными событиями. В статистическом же анализе, наоборот, вероятности и интервалы значений исследуемых параметров – обычные характеристики. В связи со сказанным при применении в аудите методов математической статистики приходится совмещать понятия из обеих областей знания.

В зарубежной аудиторской практике ориентируются на различные руководства по аудиту, разработанные в крупных аудиторских компаниях. В этих руководствах, как правило, уделяется определенное внимание конкретным методикам выборочного анализа. Характерной чертой таких руководств, по крайней мере тех, с которыми удалось ознакомиться автору настоящей работы [4, 13], является почти полное отсутствие теоретических обоснований предлагаемых методов. Обычно в руководствах указываются правила, по которым аудитор может получить необходимую информацию по принципу «черного ящика»: есть данные на входе и есть данные на выходе, а что и как преобразовывается и получается – это остается в «черном ящике». Как правило, просто даются таблицы, на пересечении строк и столбцов которых аудитор находит нужный ему параметр, без объяснения какие методики и принципы положены в основу построения таких таблиц. В данной же работе, как уже было сказано в введении, рассматриваются методы, имеющие четкую увязку с положениями математической статистики и теории выборок.

## 2. ДИСКРЕТНЫЕ И НЕПРЕРЫВНЫЕ РАСПРЕДЕЛЕНИЯ В АУДИТЕ

Аудитор на основе случайной выборки получает в свое распоряжение некоторую совокупность единиц наблюдения (документов). Каждый документ, оказавшийся в выборке, содержит сведения, ради которых, собственно, и осуществляется выборка. Аудитор, таким образом, располагает необходимой ему количественной информацией или данными о наличии или отсутствии тех или иных свойств проверяемых документов (атрибутивные выборки). Очевидно, что выборочные данные, которые попадают в руки аудитора, являются случайными величинами в том смысле, что они, в зависимости от сложившихся обстоятельств, могут принимать те или иные значения в пределах определенной шкалы. Например, суммы на счетах у наугад отобранных клиентов банка, суммы приходных ордеров, накладных и на других финансовых документах в общем случае заранее неизвестны и являются случайными величинами. В теории вероятностей случайная величина обычно определяется как переменная, принимающая в результате испытаний то или иное числовое значение в зависимости от случайного исхода испытаний.

Случайные величины делятся на два класса – непрерывные и дискретные. Стоимостные данные, с которыми в большинстве случаев имеет дело аудитор, являются практически непрерывными. В самом деле, суммы остатков на текущих счетах, размеры оплаченных чеков и т.д. могут быть любыми положительными величинами от нуля и выше (с точностью до копейки, рубля или другой денежной единицы). Случайные величины, которые имеют целочисленные значения (их можно последовательно пронумеровать), являются дискретными. Например, количество документов с нарушениями при оформлении, обнаруженными аудитором в выборке, может быть ноль, один, два и т.д.

Полученные по выборочным данным параметры (средние, доли, дисперсии) также являются случайными величинами. Например, если из одной и той же совокупности последовательно отобрать несколько выборок, то полученные выборочные средние будут разными – они варьируют по своим размерам. Их можно рассматривать как случайные величины, которые более или менее отличаются от неизвестного среднего исследуемой совокупности.

Как известно, для описания случайной величины необходимо не только указать на множество ее возможных значений, но и показать вероятности для каждого из этих значений. В статистике такое полное описание случайной величины называют теоретическим распределением.

Методы математической статистики, используемые в ряде экономических исследований, в том числе в выборочном анализе, применяют следующие виды теоретических распределений:

- биномиальное и гипергеометрическое распределение;
- распределение Пуассона;
- нормальное распределение;
- распределение Стьюдента.

Первые три из перечисленных распределений относятся к группе дискретных, остальные – к непрерывным распределениям.

Нормальное распределение играет чрезвычайно важную роль в теории и практике выборочных обследований. Некоторые сведения о нем будут приведены в следующих разделах, а сейчас немного о дискретных распределениях. Аудитор имеет дело не только со стоимостными характеристиками. Ему часто приходится оценивать по выборке качество предоставленного материала – частоту ошибок, неправильно оформленных документов и т.д. Назовем такие ошибки отклонениями, а выборки, применяемые для оценивания количества отклонений – атрибутивными.

Очевидно, что число отклонений в выборке является случайной величиной, случайной переменной. В одной выборке может оказаться два отклонения, в другой выборке из этой же совокупности – три и т.д.

Пусть случайная величина $X$ принимает дискретные значения $k$ $(k=0, 1, 2, \dots, n)$ с вероятностями:

$$P(X=k)=p_x(k).$$

По определению $0 \le p_x(k) \le 1$. Сумма вероятностей от $p_0$ до $p_m$ характеризует вероятность того, что случайная дискретная переменная не превысит некоторое заданное значение ***m***:

$$F_x = P(X \le m) = \sum_{k=0}^{m} p_x(k).$$

Функцию $F_x$ называют функцией распределения дискретной случайной величины или интегральной вероятностью.

В выборочных исследованиях в аудите для описания дискретных величин (числа отклонений в выборках) применяют два основных вида дискретных распределений: биномиальное (Deloitte & Touche) и Пуассона (Arthur Anderson & Co). В некоторых условиях более обосновано применение гипергеометрического распределения.

Теперь перейдем непосредственно к описанию каждого из упомянутых распределений и методам в аудите, основанных на этих распределениях.

# 3. МЕТОДЫ, ОСНОВАННЫЕ НА БИНОМИАЛЬНОМ РАСПРЕДЕЛЕНИИ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

## 3.1. Общая информация о биномиальном распределении

Биномиальное распределение случайной величины является одним из основополагающих в аудите, так как, по сути, большинство остальных распределений, рассматриваемых в этой работе, являются его частным случаем и выводятся из него.

Допустим, на основе прошлого опыта известно или ожидается, что вероятность случайного события (в аудиторской проверке – отклонения) равна ***p***. Вероятность противоположного события составит $(1-p)=q$. Объем выборки равен ***n***. С вероятностью $p_x(k)$ в выборке оказывается ***k*** таких документов. Указанные вероятности определяются по формуле Бернулли.

Схема Бернулли является самой распространенной вероятностной схемой повторения независимых испытаний в неизменных условиях и описывается как

$$Y_i = \begin{cases} 1, & p \\ 0, & q = 1 - p \end{cases}, \quad i = 1, \dots, n,$$

где $Y_i$ – случайная величина, которая с вероятностью ***p*** принимает значение один в случае успеха (аудитором найдено отклонение) и с противоположной вероятностью ***q*** принимает ноль в случае неудачи (аудитором не найдено отклонение).

Если мы построим случайную величину ***X***, равную

$$X = \sum_{i=1}^{n} y_i,$$

то эта величина будет иметь биномиальный закон с ***n*** степенями свободы и вероятностью успеха ***p***. ***X*** представляет собой число единиц (успехов, найденных отклонений) в последовательности $Y_1, Y_2, \dots, Y_n$.

При выводе формулы Бернулли, которая описывает биномиальное распределение, исходят из следующей модели. Пусть для начала производятся два независимых испытания (отбор только двух документов).

В этом случае получим следующие исходы:

- в обоих документах нет отклонения, $k = 0$;
- в одном документе есть отклонение, в другом нет, $k = 1$;
- в обоих документах есть отклонения, $k = 2$.

Вероятности перечисленных исходов $p_k$ в этом случае определяются следующим образом:

$$p_x(0) = q^2;\ p_x(1) = pq;\ p_x(2) = p^2.$$

В свою очередь для трех документов получим следующие вероятности:

$$p_x(0) = q^3;\ p_x(1) = pq^2;\ p_x(2) = p^2q;\ p_x(3) = p^3.$$

При обобщении полученного результата на ***n*** независимых испытаний получим формулу Бернулли, определяющую вероятность появлении ***k*** случайных событий (в аудите отклонений) из выборки, объемом ***n*** [6, 15]:

$$p_x(k) = P(X = k) = C_n^k p^k q^{n-k},\ k = 0, 1, 2, \dots, n, \quad (3.1)$$

где

***n*** – объем выборки;

***p*** – вероятность единичного отклонения.

Число сочетаний определяется как:

$$C_n^k = \frac{n!}{k!(n-k)!},$$

где

***n!*** – факториал ***n*** (вспомним: $0! = 1,\ 1! = 1,\ 2! = 1*2,\ 3! = 1*2*3$ и т.д.).

При больших ***k*** и ***n*** вычисления факториалов выполняются приближенно по формуле Стирлинга (1730 г.):

$$n! \approx n^n e^{-n}\sqrt{2\pi n}.$$

При ***n = 10*** относительная погрешность вычислений не превышает 0,83%.

Вместо (3.1) можно воспользоваться рекуррентной формулой:

$$p_x(k+1) = \frac{(n-k)p}{(k+1)q} p_x(k).$$

Распределение случайной величины, задаваемое формулой (3.1), называется биномиальным. Свое название она получила, от того, что значения $p_x(k)$ являются членами в разложении $(p+q)^n$ по формуле бинома Ньютона:

$$(p+q)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k}.$$

Поскольку $p + q = 1$, то

$$\sum_{k=0}^{n} C_n^k p^k q^{n-k} = 1.$$

Эта формула просто иллюстрирует аксиому теории вероятностей, согласно которой вероятность достоверного события равна единице.

При применении атрибутивных выборок необходимы только значения функции распределения. Они определяются путем суммирования полученных вероятностей:

$$F_x = P(X \le m) = \sum_{k=0}^{m} p_x(k) = \sum_{k=0}^{m} C_n^k p^k q^{n-k}. \quad (3.2)$$

Данная функция показывает с какой вероятностью количество отклонений ***X*** в выборке объемом ***n*** не превысит некоторого заданного значения ***m***. Математическое ожидание (среднее значение переменной ***X***) определяется весьма просто: $\mu = np$.

В табл. 3.1 приведены значения вероятностей и функции биномиального распределения для условий: $p = 4\%$, $n = 40$. Соответствующий график показан на рис. 3.1.

**Таблица 3.1**

**БИНОМИАЛЬНОЕ РАСПРЕДЕЛЕНИЕ ($p = 0,04$, $n = 40$)**

| $k$ | $p_x(k)$ | $F_x$ |
|---|---|---|
| 0 | 0,19537 | 0,19537 |
| 1 | 0,32561 | 0,52098 |
| 2 | 0,26456 | 0,78553 |
| 3 | 0,13963 | 0,92516 |
| 4 | 0,05381 | 0,97898 |
| 5 | 0,01614 | 0,99512 |
| 6 | 0,00392 | 0,99905 |
| 7 | 0,00079 | 0,99984 |
| 8 | 0,00014 | 0,99998 |
| 9 | 0,00002 | 1,00000 |

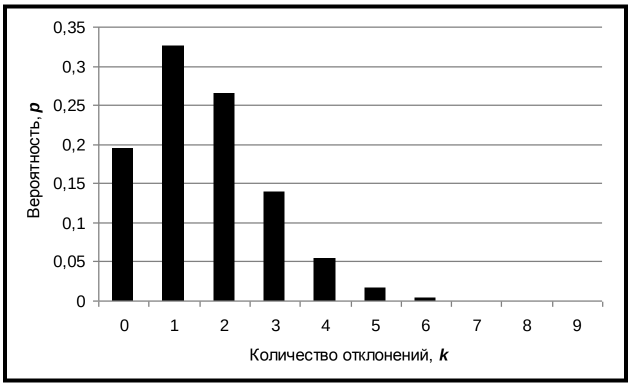


**Рис. 3.1. Функция вероятностей биномиального распределения при $p = 0,04$ и $n = 40$**

Данную таблицу и график можно интерпретировать следующим образом. В атрибутивной выборке из 40 документов и ожидаемой вероятностью одного отклонения в совокупности равной 4%, вероятность найти, например, ровно два отклонения составляет 26,5% ($p_x(2) = 0,26456$). Наиболее вероятное количество отклонений в рассматриваемой выборке равно одному ($p_x(1) = 0,32561$, т.е. аудитор с вероятностью 32,6% найдет ровно одно отклонение в выборке), что соответствует самому высокому столбику в графике. Действительно, математическое ожидание составляет 1,6 ($np = 40 \times 0,04$). Вероятность же того, что количество отклонений в выборке не превысит, например, четыре отклонения составляет 97,9% ($F_x = 0,97898$). Также можно оценить значение функции для конкретного интервала: вероятность того, что будет найдено от одного до трех отклонений составляет 73% ($\sum_{k=1}^{3} p_x(k) = 0,7298$).

Очевидно, что требование независимости испытаний при атрибутивном отборе на практике не соблюдается. Для соблюдения этого условия, строго говоря, пришлось бы возвращать каждый протестированный документ в исследуемую совокупность и заново осуществлять отбор. Причем это надо было бы сделать исключительно для того, чтобы не изменилась вероятность $p$, что лишено практического смысла. Несоблюдение требования независимости при безвозвратном отборе компенсируется большим объемом проверяемой совокупности, при котором можно считать вероятность p примерно постоянной.

Строго говоря, в расчетах значений $p_x(k)$ и $F_x$ по приведенным формулам нет крайней необходимости, так как значения функций можно найти в специальных таблицах. Хотя, как правило, соответствующие таблицы разработаны для небольших величин $n$. Эти же величины можно получить с помощью электронной таблицы Excel.

Теперь разберем непосредственно методы, основанные на биномиальном распределении, в контексте выборочного исследования в аудите.

## 3.2. Процедуры на соответствие, основанные на биномиальном распределении (атрибутивные выборки)

Один из видов аудита заключается в проверке совокупности документов с целью определения их соответствия определенным требованиям или свойствам. Например, требованию наличия всех реквизитов, печатей и подписей, требованию отсутствия арифметических ошибок. В одном из предыдущих разделов перечисленные и некоторые другие свойства были названы атрибутами. Атрибутом также может быть обязательный набор сопроводительных документов и т.п. Оговоренные перед проведением тестирования атрибуты должны присутствовать в каждом проверяемом носителе информации. Атрибуты или есть, или их нет. В аудите отсутствие атрибута называют отклонением или ошибкой. Так, если наличие разрешительной подписи рассматривается как атрибут, то ее отсутствие – как отклонение. Подобного рода тестирование ранее уже было названо «проверкой на соответствие».

Атрибутивное тестирование заключается в определении (оценивании) доли документов в совокупности с наличием отклонений. Возможна и обратная постановка задачи – определение доли документов с отсутствием соответствующего атрибута.

При большом массиве документов сплошные проверки на наличие атрибута (или его отсутствие) – весьма трудоемкие процедуры. Применение выборок в этих случаях является весьма продуктивным и во многих случаях целесообразным. Атрибутивная выборка позволяет оценить с определенной степенью надежности максимальное количество документов с наличием атрибута в проверяемой совокупности. В статистической теории подобного рода выборки обычно называют выборками для измерения качественных признаков. По своему назначению атрибутивные выборки напоминают выборки, применяемые при контроле качества продукции. Целью последних является оценка качества выпущенной партии изделий по выборочным данным [16].

Полученные в ходе атрибутивной выборки результаты говорят только о масштабах изучаемого явления – наличия атрибутов (например, при проверке обнаружено, что в 0,5% документов содержаться арифметические ошибки или 0,1% документов некомплектны и т.п.).

К атрибутивным выборкам обычно прибегают во внутреннем и внешнем аудите при тестировании документов, сопровождающих массовые хозяйственные или деловые операции:

- счета-фактуры;
- накладные;
- платежные поручения и т.д.

Необходимость в таком тестировании возникает как при проверке качества внутреннего контроля, так и по специальному запросу, например, клиента банка, наконец, при наличии сомнений у аудитора в качестве определенной группы документов, поступивших от ряда организаций.

При планировании атрибутивной выборки необходимо:

- сформулировать цель тестирования;
- определить единицу наблюдения;
- установить рамки исследуемой совокупности, причем каждая единица наблюдения (документ совокупности) должна быть одинаково доступна для отбора;
- четко определить, какое свойство является тестируемым атрибутом. Атрибут может быть единичным или комплексным (документ содержит набор реквизитов) и случай, когда отсутствует даже одно свойство или реквизит, документ рассматривается как отклонение;
- принять уровень доверительной вероятности;
- определить приемлемый верхний уровень точности оценки количества отклонений в совокупности;
- определить размер (объем) выборки;
- извлечь выборку из совокупности. Обычно используют схемы случайного (в том числе механического) или монетарного отбора документов;
- тестировать единицы наблюдения, оказавшиеся в выборке;
- определить по выборке верхний уровень точности оценки количества отклонений;
- принять решение о качестве тестируемой системы.

Возможна и другая постановка задачи – тестирование системы через риск выборки. В этом случае критерием для определения качества тестируемой системы является не превышение максимально допустимой нормы отклонений в совокупности при заданной вероятности, а превышение максимально приемлемого риска выборки – риска того, что ошибка в совокупности превысит максимально допустимый размер. И оценка риска и расчет верхнего предела ошибки осуществляются с использованием одних и тех же базовых формул статистики.

Обсудим некоторые из перечисленных этапов применения атрибутивных выборок. Предварительно вспомним, что количество документов с наличием атрибута (отклонения) в выборке является случайной величиной. Количество отклонений в выборке ***k*** может составлять ноль, одно, два и т.д. Каждое из значений ***k*** может быть сопоставлено с вероятностью соответствующего числа отклонений. Если задана величина ожидаемой вероятности отклонения в совокупности ***p***, то, приняв правило определения таких вероятностей, получим соответствующее дискретное распределение случайных величин. Такое распределение дает возможность определить вероятность каждого из возможного числа отклонений.

Как уже отмечалось ранее, атрибутивные выборки на практике применяют при решении ряда вариантов тестирования. В каждом из них основным является оценивание неизвестного количества отклонений в совокупности. Получив такую оценку, аудитор должен вынести суждение о качестве работы проверяемой системы.

Начнем с одной из наиболее распространенных постановок задачи, согласно которой аудитор по выборке определяет максимально возможное количество отклонений в совокупности. Эти величины в специальной литературе называют расчетный верхний предел точности. Обозначим этот предел как ***U***.

Обсудим теперь смысл и общую методику определения ***U***. Пусть для описания закона распределения случайных величин ***k*** (отклонений) принято биномиальное распределение. Каждому значению ***k*** в этом распределении соответствует определенная величина вероятности – см. выражение (3.1). Вспомним, что биномиальное распределение зависит от двух параметров ***p*** и ***n***.

Ожидаемая вероятность отклонения р устанавливается в виде доли документов в совокупности с наличием отклонения. Первоначальное определение параметра ***p*** может быть весьма приближенным – на основе прошлого опыта или специального выборочного обследования, даже по интуиции. На основе выборочных данных получим следующую оценку:

$$p = \frac{m}{n} 100\ \%\ ,$$

где ***m*** – количество отклонений, обнаруженных в выборке, ед.;

***n*** – объем выборки, ед.

Что касается объема выборки ***n***, то этот параметр, как известно, играет важную роль в тестировании и анализе полученных результатов. Методика определения оптимального объема выборки не рассматривается в рамках настоящей работы, так как при проведении сравнительного анализа данный параметр рассматривается как исходное условие, в зависимости от которого рекомендуется использование того или иного выборочного метода. Здесь же ограничимся одним общим замечанием – если в ходе анализа выборочных данных обнаруживается недостаточность первоначального размера выборки, то следует осуществить дополнительный отбор единиц наблюдения.

Итак, на основе выборочной оценки вероятности случайного события (отклонения) р и принятого объема выборки ***n*** определяются вероятности появления k отклонений. Сумма вероятностей от $p_0$ до $p_m$ биномиального распределения (функция распределения) составляет величину, описанную в предыдущем подразделе: см. выражение (3.2).

Функция характеризует вероятность того, что случайная дискретная переменная ***k*** не превысит некоторое заданное значение ***m***. Чем больше ***m***, тем больше слагаемых в формуле (3.2) и, следовательно, выше значение функции распределения.

Приведенные выше сведения о распределении величины ***k*** необходимы для пояснения смысла расчетного верхнего предела точности, названного выше параметром ***U***. Для этого несколько изменим постановку задачи. Пусть значение функции распределения не вычисляется, а задается аудитором. В этом случае оно играет роль доверительной вероятности, или уровня надежности. Если ***m*** в выражении (3.2) является приемлемой для аудитора величиной, то соответствующая величина вероятности $F_x$ считается аудитором достаточной для обоснования своего заключения. Она измеряется в долях единицы, а в заключении аудитора обычно указывается в процентах. Например, «с вероятностью 90% можно утверждать, что количество отклонений в совокупности не превысит…». Если доверительная вероятность в долях единицы составляет $\alpha$, то риск получить недостоверное заключение по выборке (риск выборки) равен $1 - \alpha$.

Компания «Deloitte & Touche» рекомендует устанавливать доверительную вероятность на уровне от 90% до 95% в зависимости от важности проверяемых сведений

[20]. В других зарубежных руководствах предлагается более широкий диапазон значений – от 80% до 99% [19].

Задавшись конкретным уровнем доверительной вероятности, аудитор может определить значение максимального числа отклонений $m$. Нетрудно понять, что величина $U$ есть ничто иное, как параметр $m$, соответствующий принятой доверительной вероятности $F_x$. В практических расчетах величину $U$ удобнее измерять в процентах ($U = \frac{m}{n} 100\%$).

Заметим, что на практике нет необходимости определять $U$, исходя из формулы (3.2). Для этого можно воспользоваться соответствующими таблицами.

Пусть в выборке, объем которой 200 единиц, обнаружено 12 отклонений. Соответственно

$$p = \frac{12}{200} 100\% = 6\% .$$

Допустим, аудитор установил доверительную вероятность $F_x$ на уровне 95%. Табличное значение $U$ для этого уровня вероятности и перечисленных выше условий примерно равно 8%. Таким образом, с вероятностью 95% он может утверждать, что наибольшее количество отклонений в совокупности не должно превышать 8%. Иначе говоря, вероятность того, что максимальное количество отклонений превысит 8% (риск выборки), составляет $100 - 95 = 5\%$.

Подсчитанный по выборочным данным параметр $U$ сопоставляют с «нормативом», который называют приемлемым (допустимым) уровнем отклонений, или планируемой точностью. Это наибольшее приемлемое для аудитора количество отклонений в совокупности. Иначе говоря, это то количество отклонений, с которым аудитор готов примириться при заданном уровне доверительной вероятности. Обозначим приемлемый уровень отклонений в совокупности как $T$.

Очевидно, что выбор параметра $T$ в известной степени субъективен. Естественные относительные пределы его – от 2% до 10-12%, так как считается, что вряд ли есть смысл применять выборочный контроль в том случае, когда в совокупности ожидается более 10-12% отклонений – система явно работает с большим «сбоем». С другой стороны, если предполагается, что ожидаемый уровень отклонений менее 2%, то для проверки выполнения этого условия потребуется очень большой объем выборки. Обычно этот параметр $T$ устанавливается на уровне от 3% до 7%.

Итак, полученную выборочную оценку максимального количества отклонений в совокупности $U$ сравнивают с приемлемым значением количества отклонений $T$. Если рассматриваемые параметры соотносятся как $U \le T$, то у аудитора есть основание считать, что проверяемая система контроля на соответствие функционирует нормально.

Ситуация, когда $U > T$, может служить указанием на то, что фактическая вероятность отклонения в проверяемой совокупности заметно превышает ожидаемую аудитором вероятность $p$. В этом случае аудитор может в определенных границах изменить принятые им первоначальные условия для расчета параметра $U$. Конкретно это сводится к снижению уровня доверительной вероятности и / или увеличению объема выборки. Если указанные поправки условий не достигают цели и параметр $U$ остается больше $T$, то аудитору остается признать тестируемую систему контроля плохо функционирующей, со всеми вытекающими отсюда последствиями.

Как уже было отмечено ранее, задача аудитора при применении атрибутивной выборки обычно заключается в оценивании величины верхнего предела точности $U$ для совокупности. Задавшись уровнем доверительной вероятности $F_x$ и зная объем выборки $n$ и ожидаемую вероятность отклонения $p$, определяют искомую величину $U$. Для этого используют специальные таблицы.

Показатели таких таблиц получают на основе таблиц функции биномиального распределения. Для каждого уровня доверительной вероятности разрабатывается своя таблица. Например, для $n = 100$ и $p = 0{,}05$ в таблице функции биномиального распределения находим наиболее близкие к 0,85 значения интегральной вероятности 0,76601 и 0,87024. Этим вероятностям соответствуют $m = 6\%$ и $m = 7\%$. После интерполяции получим $U = 6{,}8\%$.

Если известен размер совокупности $N$, то в табличное значение $U$ можно внести небольшое уточнение, которое в статистике называют поправкой на конечность совокупности. С учетом этой поправки верхний предел точности (максимальное возможное количество отклонений) для совокупности оценивается следующим образом:

$$U' = p + (U - p)\sqrt{1 - \frac{n}{N}} . \tag{3.3}$$

Данная поправка компенсирует неточность, вызванную пренебрежением повторного отбора, который подразумевается при биномиальном распределении, тогда как в аудите почти всегда применяется бесповторный отбор.

Рассмотрим конкретный пример. Пусть выборочная оценка ожидаемой вероятности отклонений в совокупности равна 7%, объем выборки 150 ед., доверительная вероятность принята на уровне 85%. Для $p = 7\%$, $n = 150$ и $F_x = 85\%$ находим: $U = 9\%$. Таким образом, в выборке указанного объема максимальная доля документов с отклонениями при принятом уровне доверительной вероятности не должна превышать 9%. Иными словами, если в выборке обнаружено 7% отклонений, то предельная ошибка этого параметра не превышает 2% при принятом уровне доверительной вероятности:

$$U - p = 9 - 7 = 2 .$$

Немного изменим условия. Пусть теперь известен размер совокупности: $N = 750$. Тогда с учетом поправки на конечность совокупности, используя выражение (3.3) получим:

$$U' = 7 + (9 - 7)\sqrt{1 - \frac{150}{750}} = 8{,}8\% .$$

Повышение уровня доверительной вероятности увеличивает верхний предел точности. Так, для $F_x = 90\%$ и сохранения остальных условий предыдущего примера находим $U = 9{,}4\%$. В свою очередь увеличение размера выборки снижает верхний предел точности при всех прочих равных условиях. Для варианта $p = 7\%$, $n = 200$ и $F_x = 90\%$ находим $U = 9{,}1\%$.

Теперь рассмотрим еще одну популярную постановку задачи, согласно которой аудитор оценивает риск выборки, то есть определяет вероятность того, что доля отклонений в совокупности превысит максимально допустимую норму. Приемлемый уровень отклонений в совокупности ранее был обозначен как ***T***.

Этот параметр определяется следующим образом:

$$T = \frac{S}{N},$$

где ***S*** – максимально допустимое количество отклонений в совокупности, ед.

Также исходя из условия репрезентативности выборки:

$$\frac{s}{n} \approx \frac{S}{N},$$

где ***s*** – максимально допустимое количество отклонений в выборке, ед.

Из вышеприведенных выражений получаем:

$$s = Tn.$$

Подставив ***s*** в формулу (3.2) получим следующее выражение:

$$F_x = P(X \le Tn) = \sum_{k=0}^{Tn} p_x(k) = \sum_{k=0}^{Tn} C_n^k p^k q^{n-k}.$$

В данном случае $F_x$ это вероятность того, что доля отклонений в генеральной совокупности не превысит их максимально допустимую норму (***T***), принятую аудитором. Величина $1 - F_x$ является искомым риском выборки, который сравнивается с некоторым пороговым значением и на основании этого принимается решение о качестве тестируемой системы, в частности о необходимости проведения дополнительных аудиторских процедур.

***Пример***

Аудитор проверяет 160 счетов-фактур на предмет наличия всех необходимых реквизитов ($n = 160$). Находит пять документов с отклонениями. Таким образом, $p = \frac{5}{160} = 0{,}03125 = 3\%$. Уровень существенности составляет 5%. Следовательно, максимально допустимое количество отклонений в выборке:

$$s = 0{,}05 * 160 = 8.$$

Обозначим риск выборки чрез ***r***. Тогда

$$r = 1 - \sum_{k=0}^{8} C_{160}^k 0{,}03^k 0{,}97^{160-k} = 0{,}053.$$

В итоге риск того, что доля отклонений в совокупности счетов-фактур превысит 5% (максимально приемлемую долю) составляет 5,3%. Если аудитор считает риск, например, до 10% приемлемым, то результат тестирования счетов-фактур является удовлетворительным.

## 3.3. Процедуры по существу, основанные на биномиальном распределении

При проверках же по существу аудитора интересует не количество ошибочных документов (отклонений) в генеральной совокупности, а денежная сумма ошибок. Метод, основанный на биномиальном распределении, может быть применен и в этом случае, но с определенными ограничениями.

Если генеральная совокупность однородна (отсутствуют элементы, стоимость которых резко отличается от средней) и вариация стоимости незначительна (коэффициент вариации не превышает 20-30%), то денежная оценка максимальной ошибки ***K*** в генеральной совокупности может быть получена из средней стоимости документа:

$$\bar{J} = \frac{J}{N},$$

где ***J*** (руб.) – общая сумма, проведенная по документам, составляющим генеральную совокупность:

$$K = \bar{J}\frac{N}{n} m_{bin(F_X, p, n)},$$

где $m_{bin(F_X, p, n)}$ – параметр ***m***, полученный из выражения (3.2) при заданных $F_x, p, n$, ед.

Пример. Объем генеральной совокупности составляет 800 счетов-фактур ($N = 800$). Пусть их общая стоимость (в части налога на добавленную стоимость, НДС) составляет 800 тыс. руб. ($J = 800\ 000$). Аудитор выборочно проверяет 100 из них ($n = 100$) и находит три отклонения – неправильно заполненные счета-фактуры. Доверительная вероятность принята на уровне 92%. Таким образом, расчеты будут следующими.

Средняя стоимость документа:

$$\bar{J} = \frac{800\ 000}{800} = 1\ 000 \text{ руб}.$$

Вероятность появления ошибки:

$$p = \frac{3}{100} = 0{,}03\ (3\%).$$

Максимальное количество отклонений по выборке:

$$m_{bin(0,92;0,03;100)} = 5 \text{ ед}.$$

Максимальная ошибка в генеральной совокупности:

$$K = 1\ 000 * \frac{800}{100} * 5 = 40\ 000 \text{ руб}.$$

Таким образом, аудитор может с вероятностью 92% утверждать, что суммарная ошибка в генеральной совокупности не превысит 40 тыс. руб. Впоследствии, аудитор может сравнить эту сумму с допустимым размером ошибки и принять решение о качестве тестируемых документов.

Нужно иметь в виду, что 40 тыс. руб. является максимально возможной ошибкой при заданной вероятности, но не наиболее вероятной. Наиболее вероятная ошибка ***E*** рассчитывается следующим образом:

$$E = Np\bar{J}.$$

В нашем примере, ожидаемая ошибка составляет:

$$E = 800 * 0{,}03 * 1000 = 24\ 000 \text{ руб}.$$

Выражение для нахождения максимальной ошибки в генеральной совокупности с учетом поправки на конечность совокупности можно записать следующим образом:

$$K' = E + (K - E)\sqrt{1 - \frac{n}{N}}.$$

Таким образом, скорректированная максимальная ошибка в совокупности счетов-фактур равна

$$K' = 24\ 000 + (40\ 000 - 24\ 000)\sqrt{1 - \frac{100}{800}} = 3\ 8967 \text{ руб}.$$

Рассмотренный выше метод может быть применен только в том случае, когда ошибочной является вся учетная сумма, проведенная по документу, что обычно имеет место при формальных ошибках, неправильном или безосновательном отражении операций, отражении незаконных операций и др. В других случаях (ошибки арифметические, пересчетные, в оценке, в расчетах и др.) ошибочная сумма обычно составляет какую-то часть учетной стоимости по документу или даже может превышать ее. Применение метода, базирующегося на биномиальном распределении, в этом случае не имеет под собой серьезного статистического основания.

Если генеральная совокупность неоднородна, т.е. содержит документы, стоимость которых резко (на порядок и выше) отличается от средней, то совокупность следует стратифицировать.

# 4. МЕТОДЫ, ОСНОВАННЫЕ НА ГИПЕРГЕОМЕТРИЧЕСКОМ РАСПРЕДЕЛЕНИИ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

## 4.1. Общая информация о гипергеометрическом распределении

Гипергеометрическое распределение также позволяет определить вероятности обнаружения в атрибутивной выборке ровно ***k*** случайных событий. Оно основано на той же модели, что и биномиальное распределение. Однако есть и одно отличие, которое заключается в отсутствии требования независимости испытаний и, следовательно, условия неизменности вероятности ***p***. Иначе говоря, если в биномиальном распределении предполагается неизменность вероятности наступления случайного события (хотя практически это не принимают во внимание), то в гипергеометрическом распределении она может изменяться в ходе отбора. Для расчета вероятностей $p_x(k)$ здесь необходима дополнительная информация – размер генеральной совокупности ***N***.

Например, в генеральной совокупности из ***N*** документов ***M*** документов имеют искажение. Гипергеометрическое распределение описывает вероятность того, что в выборке из ***n*** документов, вытянутых из генеральной совокупности, ровно ***k*** документов будут содержать искажение. Здесь становится очевидно, что если в биномиальном распределении предполагается повторный отбор, то есть аудитор, проверив документ, возвращает его в генеральную совокупность и у него есть шанс снова попасть в выборку, то в гипергеометрическом распределении предполагается бесповторный отбор, когда проверенный документ больше не может быть отобран, а значит вероятность найти искажение в следующем документе меняется.

Вероятность наступления ровно ***k*** случайных событий (наличие отклонений) определяется следующим образом:

$$p_x(k) = P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, \qquad (4.1)$$

где

***M*** – ожидаемое число отклонений в генеральной совокупности, по определению $M = pN$, ед.;

***p*** – вероятность случайного события до проведения испытаний или установленная по его результатам.

Приведенная формула (4.1) может трактоваться следующим образом: существует $C_N^n$ возможных выборок (без возвращения). Есть $C_M^k$ способов выбрать ***k*** документов с отклонениями и $C_{N-M}^{n-k}$ способов заполнить остаток выборки документами без отклонений.

Определим вероятность $p_x(0)$ на основе выражения (4.1):

$$p_x(0) = \frac{C_M^0 C_{N-M}^n}{C_N^n} = \frac{C_{N-M}^n}{C_N^n} = \frac{(N-M)!\,(N-n)!}{(N-M-n)!\,N!}.$$

Если вероятность $p_x(0)$ определена, то расчет вероятностей $p_x(k)$ можно осуществить по рекуррентной формуле:

$$p_x(k+1) = \frac{(n-k)(M-k)}{(k+1)(N-M-n+k+1)} p_x(k).$$

Функция распределения, как и в биномиальном распределении, определяется по общей формуле:

$$F_x = \sum_{k=0}^{m} p_x(k) = \sum_{k=0}^{m} \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}. \qquad (4.2)$$

Определим вероятности наступления случайных событий и значения функции гипергеометрического распределения для условий:

$$N = 500, \; n = 40, \; p = \frac{M}{N} = 0{,}05 .$$

Из последнего соотношения непосредственно следует $M = 0{,}05 * 500 = 25$.

Вероятность нулевого события

$$p_x(0) = \frac{C_{475}^{40}}{C_{500}^{40}} = \frac{475!*460!}{435!*500!} = 0{,}11782 .$$

По рекуррентной формуле получим значения следующих вероятностей:

$$p_x(1) = \frac{40*2}{500-25-40+1} * 0{,}11782 = 0{,}27024 ;$$

$$p_x(2) = \frac{39*24}{2*(500-25-40+1+1)} *$$

$$* 0{,}27024 = 0{,}28941 \; и \; т.д.$$

Значения $p_x(k)$ и $F_x$ для ***k*** от нуля до восьми для условий данного примера приведены в табл. 4.1.

**Таблица 4.1**

**ГИПЕРГЕОМЕТРИЧЕСКОЕ РАСПРЕДЕЛЕНИЕ**

**($N = 500,\ n = 40,\ M = 25$)**

| ***k*** | $p_x(k)$ | $F_x$ |
|---|---|---|
| 0 | 0,11782 | 0,11782 |
| 1 | 0,27024 | 0,38806 |
| 2 | 0,28941 | 0,67747 |
| 3 | 0,19250 | 0,86997 |
| 4 | 0,08923 | 0,95920 |
| 5 | 0,03066 | 0,98987 |
| 6 | 0,00811 | 0,99798 |
| 7 | 0,00169 | 0,99967 |
| 8 | 0,00028 | 0,99996 |

Подобного рода таблица (табл. 4.1) позволяет определить вероятность нахождения в выборке количества отклонений в заданном диапазоне, например от $k = 1$ до $k = 4$. В этом случае $F_x = \sum_{k=1}^{4} p_x(k) = 0{,}84138$.

Соответствующий график представлен на рис. 4.1.

Данные графика интерпретируются точно также как и в случае с биномиальным распределением.

Разработать таблицы значений $p_x(k)$ и $F_x$ для гипергеометрического распределения, аналогичные таблицам биномиального распределения, не удается, так как одним из аргументов здесь является объем генеральной совокупности. Для определения необходимых значений вероятностей и функции распределения удобно воспользоваться электронной таблицей Excel.

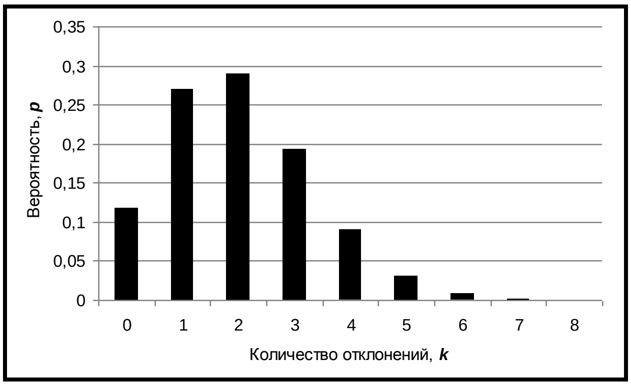


**Рис. 4.1. Функция вероятностей гипергеометрического распределения при** $N = 500$, $n = 40$, $M = 25$

## 4.2. Процедуры на соответствие, основанные на гипергеометрическом распределении

По аналогии с биномиальным распределением гипергеометрическое используется также при атрибутивных выборках в аудите, когда тестируется наличие конкретного атрибута в совокупности. В теории данное распределение должно лучше описывать вероятности отклонений, чем биномиальное, так как в нем подразумевается бесповторный отбор единиц наблюдения в выборочную совокупность, что почти всегда имеет место при аудиторских проверках. При больших объемах генеральной совокупности и сравнительно небольших выборках эти распределения хорошо аппроксимируются. Однако в некоторых ситуациях применение гипергеометрического распределения затруднено, когда, например, аудитор в силу определенных причин не знает объем генеральной совокупности. Также с учетом поправки на конечность совокупности биномиальное распределение, возможно, будет давать лучшие результаты – это вопрос сравнительного анализа, проведенного во второй главе настоящей работы.

Итак, задачи исследования в рамках процедур на соответствие при применении рассматриваемого распределения те же, что и в случае с биномиальным: определение верхнего предела точности $U$ и оценка риска выборки. Приведем конкретные формулы.

Оценка максимально возможной доли отклонений в генеральной совокупности находится следующим образом:

$$U = \frac{m_{HG(F_X, N, M, n)}}{n},$$

где $m_{HG(F_X, N, M, n)}$ – параметр $m$, выраженный из формулы (4.2) при заданных $F_x, N, M, n$, ед.

Рассмотрим пример. В рамках тестирования системы внутреннего контроля аудитор проверяет документы по операциям с денежными средствами. Проверке подлежит наличие подписи руководителя организации или уполномоченного им лица. Объем генеральной совокупности: 1 500 документов ($N = 1500$). Объем выборки: 200 документов ($n = 200$). В результате проверки было установлено, что на 6 документах отсутствует подпись. Таким образом

$$M = m\frac{N}{n} = 6 * \frac{1500}{200} = 45 \text{ ед}.$$

Аудитор принимает доверительную вероятность на уровне 95% ($F_x = 0{,}95$).

Следовательно, верхний предел точности $U$ может быть найден следующим образом:

$$U = \frac{m_{HG(0{,}95;1500;105;200)}}{200} = \frac{9}{200} = 0{,}045 \ (4{,}5\%).$$

Таким образом, аудитор с вероятностью 95% может утверждать, что доля неподписанных денежных документов в совокупности из 1,5 тыс. документов не превысит 4,5%.

Оценка риска выборки осуществляется по общей формуле:

$$r = 1 - F_X = 1 - \sum_{k=0}^{Tn} \frac{C_M^k C_{N-M}^{n-k}}{C_N^n},$$

где $T$ – уровень существенности в долях.

Возьмем исходные данные предыдущего примера. Уровень существенности составляет 5%. Аудитор хочет оценить вероятность того, что отклонения в генеральной совокупности документов могут превысить максимально допустимую норму искажения, т.е. вероятность существенной ошибки $Tn = 0{,}05 * 200 = 10$.

Таким образом, получаем:

$$r = 1 - \sum_{k=0}^{10} \frac{C_{45}^k C_{1455}^{200-k}}{C_{1500}^{200}} = 0{,}02925 \ (2{,}93\%).$$

В результате риск выборки составляет 2,93%, и если в качестве приемлемого порога принят 10%-й уровень, то аудитор может уверенно полагать, что совокупность не содержит существенных отклонений.

## 4.3. Процедуры по существу, основанные на гипергеометрическом распределении

Денежная оценка ошибки с использованием гипергеометрического распределения основана на тех же принципах, что используются при любом другом дискретном распределении в рамках атрибутивных выборок. В частности для биномиального распределения уже был описан метод определения ошибки через среднюю стоимость документа. Подобный способ применим и для гипергеометрического распределения.

Денежная оценка максимальной ошибки $K$ в генеральной совокупности может быть получена следующим образом:

$$K = \bar{J}\frac{N}{n} m_{HG(F_X, N, M, n)}.$$

Как уже было отмечено в случае биномиального распределения, подобный метод может быть применен только тогда, когда ошибочной является вся учетная сумма, проведенная по документу.

В качестве примера возьмем ситуацию из предыдущего подраздела. Аудитор проверяет наличие подписи на денежных документах. Исходные данные следующие: $N = 2000$, $n = 100$, $F_x = 90\,\%$. Общая сумма документов $J = 1\,000\,000$ руб. В результате проверки найдено четыре отклонения. Таким образом, средняя стоимость документа:

$$\bar{J} = \frac{1\,000\,000}{2\,000} = 500 \text{ руб.}.$$

Наиболее вероятное количество отклонений в генеральной совокупности:

$$M = m\frac{N}{n} = 4 * \frac{2000}{100} = 80 \text{ ед.}$$

Максимальное количество отклонений по выборке:

$$m_{HG\,(0,9;\,2000;\,80;\,100)} = 6 \text{ ед.}$$

Оценим наиболее вероятную ошибку и максимальную ошибку в генеральной совокупности при заданной доверительной вероятности.

Наиболее вероятная ошибка:

$$E = M\bar{J} = 80 * 500 = 40\,000 \text{ руб.}$$

Максимальная ошибка в совокупности:

$$K = 500 * \frac{2\,000}{100} * 6 = 60\,000 \text{ руб.}$$

Таким образом, аудитор может с 90%-й вероятностью утверждать, что ошибка в совокупности денежных документов не превысит 60 тыс. руб. При этом наиболее вероятная ошибка составляет 40 тыс. руб.

# 5. МЕТОДЫ, ОСНОВАННЫЕ НА РАСПРЕДЕЛЕНИИ ПУАССОНА

## 5.1. Общая информация о распределении Пуассона

Один из пределов биномиального распределения, представляющий практический интерес, относится к случаю, когда при неограниченном увеличении числа испытаний математическое ожидание остается постоянным:

$$np \xrightarrow{n \to \infty} \lambda = const.$$

Распределение Пуассона играет важную роль для описания редких событий в физике, теории связи, теории надежности, теории массового обслуживания и т.д. – там, где в течение определенного времени может происходить случайное число каких-то событий с очень малой вероятностью (радиоактивных распадов, телефонных вызовов, отказов оборудования, несчастных случаев и т.п.). В аудите данное распределение можно использовать тогда, когда ожидается очень небольшое количество отклонений в совокупности огромных объемов.

Выведем формулу для распределения Пуассона из биномиального распределения. Для начала приведем выражение (3.1) к виду более удобному для последующих преобразований:

$$p_x(k) = C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}. \quad (5.1)$$

Теперь введем константу $\lambda = np$. Подставив ее в формулу (5.1) и устремив $n \to \infty$, таким образом, что $p \to 0$, $q \to 1$, получаем:

$$\lim_{n\to\infty} p_x(k) = \lim_{n\to\infty}\left[\frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}\right].$$

Обозначим случайную величину, имеющую распределение Пуассона, как $P(\lambda)$. После выведения некоторых констант за скобки, а также при использовании второго замечательного предела получаем:

$$P(\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}\lim_{n\to\infty}\left[\frac{n!}{(n-k)!}\frac{1}{n^k}\left(1-\frac{\lambda}{n}\right)^{-k}\right].$$

Преобразуем предел в правой части:

$$\lim_{n\to\infty}\left[\frac{n!}{(n-k)!}\frac{1}{n^k}\left(1-\frac{\lambda}{n}\right)^{-k}\right] =$$

$$= \lim_{n\to\infty}\left[\frac{n(n-1)\ldots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^{-k}\right].$$

Так как $p \to 0$ и $k \ll n$, то:

$$\lim_{n\to\infty} n(n-1)\ldots(n-m+1) = n^k,$$

а значит

$$\lim_{n\to\infty}\left[\frac{n(n-1)\ldots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^{-k}\right] = 1.$$

Таким образом, получаем окончательную формулу для распределения Пуассона:

$$P(\lambda) = p_x(k) = \frac{\lambda^k}{k!}e^{-\lambda}. \quad (5.2)$$

При больших размерах тестируемой совокупности (***N***) распределение Пуассона хорошо аппроксимируется с гипергеометрическим распределением. Как уже было показано выше, распределение задается только одним параметром $\lambda$ (среднее число случаев в выборке).

Вероятности $p_x(k)$ могут быть определены и по рекуррентной формуле:

$$p_x(k+1) = \frac{\lambda}{K+1}p_x(k).$$

Функция распределения:

$$F_x = P(X \le m) = \sum_{k=0}^{m} p_x(k) = \sum_{k=0}^{m}\frac{\lambda^k}{k!}e^{-\lambda}. \quad (5.3)$$

Значения функции распределения Пуассона приводятся в соответствующих справочниках [12]. Их нетрудно определить и по электронной таблице Excel. Как уже было отмечено, в выборочных обследованиях к этому распределению обычно прибегают при невысоких значениях вероятности ***p***.

В демонстративных целях определим вероятности для условия $\lambda = 1{,}2$ и функцию распределения Пуассона для $m = 7$. По формуле (5.2) находим:

$$p_x(0) = \frac{1{,}2^0}{0!}e^{-1,2} = 0{,}30119;$$

$$p_x(1) = \frac{1{,}2^1}{1!}e^{-1,2} = 0{,}36143 \text{ и т.д.}$$

В табл. 5.1 приведены значения вероятностей $p_x(k)$ и функции распределения $F_x$ для ***k*** от нуля до семи.

Соответствующий график представлен на рис. 5.1.

**Таблица 5.1**

**РАСПРЕДЕЛЕНИЕ ПУАССОНА ($\lambda = 1{,}2$)**

| $k$ | $p_x(k)$ | $F_x$ |
|---|---|---|
| 0 | 0,30119 | 0,30119 |
| 1 | 0,36143 | 0,66263 |
| 2 | 0,21686 | 0,87949 |
| 3 | 0,08674 | 0,96623 |
| 4 | 0,02602 | 0,99225 |
| 5 | 0,00625 | 0,99850 |
| 6 | 0,00125 | 0,99975 |
| 7 | 0,00021 | 0,99996 |

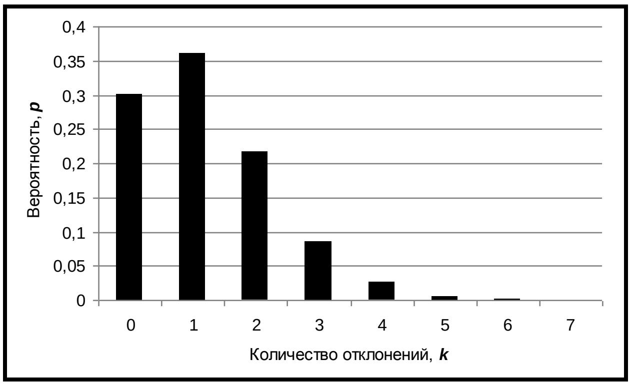


**Рис. 5.1. Функция вероятностей распределения Пуассона при $\lambda = 1{,}2$**

И график, и таблица интерпретируются точно так же как и в случае биномиального или гипергеометрического распределения. На графике видно, что самый высокий столбик принадлежит $k = 1$, что соответствует математическому ожиданию, которое в случае распределения Пуассона равно $\lambda$ ($\lambda = 1{,}2$ в нашем примере). Вероятность нахождения ровно 2-х отклонений в выборке составляет 21,7% ($p_x(2) = 0{,}21686$). Вероятность найти не более двух отклонений в выборочной совокупности – 88% ($F_x = 0{,}87949$). Вероятность найти от двух до трех отклонений – 30,36% ($\sum_{k=2}^{3} p_x(k) = 0{,}3036$).

## 5.2. Процедуры на соответствие, основанные на распределении Пуассона

Распределение Пуассона в аудите может применяться в рамках атрибутивной проверки по аналогии с двумя другими распределениями, уже описанными в данной работе. Если ожидается небольшое количество отклонений по прошлому опыту аудитора или его интуиции при проверке достаточно большой совокупности таким образом, что вероятность нахождения отклонения в конкретно взятом документе мала, то методы, основанные на распределении Пуассона, статистически оправданы. Основные задачи при использовании этого распределения такие же, как и в случае биномиального или гипергеометрического распределения:

- определение максимального количества отклонений в совокупности при заданной доверительной вероятности;
- оценка риска выборки с учетом уровня существенности.

Перейдем сразу к формулам. Верхний предел точности определяется следующим образом:

$$U = \frac{m_{P(F_X,\lambda)}}{n},$$

где $m_{P(F_X,\lambda)}$ – параметр $m$, выраженный из формулы (5.3) при заданных $F_x$, $\lambda$, ед.

Для демонстрации рассмотрим данную формулу на примере. Аудитор проверяет порядок заполнения счетов-фактур. Из 200 проверенных документов у двух отсутствует порядковый номер и дата выписки. Аудитор хочет найти максимально возможную долю отклонений в генеральной совокупности при доверительной вероятности в 95%. Так как объем выборки велик, а вероятность отклонения в документе невысока ($p = \frac{2}{200} = 0{,}01$), можно применить распределение Пуассона. Таким образом, имеем следующие данные: $n = 200$, $\lambda = 2$, $F_x = 0{,}95$.

Максимальное количество отклонений по выборке:

$$m_{P(0,95;2)} = 4 \text{ ед.}$$

Верхняя граница доли отклонений в совокупности:

$$U = \frac{4}{200} = 0{,}02\ (2\%).$$

По результатам проверки аудитор может утверждать с вероятностью 95%, что доля отклонений в генеральной совокупности счетов-фактур не превысит 2%.

Рассмотрим оценку риска выборки с применением распределения Пуассона. Риск выборки в данном случае определяется следующим образом:

$$r = 1 - \sum_{k=0}^{Tn} \frac{\lambda^k}{k!} e^{-\lambda}.$$

Взяв данные предыдущего примера и задавшись уровнем существенности в 3%, рассчитаем риск выборки. При $Tn = 0{,}03 * 200 = 6$ получаем:

$$r = 1 - \sum_{k=0}^{6} \frac{2^k}{k!} e^{-2} = 1 - 0{,}99547 = 0{,}00453\ (0{,}45\%).$$

Таким образом, вероятность того, что доля счетов фактур с отклонением в генеральной совокупности превысит максимально допустимую норму искажения (3%), составляет 0,45%. Очевидно, что такой риск выборки является более чем удовлетворительным.

## 5.3. Процедуры по существу, основанные на распределении Пуассона

В процедурах по существу, основанных на распределении Пуассона, используется тот же механизм стоимостного определения ошибки, что и в двух уже описанных случаях при биномиальном и гипергеометрическом распределении. Денежную оценку верхнего предела ошибки при заданной вероятности можно осуществить следующим образом:

$$K = \bar{J}\frac{N}{n} m_{P(F_X,\lambda)}.$$

Как и ранее, в данном методе подразумевается строгая связь между стоимостью документа и ошибкой: наличие отклонения делает всю сумму документа ошибочной. Ошибка не может составлять часть от суммы по документу – только целиком, иначе распределение Пуассона, как и любое другое дискретное распределение, описанное в данной работе, не применимо.

Найдем максимальную ошибку в совокупности документов со следующими условиями: из 150 проверенных документов найдено три отклонения. Доверительная вероятность установлена на уровне 90%. Генеральная совокупность: 1 000 документов с общей суммой в 2,1 млн. руб. Следовательно, $n = 150$, $\lambda = 3$, $F_x = 0{,}9$, $N = 1000$.

Средняя стоимость документа:

$$\bar{J} = \frac{2100000}{1000} = 2100 \ \text{руб}.$$

Максимальное количество отклонений по выборке при заданной вероятности:

$$m_{P(0,9;3)} = 5 \ \text{ед}.$$

Таким образом, можем рассчитать стоимостную величину ошибки. Максимальная стоимостная ошибка в генеральной совокупности равна

$$K = 2100 * \frac{1000}{150} * 5 = 70\ 000 \ \text{руб}.$$

В результате с вероятностью 90% ошибка в генеральной совокупности документов не должна превысить 70 тыс. руб.

Дополнительно оценим наиболее вероятную ошибку в совокупности:

$$E = \lambda \frac{N}{n} \bar{J} = 3 * \frac{1\ 000}{150} * 2100 = 42\ 000 \ \text{руб}.$$

Разница между максимальной ошибкой при заданной доверительной вероятности и наиболее вероятной ошибкой представляет собой предельную ошибку выборки:

$$\Delta = 70\ 000 - 42\ 000 = 28\ 000 \ \text{руб}.$$

# 6. МЕТОДЫ, ОСНОВАННЫЕ НА НОРМАЛЬНОМ РАСПРЕДЕЛЕНИИ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

## 6.1. Общая информация о нормальном распределении

Нормальное распределение занимает одно из самых важных мест не только в теории и практике выборочных исследований, но и во многих других областях науки, например, в математической статистике и статистической физике. Данное обстоятельство объясняется тем, что большинство процессов в нашем мире описываются законом нормального распределения. Оно часто встречается в природе. Например, следующие случайные величины хорошо моделируются нормальным распределением:

- отклонение при стрельбе;
- погрешности измерений (однако, погрешности некоторых измерительных приборов имеют не нормальные распределения);
- некоторые характеристики живых организмов в популяции.

Важное значение нормального распределения вытекает из центральной предельной теоремы теории вероятностей. Если результат наблюдения является суммой многих случайных слабо взаимозависимых величин, каждая из которых вносит малый вклад относительно общей суммы, то при увеличении числа слагаемых распределение центрированного и нормированного результата стремится к нормальному. Этот закон теории вероятностей имеет следствием широкое распространение нормального распределения, что и стало одной из причин его наименования.

Нормальное распределение характеризует распределение непрерывной случайной величины $x$, т.е. является непрерывным распределением в отличие от всех ранее рассмотренных дискретных распределений. Параметрами нормального распределения являются среднее значение ($\bar{x}$) и дисперсия ($\sigma^2$). Обычно это распределение обозначается как $N(\bar{x}, \sigma)$ или $X \sim N(\bar{x}, \sigma)$. Функция симметрична относительно $\bar{x}$.

Нормальное распределение, как и распределение Пуассона, является предельным случаем биномиального. Однако при нормальном распределении в отличие от распределения Пуассона

$$np \xrightarrow{n \to \infty} \infty.$$

Плотность вероятности выводится из биномиального распределения при помощи уже упомянутой в этой работе формулы Стирлинга и разложения в ряд Тейлора. Функция плотности нормального распределения выглядит следующим образом:

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

где

$\mu$ – математическое ожидание, медиана и мода распределения;

$\sigma$ – стандартное отклонение ($\sigma^2$ – дисперсия) распределения.

Кривая нормального распределения $N(\bar{x}, \sigma)$ показана на рис. 6.1. Площадь под кривой в пределах $\bar{x} \pm 3\sigma$ чуть меньше единицы. Это называется правило 3-х сигм.

На практике обычно применяют нормированное (стандартизированное) нормальное распределение $N(0,1)$, которое не зависит от масштабов измерения случайной величины $x$.

Вместо $x$ здесь фигурирует нормированное по $\sigma$ отклонения $x$ от среднего $\bar{x}$, которые обозначим $z$:

$$z = \frac{x - \bar{x}}{\sigma}. \tag{6.1}$$

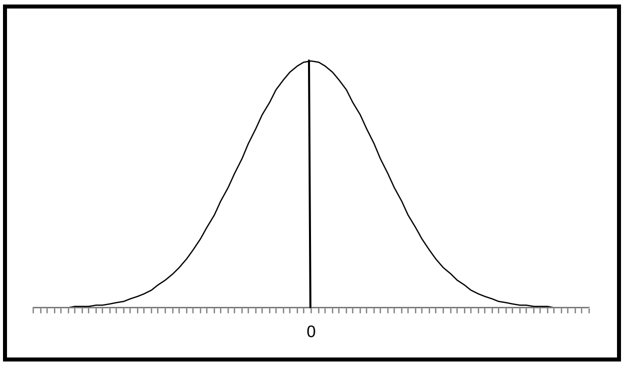


**Рис. 6.1. Кривая нормального распределения**

Величина $z$ также является случайной. Заменим в выражении (6.1) значения средней и стандартного отклонения в совокупности на их выборочные оценки. Получим:

$$z = \frac{x - \tilde{x}}{\sigma}, \qquad (6.2)$$

где

$\tilde{x}$ – выборочное среднее;

$\sigma$ – стандартное отклонение по выборке.

Из (6.2) легко найти абсолютный размер отклонения $x$ от средней как функцию $z$:

$$x - \tilde{x} = z\sigma .$$

Например, для $x = 25$ при условии, что $\tilde{x} = 40$ и $\sigma = 10$, нормированное отклонение составит:

$$z = \frac{25 - 40}{10} = -1{,}5 .$$

Кривая стандартного нормального распределения $N(0{,}1)$ показана на рис. 6.2. Кривая симметрична относительно $z = 0$. Для $z$ в пределах от $-\infty$ до $+\infty$ площадь под кривой распределения равна единице, а для $z$ в пределах от -3 до +3 эта площадь чуть меньше единицы (0,997).

Функция стандартного нормального распределения определяется как:

$$F_x(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{y^2}{2}} dy . \qquad (6.3)$$

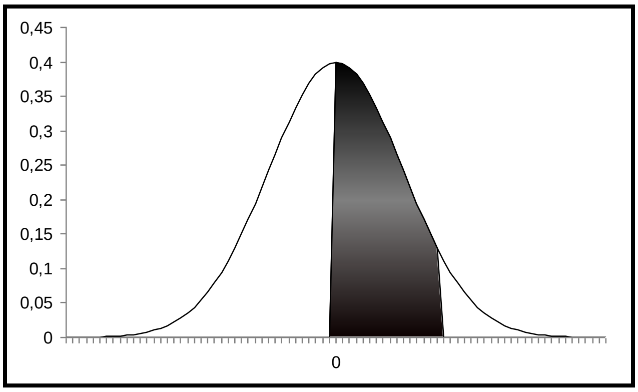


**Рис. 6.2. Кривая стандартного нормального распределения с односторонним интервалом**

Величина функции распределения равна площади под кривой в интервале значений от $-\infty$ до $z$. В свою очередь площадь под кривой распределения характеризует вероятность того, что случайная переменная находится в указанном интервале. Именно это свойство кривой стандартного нормального распределения используется в анализе выборочных данных. Практически нет необходимости прибегать к формуле (6.3), так как соответствующие значения функции табулированы. Значения функции нормального распределения можно получить и по электронной таблице Excel.

В практических расчетах возможны следующие варианты определения вероятности на основе указанных таблиц в зависимости от заданных значений случайной переменной:

не превышает заданного значения $z$, то есть находится в пределах от $-\infty$ до $z$ (рис. 6.2). Это самый применяемый вариант в аудите в связи с использованием понятия существенности и определенной на его основе границы – допустимого размера ошибки.

- находится в пределах $\pm z$, то есть интервал симметричен относительно начала координат (рис. 6.3);
- находится в пределах от $z_1$ до $z_2$ (рис. 6.4).

Значение функции при двухстороннем симметричном интервале получается следующим образом:

$$P(X \in |-z, +z|) = F_x(z) - F_x(-z) = 2(F_x(z) - 0{,}5) .$$

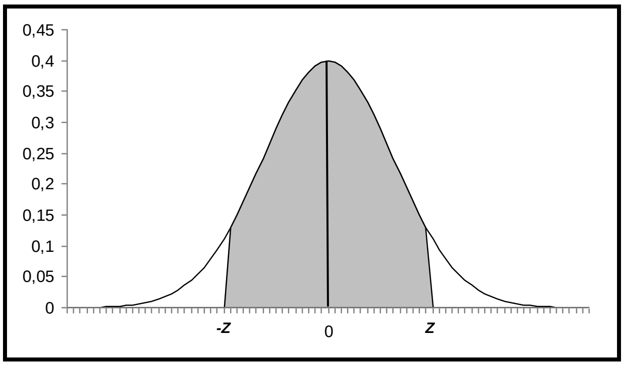


**Рис. 6.3. Кривая стандартного нормального распределения с двухсторонним симметричным интервалом**

Что касается произвольного интервала, то соответствующее значение получим как разность

$$P(X \in |z_1, z_2|) = F_x(z_2) - F_x(z_1), \; -\infty < z_1 < z_2 < +\infty .$$

В силу симметричности распределения

$$F_x(-z) = 1 - F_x(z).$$

Рассмотрим некоторые примеры использования функции нормального распределения.

Определим вероятности (площади под кривой стандартного нормального распределения) для следующих вариантов значений нормированного отклонения:

- не превышает 0,65;
- находится в пределах $\pm 0{,}25$;
- находится в пределах от 0,1 до 0,25.

По приложению 2 находим значение функции распределения для первого варианта: $F_x(0{,}65) = 0{,}742$. Иначе говоря, вероятность того, что нормированное отклонение случайной величины не превысит величину $z = 0{,}65$, немного превышает 74%.

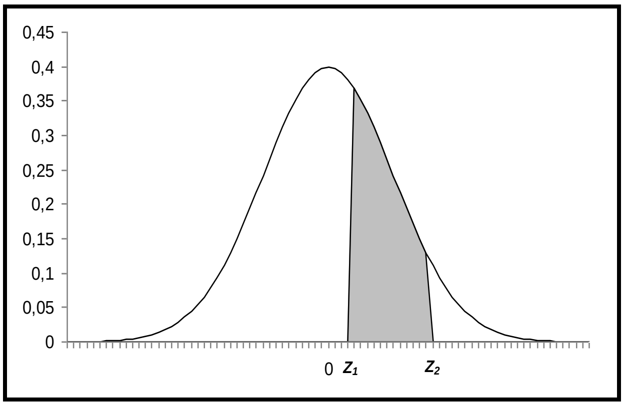


**Рис. 6.4. Кривая стандартного нормального распределения с двухсторонним произвольным интервалом**

Для второго варианта получим:

$$P(X \in |-0,25;+0,25|) = 2(F_x(0,25) - 0,5) = 0,198 \ .$$

То есть вероятность указанного события порядка 20%.

Наконец, для третьего варианта получим вероятность чуть ниже 6%:

$$P(X \in |0,1;0,25|) = F_x(0,25) - F_x(0,1) =$$
$$= 0,5987 - 0,5398 = 0,0589 \ .$$

Разберем еще один пример. Известно, что случайная величина $X$ имеет нормальное распределение. Среднее значение в выборке – 200, стандартное отклонение – 100. Необходимо определить вероятность того, что случайная величина находится в интервале150-400.

Нормируем значения:

$$z = \frac{x - 200}{70} \ .$$

Граничные значения $z$ для указанного диапазона $X$:

$$z_1 = \frac{150 - 200}{100} = -0,5 \ ;$$

$$z_2 = \frac{400 - 200}{100} = 2 \ .$$

Искомая вероятность равна 69%:

$$P(X \in |-0,5;2|) = F_x(2) - (1 - F_x(0,5)) =$$
$$= 0,9972 - 0,3085 = 0,6887 \ .$$

Как было показано выше, нормированное нормальное распределение позволяет определить вероятность в зависимости от заданного интервала значений.

Аудитору часто приходится оценить по выборочным данным неизвестную сумму стоимостных характеристик для тестируемой совокупности. Эта задача равнозначна оцениванию неизвестного размера средней. Для этого аудитор отбирает из совокупности $n$ документов (единиц наблюдения) и определяет значение выборочной суммы и средней. Допустим, что он повторяет отбор. Тогда, как правило, он получит другие значения суммы и средней. Для тестируемой совокупности эти величины остаются неизвестными. Для их оценивания необходимо располагать соответствующими выборочными распределениями.

Как утверждает центральная предельная теорема, распределение сумм независимых случайных величин асимптотически приближается к нормальному распределению. Иначе говоря, каким бы ни было распределение исходных данных, с увеличением объема выборки распределение выборочных сумм (а следовательно, и средних) стремится к нормальному [5]. Это свойство позволяет применять нормальное распределение при определении интервальных оценок параметров для совокупности.

## 6.2. Процедуры по существу. Оценивание средней и суммы стоимостных характеристик совокупности

До этого момента в данной работе рассматривались процедуры на соответствие и процедуры по существу, основанные на атрибутивных выборках. Теперь перейдем к проверкам по существу, где случайной величиной является не атрибут, а числовая характеристика – средняя сумма ошибки или любой другой стоимостный показатель. Для начала рассмотрим вид проверки по существу, который заключается в оценивании среднего размера какого-либо стоимостного показателя. Например, среднего остатка на счете, средней суммы просроченного кредита и т.п. На основе выборочных средних легко оценить общие суммы соответствующих показателей в совокупности.

При оценивании средней и суммы возможны два варианта. Согласно первому правильность зарегистрированных стоимостных показателей документов не проверяется. По второму варианту выборочные стоимостные характеристики документов тестируются аудитором. Обнаруженные отклонения (ошибки) статистически обобщаются в виде соответствующих средних или сумм.

Необходимость оценивания по выборке средних и сумм стоимостных показателей для совокупности возникает в ряде ситуаций. Например, когда эти данные по каким-либо причинам отсутствуют, однако, они нужны аудитору для анализа, для различных сопоставлений. Нельзя исключить случаи, когда необходимо проверить правильность предоставленной для проверки итоговой суммы, не обращаясь ко всем документам совокупности.

Оценивание средней и суммы по выборке является стандартной и наиболее простой постановкой задачи определения неизвестных параметров совокупности. Ее обсуждение целесообразно в методическом плане до рассмотрения более важной для аудита специальной постановки задачи. Последняя предполагает предварительное тестирование стоимостных характеристик, а также оценивание абсолютных и относительных ошибок регистрации и определение их влияния на суммарные стоимостные показатели в совокупности. Соответствующие методики будут обсуждаться позже.

При оценивании параметров совокупности по данным выборок получают два вида оценок: точечные и интервальные. В качестве точечной оценки средней для совокупности применяют выборочную среднюю арифметическую.

Это величина определяется следующим образом:

$$\tilde{x} = \frac{\sum_{i=1}^{n} x_i}{n},$$

где

$x_i$ – стоимостная характеристика $i$-го документа, руб.;

$n$ – объем выборки, ед.

Оценивание среднего размера стоимостной характеристики документа является первым шагом при оценивании суммы. После того как выборочная средняя определена, нетрудно найти точечную оценку суммы для совокупности. Для этого распространим значение полученной средней на всю совокупность:

$$\tilde{Q} = \tilde{x}N,$$

где

$\tilde{Q}$ – точечная оценка суммы для совокупности, руб.;

$N$ – общий размер тестируемой совокупности, ед.

Точечную оценку суммы можно получить и на основе выборочного значения суммы $\sum_{i=1}^{n} x_i$. Естественно полагать, что точечная оценка суммы для совокупности в $\frac{N}{n}$ раз больше, чем сумма, полученная по выборке:

$$\tilde{Q} = \sum_{i=1}^{n} x_i \frac{N}{n} = \tilde{x}N.$$

При получении точечной оценки средней не учитывается один важный момент – оценка содержит некоторую ошибку, связанную с тем, что при расчете учитывалась только часть документов, а не вся совокупность. В одной выборке средняя имеет одно значение, в другой из этой же совокупности – другое. В связи с этим следует ожидать, что действительная средняя тестируемой совокупности более или менее отличается от полученной выборочной средней, но находится в ее окрестности. Таким образом, задача заключается в определении границ интервалов, в которых находятся средняя и сумма для совокупности. Основной для таких интервалов всегда являются соответствующие точечные оценки.

Границы интервала для средней совокупности находятся следующим образом:

$$\bar{x} = \tilde{x} \pm \Delta,$$

где $\Delta$ – предельная ошибка выборочной средней.

В свою очередь границы интервала для суммы стоимостных показателей совокупности находятся как

$$Q = N(\tilde{x} \pm \Delta) = \tilde{Q} \pm N\Delta.$$

Как следует из вышеприведенных формул, отклонения от точечных оценок ($\Delta$ и $N\Delta$) симметричны относительно центральных значений в соответствующих интервалах.

Если расчет $\tilde{x}$ не требует комментариев, то метод определения $\Delta$, вероятно, нуждается в них. Начнем с того, что выборочная средняя является случайной величиной. Ее значение зависит от случая – от того какие документы попадут в выборку. Изменчивость, или колеблемость выборочной средней измеряется, как известно, с помощью стандартной ошибки средней. Обозначим эту величину как $\mu$. Находится она следующим образом:

$$\mu = \frac{\sigma}{\sqrt{n}},$$

где

$\sigma$ – стандартное (среднеквадратическое) отклонение в выборке;

$n$ – объем выборки, ед.

Стандартное отклонение находится по выборочным данным следующим образом:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \tilde{x})^2}{n-1}}.$$

Поскольку $\sum_{i=1}^{n}(x_i - \tilde{x})^2 = \sum_{i=1}^{n} x_i^2 - n\tilde{x}^2$, то расчет $\sigma$ можно осуществить, применив следующую рабочую формулу:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \tilde{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - n\tilde{x}^2}{n-1}}.$$

Вернемся к предельной ошибке выборочной средней или предельной ошибке выборки $\Delta$. Эта величина характеризует наибольшее отклонение от выборочной средней, соответствующее принятому уровню надежности. Размер предельной ошибки в конченом счете зависит от двух факторов – степени изменчивости исходных данных и принятого уровня надежности оценки.

Влияние первого фактора измеряется стандартным отклонением средней $\mu$. Чем больше колеблемость данных в выборке, тем, естественно, больше предельная ошибка. Аналогично можно сказать и о влиянии второго фактора – уровня надежности, или доверительной вероятности, с которым мы уже познакомились в предыдущих разделах. Уровень доверительной вероятности устанавливает аудитор в зависимости от важности оцениваемого параметра для аудиторского заключения.

Согласно центральной предельной теореме, о которой говорилось ранее, распределение выборочной средней следует нормальному закону. Таким образом, предельная ошибка определяется на основе нормированного нормального распределения в зависимости от заданного уровня доверительной вероятности.

С учетом влияния двух факторов, о которых говорилось выше, предельная ошибка выборки определяется следующим образом:

$$\Delta = z\mu,$$

где $z$ – нормированное отклонение от средней (предел интеграла Лапласа).

Окончательно для предельной ошибки выборочной средней имеем следующее выражение:

$$\Delta = z\sqrt{\frac{\sum_{i=1}^{n}(x_i - \tilde{x})^2}{(n-1)n}}.$$

Корень всегда берется с положительным знаком.

Из вышеприведенной формулы следует, что чем больше объем выборки, тем меньше ошибка выборки при всех прочих равных условиях.

**Таблица 6.1**

**ВЕРОЯТНОСТИ СТАНДАРТНОГО НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ДЛЯ ПАРАМЕТРА *Z***

| $p$ | 0,85 | 0,9 | 0,95 | 0,99 |
|---|---|---|---|---|
| $\alpha$ | 1,03 | 1,28 | 1,64 | 2,33 |
| $\alpha/2$ | 1,15 | 1,64 | 1,96 | 2,57 |

На практике в аудиторских выборочных обследованиях часто применяют $z = 2$, что соответствует доверительной вероятности 95,4% (см. табл. 6.1).

### *Пример*

Тестируется совокупность счетов-фактур. Необходимо выборочным путем определить их общую сумму. Количество счетов в совокупности 2 тыс., объем выборки 120 ед. В табл. 6.2 приведены значения стоимостных показателей (руб.) счетов-фактур (первые и последние семь счетов, оказавшихся в выборке) и суммы, которые требуются для расчета $\sigma$ .

Точечная оценка среднего значения для совокупности (выборочная средняя):

$$\tilde{x} = \frac{165\ 445}{120} = 1\ 379 \text{ руб}.$$

**Таблица 6.2**

**РАСЧЕТ СУММ, НЕОБХОДИМЫХ ДЛЯ ОПРЕДЕЛЕНИЯ СТАНДАРТНОГО ОТКЛОНЕНИЯ**

| $i$ | $x_i$ | $x_i^2$ |
|---|---|---|
| 1 | 1 205 | 1 452 025 |
| 2 | 92 | 8 464 |
| 3 | 910 | 828 100 |
| 4 | 3 150 | 9 922 500 |
| 5 | 2 546 | 6 482 116 |
| 6 | 1 293 | 1 671 849 |
| 7 | 315 | 99 225 |
| … | … | … |
| 114 | 136 | 18 496 |
| 115 | 1 809 | 3 272 481 |
| 116 | 1 521 | 2 313 441 |
| 117 | 837 | 700 569 |
| 118 | 2 100 | 4 410 000 |
| 119 | 3 123 | 9 753 129 |
| 120 | 128 | 16 384 |
| Сумма | 165 445 | 350 181 362 |

Точечная оценка суммы для совокупности:

$$\tilde{Q} = 1\ 379 * 2\ 000 = 2\ 758\ 000 \text{ руб}.$$

Перейдем к интервальному оцениванию. Определим стандартное отклонение и стандартную ошибку средней. Используя известные формулы, получим:

$$\sigma = \sqrt{\frac{120 * 350\ 181\ 362 - 165\ 445^2}{119 * 120}} = 1\ 013 \text{ руб}.;$$

$$\mu = \frac{1013}{\sqrt{120}} = 92 \text{ руб}.$$

Пусть доверительная вероятность составляет 90% ($t = 1{,}64$). Тогда с вероятностью 90% можно утверждать, что средний размер счета в совокупности находится в пределах:

$$\bar{x} = 1\ 379 \pm 1{,}64 * 92 = 1\ 379 \pm 152 \text{ руб}.$$

А общая сумма счетов составляет:

$$Q = 2000 * (1379 \pm 152) \text{ руб}.$$

Таким образом, с вероятностью 90% можно утверждать, что общая сумма счетов определяется пределами:

$$Q_L = 2\ 000 * 1227 = 2\ 454\ 000 \text{ руб}.;$$

$$Q_U = 2\ 000 * 1\ 531 = 3\ 062\ 000 \text{ руб}.$$

Приведенные выше формулы для расчета $\sigma$ , строго говоря, получены для повторного отбора, который в принципе вряд ли уместен в аудите. Применение указанных формул при бесповторном отборе приводит к некоторым смещениям в получаемых оценках. Впрочем, при небольшой величине соотношения $\frac{n}{N}$ этими смещениями вполне можно пренебречь. Для бесповторного отбора оценку $\mu$ получим по формуле:

$$\mu = \sqrt{\frac{\sigma^2}{n}} * \sqrt{1 - \frac{n}{N}} = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)}.$$

С множителем $\sqrt{1 - \frac{n}{N}}$ мы уже встречались в предыдущих разделах. Там он был назван поправкой на конечность совокупности (имеется в виду ограниченность совокупности, из которой произведена выборка) или кратко ПКС. При достаточно небольшом относительном объеме выборки эта величина близка к единице.

Продолжим предыдущий пример и оценим ошибку средней теперь по формуле для бесповторного отбора. По условиям примера $N = 2000$, $n = 120$, тогда ПКС составит всего:

$$\sqrt{1 - \frac{120}{2000}} = 0{,}97.$$

Таким образом, с учетом ПКС:

$$\sigma = 1013 * 0{,}97 = 983 \text{ руб}.;$$

$$\mu = 92 * 0{,}97 = 89 \text{ руб}.;$$

$$\bar{x} = 1\ 379 \pm 1{,}64 * 89 = 1\ 379 \pm 146 \text{ руб}.;$$

$$Q = 2\ 000\ (1\ 379 \pm 146);$$

$$Q_L = 2\ 000 * 1\ 233 = 2\ 466\ 000 \text{ руб}.;$$

$$Q_U = 2\ 000 * 1\ 525 = 3\ 050\ 000 \text{ руб}.$$

Как видим, учет ПКС практически не изменил значения рассчитываемых параметров по сравнению с полученными ранее результатами.

Выше внимание было сосредоточено на определении предельной ошибки выборочной средней. Для расчета последней необходимо иметь стандартную ошибку этой средней. Практически ничего не изменится, если расчет базировать на стандартной ошибке выборочной суммы:

$$\mu_\Sigma = \frac{N\sigma}{\sqrt{n}}.$$

Как видим, стандартная ошибка выборочной суммы в $N$ раз больше стандартной ошибки выборочной средней.

Перейдем к оцениванию средней и суммы стоимостных характеристик совокупности при расслоенном отборе единиц наблюдения. Но для начала рассмотрим общие сведения о стратифицированном (расслоенном) отборе. Согласно этому способу обследуемая совокупность первоначально делится на слои (группы, страты). Каждый слой содержит более или менее однородные в каком-либо отношении единицы наблюдения. Например, группы документов по источникам поступлений, по размеру сумм, по срокам задолженности и т.п. Затем из каждой группы в случайном порядке (собственно случайный или механический отбор) производится извлечение выборочных единиц наблюдения. В аудите в основном имеют дело с денежными показателями. Соответственно совокупность делится на интервалы или слои по величине стоимости показателей.

Число слоев, границы между ними и распределение общего объема выборки между слоями устанавливается аудитором. Переход от собственно случайного к расслоенному отбору позволяет учитывать наиболее значимые для аудита слои документов.

На практике применяют несколько видов расслоения. Рассмотрим некоторые из них.

- Пропорциональное расслоение (размещение) выборки по слоям. Иначе говоря, доля отбора одинакова в каждом слое. Например, по 10% от количества документов в слое. Таким образом, количество документов, извлеченных в выборку из каждого слоя совокупности, пропорционально размеру слоя.
- Непропорциональное расслоение по слоям. Доли отбора здесь изменяются от слоя к слою. Непропорциональное расслоение реализуется различными способами. Наиболее простым является отбор равных долей выборки из каждого слоя. Например, при пяти слоях из каждого слоя извлекается одна пятая установленного общего объема выборки.

Важным частным случаем непропорционального расслоения является так называемое оптимальное расслоение выборочных единиц наблюдения по слоям. Такой метод позволяет во многих случаях получить наибольший объем информации при фиксированных затратах труда на извлечение выборки. Оптимальным является расслоение объема выборки пропорционально взвешенным стандартным отклонениям по слоям. В качестве весов здесь берут доли слоев в совокупности.

Также целесообразно здесь упомянуть о двухслойной выборке. Это упрощенный вариант расслоенного отбора (отбор с двумя слоями), применяемый, вероятно, только в аудите. Порядок формирования выборки сводится к следующему. Документы генеральной совокупности, подлежащие выборочной проверке, распределяются по двум слоям. Учитывая то, что распределение единиц наблюдения по стоимостным показателям обычно является заметно асимметричным (преобладание документов с относительно небольшими стоимостными характеристиками), граница между слоями определяется как удвоенная средняя стоимостная характеристика документа [18].

Расслоение, естественно, связано с некоторыми усложнениями, как при извлечении выборки, так и при ее анализе. К тому же здесь необходима некоторая дополнительная информация. Однако почти всегда расслоение приводит к повышению точности полученных результатов по сравнению с оценками, полученными без расслоения.

Применение стратифицированного отбора, как правило, сокращает стандартное отклонение выборочной средней (суммы) и тем самым повышает точность оценивания без увеличения объема выборки и при одной и той же надежности. Иначе говоря, расслоение сокращает интервалы оценок параметров совокупности. Заметим, что стандартное отклонение наиболее заметно уменьшается при увеличении количества слоев до пяти. Далее эффект расслоения растет не так быстро. За пределами 20 слоев этот эффект пренебрежимо мал. Сказанное объясняет, почему в аудите часто применяют двухслойную выборку – затраты минимальные, а эффект ощутимый.

Аудитор, решив применить расслоение тестируемой совокупности, должен последовательно определить число слоев, границы между слоями, общий объем выборки и распределить этот объем между слоями.

Число слоев может быть установлено аудитором на основе собственного понимания структуры тестируемой совокупности, опыта или интуиции, иными словами, достаточно произвольно. Соответственно устанавливаются и границы интервалов. В то же время, если единицы наблюдения ранжированы по величине стоимостных характеристик документов и установлено число слоев, то расслоение можно осуществить исходя из условия примерного равенства сумм в каждом слое [17].

После того как установлено число слоев и границы интервалов, необходимо общее количество единиц наблюдения в выборке (объем выборки) распределить по слоям. Как уже отмечалось ранее, распределить это объем можно пропорционально количеству документов в слое и не пропорционально.

Цели, порядок и методики оценивания в принципе остаются такими же, что и для выборки без расслоения – получение точечных и интервальных оценок средней и суммы стоимостных характеристик.

Рассмотрим отбор единиц наблюдения непропорциональный размеру слоя в совокупности. Оценка средней для совокупности находится как средняя взвешенная из выборочных средних по слоям. Причем в качестве весов берутся количества документов в каждом из них.

Пусть общий размер выборки распределен по ***m*** слоям. Тогда средняя для слоя ***k*** составит:

$$\tilde{x}_k = \frac{\sum_{i=1}^{n_k} x_i}{n_k},$$

где

$x_i$ – величина стоимостной характеристики ***i***-й выборочной единицы наблюдения, руб.;

$n_k$ – количество документов в выборке из слоя ***k***, ед.

Общая средняя для выборок по всем слоям:

$$\tilde{x}_s = \frac{\sum_{k=1}^{m} \tilde{x}_k N_k}{N}, \tag{6.4}$$

где

$N_k$ – количество документов в слое ***k***, ед.;

***N*** – размер совокупности, ед.

По определению $N = \sum_{k=1}^{m} N_k$ и $N = \sum_{k=1}^{m} n_k$ .

Параметр $\tilde{x}_s$ рассматривается как точечная оценка средней для совокупности. Заметим, что в общем случае средняя $\tilde{x}_s$ отличается от выборочной средней с весами, равными размерам выборок по слоям. Обозначим последнюю как $\tilde{x}_{sk}$ :

$$\tilde{x}_{sk} = \frac{\sum_{k=1}^{m} \tilde{x}_k n_k}{n}. \tag{6.5}$$

Известно, что дисперсия суммы независимых случайных величин (в нашем случае – выборочных средних по слоям) равна сумме дисперсий. Отсюда следует, что стандартная ошибка выборочной средней с учетом размеров слоев определяется в виде средней взвешенной стандартных ошибок по слоям:

$$\mu = \sqrt{\sum_{k=1}^{m} \frac{\sigma_k^2}{n_k} \left(\frac{N_k}{N}\right)^2} = \frac{1}{N}\sqrt{\sum_{k=1}^{m} \frac{\sigma_k^2}{n_k} N_k^2}. \tag{6.6}$$

Стандартное отклонение выборочных данных из слоя ***k***:

$$\sigma_k = \sqrt{\frac{\sum_{i=1}^{n_k} (x_i - \tilde{x}_k)^2}{n_k - 1}}.$$

В свою очередь стандартную ошибку выборочной средней для бесповторного отбора (с учетом ПКС) находим по формуле:

$$\mu = \frac{1}{N}\sqrt{\sum_{k=1}^{m}\frac{\sigma_k^2}{n_k}N_k^2(1-\frac{n_k}{N_k})}.$$

Разберем пример. Размер совокупности $N = 3700$. Массив данных разделен на четыре слоя по величине стоимостной характеристики. Необходимо по выборочным данным найти точечные и интервальные оценки суммы стоимостных показателей.

Допустим, что аудитор располагает некоторыми оценками стандартных отклонений по слоям. Из каждого слоя произведена выборка. Общий объем выборки $n = 200$. Распределение объема выборки по слоям ($n_k$) произведено с учетом стоимостного фактора.

Размеры слоев по числу документов, объемы выборки по слоям и другие данные, необходимые для расчета выборочной средней $x_s$ и стандартной ошибки средней $\mu$, показаны в табл. 6.3.

Точечная оценка средней и суммы для совокупности:

$$\tilde{x}_s = \frac{243\ 600}{3\ 700} = 65,8 \text{ тыс. руб.}$$

**Таблица 6.3**

**Расчет величин необходимых для определения выборочной средней и стандартной ошибки средней**

*Тыс. руб.*

| Слои | $N_k$ | $n_k$ | $\tilde{x}_k$ | $\tilde{x}_k N_k$ | $\sigma_k$ | $\frac{\sigma_k}{N_k}N_k^2$ |
|---|---|---|---|---|---|---|
| 0,1-20 | 2 000 | 12 | 9,8 | 19 600 | 11 | $40,333 * 10^6$ |
| 20-100 | 1 000 | 35 | 68 | 68 000 | 45 | $57,857 * 10^6$ |
| 100-400 | 500 | 72 | 236 | 11 800 | 97 | $32,67 * 10^6$ |
| 400-1000 | 200 | 81 | 721 | 144 200 | 152 | $45,637 * 10^6$ |
| Всего | 3 700 | 200 | - | 243 600 | - | $176,797 * 10^6$ |

$$\tilde{Q} = \sum_{k=1}^{m}\tilde{x}_k N_k = 243\ 600 \text{ тыс. руб.}$$

Перейдем к определению стандартной ошибки средней. Согласно формуле (6.6) получаем:

$$\mu = \frac{1}{3700}\sqrt{176,797 * 10^6} = 3,6 \text{ тыс. руб.}$$

Пусть доверительная вероятность равна 90%, тогда:

$$\Delta = z\mu = 1,64 * 3,6 = 5,9;$$

$$Q = 243600 \pm 5,9 * 3700 \text{ тыс. руб.}$$

Теперь есть основание утверждать, что оценка общей суммы для совокупности находится в пределах от 222 до 265 млн. руб.

Как уже отмечалось выше, преимущество стратифицированного отбора заключается в более точной оценке предельной ошибки выборки. Сказанное можно продемонстрировать даже на примере двухслойной выборки, о которой говорилось ранее.

Необходимо выборочным путем определить среднюю величину стоимостных показателей в совокупности документов. Общее количество документов равно 1 500. Случайным образом отобрано без расслоения 150 документов (выборка 1).

В порядке эксперимента выборочное исследование было повторено. Теперь документы были разделены на две группы. В первую вошли документы со стоимостями до 100 тыс. руб., во вторую – превышающие эту границу. Из первой группы отобрано 50 ед. (выборка 2, слой 1), из второй – 100 ед. (выборка 2, слой 2). Таким образом, в первом варианте применен нестратифицированный отбор, во втором – двухслойная выборка. Сравним результаты оценивания, сопоставив только предельные ошибки выборочной средней, полученные по двум вариантам.

Выборочные средние и стандартные отклонения (без ПКС) и другие данные, необходимые для получения соответствующих оценок для двух вариантов условий, показаны в табл. 6.4.

**Таблица 6.4**

**ДАННЫЕ ДЛЯ РАСЧЕТА ТОЧЕЧНЫХ И ИНТЕРВАЛЬНЫХ ОЦЕНОК**

| Показатель | Выборка 1 | Выборка 2 слой 1 | Выборка 2 слой 2 |
|---|---|---|---|
| $N_k$ | - | 950 | 550 |
| $n_k$ | 150 | 50 | 100 |
| $\tilde{x}_k$ | 220,4 | 54,5 | 315,6 |
| $\sigma, \sigma_k$ | 218 | 108 | 185 |

Определим интервал для средней по данным выборки 1 при условии, что доверительная вероятность равна 90% ($z = 1,64$). В итоге получаем:

$$\mu_1 = \frac{218}{\sqrt{150}} = 17,8;$$

$$\Delta_1 = z\mu_1 = 1,64 * 17,8 = 29,2 \text{ тыс. руб.}$$

$$\bar{x} = 220,4 \pm 29,2 \text{ тыс. руб.}$$

Таким образом, средняя для совокупности находится в пределах от 191,2 до 249,6 тыс. руб.

Перейдем к двухслойной выборке. Общая средняя (точечная оценка для совокупности):

$$\tilde{x}_s = \frac{\sum_{k=1}^{m}\tilde{x}_k N_k}{N} =$$

$$= \frac{54,5 * 50 + 315,6 * 100}{150} = 228,6 \text{ тыс. руб.}$$

Стандартные ошибки средней при расслоении на два слоя:

$$\mu_2 = \frac{1}{1500}\sqrt{\frac{108^2 * 950^2}{50} + \frac{185^2 * 550^2}{100}} =$$

$$= 11,8 \text{ тыс. руб.}$$

Предельная ошибка:

$$\Delta_2 = z\mu_2 = 1,64 * 11,8 = 19,3 \text{ тыс. руб.}$$

Средняя для совокупности лежит в интервале $228,6 \pm 19,3$, а именно от 209,3 до 247,9 тыс. руб. Интервал здесь заметно уже, чем в предыдущем варианте.

Рассмотрим теперь вариант отбора, пропорциональный размерам слоев в совокупности. Оценку средней для совокупности можно получить по формулам (6.4) и (6.5).

Очевидно, что результат будет одинаковым, так как при пропорциональном отборе:

$$\frac{n_k}{n} = \frac{N_k}{N},$$

отсюда

$$n_k = n\frac{N_k}{N}.$$

Не будем рассматривать во всех деталях порядок расчета границ интервала для средней и суммы. Вполне достаточно показать методику определения стандартной ошибки выборки без учета ПКС.

Для этого подставим значение $n_k$ в формулу (6.6):

$$\mu = \frac{1}{N}\sqrt{\sum_{k=1}^{m}\frac{\sigma_k^2}{n_k}N_k^2} = \sqrt{\frac{1}{n}\sum_{k=1}^{m}\sigma_k^2\frac{N_k}{N}} = \sqrt{\frac{\tilde{\sigma}^2}{n}},$$

где

$\tilde{\sigma}^2$ – средняя взвешенная величина из выборочных дисперсий по слоям;

$\sigma_k$ – стандартное отклонение в слое ***k***, руб.

Аналогичная по содержанию величина, но для бесповторного отбора (с учетом ПКС), находится следующим образом:

$$\mu = \sqrt{\frac{\tilde{\sigma}^2}{n}\left(1 - \frac{n}{N}\right)}.$$

## 6.3. Процедуры по существу (оценивание абсолютной ошибки)

Ранее обсуждавшиеся атрибутивные выборки позволяют оценивать количество документов в совокупности, содержащих те или иные отклонения, в том числе ошибки оформления. Вспомним, что к отклонениям могут быть отнесены и различного рода погрешности в размерах стоимостных показателей (тест на соответствие). Причем размер самих ошибок во внимание не принимался. Во многих случаях подобного рода контроль является недостаточным для аудитора, поскольку для анализа ему необходимы оценки обобщенного размера стоимостных ошибок для совокупности (тест по существу). Для решения такой задачи воспользуемся методикой точечного и интервального оценивания средних и сумм стоимостных показателей для совокупности, которая рассмотрена в предыдущем подразделе. В связи с этим несколько изменим постановку задачи и распространим методику оценивания не на сами показатели, а на обнаруженные в ходе тестирования отклонения от реальных или действительных данных. Такие данные получает аудитор при проверке и корректировке. Далее отклонения от реальных данных назовем ошибками.

Для аудитора представляет безусловный интерес оценка суммы ошибок для совокупности, а иногда и оценка среднего их размера. Эти параметры дают возможность характеризовать качество представленной для тестирования документации и определить действительные размеры совокупности по стоимости с учетом ошибок, обнаруженных аудитором. Для такого тестирования применяют два метода, которые условно назовем оценивание разности (через абсолютную ошибку) и оценивание отношения (через относительную ошибку). В первом случае в качестве исходных данных рассматриваются абсолютные ошибки, обнаруженные в представленных документах, во втором – относительные ошибки. В данном подразделе обсуждается первый метод, в следующем подразделе – второй.

К оцениванию размера ошибок на основе выборок аудитор прибегает при необходимости проверить массив первичных документов и при ожидании незначительных ошибок. Например, счетов-фактур, накладных, различных документов, подтверждающих оплату товаров с указанием цены и стоимости, и т.п.

Обсудим методики получения точечных и интервальных оценок. Точечные оценки, как известно, служат базой при определении интервальных оценок средней абсолютной ошибки и суммарной ошибки для совокупности. По-видимому, для аудитора наибольший интерес представляет последняя величина. Располагая такой оценкой, аудитор может с определенной степенью надежности скорректировать представленные ему данные, сформулировать определенные выводы для своего заключения.

Для записи формул, применяемых при точечном и интервальном оценивании, примем следующие обозначения:

$x_i$ – величина стоимостной характеристики в ***i***-й выборочной единице наблюдения (документе), руб.;

$y_i$ – скорректированный после тестирования (реальный) стоимостной показатель ***i***-й единицы наблюдения, руб.;

$e_i$ – обнаруженная аудитором величина абсолютной ошибки в ***i***-м документе, руб.;

$\tilde{e}$ – среднее значение ошибки в выборке, руб.;

***n*** – объем выборки, ед.;

***N*** – размер генеральной совокупности, из которой произведена выборка, ед.;

***E*** – точечная оценка суммы ошибок для совокупности, руб.;

$\bar{e}$ – интервальная оценка средней ошибки для совокупности, руб.;

***K*** – интервальная оценка сумм ошибок для совокупности, руб.

Величину абсолютной ошибки стоимостной характеристики ***i***-й единицы наблюдения в выборке получим как разность:

$$e_i = x_i - y_i.$$

Ошибка может быть нулевой, положительной или отрицательной, тогда как $x_i$ и $y_i$ только положительные величины.

Среднее значение абсолютной ошибки в расчете на единицу наблюдения в выборке:

$$\tilde{e} = \frac{\sum_{i=1}^{n} e_i}{n}.$$

Среднее значение абсолютной ошибки рассматривается как точечная оценка средней ошибки для совокупности. Экстраполируем полученное значение средней на весь объем совокупности. Получим точечную оценку суммы ошибок для совокупности:

$$E = \tilde{e}N.$$

Если ошибки случайны, то следует ожидать, что их сумма и среднее значение близки к нулю, поскольку положительные ошибки в какой-то мере компенсируются отрицательными. Заметное преобладание положительных (или наоборот отрицательных) ошибок может свидетельствовать о приписках и других преднамеренных искажениях в представленных документах.

При полном совпадении по суммам ошибок с противоположными знаками (разумеется, это крайний, но теоретически возможный случай) имеем: $\sum_{i=1}^{n} e_i = 0$ и $\tilde{e} = 0$. В этой или близкой ситуации точечная оценка средней абсолютной ошибки теряет свой смысл, но только в качестве индикатора наличия и размера ошибок. В самом деле, нулевая или близкая к нулевой точечная оценка вовсе не означает, что тестируемые документы не содержат ошибок. Однако сказанное не относится к интервальным оценкам. Также аудитор может исследовать абсолютный размер ошибок без учета знака, взяв разницу между стоимостной характеристикой и скорректированной суммой по модулю:

$$e_{i\ abs} = |x_i - y_i|.$$

Следовательно, среднее значение ошибки в данном случае определяется следующим образом:

$$\tilde{e}_{abs} = \frac{\sum_{i=1}^{n} e_{i\ abs}}{n}.$$

Перейдем к методике интервального оценивания. Средняя ошибка для совокупности лежит в интервале:

$$\bar{e} = \tilde{e} \pm \Delta,$$

где $\Delta$ – предельная ошибка выборочной средней для заданного уровня надежности или интервал точности (если $\tilde{e} = 0$, то $\bar{e} = \pm\Delta$), руб.

Как было показано ранее, выборочные средние распределяются по нормальному закону или близко к нему. Соответственно для определения доверительных границ следует воспользоваться квантилями нормального распределения ***z***.

Предельную ошибку находим стандартным путем:

$$\Delta = z\mu = z\frac{\sigma}{\sqrt{n}},$$

где

$\sigma$ – среднеквадратическое отклонение в выборке, руб.;

$\mu$ – стандартная ошибка выборочной средней, руб.;

***n*** – объем выборки, ед.

С учетом ПКС:

$$\Delta = z\frac{\sigma}{\sqrt{n}}\sqrt{1-\frac{n}{N}}.$$

С увеличением объема выборки размер предельной ошибки уменьшается.

Формула среднеквадратического отклонения уже была приведена ранее в предыдущих разделах. Перепишем ее с учетом нового обозначения стоимостной характеристики – ошибки:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(e_i - \tilde{e})^2}{n-1}}.$$

Границы интервальной оценки суммы ошибок для совокупности находятся элементарно:

$$K = (\tilde{e} \pm \Delta)N = E \pm \Delta N.$$

Как видим, с увеличением размера совокупности растет интервал для суммы ошибок. Если стандартная ошибка средней определяется без учета ПКС, то размер интервала строго пропорционален объему совокупности.

Остановимся на одном частном случае, который, впрочем, нельзя исключить. Пусть все ошибки одинаковы по размеру и имеют один знак (например, одинаковые приписки к цене). В этом случае среднеквадратическое отклонение равно нулю. Соответственно $\Delta = 0$ и, следовательно, интервала для оценки средней нет: $\bar{e} = \tilde{e}$.

В случае исследования абсолютных ошибок без учета знака, получаем следующие оценки:

$$\sigma_{abs} = \sqrt{\frac{\sum_{i=1}^{n}(e_{i\ abs} - \tilde{e}_{abs})^2}{n-1}};$$

$$\Delta_{abs} = z\frac{\sigma_{abs}}{\sqrt{n}} \text{ или } \Delta_{abs} = z\frac{\sigma_{abs}}{\sqrt{n}}\sqrt{1-\frac{n}{N}};$$

$$K_{abs} = (\tilde{e}_{abs} \pm \Delta_{abs})N = E_{abs} \pm \Delta_{abs}N.$$

Рассмотрим очередной пример. Общий объем документов в совокупности равен 2 000. Для проверки случайным образом отобрано 200 документов. Тестирование показало, что в 15 документах выборки обнаружены ошибки. Причем имели место как завышения, так и приуменьшения. Результаты корректировки в тыс. руб. ($y_i$) и остальные данные показаны в табл. 6.5.

Ошибки имеют разные знаки. Как показано в таблице, наибольшая по абсолютной величине отрицательная ошибка равна -0,3, а наибольшая положительная 0,4. Сумма положительных ошибок немного больше, чем отрицательных (соответственно 1,7 и -1,5), что может служить свидетельством об отсутствии преднамеренных ошибок. Имеет место частичное взаимное их погашение, что уменьшает их влияние на размер суммарной ошибки для совокупности.

**Таблица 6.5**

**ДАННЫЕ ДЛЯ ОЦЕНКИ ОШИБОК**

| $i$ | $x_i$ | $y_i$ | $e_i$ | $e_i^2$ |
|---|---|---|---|---|
| 1 | 3 | 2,8 | 0,2 | 0,04 |
| 2 | 5,4 | 5,6 | -0,2 | 0,04 |
| 3 | 1,7 | 1,8 | -0,1 | 0,01 |
| 4 | 10,9 | 10,8 | 0,1 | 0,01 |
| 5 | 9,9 | 9,5 | 0,4 | 0,16 |
| 6 | 3,7 | 3,4 | 0,3 | 0,09 |
| 7 | 16 | 16,1 | -0,1 | 0,01 |
| 8 | 8,2 | 8 | 0,2 | 0,04 |
| 9 | 7,5 | 7,6 | -0,1 | 0,01 |
| 10 | 12,3 | 12,6 | -0,3 | 0,09 |
| 11 | 6,4 | 6 | 0,4 | 0,16 |
| 12 | 7,8 | 8 | -0,2 | 0,04 |
| 13 | 2,1 | 2,3 | -0,2 | 0,04 |
| 14 | 3,3 | 3,2 | 0,1 | 0,01 |
| 15 | 9 | 9,3 | -0,3 | 0,09 |
| Итого | 107,2 | 107 | 0,2 | 0,84 |

По данным таблицы находим среднюю ошибку (точечную оценку абсолютной ошибки для совокупности):

$$\tilde{e} = \frac{0,2}{200} = 0,001 \text{ тыс. руб.}$$

Иначе говоря, на одну единицу наблюдения в выборке приходится ошибка равная 1 руб. Экстраполируя полученную среднюю на все 2 000 документов, получим точечную оценку суммы ошибок для совокупности:

$$E = 0,001 * 2000 = 2 \text{ тыс. руб.}$$

Перейдем к определению более важных для аудитора интервальных оценок. Начнем с расчета стандартной и предельной ошибок. Для этого найдем сумму квадратов отклонений: $\sum_{i=1}^{n} e_i^2 = 0,84$ . Определим среднеквадратическое отклонение:

$$\sigma = \sqrt{\frac{0,84 - 200 * 0,001^2}{200 - 1}} = 0,065 \text{ тыс. руб.}$$

Примем доверительную вероятность на уровне 90%, тогда $z = 1,64$ . В этом случае предельная ошибка средней составит:

$$\Delta = 1,64 * \frac{0,065}{\sqrt{200}} = 0,0075 \text{ тыс. руб.}$$

Интервал для средней ошибки в рублях в расчете на один документ составляет $\bar{e} = 1 \pm 7,5$ руб., иначе говоря, ошибка средней для совокупности при принятом уровне доверительной вероятности лежит в пределах от -6,5 до 8,5 руб.

В свою очередь интервал для суммы ошибок во всех документах тестируемой совокупности определяется как:

$$K = 2 \pm 0,0075 * 2000 \text{ тыс. руб.}$$

Аудитор с вероятностью 90% может утверждать, что суммарная ошибка находится в пределах от –13 до 17 тыс. руб. У аудитора теперь есть объективное основание для суждения о представленных ему документах.

Если приведенные выше параметры определить с учетом ПКС, то получим:

$$\Delta = 1,64 * \frac{0,065}{\sqrt{200}} * \sqrt{1 - \frac{200}{2000}} = 0,007 \text{ тыс. руб.};$$

$$\bar{e} = 1 \pm 7 \text{ руб.}$$

Окончательно получаем следующий интервал:

$$K = 2 \pm 0,007 * 2000 \text{ тыс. руб.}$$

Полученные интервалы немного уже, чем найденные ранее. Так, интервал для суммы ошибок с учетом ПКС теперь составляет от –12 до 16 тыс. руб.

Если аудитора интересует только верхняя граница, которую с заданной вероятностью ошибка в генеральной совокупности не превысит, то при расчете предельной ошибки нужно использовать односторонний параметр нормального распределения. Введем новые обозначения. Отныне двухсторонний параметр нормального распределения будет обозначаться как $z_{\alpha/2}$ , а односторонний – $z_\alpha$ . Если же, исходя из контекста выражения, критерий оценки (односторонний или двухсторонний) не имеет значения, то будет использоваться просто $z$. Использование одностороннего параметра в аудите более распространено в силу того, что аудитору чаще всего требуется найти только верхнюю границу ошибки, которая сравнивается с допустимым размером (уровнем существенности). Вероятности при одностороннем интервале и двухстороннем связаны следующей зависимостью:

$$F_X(z_{\alpha/2}) = 2F_X(z_\alpha) - 1.$$

По таблицам нормального распределения для одностороннего интервала находим соответствующую величину параметра для заданной в нашем примере доверительной вероятности. При уровне надежности 90% величина $z_\alpha$ равна 1,28. Таким образом, оценки предельных ошибок по условиям предыдущего примера будут следующими.

Без учета ПКС:

$$\Delta = 1,28 * \frac{0,065}{\sqrt{200}} = 0,0059 \text{ тыс. руб.}$$

С учетом ПКС:

$$\Delta = 1,28 * \frac{0,065}{\sqrt{200}} * \sqrt{1 - \frac{200}{2\,000}} = 0,0055 \text{ тыс. руб.}$$

Соответственно верхние границы можно найти следующим образом.

Без учета ПКС:

$$K = 2 + 0,0059 * 2000 = 13,8 \text{ тыс. руб.}$$

С учетом ПКС:

$$K = 2 + 0,0055 * 2000 = 13 \text{ тыс. руб.}$$

Как видим, верхняя граница при одностороннем интервале ниже, чем при двухстороннем, что естественно для аудитора лучше. Большая точность верхнего предела вызвана максимальным снижением точности нижнего предела (он вообще ничем неограничен).

Также можем оценить абсолютные ошибки без учета знака по условиям того же примера. Найдем соответствующие оценки для одностороннего интервала с учетом ПКС.

Средняя ошибка (ошибки в документах суммируются по модулю):

$$\tilde{e} = \frac{3,2}{200} = 0,016 \text{ тыс. руб.}$$

Точечная сумма ошибок в совокупности:

$$E = 0,016 * 2000 = 32 \text{ тыс. руб.}$$

Среднеквадратическое отклонение:

$$\sigma = \sqrt{\frac{0,84 - 200 * 0,016^2}{200 - 1}} = 0,063 \text{ тыс. руб.}$$

Предельная ошибка выборки:

$$\Delta = 1,28 * \frac{0,063}{\sqrt{200}} * \sqrt{1 - \frac{200}{2000}} = 0,0054 \text{ тыс. руб.}$$

Верхняя граница для средней ошибки в совокупности:

$$\bar{e} = 16 + 5,4 = 21,4 \text{ руб.}$$

Верхняя граница для суммы абсолютных ошибок в совокупности:

$$K = 32 + 0,0054 * 2000 = 42,8 \text{ тыс. руб.}$$

Верхний предел при оценке ошибок по модулю естественно намного выше, чем с учетом знака. При этом среднеквадратическое отклонение всегда ниже при оценке по модулю, так как уменьшается размах вариации.

Рассмотрим теперь оценку риска выборки. Как известно, риск выборки может быть получен из выражения для верхней границы доверительного интервала. Выражая величину $z_\alpha$ и заменяя границу на допустимый размер ошибки, получаем

$$z_\alpha = \frac{\bar{s} - \tilde{e}}{\mu},$$

где $\bar{s}$ – допустимый размер ошибки в расчете на один документ совокупности, руб.

Посчитаем риск выборки для примера выше. Предположим сумма, проведенная по всем документам, составляет 2 млн. руб. Существенность установлена на уровне 1%. Таким образом, получаем

$$\bar{s} = \frac{2\ 000\ 000}{2\ 000} * 0{,}01 = 10 \text{ руб}.$$

Теперь можем найти параметр $z_\alpha$ для последующей оценки риска. С учетом знака при определении ошибок и поправки на конечность совокупности находим:

$$\mu = \frac{65}{\sqrt{200}} * \sqrt{1 - \frac{200}{2\ 000}} = 4{,}3 \text{ руб};$$

$$z_\alpha = \frac{10 - 1}{4{,}3} = 2{,}09.$$

По полученному значению $z_\alpha$ из таблиц нормального распределения находим значение риска выборки.

Риск выборки $r$ рассчитывается по следующей формуле:

$$r = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}} dx.$$

Для $z_\alpha$ равному 2,09 риск выборки составляет 1,82%. Таким образом, вероятность того, что абсолютная ошибка в рублях превысит допустимый размер ошибки в генеральной совокупности, не превышает 2%, что безусловно является низким уровнем риска.

Стратифицированный отбор при оценивании абсолютных ошибок аналогичен тому, что был рассмотрен в подразделе 6.2. Соответственно вместо $x_i, \tilde{x}_k, \tilde{x}_s, \tilde{x}_{sk}$ используются $e_i, \tilde{e}_k, \tilde{e}_s, \tilde{e}_{sk}$.

## 6.4. Процедуры по существу (оценивание относительной ошибки)

В предыдущем подразделе обсуждались методики оценивания ошибок и различного рода искажений, определяемых как разности показанных в документах данных и этих же показателей после их корректировки аудитором. Возможен и другой вариант измерения ошибок – в виде относительных величин. Обобщающей характеристикой таких ошибок является оценка доли суммарной ошибки в общей стоимости. Оценка доли позволяет найти и сумму абсолютных ошибок для совокупности. В статистической литературе такого рода выборочные оценки относительных величин называют оценками по отношению.

Итак, вместо абсолютных разностей $e_i$ определим их отношения для каждой единицы наблюдения в выборке:

$$a_i = \frac{e_i}{x_i}.$$

Введем несколько новых обозначений:

$\tilde{a}$ – выборочная оценка относительной ошибки;

$\tilde{x}$ – среднее значение стоимостной характеристики в выборке, руб.;

$\bar{X}$ – среднее значение стоимостной характеристики в совокупности, руб.;

$Q$ – сумма стоимостных показателей совокупности, руб.;

$E$ – точечная оценка суммы ошибок для совокупности, руб.;

$\bar{a}$ – интервальная оценка относительной ошибки для совокупности;

$K$ – интервальная оценка суммы ошибок для совокупности, руб.

Обобщенной выборочной характеристикой размера относительных ошибок является величина $\tilde{a}$, которую примем в качестве точечной оценки относительной ошибки для совокупности. Определим ее следующим образом:

$$\tilde{a} = \frac{\sum_{i=1}^{n} e_i}{\sum_{i=1}^{n} x_i} = \frac{\sum_{i=1}^{n} e_i / n}{\sum_{i=1}^{n} x_i / n} = \frac{\tilde{e}}{\tilde{x}}.$$

Оценку по отношению $\tilde{a}$ можно применить для определения точечной оценки абсолютной суммы ошибок для совокупности. Для этого необходима общая сумма стоимостных показателей. Искомая величина составит:

$$E = \tilde{a}Q = \tilde{a}\bar{X}N.$$

Перейдем к интервальным оценкам относительной ошибки и суммы ошибок для совокупности. Прежде всего необходимо определить стандартную ошибку оценки по отношению, а затем предельную ошибку выборки. Однако величина $\tilde{a}$ определяется как отношение двух выборочных средних – $\tilde{e}$ и $\tilde{x}$, каждое из которых имеет свое распределение. В силу сказанного для расчета $\sigma$ нельзя воспользоваться рабочими формулами, которые получены для средней, а не отношения двух средних. Оказывается, что точной формулы для расчета стандартного отклонения по отношению нет.

Приходится пользоваться приближенной формулой [7]:

$$\sigma = \frac{1}{\bar{X}} \sqrt{\frac{\sum_{i=1}^{n} e_i^2 - 2\tilde{a}\sum_{i=1}^{n} e_i x_i + \tilde{a}^2 \sum_{i=1}^{n} x_i^2}{n-1}},$$

А с учетом ПКС имеем:

$$\sigma = \frac{1}{\bar{X}} \sqrt{\frac{\sum_{i=1}^{n} e_i^2 - 2\tilde{a}\sum_{i=1}^{n} e_i x_i + \tilde{a}^2 \sum_{i=1}^{n} x_i^2}{n-1}\left(1 - \frac{n}{N}\right)}.$$

Предельную ошибку выборки находим по стандартной формуле. В итоге получим границы интервальной оценки по отношению:

$$\bar{a} = \tilde{a} \pm \Delta.$$

Экстраполируем эти границы на всю сумму стоимостных характеристик совокупности. Получим интервал для оценки абсолютной суммы ошибок в совокупности:

$$K = (\tilde{a} \pm \Delta)Q = E \pm \Delta\bar{X}N.$$

Как уже отмечалось, в определенных условиях $\sum_{i=1}^{n} e_i = 0$ соответственно $\tilde{a}$ и $E = 0$. Нулевая точечная оценка вовсе не указывает на отсутствие ошибок в совокупности. Поскольку $\Delta \neq 0$, то интервальные оценки в этой ситуации отличаются от нуля: $\bar{a} = \pm\Delta$, $K = \pm\Delta Q$.

На первый взгляд представляется, что оценивание суммарной ошибки на основе абсолютных разностей дает близкие результаты, что и оценивание по отношению. Однако это не так. Дело в том, что методика, базирующаяся на абсолютных разностях, не принимает во внимание такой важный фактор, как стоимостные характеристики документов. Вместе с тем, как правило, абсолютные размеры ошибок коррелируют с этими характеристиками.

Воспользуемся данными последнего примера предыдущего подраздела ($N = 2000$, $n = 200$) и определим точечные и интервальные оценки по отношению и оценки абсолютной суммы ошибок в совокупности. Исходные и необходимые для расчетов данные представлены в табл. 6.6.

**Таблица 6.6**

**ОТНОСИТЕЛЬНЫЕ ОШИБКИ В СТОИМОСТНЫХ ПОКАЗАТЕЛЯХ ВЫБОРКИ**

| $i$ | $x_i$ | $y_i$ | $e_i$ | $e_i^2$ | $a_i$ | $e_i x_i$ | $x_i^2$ |
|---|---|---|---|---|---|---|---|
| 1 | 3 | 2,8 | 0,2 | 0,04 | 0,067 | 0,6 | 9 |
| 2 | 5,4 | 5,6 | -0,2 | 0,04 | -0,037 | -1,08 | 29,2 |
| 3 | 1,7 | 1,8 | -0,1 | 0,01 | -0,059 | -0,17 | 2,9 |
| 4 | 10,9 | 10,8 | 0,1 | 0,01 | 0,009 | 1,09 | 118,8 |
| 5 | 9,9 | 9,5 | 0,4 | 0,16 | 0,04 | 3,96 | 98 |
| 6 | 3,7 | 3,4 | 0,3 | 0,09 | 0,081 | 1,11 | 13,7 |
| 7 | 16 | 16,1 | -0,1 | 0,01 | -0,006 | -1,6 | 256 |
| 8 | 8,2 | 8 | 0,2 | 0,04 | 0,024 | 1,64 | 67,2 |
| 9 | 7,5 | 7,6 | -0,1 | 0,01 | -0,013 | -0,75 | 56,3 |
| 10 | 12,3 | 12,6 | -0,3 | 0,09 | -0,024 | -3,69 | 151,3 |
| 11 | 6,4 | 6 | 0,4 | 0,16 | 0,063 | 2,56 | 41 |
| 12 | 7,8 | 8 | -0,2 | 0,04 | -0,026 | -1,56 | 60,8 |
| 13 | 2,1 | 2,3 | -0,2 | 0,04 | -0,095 | -0,42 | 4,4 |
| 14 | 3,3 | 3,2 | 0,1 | 0,01 | 0,030 | 0,33 | 10,9 |
| 15 | 9 | 9,3 | -0,3 | 0,09 | -0,033 | -2,7 | 81 |
| Итого | 107,2 | 107 | 0,2 | 0,84 | - | -0,68 | 1 000,4 |

Пусть общая сумма стоимостных характеристик тестируемых документов совокупности равна 13 млн. руб. Средняя для совокупности составит:

$$\bar{X} = \frac{Q}{N} = \frac{13000}{2000} = 6{,}5 \text{ тыс. руб.}$$

В 15 из 200 документов выборки обнаружены различного рода искажения. Как показано в столбце $a_i$ таблицы, относительные ошибки находятся в интервале от – 0,095 до 0,081. Относительная ошибка в выборке или точечная оценка по отношению для совокупности составит:

$$\bar{a} = \frac{0{,}2}{107{,}2} = 0{,}0019 .$$

Обобщенная относительная ошибка в величине стоимостного показателя равна 0,19%. Иначе говоря, на каждую тысячу рублей стоимостной характеристики приходится в среднем ошибка, которая чуть меньше двух рублей. Распространим эту оценку на всю суммарную стоимость документов в совокупности:

$$E = 0{,}0019 * 13\ 000 = 24{,}7 \text{ тыс. руб.}$$

Перейдем к определению интервальных оценок. Рассчитаем стандартное отклонение выборочной доли. По табл. 6.6 находим необходимые для расчета суммы:

$$\sum_{i=1}^{n} e_i^2 = 0{,}84,\quad \sum_{i=1}^{n} e_i x_i = -0{,}68,\quad \sum_{i=1}^{n} x_i^2 = 1000{,}4 .$$

Стандартное отклонение без учета ПКС:

$$\sigma = \frac{1}{6{,}5}\sqrt{\frac{(0{,}84 - 2 * 0{,}0019 * (-0{,}68) + 0{,}0019^2 * 1000{,}4)}{199}} = 0{,}01 .$$

Предельная ошибка выборки при доверительной вероятности 90% равна:

$$\Delta = z\frac{\sigma}{\sqrt{200}} = 1{,}64\frac{0{,}01}{\sqrt{200}} = 0{,}0012 .$$

Интервал оценки по отношению для совокупности составит:

$$\bar{a} = \tilde{a} \pm \Delta = 0{,}0019 \pm 0{,}0012 ,$$

т.е. от 0,07% до 0,31%.

Интервал для суммы ошибок в совокупности:

$$K = 24{,}7 \pm 0{,}0012 * 6{,}5 * 2000 ,$$

то есть от 9,1 тыс. до 40,3 тыс. руб.

С учетом ПКС последовательно получим:

$$\sigma = 0{,}0095 ;$$

$$\Delta = 0{,}0011 ;$$

$$\bar{a} = 0{,}0019 \pm 0{,}0011 ;$$

$$K = 24{,}7 \pm 0{,}0011 * 6{,}5 * 2000 .$$

Сравним результаты точечного и интервального оценивания сумм ошибок для совокупности, полученные двумя методами (на основе абсолютных разностей и на основе отношений).

Для точечной оценки по первому методу была получена оценка, равная 2 тыс. руб., по второму – 24,7 тыс. руб. Разница, как видим, существенная. Размеры интервалов сумм абсолютных ошибок для совокупности примерно одинаковы (30 и 31 тыс. руб.). Интервалы смещены относительно друг друга: от –13 до 17 тыс. руб. по первому методу и от 9,1 до 40,3 тыс. руб. по второму.

Рассмотренный выше метод оценивания ошибок по отношению не является единственно возможным. В практике зарубежного аудита [17] иногда вместо соотношения $a_i = \frac{e_i}{x_i}$ объектом анализа является выборочное соотношение $b_i = \frac{y_i}{x_i}$. Обобщающая характеристика для выборки или точечная оценка для совокупности находится как отношение:

$$\tilde{b} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{\tilde{y}}{\tilde{x}} .$$

Как и раньше, $y_i$ означает уточненное аудитором значение стоимостной характеристики.

Между двумя выборочными характеристиками относительных ошибок существует функциональная зависимость. Нетрудно доказать, что

$$\tilde{a} = \frac{\sum_{i=1}^{n} (x_i - y_i)}{\sum_{i=1}^{n} x_i} = 1 - \tilde{b} .$$

Стандартное отклонение отношения находим по приближенной формуле:

$$\sigma = \frac{1}{X}\sqrt{\frac{\sum_{i=1}^{n} y_i^2 - 2\tilde{b}\sum_{i=1}^{n} y_i x_i + \tilde{b}^2\sum_{i=1}^{n} x_i^2}{n-1}}.$$

Предельная ошибка выборочного отношения определяется стандартным путем. Точечная и интервальная оценки суммы ошибок составляют:

$$E = (1 - \tilde{b})Q;$$

$$K = [1 - (\tilde{b} \pm \Delta)]Q.$$

Таким образом, принципиально нового здесь ничего нет. Немного меняется лишь порядок расчета. Результат должен быть таким же, что и при применении $\tilde{a}$. Разница в результатах может быть только следствием округлений. В силу сказанного не будем больше останавливаться на рассмотренной выше методике.

Риск выборки также можно определить с помощью метода оценивания по отношению. В данном случае риск будет отражать вероятность того, что доля искажения в стоимостной характеристике документа превысит максимально допустимую величину.

Формула для расчета величины $z_\alpha$ необходимой для определения риска выборки выглядит следующим образом:

$$z_\alpha = \frac{s - \tilde{a}}{\mu},$$

где

$s$ – максимально допустимая доля искажения стоимости документа (уровень существенности);

$\mu$ – стандартная ошибка выборочной доли по отношению, руб.

Когда известна величина $z_\alpha$, риск выборки рассчитывается точно так же, как это было показано в предыдущих подразделах, посвященных нормальному распределению. В связи с этим пропустим демонстрацию определения риска на примере.

Дальнейшее развитие методики оценивания по отношению заключается в применении расслоенного отбора. Однако для реализации такой методики необходима дополнительная информация, что в рамках аудиторской проверки, скорей всего, сопряжено с определенными трудностями, а часто и вовсе представляется невозможным.

## 6.5. Процедуры по существу (монетарный метод)

Цель монетарной выборки остается такой же, что и при собственно-случайном и расслоенном отборе – расчет суммарной ошибки для совокупности (ошибка в сумме остатков на счетах, в суммарной стоимости товарных документов и т.д.) в виде точечной и интервальной оценок. Методика получения необходимых статистических параметров в принципе не изменяется.

При монетарной выборке документы отбирают случайным образом. Однако принцип равной вероятности здесь применяется не к документам (единицам наблюдения), а к отдельным денежным единицам. В результате чем больше стоимостная характеристика документа, тем больше вероятность для него попасть в выборку. Например, если стоимость, указанная в одном документе совокупности, равна 20 тыс. руб., а в другом – 2 тыс. руб., то вероятность оказаться в выборке у первого документа при монетарной выборке в 10 раз выше, чем у второго. При собственно случайном отборе эти вероятности должны быть одинаковыми. Очевидно, что при монетарном отборе внимание аудитора в большей мере сосредотачивается на документах с более высокими стоимостными характеристиками. Соответственно уменьшается риск пропустить значительную погрешность. Более того, такой метод отбора увеличивает охват по стоимости тестируемой совокупности без увеличения размера выборки.

Перейдем к оцениванию суммарной ошибки для совокупности на основе данных монетарной выборки в виде относительных ошибок, обнаруженных аудитором. Что касается использования абсолютных ошибок, то в литературе в этих случаях иногда предлагается применить методику определения точечных и интервальных оценок для совокупности, которая была рассмотрена в подразделе 6.3. С этим трудно согласиться. Дело в том, что выборочная средняя абсолютная ошибка при монетарном отборе оказывается смещенной (завышенной), так как она определяется в расчете на единицу наблюдения, а в выборке преобладают документы с высокими стоимостями. Причем следует ожидать, что имеет место корреляция размера абсолютной ошибки и стоимостной характеристики документа – чем выше стоимостная характеристика, тем больше абсолютная ошибка. В итоге получаемые оценки сумм ошибок для совокупности будут завышенными. Попутно заметим, что монетарный отбор приводит к смещению и при оценивании суммарной стоимости для совокупности. В связи со сказанным остановимся только на методике оценивания на основе относительных ошибок.

Пусть, как и раньше, $x_i$ и $y_i$ означают стоимостные характеристики документа в выборке – показанные в документе, и стоимости после корректировки аудитором. Обнаруженная абсолютная ошибка в документе равна $e_i = x_i - y_i$, обе переменные больше нуля. Ошибка может быть равна нулю, больше или меньше нуля. Ошибка в расчете на один рубль стоимостной характеристики документа составляет:

$$h_i = \frac{e_i}{x_i}.$$

По выборочным данным находим значения $h_i$ и затем их обобщенную оценку:

$$\tilde{h} = \frac{\sum_{i=1}^{n} h_i}{n}.$$

Эту величину можно трактовать как среднюю ошибку в расчете на один рубль стоимостной характеристики единицы наблюдения. В связи с этим ею можно воспользоваться для получения точечной оценки суммы ошибок для совокупности:

$$E = \tilde{h}Q.$$

Определим теперь стандартное отклонение выборочной средней без ПКС. Для этого воспользуемся уже известной рабочей формулой, где заменим $x_i$ на $h_i$:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} h_i^2 - nh^2}{n-1}}.$$

Предполагается, что выборочная средняя имеет распределение, близкое к нормальному. С учетом этого условия предельную ошибку выборки для средней находим стандартным путем:

$$\Delta = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

Предельная ошибка для суммы составит:

$$\Delta_s = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} Q.$$

Интервальная оценка суммарной ошибки для совокупности:

$$K = E \pm \Delta_s.$$

Разберем пример. На основе монетарной выборки из совокупности отобрано 100 документов. В 15 из них обнаружены ошибки в стоимостных показателях. Суммарная стоимость всех документов совокупности равна 2 млн. руб. Выборочные данные и показатели, необходимые для расчета предельной ошибки, приведены в табл. 6.7.

По данным таблицы находим средний размер обнаруженных в выборке ошибок:

$$\bar{h} = \frac{0,4099}{100} = 0,0041.$$

Точечная оценка суммы ошибок для совокупности:

$$E = 0,0041 * 2000 = 8,2 \text{ тыс. руб.}$$

Стандартное отклонение в выборке:

$$\sigma = \sqrt{\frac{0,0459 - 100 * 0,0041^2}{100 - 1}} = 0,0211.$$

**Таблица 6.7**

**ДАННЫЕ ДЛЯ РАСЧЕТА ПРЕДЕЛЬНОЙ ОШИБКИ МОНЕТАРНЫМ МЕТОДОМ**

| i | $x_i$ | $y_i$ | $e_i$ | $h_i$ | $h_i^2$ |
|---|---|---|---|---|---|
| 1 | 13,2 | 12,4 | 0,8 | 0,0606 | 0,0037 |
| 2 | 25,4 | 23,6 | 1,8 | 0,0709 | 0,005 |
| 3 | 11,7 | 12 | -0,3 | -0,0256 | 0,0007 |
| 4 | 10,9 | 10,3 | 0,6 | 0,055 | 0,003 |
| 5 | 2,9 | 2,8 | 0,1 | 0,0345 | 0,0012 |
| 6 | 3,7 | 3,4 | 0,3 | 0,0811 | 0,0066 |
| 7 | 17,9 | 16,1 | 1,8 | 0,1006 | 0,0101 |
| 8 | 18,2 | 17,4 | 0,8 | 0,044 | 0,0019 |
| 9 | 7,5 | 7,6 | -0,1 | -0,0133 | 0,0002 |
| 10 | 12,3 | 12,6 | -0,3 | -0,0244 | 0,0006 |
| 11 | 6,4 | 6,2 | 0,2 | 0,0313 | 0,001 |
| 12 | 17,8 | 17,1 | 0,7 | 0,0393 | 0,0015 |
| 13 | 2,1 | 2,3 | -0,2 | -0,0952 | 0,0091 |
| 14 | 3,3 | 3,2 | 0,1 | 0,0303 | 0,0009 |
| 15 | 9,5 | 9,3 | 0,2 | 0,0211 | 0,0004 |
| Итого | 162,8 | 156,3 | 6,5 | 0,4099 | 0,0459 |

Предельная ошибка средней и суммы ($z_{\alpha/2} = 1,64$):

$$\Delta = 1,64 \frac{0,0211}{\sqrt{100}} = 0,00346;$$

$$\Delta_s = 0,00346 * 2000 = 6,92 \text{ тыс. руб.}$$

Интервальная оценка суммы ошибок для совокупности:

$$K = 8,2 \pm 6,92 \text{ тыс. руб.}$$

Таким образом, аудитор, основываясь на выборке, с доверительной вероятностью 90% может утверждать, что в проверенной им совокупности суммарная ошибка лежит в широких пределах от 1,28 до 15,12 тыс. руб. Сузить интервал можно за счет снижения уровня доверительной вероятности. Если последнюю принять на уровне 85%, то интервал будет находиться в пределах 2,12-14,28 тыс. руб.

Теперь рассчитаем риск выборки для примера выше. Формула для расчета величины $z_\alpha$ в случае монетарного отбора выглядит следующим образом:

$$z_\alpha = \frac{s - \bar{h}}{\mu},$$

где $s$ – максимально допустимая доля искажения рубля (уровень существенности в долях).

Предположим, что аудитор установил существенность на уровне 1%. Тогда согласно уже рассчитанным выше данным получаем:

$$z_\alpha = \frac{0,01 - 0,0041}{\frac{0,0211}{\sqrt{100}}} = 2,8.$$

По таблицам нормального распределения находим, что данной величине $z_\alpha$ соответствует риск выборки равный 0,26%.

Рассмотрим случай стратифицированного непропорционального отбора. Формула средней для слоя $k$ в данном случае выглядит следующим образом:

$$\tilde{h}_k = \frac{\sum_{i=1}^{n_k} h_i}{n_k}.$$

Общая средняя для выборок по всем слоям:

$$\tilde{h}_s = \frac{\sum_{k=1}^{m} \tilde{h}_k N_k}{N}.$$

Стандартное отклонение выборочных данных из слоя $k$:

$$\sigma_k = \sqrt{\frac{\sum_{i=1}^{n_k} (h_i - \tilde{h}_k)^2}{n_k - 1}}.$$

В конечном итоге выражение для доверительного интервала суммы ошибок в совокупности выглядит следующим образом (стандартная ошибка с учетом ПКС):

$$K = E \pm \Delta_s = Q\tilde{h}_s \pm z_{\alpha/2} \frac{1}{N} \sqrt{\sum_{k=1}^{m} \frac{\sigma_k^2}{n_k} N_k^2 \left(1 - \frac{n_k}{N_k}\right)}.$$

Судя по ряду зарубежных источников [1; 2; 4; 14], монетарный метод достаточно популярен (во всяком случае, в литературе по аудиту).

Между тем, анализ показывает следующее.

В случае, когда размер ошибки не связан со стоимостью документа (например, ошибки в авансовых отчетах), применение монетарного метода не имеет достаточного статистического основания, поскольку отсутствуют какие-либо данные о характере распределения относительной ошибки $h_i = \frac{e_i}{x_i}$, которая используется в «монетарном» методе при определении точечной оценки ошибки генеральной совокупности. Очевидно, в подобном случае более оправданным является применение методики оценивания через абсолютную ошибку.

В случае, когда размер ошибки связан со стоимостью документа (например, формальные ошибки в обязательных реквизитах счетов-фактур), размер относительной ошибки $h_i = 1$ (так как $e_i = x_i$). Тогда наряду с монетарным методом для проверок «по существу» может быть применена и методика оценивания суммы ошибок через долю отклонений, о которой речь пойдет в последующих подразделах. Это сводит к нулю преимущества монетарного метода вследствие большей сложности формирования выборки при его применении. Применение же монетарного метода при неоднородной стоимости документов генеральной совокупности (коэффициент вариации более 30%), как показывает расчетный анализ, проведенный в [11], может привести к недопустимо высокой погрешности.

## 6.6. Процедуры на соответствие. Оценивание доли отклонений

К оцениванию отношений (долей) можно подойти, использовав методику, применяемую при оценивании по выборке средней для совокупности. Предварительно сделаем в методике оценивания некоторые уточнения. Разделим единицы наблюдения (документы) совокупности на две группы. Пусть в первой из них документы обладают некоторыми свойствами. Например, содержат приписки, ошибки, неправильно оформлены и т.д. Подобного рода случаи в предыдущих разделах были названы отклонениями. Во второй группе документов такие свойства не наблюдаются. Необходимо выборочным путем оценить для тестируемой совокупности долю документов с отклонениями. Иначе говоря, это обычная проверка на соответствие, с которой мы сталкиваемся при атрибутивных выборках.

Для начала положим, что оценивание производится без расслоения совокупности. Пусть количество документов первой группы в совокупности равно ***M***. Величина ***M*** неизвестна и ее необходимо оценить по выборке.

Доля документов этой группы в совокупности соответствует отношению:

$$P = \frac{M}{N},$$

где ***N*** – размер совокупности.

Для применения методики, обсуждавшейся в предыдущем параграфе, положим, что если документ относится к первой группе (содержит отклонение), случайная переменная равна единице, $x_i = 1$, в противоположном случае $x_i = 0$. Таким образом, точечная оценка доли в совокупности, полученная по выборке, равна:

$$p = \tilde{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{m}{n},$$

где

***m*** – количество отклонений в выборке, ед.;

***n*** – объем выборки, ед.

Итак, доля документов с отклонениями в выборке равна p, а без отклонений $q = 1 - p$. Точечная оценка количества документов с отклонениями в совокупности составит:

$$M = pN.$$

Для получения интервальной оценки доли в совокупности необходимо определить стандартную и предельную ошибку выборочной средней при условии, что $x_i$ равно одному или нулю. Для этого воспользуемся одной рабочей формулой, в которой сумма квадратов отклонений определяется как $\sum_{i=1}^{n} x_i^2 - n\tilde{x}^2$. Рассмотри элементы этой разности.

Поскольку в оговоренных условиях:

$$\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i = a = np \quad \text{И} \quad n\tilde{x}^2 = n\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) = np^2,$$

то окончательно получаем:

$$\sum_{i=1}^{n} x_i^2 - n\tilde{x}^2 = np - np^2 = np(1-p) = npq.$$

В итоге стандартное отклонение определяется как

$$\sigma = \sqrt{\frac{npq}{n-1}}.$$

В свою очередь стандартная ошибка доли с учетом поправки на конечность совокупности (ПКС) составит:

$$\mu = \sqrt{\frac{pq}{n}\left(\frac{N-n}{N-1}\right)} \cong \sqrt{\frac{pq}{n}\left(1 - \frac{n}{N}\right)}.$$

При большом размере совокупности и относительно небольшом объеме выборки поправкой на конечность совокупности можно пренебречь. Без ПКС получаем:

$$\mu = \sqrt{\frac{pq}{n}}.$$

Предельную ошибку отношения находим стандартным путем:

$$\Delta = z\mu.$$

Границы интервалов оценок для отношения ($\bar{p}$) и количества документов с отклонениями в совокупности (***D***) определяем как

$$\bar{p} = p \pm z_{\alpha/2}\mu;$$

$$D = N(p \pm z_{\alpha/2}\mu) = M \pm z_{\alpha/2}\mu N. \qquad (6.7)$$

Пример. Из совокупности, содержащей 3 тыс. документов, случайным образом отобрано 200 ед. Оказалось, что 18 из них имеют отклонения. Итак: $N = 3000$, $n = 200$, $a = 18$. Необходимо по выборке оценить количество отклонений в тестируемой совокупности.

Точечные оценки доли и количества документов с отклонениями составят:

$$p = \frac{18}{200} = 0{,}09;$$

$$M = 0{,}09 * 3000 = 270.$$

Для получения интервальных оценок определим стандартную ошибку доли (с учетом ПКС):

$$\mu = \sqrt{\frac{0{,}09 * 0{,}91}{200} * \frac{3\,000 - 200}{3\,000 - 1}} = 0{,}0195.$$

Без учета ПКС стандартная ошибка доли чуть больше 0,02.

Интервальную оценку доли найдем при уровне доверительной вероятности 85% ($z_{\alpha/2} = 1{,}44$):

$$\bar{p} = 0{,}09 \pm 1{,}44 * 0{,}0195 = 0{,}09 \pm 0{,}028.$$

В свою очередь интервал для количества отклонений составит:

$$D = 270 \pm 1,44 * 0,0195 * 3000 = 270 \pm 84.$$

Таким образом, аудитор с вероятностью 85% может утверждать, что в тестируемой совокупности содержится от 6,2% до 11,8% документов с отклонениями. Соответственно, в совокупности ожидается от 186 до 354 таких документов.

Теперь рассмотрим ситуацию, когда, задавшись уровнем существенности, аудитора интересует риск выборки. Для этого возьмем выражение для верхней границы интервала оценки для отношения, выразим из нее предел интеграла Лапласа (величину $z_\alpha$) и заменим границу на максимально допустимую долю отклонений (уровень существенности). В итоге получим:

$$z_\alpha = \frac{s - p}{\mu},$$

где $s$ – принятый уровень существенности (максимально допустимая доля отклонений).

Возьмем пример, рассмотренный выше, и предположим, что аудитор установил существенность на уровне 10%. Тогда величину $z$ можно найти следующим образом:

$$z_\alpha = \frac{0,1 - 0,09}{0,0195} = 0,5128.$$

По таблицам нормального распределения находим, что данной величине соответствует риск выборки равный 30%, что безусловно является очень высоким показателем. Это предполагает, что с вероятностью 30% искажение в генеральной совокупности превысит заданный уровень существенности. Таким образом, аудитору требуется провести дополнительные аудиторские процедуры по снижению риска выборки или принять решение о том, что проверяемая система внутреннего контроля функционирует неэффективно.

## 6.7. Процедуры по существу. Оценивание суммы ошибок через долю отклонений

Оценивание суммы ошибок с использованием атрибутивных выборок, как уже известно, возможно только при условии, что наличие атрибута (отклонения) делает всю сумму документа ошибочной. Данное условие позволяет использовать среднюю стоимость документа для обобщения ошибок в генеральной совокупности. Для этого средняя стоимость умножается на количество ошибочных документов, найденное посредством методики, описанной в предыдущем подразделе, где рассматривались процедуры на соответствие, основанные на нормальном распределении. Таким образом, формула (6.7) для нахождения интервала ошибочных документов модифицируется следующим образом:

$$K = N\bar{J}(p \pm z_{\alpha/2}\mu) = M\bar{J} \pm z_{\alpha/2}\mu N\bar{J} = E \pm \Delta N\bar{J},$$

где $\bar{J}$ – средняя стоимость документа, руб.

Все остальные параметры вышеприведенной формулы ($p$, $z_{\alpha/2}$ и $\mu$) определяется точно так же, как это было уже описано в предыдущем подразделе.

Сразу разберем пример. Аудитор проверяет совокупность счетов-фактур. Объем совокупности – 1 000 документов. Для выборочного тестирования случайным образом отобрано 100 документов. В двух документах найдены несоответствия.

Таким образом, точечная оценка доли составит:

$$p = \frac{4}{100} = 0,04.$$

Точечная оценка количества документов с отклонениями:

$$M = 0,04 * 1000 = 40 \text{ ед}.$$

В свою очередь стандартная ошибка доли с учетом ПКС будет равна:

$$\mu = \sqrt{\frac{pq}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0,04 * 0,96}{100} * \left(1 - \frac{100}{1000}\right)} = 0,019.$$

Доверительная вероятность установлена на уровне 95% ($z_{\alpha/2} = 2$). Соответственно предельная ошибка выборки находится следующим образом:

$$\Delta = 2 * 0,019 = 0,038.$$

Предположим, общая стоимость всех документов – 1,5 млн. руб. Средняя стоимость одного документа:

$$\bar{J} = \frac{1\,500\,000}{1\,000} = 1\,500 \text{ руб}.$$

Теперь мы имеем все данные, чтобы посчитать доверительный интервал для суммы ошибок в совокупности:

$$K = E \pm \Delta N\bar{J} = 40 * 1\,500 \pm 0,038 * $$
$$* 1\,000 * 1\,500 = 60\,000 \pm 57\,000.$$

Таким образом, аудитор с вероятностью 95% может утверждать, что ошибка в генеральной совокупности лежит в интервале от 3 до 117 тыс. руб.

Можно рассмотреть вариант с односторонним критерием оценки. Пусть аудитора интересует только верхняя граница, которую ошибка в 95% случаях не должна превысить. Тогда $z_\alpha = 1,65$ и предельная ошибка выборки составит:

$$\Delta = 1,64 * 0,019 = 0,031.$$

Следовательно, верхний предел ошибки составит:

$$K = E + \Delta N\bar{J} = 60\,000 + 0,031 * $$
$$* 1\,000 * 1\,500 = 106\,500 \text{ руб}.$$

## 6.8. Процедуры по существу (оценивание суммы ошибок через долю отклонений с учетом дисперсии)

Известно, что использование классических методов возможно лишь при однородной стоимости элементов генеральной совокупности. В [9], например, показано, что при значениях коэффициента вариации стоимости элементов генеральной совокупности, превышающих 0,2 ÷ 0,3, погрешность может быть весьма существенной. В подобном случае все без исключения литературные источники, в которых рассматриваются вопросы статистических выборочных проверок, указывают, что генеральную совокупность следует стратифицировать по стоимости элементов, в результате чего может быть достигнута однородность каждой страты.

Следует отметить, что подобная рекомендация далеко не всегда бывает эффективной, поскольку, как показывает практика, в большинстве случаев для достижения однородности элементов генеральную совокупность приходится стратифицировать несколько раз, что усложняет и формирование выборок и оценку результатов. Рассмотрим это на наглядном примере. Пусть некая генеральная совокупность состоит из 10 элементов, значения которых составляют 10; 20; 30; 40; 50; 60; 70; 80; 90; 100. Простейшие вычисления показывают, что для данной генеральной совокупности генеральная средняя равна 55, генеральная дисперсия – 825, среднеквадратическое отклонение – 28,7. Коэффициент вариации, таким образом, составит $28{,}7 / 55 = 0{,}52\ (52\ \%)$. Поскольку это значение превосходит рекомендуемое, то разбиваем генеральную совокупность на две страты (со значениями элементов в первой страте 10; …50 и во второй – 60; …100). Для этих страт получаем значения коэффициентов вариации, равные 0,47 (47%), что заставляет нас продолжить процесс стратификации.

Альтернативой может быть метод, который позволяет оценить ожидаемую ошибку в генеральной совокупности и оценить риск выборки с учетом дисперсии стоимости элементов генеральной совокупности, что позволяет избежать необходимости ее стратификации. Подобному методу присущи те же недостатки, что и любому методу, основанному на атрибутивных выборках при использовании в процедурах по существу – ограниченное применение в силу обязательной связи между стоимостью документа и величиной ошибки.

Рассмотрим данный метод. Исходные данные: имеется генеральная совокупность неких элементов. Элементами могут быть либо отраженные в регистре бухгалтерского учета операции, относящиеся к обороту какого-либо счета (например, расходы, отраженные проводками Д-т 20 – К-т 60); либо первичные документы, подлежащие отражению на каком-либо счете (например, полученные от поставщиков товаров, работ, услуг счета-фактуры, подлежащие отражению проводками Д-т 68 – К-т 19). В генеральной совокупности может находиться какое-то количество «отмеченных» элементов (элементов, содержащих искажения). Будем исходить из того, что при наличии в элементе искажения ошибочной будет являться вся сумма, отраженная в учете в соответствии с данным элементом. Подобные случаи, когда ошибочной является вся учетная сумма, проведенная по первичному документу, обычно имеют место при формальных ошибках (например, в счетах-фактурах), неправильном или безосновательном отражении операций, отражении незаконных операций и др.

Пусть $N$ (в натуральных единицах) – объем генеральной совокупности элементов (операций либо первичных документов, относящихся к обороту счета бухгалтерского учета – накладных, счетов-фактур и т.п.). Пусть $J$ (руб.) – суммарная стоимость элементов, составляющих генеральную совокупность, тогда

$$J = \sum_{i=1}^{N} j_i ,$$

где $j_i$ – стоимость $i$-го элемента генеральной совокупности, руб.

Предлагаемый метод основан на том, что каждый элемент генеральной совокупности имеет два признака случайности (количественный и качественный):

- размер (стоимость, руб.);
- отмеченность (наличие искажений).

Известным образом определим наиболее вероятное количество отмеченных элементов (элементов, содержащих искажения) в генеральной совокупности. Для этого сформируем случайную выборку элементов объемом $n$ и проверим ее.

Пусть $m$ – количество отмеченных элементов (элементов, содержащих искажения) в выборке.

Тогда $p = \frac{m}{n}$ – относительная частота появления элементов, содержащих искажения, в выборке объемом $n$ или точечная оценка доли согласно терминологии принятой в предыдущих подразделах.

Наиболее вероятное количество элементов, содержащих искажения, в генеральной совокупности (обозначим его $M$), как известно, составит:

$$M = pN .$$

Поскольку мы априорно исходим из случайности распределения отмеченных элементов (элементов, содержащих искажения) в генеральной совокупности, то совокупность объемом $M$ элементов, содержащих искажения, применительно к количественному признаку можно рассматривать как случайную выборку. Обозначим выборочную среднюю выборки объемом $M$ через $\bar{k}$ (руб.).

Тогда наиболее вероятное значение суммарной стоимости элементов, составляющих выборку объемом $M$, составит:

$$E = M\bar{k} .$$

Очевидно, что $E$ – искомая ожидаемая ошибка генеральной совокупности (руб.) или точечная оценка ошибки в совокупности. Подобная формула оценки ожидаемой ошибки приведена в [8], но с учетом применения исключительно в однородной генеральной совокупности.

Для того, чтобы найти значение $\bar{k}$, вспомним, что наиболее вероятным значением генеральной средней является значение выборочной средней. Справедливо и обратное утверждение: наиболее вероятным значением выборочной средней является значение генеральной средней. Тогда в качестве наиболее вероятного значения выборочной средней $\bar{k}$ может быть принято известное нам значение генеральной средней:

$$\bar{J}\ \ (\bar{k} = \bar{J}),$$

где

$$\bar{J} = \frac{J}{N} = \frac{\sum_{i=1}^{N} j_i}{N} .$$

Для определения границ ошибки и риска выборки необходимо оценить дисперсию $\sigma_E^2$ случайной величины $E$, являющейся произведением двух случайных величин ($M$ и $\bar{k}$). Для этого оценим дисперсии указанных сомножителей.

Из теории вероятности известно, что дисперсия $\sigma_p^2$ ошибки доли может быть оценена из выражения:

$$\sigma_p^2 = \frac{pq}{n} .$$

Поскольку $N$ – величина постоянная, то оценка дисперсии случайной величины $M$ составит:

$$\sigma_M^2 = N^2 \sigma_p^2 .$$

Дисперсия выборочной средней $\sigma_{\bar{k}}^2$ может быть оценена по выборочной дисперсии $\sigma_k^2$:

$$\sigma^2_{\bar{k}} = \frac{\sigma^2_k}{M}.$$

Выборочная дисперсия $\sigma^2_k$, в свою очередь, может быть оценена по генеральной дисперсии $\sigma^2_J$ (смещенностью оценки пренебрегаем):

$$\sigma^2_k = \sigma^2_J.$$

Тогда для оценки дисперсии выборочной средней $\sigma^2_{\bar{k}}$ получаем выражение:

$$\sigma^2_{\bar{k}} = \frac{\sigma^2_J}{M}.$$

Генеральная дисперсия $\sigma^2_J$ в полученном выражении может быть найдена известным образом:

$$\sigma^2_J = \frac{\sum_{i=1}^{N}(J_i - \bar{J})^2}{N}.$$

Итак, случайная величина ***E*** является произведением двух независимых случайных величин (***M*** и $\bar{k}$), оценки дисперсий которых нам известны. Получим выражение для дисперсии произведения двух независимых случайных величин.

Как известно дисперсию случайной величины можно найти по следующей формуле:

$$\sigma^2_E = M_0(E^2) - M^2_0(E),$$

где $M_0(E^2)$ и $M_0(E)$ – математические ожидания случайных величин $E^2$ и $E$ соответственно.

Учитывая то, что математическое ожидание произведения двух независимых случайных величин равно произведению математических ожиданий этих величин, получаем:

$$\sigma^2_E = M_0(\bar{k}^2 M^2) - M^2_0(\bar{k}M) =$$

$$= M_0(\bar{k}^2)M_0(M^2) - M^2_0(\bar{k})M^2_0(M).$$

По определению дисперсии:

$$M_0(\bar{k}^2) = \sigma^2_{\bar{k}} + M^2_0(\bar{k})$$

и

$$M_0(M^2) = \sigma^2_M + M^2_0(M).$$

Теперь, подставляя данные выражения в выражение для $\sigma^2_E$, получаем:

$$\sigma^2_E = (\sigma^2_{\bar{k}} + M^2_0(\bar{k}))(\sigma^2_M + M^2_0(M)) - M^2_0(\bar{k})M^2_0(M).$$

Раскрывая скобки, получаем следующее выражение:

$$\sigma^2_E = \sigma^2_{\bar{k}}\sigma^2_M + \sigma^2_{\bar{k}}M^2_0(M) + M^2_0(\bar{k})\sigma^2_M +$$

$$+ M^2_0(\bar{k})M^2_0(M) - M^2_0(\bar{k})M^2_0(M) =$$

$$= \sigma^2_M\sigma^2_{\bar{k}} + \sigma^2_M M^2_0(\bar{k}) + \sigma^2_{\bar{k}}M^2_0(M).$$

Так как $\bar{k}$ представляет собой среднее арифметическое генеральной совокупности, то

$$M_0(\bar{k}) = \bar{k}.$$

Величина ***M*** является лишь наиболее вероятным количеством ошибочных элементов в генеральной совокупности, и поэтому мы можем получить только приближенную оценку ее математического ожидания:

$$M_0(M) \cong M.$$

Таким образом, оценка дисперсии случайной величины **E** может быть получена из зависимости:

$$\sigma^2_E = \sigma^2_M\sigma^2_{\bar{k}} + \sigma^2_M\bar{k}^2 + \sigma^2_{\bar{k}}M^2.$$

Среднеквадратическое отклонение случайной величины ***E***:

$$\sigma_E = \sqrt{\sigma^2_E}.$$

В итоге доверительный интервал для суммы ошибок в генеральной совокупности в случае двухсторонней оценки определяется следующим образом:

$$K = E \pm z_{\alpha/2}\sigma_E.$$

Оценку риска выборки произведем из следующих соображений. Риск выборки – это вероятность того, что ожидаемая ошибка (случайная величина ***E***) превысит уровень существенности ***S***, «применяемый» для данной генеральной совокупности (о «применяемом» уровне существенности см. в [10]).

Приравняем верхнюю границу доверительного интервала для случайной величины ***E*** (используется только односторонняя оценка) применяемому уровню существенности ***S***:

$$S = E + z_\alpha\sigma_E,$$

откуда получаем:

$$z_\alpha = \frac{S - E}{\sigma_E}.$$

Рассмотрим применение предложенного метода на примере и рассчитаем только риск выборки.

Аудитор проверяет правомерность предъявления налога на добавленную стоимость (НДС) к вычету. Объем генеральной совокупности ***N = 1 000*** счетов-фактур. Дебетовый оборот счета 68 в корреспонденции со счетом 19 – ***J = 3 000*** тыс. рублей. Применяемый уровень существенности ***s = 5%*** (тогда в рублях ***S = 150*** тыс. руб.). Объем выборки ***n = 100*** счетов-фактур. В выборке обнаружены два недостоверных счета-фактуры (***m = 2***).

Генеральная средняя:

$$\bar{J} = \frac{J}{N} = \frac{3\ 000\ 000}{1\ 000} = 3\ 000 \text{ руб}.$$

Генеральная дисперсия:

$$\sigma^2_J = \frac{\sum_{i=1}^{N}(J_i - \bar{J})^2}{N} = 2\ 500\ 000.$$

Точечная оценка доли:

$$p = \frac{m}{n} = \frac{2}{100} = 0{,}02.$$

Наиболее вероятное количество ошибок в генеральной совокупности:

$$M = Np = 1\ 000 * 0{,}02 = 20.$$

Дисперсия величины ***M***:

$$\sigma^2_M = N^2\frac{pq}{n} = 1\ 000^2 * \frac{0{,}02 * (1 - 0{,}02)}{100} = 196.$$

Дисперсия выборочной средней:

$$\sigma^2_{\bar{k}} = \frac{\sigma^2_J}{M} = \frac{2\ 500\ 000}{20} = 125\ 000.$$

Ожидаемая ошибка:

$$E = M\bar{k} = 20 * 3\ 000 = 60\ 000 \text{ руб}.$$

Дисперсия ожидаемой ошибки ***K***:

$$\sigma_E^2 = \sigma_M^2 \sigma_k^2 + \sigma_M^2 \bar{k}^2 + \sigma_k^2 M^2 =$$

$$= 196 * 125\ 000 + 196 * 3\ 000^2 +$$

$$+ 125\ 000 * 20^2 = 1\ 838\ 500\ 000 .$$

Среднеквадратическое отклонение:

$$\sigma_E = \sqrt{\sigma_E^2} = \sqrt{1\ 838\ 500\ 000} = 42\ 878 \text{ руб} .$$

Предел интеграла Лапласа:

$$z_\alpha = \frac{S - E}{\sigma_E} = \frac{150\ 000 - 60\ 000}{42\ 878} = 2,1 .$$

По таблицам нормального распределения находим: для $z_\alpha = 2,1$ риск выборки составляет 1,8%.

В случае, если риск выборки с точки зрения аудитора окажется чрезмерно высоким, может быть принято решение о расширении объема выборки или компенсации высокого риска за счет снижения риска, не связанного с выборкой, например, посредством привлечения более опытных аудиторов и т.п. (напомним, что аудиторский риск на уровне сальдо и оборотов по счетам состоит из риска выборки и риска, не связанного с выборкой). Рассмотрим на примере подобный случай.

Возьмем исходные данные предыдущего примера, но предположим, что применяемый уровень существенности ***s = 4%*** (в рублях ***S = 120*** тыс. руб.). Тогда предел интеграла Лапласа:

$$z_\alpha = \frac{S - E}{\sigma_E} = \frac{120\ 000 - 60\ 000}{42\ 878} = 1,4 .$$

При таком значении $z_\alpha$ риск выборки составит 8,1%, что по мнению аудитора может оказаться неприемлемым. Он принимает решение вдвое снизить риск посредством увеличения объема выборки.

Получим выражение для оценки требуемого объема выборки, полагая, что любая выборочная дисперсия приблизительно равна генеральной дисперсии и поэтому изменение объема выборки не оказывает существенного влияния на ее оценку. Таким образом, требуется выразить ***n***, которое содержится в знаменателе $\sigma_M^2$. В качестве исходной зависимости возьмем верхнюю границу доверительного интервала

$$S = E + z_\alpha \sigma_E ,$$

выразим среднеквадратическое отклонение

$$\sigma_E = \frac{S - E}{z_\alpha} ,$$

возведем обе части выражения в квадрат, чтобы получить дисперсию, и подставим в левую часть выражение, полученное выше для дисперсии произведения двух независимых случайных величин:

$$\sigma_M^2 \sigma_k^2 + \sigma_M^2 \bar{k}^2 + \sigma_k^2 M^2 = \frac{(S - E)^2}{z_\alpha^2} .$$

Вынесем за скобки величину $\sigma_M^2$, в знаменателе которой содержится требуемый параметр ***n***, а оставшееся слагаемое перенесем в правую часть:

$$\sigma_M^2 (\sigma_k^2 + \bar{k}^2) = \frac{(S - E)^2}{z_\alpha^2} - \sigma_k^2 M^2 .$$

Раскроем дисперсию величины ***M*** в левой части и оставшейся множитель перенесем в правую часть:

$$N^2 \frac{p(1 - p)}{n} = \frac{(S - E)^2}{z_\alpha^2 (\sigma_k^2 + \bar{k}^2)} - \frac{\sigma_k^2 M^2}{\sigma_k^2 + \bar{k}^2} ,$$

откуда получаем:

$$n = \frac{N^2 (\sigma_k^2 + \bar{k}^2)(p - p^2)}{\frac{(S - E)^2}{z_\alpha^2} - \sigma_k^2 M^2} .$$

Теперь подставим уже известные из первого примера значения параметров в вышеприведенную формулу, принимая $z_\alpha$, исходя из требуемого риска выборки ($z_\alpha = 1,75$ при риске выборки в 4%):

$$n = \frac{1000^2 * (125\ 000 + 3\ 000^2) * (0,02 - 0,02^2)}{\frac{(120\ 000 - 60\ 000)^2}{1,75^2} - 125\ 000 * 20^2} \approx 160 .$$

Таким образом, аудитору понадобится произвести добор:

$$\Delta_n = 160 - 100 = 60 .$$

Из всего вышесказанного следует, что в данном примере для снижения риска выборки с 8,1% до минимально приемлемого в 4%, аудитору необходимо сделать дополнительную выборку 60 счетов-фактур из генеральной совокупности.

# 7. МЕТОДЫ, ОСНОВАННЫЕ НА РАСПРЕДЕЛЕНИИ СТЬЮДЕНТА

## 7.1. Общая информация о распределении Стьюдента

Рассмотрим еще одно непрерывное распределение, которое может применяться в аудите в рамках выборочного исследования.

Анализируя случайные отклонения выборочной средней ***x*** от истинного среднего значения исследуемой случайной величины ***Y***, английский статистик В. Госсет (писавший под псевдонимом Стьюдент) получил следующий результат. Пусть $Y_0, Y_1, ..., Y_n$ – независимые стандартные нормальные случайные величины, такие что $Y_i \sim N(0,1),\ i = 0,1, ..., n$.

Тогда плотность распределения случайной величины ***t***, где:

$$t = \frac{Y_0}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} Y_i^2}}$$

описывается функцией:

$$f_t(y) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{y^2}{n}\right)^{-\frac{n+1}{2}} ,$$

где $\Gamma$ – гамма функция Эйлера.

Распределение получило название распределения Стьюдента с ***n*** степенями свободы. Пишут $t \sim t(n)$. Распределение абсолютно непрерывно. Функция плотности распределения Стьюдента не зависит от дисперсии $\sigma^2$ случайных величин ***Y***, кроме того является унимодальной и симметричной относительно точки $y = 0$.

Если $n \to \infty$, то функция плотности распределения Стьюдента стремится к функции плотности нормального распределения. На практике полагают, что распределение Стьюдента можно заменить нормальным при $n > 30$ .

Распределение Стьюдента используется в статистике для точечного оценивания, построения доверительных интервалов и тестирования гипотез, касающихся неизвестного среднего статистической выборки из нормального распределения. Также данное распределение может использоваться и в аудите при малых выборках.

## 7.2. Процедуры на соответствие и по существу, основанные на распределении Стьюдента (малые выборки)

Далеко не всегда аудитор может осуществить выборку в полном объеме. В силу разных причин, например нехватки времени, он ограничивается небольшим объемом выборки – менее 20 единиц. Это так называемые малые выборки. Аудитор и при анализе малых выборок в ряде ситуаций может быстро получить полезную информацию.

Как уже отмечалось ранее, согласно центральной предельной теореме, выборочные средние имеют нормальное или близкое к нему распределение. Однако это утверждение, справедливое для больших выборок, не относится к выборочным средним, извлеченным из малых выборок. Выходом является применение другого распределения, а именно – распределения нормированных отклонений выборочных средних от средней генеральной совокупности. Как было отмечено в предыдущем подразделе, это распределение получило название ***t***-распределение Стьюдента [3]. Распределение получено для случайной выборки из нормально распределенной генеральной совокупности. Оно зависит только от одного параметра – числа степеней свободы ***v***:

$$v = n - 1,$$

где ***n*** – размер выборки, ед.

Величина ***t*** находится как отношение разности выборочной и генеральной средней к стандартной ошибке выборки:

$$t = \frac{\tilde{x} - \bar{x}}{\mu}, \qquad (7.1)$$

где

$\tilde{x}$ – выборочная средняя;

$\bar{x}$ – средняя для совокупности (генеральная средняя);

$\mu$ – стандартная ошибка выборки.

Стандартная ошибка выборки находится как:

$$\mu = \frac{\sigma}{\sqrt{n}}.$$

Выборочное стандартное отклонение $\sigma$ определяется известным образом.

Как видим, величина ***t*** характеризует отклонение выборочной средней от средней совокупности, нормированное по стандартной ошибке выборки.

Распределение Стьюдента симметрично относительно нуля, как и нормальное распределение, но более пологое. Как уже было сказано в предыдущем подразделе, с ростом числа степеней свободы оно приближается к нормальному. Площадь под кривой распределения равна единице (или 100%). Площадь под кривой для значений, превышающих ***t*** (или меньше – ***t***), характеризует вероятность превзойти указанное значение. Обозначим эту вероятность как ***r*** (риск выборки).

Вернемся к формуле (7.1) и решим ее относительно средней генеральной совокупности:

$$\bar{x} = \tilde{x} \pm t\mu. \qquad (7.2)$$

Как видим, приведенная формула позволяет оценить границы для неизвестной средней генеральной совокупности. Границы определяются таким же способом, что и при оценивании интервалов на основе большой выборки. Соответствующие методики обсуждались в предыдущем разделе, посвященному нормальному распределению. Различие состоит только в том, что здесь вместо квантиля ***z*** нормального распределения используется величина ***t***.

С доверительной вероятностью, равной $1 - r$ , можно утверждать, что неизвестная генеральная средняя находится в пределах, определяемых формулой (7.2). Произведение $t\mu$ в этой формуле характеризует предельную ошибку выборочной средней. Чем больше доверительная вероятность, тем больше ***t*** и соответственно предельная ошибка выборки. Уровень доверительной вероятности можно указывать при записи величины ***t***. Например, для $r = 0{,}05$ запишем $t_{95}$ и т.д.

Разберем небольшой пример. Необходимо на основе выборки быстро оценить пределы, в которых находится средняя величина для совокупности чеков.

Для оценивания случайным образом отобрано 18 чеков со следующими стоимостными характеристиками (суммами).

**Таблица 7.1**

**СУММЫ ЧЕКОВ, ПОПАВШИХ В ВЫБОРКУ**

*Тыс. руб.*

| № | Сумма чека | № п/п | Сумма чека |
|---|---|---|---|
| 1 | 40,9 | 10 | 32,7 |
| 2 | 29,3 | 11 | 48,9 |
| 3 | 50,5 | 12 | 38,7 |
| 4 | 45,9 | 13 | 28,9 |
| 5 | 50,6 | 14 | 43,7 |
| 6 | 14,6 | 15 | 64,9 |
| 7 | 41,4 | 16 | 35,6 |
| 8 | 18,5 | 17 | 56,1 |
| 9 | 20,6 | 18 | 65,7 |

По приведенным данным число степеней свободы равно 17, выборочная средняя 40,42 тыс. руб., стандартное отклонение 14,72 тыс. руб. Стандартная ошибка выборочной средней

$$\mu = \frac{14{,}73}{\sqrt{18}} = 3{,}47 \text{ тыс. руб.}$$

По таблице ***t***-распределения Стьюдента для 17 степеней свободы получим следующие значения ***t*** для доверительных вероятностей 85, 90, 95 и 99% (с округлением до третьего знака):

$$t_{85} = 1{,}508;\ t_{90} = 1{,}74;\ t_{95} = 2{,}101;\ t_{99} = 2{,}898..$$

Для принятых уровней доверительных вероятностей оценим величины интервалов, в которых следует ожидать среднее значение суммы чеков совокупности (тыс. руб.):

$$\bar{x} = 40{,}42 \pm 1{,}508 * 3{,}47 = 35{,}2 \div 45{,}7;$$

$$\bar{x} = 40{,}42 \pm 1{,}74 * 3{,}47 = 34{,}4 \div 46{,}5;$$

$$\bar{x} = 40{,}42 \pm 2{,}101 * 3{,}47 = 33{,}1 \div 47{,}4;$$

$$\bar{x} = 40{,}42 \pm 2{,}898 * 3{,}47 = 30{,}4 \div 50{,}5.$$

Как следует из полученных ответов, рост доверительной вероятности с 85% до 99% вдвое увеличил размеры интервалов. Решение этой же задачи, но на основе большой выборки, значительно сузит искомые интервалы.

В целом можно сказать, что абсолютно все методы, основанные на нормальном распределении, включая оценивание абсолютной ошибки, монетарный метод (со стратификацией или без нее) и т.д., могут применяться с использованием распределения Стьюдента. Как было отмечено выше, для этого достаточно заменить в выражении для доверительного интервала (а также производных формулах) величину $z$ на $t$. Расчет остальных параметров остается неизменным. В силу сказанного нет смысла заново писать все уже известные формулы и приводить примеры их использования. С точки зрения аудита целесообразность применения распределения Стьюдента вместо нормального распределения обусловлена только малыми выборками.

## 8. Выводы по аналитическому обзору статистических выборочных методов в аудите

Подводя итог, можно сказать следующее. Существует огромное разнообразие статистических методов, которые могут применяться в аудите в рамках выборочного исследования. Часть из них имеет совершенно разные задачи и сравнению не подлежит, например, любые методы, основанные на дискретных распределениях в рамках процедур по существу, не могут сравниваться с методами, основанными на непрерывных распределениях в аналогичных процедурах, так как имеются определенные ограничения использования первых в подобных процедурах. Другая часть методов допускает сравнение в различных условиях, однако имеются теоретические предпосылки, в силу которых использование одних методов является более эффективным, чем других в конкретной ситуации. Например, сравнение методов, основанных на нормальном распределении, с методами, основанными на распределении Стьюдента в условиях малой выборки. Или сравнение методов, использующих гипергеометрическое распределение, с методами, которые основываются на биномиальном распределении, в условиях бесповторного отбора. Теоретические предпосылки подлежат практической проверке. Есть группа методов, которая не только допускает прямое сравнение, но и не имеет ярко выраженных теоретических предпосылок для преобладания одних методов над другими с точки зрения эффективности применения. Например, неясно какой метод, использующий нормальное распределение, дает наилучший результат при оценке совокупной суммы ошибок в совокупности – оценивание абсолютной ошибки напрямую, через относительную ошибку или монетарный метод.

Вышеобозначенные обстоятельства делают актуальным сравнительный анализ статистических методов в рамках выборочного аудита. Безусловно эффективность применения того или иного метода должна быть подтверждена на практике.

## Литература


1. Адамс Р. Основы аудита [Текст] / Р. Адамс. – М. : Аудит-ЮНИТИ, 1995. – 398 с.
2. Аренс А. Аудит [Текст] / А. Аренс, Дж. Лоббек. – М. : Финансы и статистика, 1995. – 560 с.
3. Вайнберг Дж. Статистика [Текст] / Дж. Вайнберг, Дж. Шумекер. − М. : Статистика, 1979. – 389 с.
4. Дефлиз Ф.Л. и др. Аудит Монтгомери [Текст] / Ф.Л. Дефлиз, Г.Р. Дженик, В.М. О'Рейли. – М. : Аудит-ЮНИТИ, 1997. – 542 с.
5. Джессен Р. Методы статистических обследований [Текст] / Р. Дженсен. − М. : Финансы и статистика, 1985. – 479 с.
6. Кендалл М. Теория распределений [Текст] / М. Кендалл, А. Стюарт. − М. : Наука, 1966. – 588 с.
7. Кокрен У. Методы выборочных исследований [Текст] / У. Кокрен. – М. : Статистика, 1976. – 440 с.
8. Кочинев Ю.Ю. Аудит. Теория и практика [Текст] / Ю.Ю. Кочинев. – СПб. : Питер, 2010. – 448 с.
9. Кочинев Ю.Ю. Выборочная проверка с помощью «монетарного» метода: возможная погрешность и границы применения [Текст] / Ю.Ю. Кочинев // Аудиторские ведомости. − 2009. – №4. – С. 60-66.
10. Кочинев Ю.Ю. Международные стандарты аудита: оценка существенности [Текст] / Ю.Ю. Кочинев // Аудиторские ведомости. − 2010. – №12. – С. 3-7.
11. Кочинев Ю.Ю. Моделирование и автоматизация аудита [Текст] : монография / Ю.Ю. Кочинев. – СПб. : СПбГПУ, 2006. – 157 с.
12. Ликеш И. Основные таблицы математической статистики [Текст] / И. Ликеш, Й. Ляга. − М. : Финансы и статистика, 1985. – 356 с.
13. Робертсон Дж. Аудит [Текст] / Дж. Роберсон. – М. : Контакт, 1993. – 496 с.
14. Романов А.И. Автоматизация аудита [Текст] / А.И. Романов, Е.Е. Одинцов. – М. : ЮНИТИ, 1999. – 336 с.
15. Феллер В. Введение в теорию вероятностей и ее приложение [Текст] / В. Феллер. − М. : Мир, 1964. – 1276 с.
16. Шторм Р. Теория вероятностей. Математическая статистика. Статистический контроль качества [Текст] / Р. Шторм. − М. : Мир, 1970. – 380 с.
17. Arens A.A., Loebbecke J.K. Application of statistical sampling to auditing. N.J. : Prentice Hall, 1981.
18. DRT International. The DRT audit sampling. A guide to applying The DRT International audit approach. N.Y., 1991. Pp. 42-73.
19. Giliberti A. Bank internal audit manual. N.Y. : Warren, Gorham & Lamont, 1997. Vol. 2. Pp. 65-13/
20. Touche Ross International. The touch ross audit sampling manual. U.K. Edition. 1982. P. 18.



## Ключевые слова

Аудиторская выборка; риск выборки; стратификация; доверительный интервал; выборочная средняя; нормальное распределение; процедуры на соответствие; процедуры по существу.


*Логиненков Алексей Владимирович*

## РЕЦЕНЗИЯ

В настоящее время в практике аудита известно применение статистических методов выборочных проверок. Любая аудиторская компания вынуждена применять статистические методы при проверке более или менее крупного клиента, так как сплошной аудит неэффективен ни с точки зрения трудозатрат аудитора, ни с точки зрения финансовых затрат самого клиента. На данный момент в статистике описано немало методов выборочного наблюдения, однако их применение в контексте специфических задач аудита недостаточно полно освящено в литературе по аудиту, что делает актуальным тему рецензируемой статьи.

В рецензируемой статье проведен обзор статистических выборочных методов, применяемых в аудите. Произведена классификация и систематизация способов оценки результатов выборочного исследования в рамках аудита с увязкой к общим теоретическим положениям математической статистики.

Материал, представленный в работе, обладает научной новизной и дает почву для дальнейших исследований в области теории выборочного наблюдения в аудите. В частности, наиболее полный обзор и классификация методов дают возможность в будущем провести сравнительный анализ эффективности того или иного метода в зависимости от исходных условий и предпосылок выборочного аудита.

Полагаю, что работа может быть рекомендована к опубликованию в журнале «Аудит и финансовый анализ».

*Кочинев Ю.Ю., д.э.н., проф. доцент, с.н.с., Санкт-петербургский государственный политехнический университет.*

[Перейти на Главное МЕНЮ](#)
[Вернуться к СОДЕРЖАНИЮ](#)