5. ОЦЕНКА ГЕНЕРАЛЬНЫХ ПАРАМЕТРОВ

5.1. Введение

Материал, содержащийся в предыдущих главах, можно рассматривать как минимальный набор сведений, необходимых для использования основных статистических методов, объединенных в теории статистических выводов.

Перейдем теперь к рассмотрению этих методов. Для этого необходимо определить их место в рамках единого подхода к решению конкретных задач статистических исследований в области правоведения.

Основная задача, решаемая с помощью методов математической статистики, — получение информации о закономерностях изменения изучаемого признака для большой совокупности объектов исследования, объединенных по этому признаку. В терминах математической статистики это означает, что делаются выводы о свойствах генеральной совокупности.

В гл. 4 определено, что для описания генеральной совокупности используются математические модели теории вероятностей. Исчерпывающую информацию о генеральной совокупности дает распределение вероятностей. Чаще всего используется модель нормально распределенной генеральной совокупности. И в этом случае распределение полностью определено всего двумя параметрами: средним значением (математическим ожиданием) и стандартным отклонением.

Следовательно, чтобы полностью описать нормальную генеральную совокупность, нужно знать значения двух генеральных параметров: среднего значения и стандартного отклонения. Так, если интерес вызывают данные о количестве правонарушений, то это среднее количестве всех правонарушений и стандартное отклонение. Эти параметры неизвестны и предположительно находятся в каких-то пределах. Единственное, что можно сделать, чтобы их определить — это провести эксперимент. Эксперимент для всей генеральной совокупности нереализуем или неоправдан, поэтому применяется выборочный метод.

На основании данных, полученных по выборке, делается вывод относительно всей генеральной совокупности. Используемые для этого методы теории статистических выводов обычно подразделяются на два класса: оценка параметров и проверка гипотез.

Задача оценки параметров состоит в получении наилучших в определенном смысле оценок параметров распределения генеральной совокупности на основании выборочных данных.

Проверка гипотез охватывает методы использования выборочных данных для проверки предположений относительно распределения и параметров распределения генеральной совокупности, которые делаются до получения выборочных данных.

В данной главе будут рассмотрены основные положения теории оценок.

5.2. Случайная выборка из генеральной совокупности

Чтобы по выборке можно было делать выводы о свойствах всей генеральной совокупности, она должна быть представительной (репрезентативной). Это обеспечивается в тех ситуациях, когда выборка является случайной. Модель случайной выборки предъявляет к ней следующие требования:

1) каждый из объектов, составляющих генеральную совокупность, должен иметь одинаковую вероятность быть представленным в выборке;

2) все n измерений, образующих выборку, должны быть независимыми, т. е. результаты каждого измерения не должны зависеть от предыдущих измерений.

Существует два основных метода отбора объектов из генеральной совокупности в выборку: повторный и бесповторный.

При повторном отборе каждый объект после измерения значения признака возвращается в генеральную совокупность. При этом состояние генеральной совокупности перед каждым новым измерением восстанавливается и требование независимости всегда выполняется.

При бесповторном отборе после измерения объект не возвращается в генеральную совокупность. В этом случае соотношение значений признака в оставшейся части генеральной совокупности меняется, и, следовательно, проводимые измерения не являются независимыми, т. е. бесповоротный отбор не является случайным. На практике бесповоротный отбор используется чаще. Когда проводится измерение каких-то признаков, относящихся, например, к преступникам, выборка составляется таким образом, что после того, как очередной человек принял участие в измерениях, он уже не участвует в следующих измерениях.

Но, как правило, можно считать, что объем генеральной совокупности настолько велик, что при исключении из нее относительно малого числа единиц, составляющих выборку, состояние генеральной совокупности практически не меняется. При бесконечной генеральной совокупности различие между повторным и бесповторным отбором исчезает.

вторным и бесповторным отбором исчезает.

На практике используется несколько способов получения случайных выборок:

1. собственно случайная,
2. механический отбор.
3. типический отбор.
4. серийный отбор.

При проведении выборочных исследований предполагается, что выборка является однородной. Это означает, что она получена из одной генеральной совокупности, т. е. в исходной совокупности отсутствуют объекты, резко выделяющиеся по значениям изучаемого признака. Предположение об однородности выборки на практике обычно основывается на предварительном изучении условий эксперимента. Так, обычно есть уверенность в том, что полученные выборочные данные о количестве правонарушений представляют собой результаты измерений для одинаковых по численности городов.

5.3. Точечные оценки

5.3.1. Определения и требования к оценкам

Под термином "оценка" в теории оценок понимаются как сами значения параметров генеральной совокупности, полученные по выборке, так и процесс получения этих значений, т. е. правило, по которому они получены.

Оценки подразделяются на два класса; точечные и интервальные.

Точечные оценки представляют собой определенные значения параметров генеральной совокупности, полученные по выборочным данным. Эти значения должны быть максимально близки к значениям соответствующих параметров генеральной совокупности, которые являются истинными значениями оцениваемых параметров.

При формировании интервальных оценок определяют границы интервалов, между которыми с большой вероятностью находятся истинные значения параметров.

Начнем с точечных оценок и рассмотрим оценку произвольного параметра (среднего, дисперсии или какого-то другого) генеральной совокупности, который обозначим α. Оценивая параметр α по выборке, находим такую величину α_B, которую принимаем за точечную оценку параметра α. Естественно, при этом стремимся, чтобы оценка была в *определенном смысле* наилучшей, поэтому к ней предъявляется ряд требований:

1. ***Состоятельность***. Точечная оценка α_B называется состоятельной, если при неограниченном увеличении объема выборки ($n \to \infty$) она стремится к истинному значению параметра α.

В математической статистике показывается, что состоятельной оценкой генерального среднего значения μ, является выборочное среднее арифметическое $\bar{x}_B$, а состоятельной оценкой генеральной дисперсии σ^2 — выборочная дисперсия σ_B^2. Методы вычисления этих выборочных характеристик были рассмотрены в гл. 3.

2. ***Несмещенность***. Оценка α_B называется несмещенной, если она не содержит систематической ошибки, т. е. среднее значение оценки, определенное по многократно повторенной выборке объема *n* из одной и той же генеральной совокупности, стремится к истинному значению соответствующего генерального параметра α.

Выборочное среднее арифметическое $\bar{x}_B$ является несмещенной оценкой генерального среднего μ.

Несмещенной оценкой генеральной дисперсии σ^2 является *исправленная выборочная дисперсия*, вычисляемая по формуле:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x}_B)^2}{n-1} \text{ для несгруппированных данных,}$$

$$S^2 = \frac{\sum_{i=1}^{k}(x_i - \bar{x}_B)^2 n_i}{n-1} \text{ для сгруппированных данных,}$$

Замечание 1

Одним из свойств выборочного среднего арифметического является то, что сумма квадратов отклонений значений признака от среднего арифметического меньше, чем сумма квадратов отклонений от любой другой величины (в том числе и от генерального среднего μ), т.е. $\sum(x_i - \bar{x}_B)^2 < \sum(x_i - \mu)^2$ для любой выборки. Поэтому вычисление оценки дисперсии по формуле $S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_B)^2$ будет содержать систематическую ошибку, и такая оценка будет смещенной.

Можно показать, что если использовать $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_B)^2$, то она будет несмещенной, т.е. при неограниченном повторении выборки из генеральной совокупности и усреднении выборочной дисперсии, полученной на основании этой формулы, по всем выборкам получается истинное значение генеральной дисперсии.

3. ***Эффективность.*** Несмещенная оценка является эффективной, если она имеет наименьшую дисперсию по сравнению с другими несмещенными оценками того же параметра генеральной совокупности.

Это надо понимать так: полученные по выборке оценки $\bar{x}_B$ и S^2 — случайные величины, так как случайны сами выборочные значения. Поэтому можно говорить о математическом ожидании и дисперсии оценок $\bar{x}_B$ и S^2. Эффективность этих оценок означает, что их дисперсии $D(\bar{x}_B)$ и $D(S^2)$ меньше дисперсий любых других несмещенных оценок среднего значения и дисперсии генеральной совокупности.

Итак, наилучшими в указанном смысле оценками генерального среднего значения и генеральной дисперсии являются выборочные характеристики $\bar{x}_B$, S^2.

5.3.2. Стандартная ошибка среднего арифметического

Оценки $\bar{x}_B$ и S^2, полученные по выборке, естественно не совпадают с истинными значениями параметров μ и σ^2 генеральной совокупности. Экспериментально проверить это утверждение невозможно, поскольку не известны истинные значения этих параметров. Но если брать повторные выборки из одной и той же генеральной совокупности с параметрами μ и σ^2 и каждый раз вычислять их оценки $\bar{x}_B$ и S^2, то окажется, что эти оценки для разных выборок не совпадают, хотя все это из одних и тех же генеральных параметров.

Отклонения оценок генеральных параметров от истинных значений этих параметров называются статистическими ошибками, или ошибками репрезентативности. Их происхождение не имеет ничего общего с ошибками измерения, а возникают они только потому, что не все объекты генеральной совокупности представлены в выборке.

Величины статистических ошибок оценивают по среднему квадратическому (стандартному) отклонению выборочных характеристик. Здесь рассматривается только стандартное отклонение выборочного среднего арифметического.

Если взять очень много независимых выборок объема *n* из одной и той же генеральной совокупности и определить для каждой из них среднее арифметическое, то окажется, что полученные средние арифметические варьируют вокруг своего среднего значения (равного μ) в $\sqrt{n}$ раз меньше, чем отдельные варианты выборки. Т.е. стандартное отклонение выборочного среднего арифметического будет равно

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

где σ — стандартное отклонение генеральной совокупности.

В качестве оценки стандартного отклонения выборочного среднего используется величина

$$S_{\bar{x}} = \frac{S}{\sqrt{n}} \qquad (5.1)$$

называемая стандартной ошибкой среднего арифметического. В формуле (5.1) S — выборочное стандартное отклонение $S = \sqrt{S^2}$.

Величина $S_{\bar{x}}$ показывает, какая ошибка в среднем допускается, если использовать вместо генерального среднего μ его выборочную оценку $\bar{x}_B$. Поэтому вычисленное среднее арифметическое часто указывают в виде

$$\bar{x}_B \pm S_{\bar{x}}$$

чтобы оценить точность оценки x.

Из формулы (5.1) видно, как зависит стандартная ошибка $S_{\overline{x}}$ от объема выборки *n*: с увеличением объема выборки n стандартная ошибка $S_{\overline{x}}$ уменьшается пропорционально корню квадратному из *n*.

Замечание.

Теперь можно вернуться к вопросу, который был оставлен открытым при вычислении выборочных характеристик с такой точностью нужно вычислять выборочные характеристики?

Как мы только что убедились, при ограниченном объеме выборки *n* истинное значение генерального среднего μ не может быть определено сколь угодно точно, поэтому при вычислении $\overline{x}_B$ оставлять большое число значащих цифр не имеет смысла. Существует эмпирическое правило, согласно которому в окончательном результате положение последней значащей цифры должно соответствовать положению первой значащей цифры в величине $S_{\overline{x}}$/3. Чтобы избежать накопления ошибок, связанных с округлением, промежуточные результаты нужно вычислять с точностью на один порядок больше, чем точность окончательных результатов.

5.4. Интервальные оценки

По известной величине выборочной характеристики ($\overline{x}_B$ или S^2 и др.) можно определить интервал, в котором с той или иной вероятностью определяется значение параметра генеральной совокупности, оцениваемого по этой выборочной характеристике.

Вероятности, признанные достаточными для того, чтобы уверенно судить о генеральных параметрах на основании выборочных характеристик, называются доверительными.

Обычно в качестве доверительных вероятностей выбирают значения 0,95, 0,99 или 0,999 (их принято выражать в процентах). Перечисленным значениям соответствуют 95, 99 и 99,9 %. Выбор той или иной доверительной вероятности производится исследователем исходя из практических соображений о той ответственности, с какой делаются выводы о генеральных параметрах.

Замечание 3

Как правило, в научных исследованиях в области правоведения считается достаточной доверительная вероятность 0,95 (95%). В некоторых случаях, когда уточняются результаты предыдущих исследований или когда выводы, сделанные в данном исследовании, связаны с большой ответственностью (например, предлагается в корне пересмотреть результаты следствия), применяются более высокие уровни доверительной вероятности: 99 или 99,9 %.

Интервал, в котором с заданной доверительной вероятностью находится оцениваемый генеральный параметр, называется *доверительным интервалом*.

В соответствии с доверительными вероятностями на практике используются 95-, 99-, 99,9-процентные доверительные интервалы.

В литературе по математической статистике обычно говорят о $100\cdot(1-\alpha)\%$ -процентном доверительном интервале, где (1–α) — *доверительная вероятность*, а α — некоторое малое число (α = 0,05; 0,01; 0,001), задающее вероятность того, что оцениваемый генеральный параметр выходит, за границы доверительного интервала.

Теперь рассмотрим формирование доверительного интервала для среднего (математического ожидания) n, нормально распределенной генеральной совокупности. Пронормируем значение среднего арифметического $\overline{x}_B$, найденного по выборке объема *n* из этой генеральной совокупности, по формуле:

$$t = \frac{\overline{x}_B - \mu}{S_{\overline{x}}}$$

где μ — оцениваемый параметр — среднее значение генеральной совокупности; $S_{\overline{x}}$ — стандартная ошибка выборочного среднего арифметического.

Величина t имеет Т-распределение Стьюдента (определенное в гл. 4) с v=n–1 степенями свободы.

Необходимо определить доверительный интервал, в котором с доверительной вероятностью $100\cdot(1-\alpha)\%$ находится истинное значение оцениваемого параметра μ. Для этого задается значение α (например, 0,05). Доверительная вероятность будет соответствовать площади под кривой Т-распределения Стьюдента, заключенной между точками $-t_\alpha$ и t_α (рис. 5.1). Следовательно, доверительный интервал можно записать как

$$-t_\alpha \le \frac{\overline{x}_B - \mu}{S_{\overline{x}}} \le t_\alpha$$

Преобразуем это выражение к виду

$$\bar{x}_B - t_\alpha S_{\bar{x}} \le \mu \le \bar{x}_B + t_\alpha S_{\bar{x}} \qquad (5.2)$$

Это и есть стандартная форма записи доверительного интервала.

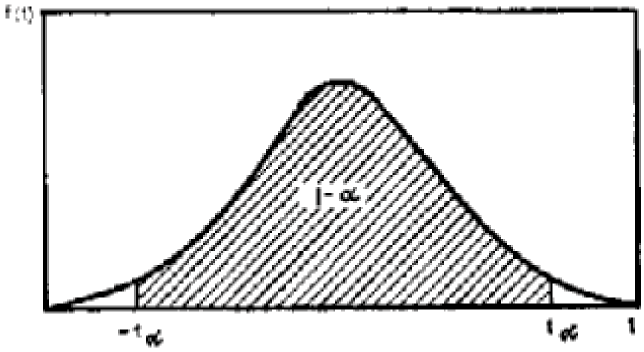

Рис. 5.1. К определению доверительного интервала

Учитывая формулу (5.1) приходим к окончательному выражению:

$$\bar{x}_B - t_\alpha \frac{S}{\sqrt{n}} \le \mu \le \bar{x}_B + t_\alpha \frac{S}{\sqrt{n}}$$

т. е. истинное значение μ с вероятностью 100 (1 — α)% лежит в границах $\bar{x}_B - t_\alpha \frac{S}{\sqrt{n}}$ и $\bar{x}_B + t_\alpha \frac{S}{\sqrt{n}}$.

Значения t_α для стандартных значений α (0,05, 0,01 и 0,001) и различных значений параметра ν t-распределения (ν = n – 1) приведены в специальных таблицах.

Чтобы найти границы доверительного интервала для среднего значения генеральной совокупности, действуем в следующем порядке:

1) по полученной выборке объема n вычисляем среднее арифметическое $\bar{x}_B$ и стандартное отклонение S. Методы вычислений рассмотрены в гл. 3;

2) задаемся доверительной вероятностью $1-\alpha$ (например, 0,95) исходя из целей исследования;

3) по таблице Т-распределения Стьюдента находим граничные значения t_α. В силу симметричности Т-распределения достаточно знать только положительное значение t_α. Например, если объем выборки n=12, то число степеней свободы Т-распределения ν=12–1=11, и по таблице Т-распределения определяем для α = 0,05 значение $t_{0,05}$ = 2,20;

4) находим границы доверительного интервала по формуле (5.3). Для α=0,05 и n=12:

$$\bar{x}_B - 2{,}2\frac{S}{\sqrt{12}} \le \mu \le \bar{x}_B + 2{,}2\frac{S}{\sqrt{12}}$$

Как было отмечено в гл. 4, при больших объемах выборки (практически при $n \ge 30$) t-распределение Стьюдента переходит в нормальное. Поэтому для определения границ доверительного интервала для μ при больших объемах выборки можно пользоваться таблицами нормированного нормального распределения (табл. 1 Приложения).

Доверительный интервал для μ при $n \ge 30$ записывается в следующем виде:

$$\bar{x} - u_\alpha \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + u_\alpha \frac{S}{\sqrt{n}} \qquad (5.4)$$

где u_α — процентные точки нормированного нормального распределения, определяемые по табл. 1 Приложения.

Для стандартных доверительных вероятностей (95, 99, 99,9%) значения u_α приведены в таблице 5.2.

Таблица 5.2

Значения u_a для стандартных доверительных вероятностей

α	1- α	u_a
0,05	0,95	1,96
0,01	0,99	2,58
0,001	0,999	3,28

Чтобы найти доверительный интервал для среднего значения генеральной совокупности при больших объемах выборки ($n \geq 30$), поступаем следующим образом:

1. По выборочным данным находим среднее арифметическое $\bar{x}_B$ и стандартное отклонение S, как показано в гл. 3.

2. Задаемся доверительной вероятностью $1-\alpha$ (например, 0,95).

3. По табл. 5.2 находим значение u_a, соответствующее заданной доверительной вероятности ($u_{0,05}$ = 1,96).

4. Определяем границы доверительного интервала по формуле (5.4). Для α = 0,05 получаем:

$$\bar{x}_B - 1{,}96\frac{S}{\sqrt{n}} \leq \mu \leq \bar{x}_B + 1{,}96\frac{S}{\sqrt{n}}$$

Как видно из сравнения найденного доверительного интервала с доверительным интервалом, полученный выше по Т-распределению, при малых объемах выборки границы первого интервала шире ($t_{0,05}$ = 2,20, а $u_{0,05}$ = 1,96). Это понятно из простых физических соображений: при малом объеме выборки получается меньше информации о свойствах генеральной совокупности.

Сделаем еще одно замечание по поводу доверительных интервалов.

Среднее значение (μ, генеральной совокупности является хотя и неизвестным, но фиксированным параметром, а границы доверительного интервала, полученные по случайной выборке объема *n*, будут также случайными величинами. Когда говорится о 95-процентной доверительной вероятности, это означает, что примерно в 95 % случаев фиксированное, но неизвестное значение μ, окажется в границах доверительного интервала.

Образная трактовка доверительных интервалов приведена в книге «Статистика и планирование эксперимента в технике и науке»[*]. «Доверительный интервал и связанные с ним понятия похожи на то, с чем мы сталкиваемся при игре с набрасыванием подковы на кол. Кол здесь играет роль оцениваемого параметра (его положение никогда не изменяется)... Подкова выступает в роли доверительного интервала. Если при 100 набрасываниях подковы удается в среднем 90 раз набросить ее на кол, то имеется 90 %-ная гарантия (или уровень доверия) набросить подкову на кол. Доверительный интервал, подобно подкове, меняет свое положение. При любом броске (или при построении некоторой интервальной оценки) кол (или параметр) может как попасть внутрь подковы (интервала), так и оказаться вне ее. Таким образом, делается вероятностное утверждение относительно переменных величин, характеризующих положение подковы».

Оценку параметра μ, найденную в форме доверительного интервала, часто записывают в виде $\bar{x}_B \pm t_\alpha S/\sqrt{n}$. Чтобы избежать неоднозначности в толковании результатов (перепутывания с записью результата как $\bar{x}_B \pm S_{\bar{x}}$), запись доверительного интервала необходимо сопровождать пояснением. Например 95 %-ный доверительный интервал для среднего роста (174см ± 1,3 см).

5.5. Определение необходимого объема выборки для получения оценок заданной точности

Обычно исследователя интересует вопрос: какой минимальный объем выборки необходим для того, чтобы оценка (чаще всего выборочное среднее арифметическое $\bar{x}$) отличалась от истинного значения среднего значения генеральной совокупности не более чем на заданную величину?

Ответить на этот вопрос можно, если ввести доверительную вероятность и выбрать объем выборки n таким образом, чтобы доверительный интервал имел заданный размер.

Если генеральная совокупность предполагается нормально распределенной и ее дисперсия σ^2 известна, то доверительный интервал для среднего значения μ записывается следующим образом:

$$\bar{x}_B - u_\alpha\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x}_B + u_\alpha\frac{\sigma}{\sqrt{n}}$$

где u_a для стандартных доверительных вероятностей определены в табл. 5.2.

Пусть требуется, чтобы выборочное среднее $\bar{x}_B$ отличалось от генерального μ, не более чем на за-

[*] Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке. — М.: Мир, 1980, с. 228—229.

данную величину d. Это означает, что половина ширины доверительного интервала должна быть равна d, т. е. половина от

$$\left(\bar{x}_B + u_\alpha \frac{\sigma}{\sqrt{n}}\right) - \left(\bar{x}_B - u_\alpha \frac{\sigma}{\sqrt{n}}\right) = 2u_\alpha \frac{\sigma}{\sqrt{n}}$$

должна равняться d:

$$u_\alpha \frac{\sigma}{\sqrt{n}} = d$$

Отсюда требуемый объем выборки определяется следующим образом:

$$n = \left(\frac{u_\alpha \sigma}{d}\right)^2 \qquad (5.5)$$

Истинное значение параметра σ генеральной совокупности обычно неизвестно, но при больших объемах выборки ($n \geq 30$) можно использовать его выборочную оценку S. Тогда

$$n = \left(\frac{u_\alpha S}{d}\right)^2 \qquad (5.6)$$