РАЗДЕЛ 2. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Математическая статистика изучает закономерности, присущие массовым случайным явлениям. Основные задачи математической статистики состоят в разработке методов сбора и группировки статистических данных, а также в последующем анализе полученных статистических данных. Этот анализ включает определение законов распределения случайных величин, определение неизвестных параметров распределения, а также связей между случайными величинами.

Глава 7. Выборка и ее распределение

7.1. Выборочный метод. Основные понятия

В математической статистике вводятся понятия генеральной и выборочной совокупности. **Генеральной совокупностью** называется совокупность всех однородных объектов, подлежащих изучению. **Выборочной совокупностью**, или **выборкой** называется совокупность объектов случайно отобранных из генеральной совокупности. **Объемом** выборки называется число ее объектов. Например, если из 100000 изготовленных деталей для обследования случайным образом отобрано 100, то объем генеральной совокупности $N = 100000$, объем выборки $n = 100$. Возможных значений случайной величины в генеральной совокупности может быть бесконечное множество, например, в некотором пункте измеряется количество выпавших осадков. Это непрерывная случайная величина; выборку также можно проводить любое количество раз.

Выборка бывает **повторной** (с возвращением исследуемого объекта в генеральную совокупность) и **бесповторной**, при которой отобранный объект в генеральную совокупность не возвращается.

Выборка должна быть **репрезентативной** (представительной), т. е. она должна правильно отображать все свойства и характеристики объектов генеральной совокупности.

Пусть в результате независимых испытаний, проведенных в одинаковых условиях, из генеральной совокупности данных, выбраны числовые значения x_1, x_2, … x_n, характеризующие некоторый признак (здесь n - объем выборки). Последовательность наблюдаемых значений x_i ($i = 1, 2, \ldots, n$), записанных в возрастающем порядке, называется **дискретным вариационным рядом**, а элементы этой последовательности x_i называют **вариантами**. Если среди вариант есть равные значения, тогда дискретный вариационный ряд записывают в виде таблицы

x_i	x_1	x_2	…	x_n
n_i	n_1	n_2	…	n_k

где n_i – частота появления значения x_i; при этом должно выполняться условие
$$n_1 + n_2 + \ldots + n_k = n.$$

Относительной частотой P_i^* варианты x_i называется отношение ее частоты к объему выборки:

$$P_i^* = \frac{n_i}{n}.$$

Очевидно, $\sum_{i=1}^{k} P_i^* = 1$,

т. е. сумма относительных частот всех вариант равна единице.

Пример. В поликлинике подсчитывали количество посетителей X, обратившихся в регистратуру с 8-00 до 9-00 в течение 30 дней. Были получены следующие результаты:
100, 100, 120, 75, 75, 60, 120, 100, 70, 70, 65, 60, 100, 120, 120, 65, 60, 70, 70, 75, 100, 100, 70, 70, 75, 100, 100, 75, 70, 65.
Требуется составить ряд распределения частот.
Решение. Вначале составим вариационный ряд, записав результаты наблюдений в возрастающем порядке:
60, 60, 60, 65, 65, 65, 70, 70, 70, 70, 70, 70, 70, 75, 75, 75, 75, 75, 100, 100, 100, 100, 100, 100, 100, 100, 120, 120, 120, 120.
В данном вариационном ряде шесть различных вариант (шесть групп). Для каждой варианты подсчитаем ее частоту (количество повторений) и соответствующую относительную частоту. Все результаты запишем в таблицу

номер группы	i	1	2	3	4	5	6
число обращений в регистратуру	x_i	60	65	70	75	100	120
частота	n_i	3	3	7	5	8	4
относительная частота	P_i^*	3/30	3/30	7/30	5/30	8/30	4/30

Контроль: $\sum_{i=1}^{6} P_i^* = 1$

7.2. Полигон и гистограмма

В целях наглядности статистическое распределение изображается графически. **Полигоном** частот называют ломаную, отрезки которой соединяют точки (x_1, n_1), (x_2, n_2),…, (x_n, n_n), где x_i – варианта, n_i – частота.

Аналогично строится полигон относительных частот, т. е. ломаная, отрезки которой последовательно соединяют точки (x_1, P_1^*), (x_2, P_2^*),…, (x_n, P_n^*). На рис. 28 изображен полигон частот распределения рассмотренного примера.

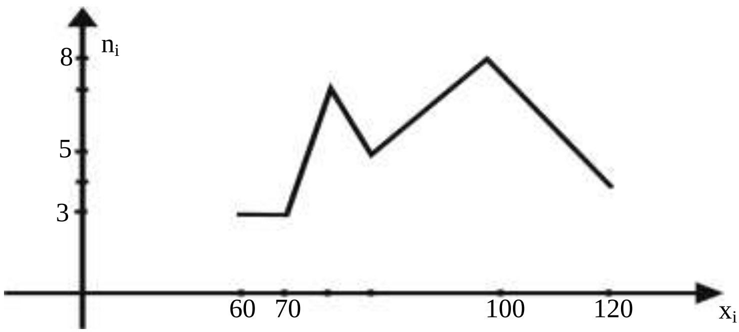

Рис. 28. Полигон частот

Кроме дискретных вариационных рядов рассматриваются **интервальные вариационные ряды**, в которых значения признака могут меняться непрерывно или число значений признака велико. Для построения такого ряда промежуток изменения признака разбивается на ряд отдельных интервалов и подсчитывается количество значений величины в каждом из них. Если все интервалы имеют одинаковую длину, то число интервалов в случае нормально распределенной совокупности рекомендуется находить по формуле Стерджесса

$$K = 1 + 3{,}322 \lg n,$$

обычно, $6 \le K \le 12$. Длина частичного интервала определяется по формуле

$$h = (X_{max} - X_{min})/K,$$

где $X_{max} - X_{min}$ – разность между наибольшим и наименьшим значениями признака, называемая размахом вариации.

Пример. Пусть дан ряд распределения результатов тестирования (в баллах) для 60 студентов:

12	6	8	10	11	7	10	12	8	7	7	6	7	8	6	11	9
9	10	11	9	10	7	8	8	8	11	9	8	7	5	9	7	7
9	8	7	4	7	5	5	10	7	7	5	8	10	10	15	10	10
11	6	11	14	13	6	11	12	6								

Построить интервальный вариационный ряд.

Решение. Количество интервалов найдем по формуле Стерджесса

$$K = 1 + 3{,}322 \lg 60 \approx 6{,}907;\ K = 7.$$

Найдем длину частичного интервала

$$h = (X_{\max} - X_{\min})/K = (15 - 4)/7 \approx 1{,}6.$$

Разобьем интервал изменения X (количества баллов) на 7 интервалов с шагом h = 1,6 и подсчитаем количество студентов в каждом интервале. Результаты вычислений запишем в таблицу

Частичные интервалы (x_{i-1}, n_i)	Число студентов в группе, n_i	Относительная частота $P_i^* = \frac{n_i}{n}$
4 – 5,6	5	5/60
5,61 – 7,2	17	17/60
7,21 – 8,8	9	9/60
8,81 – 10,4	15	15/60
10,41 – 12,0	10	10/60
12,01 – 13,6	1	1/60
13,61 – 15,2	3	3/60
сумма	60	1

Если статистическое распределение задано перечнем интервалов и соответствующих им частот, то строят гистограмму частот. **Гистограммой частот** называется ступенчатая фигура состоящая из прямоугольников с основаниями $h_i = X_i - X_{i-1}$ и высотами n_i/h_i. Если интервалы имеют одинаковую длину, то прямоугольники имеют одно и тоже основание h, их высоты равны n_i/h.

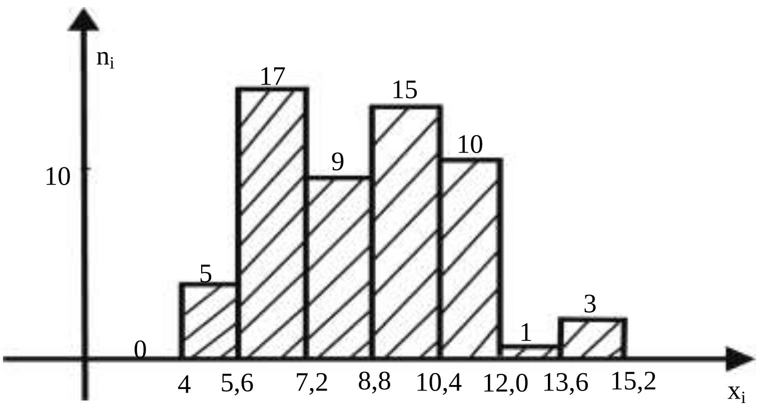

Рис.29. Гистограмма частот

На оси абцисс откладывают частичные интервалы длины h и над каждым интервалом строят прямоугольник.

Замечание. Гистограмма относительных частот является аналогом дифференциальной функции случайной величины. На рис. 29 изображена гистограмма частот распределения объема $n = 60$ заданного в таблице.

7.3. Эмпирическая функция распределения

Эмпирической функцией распределения, или **функцией распределения выборки**, называется функция, определяющая для каждого значения X относительную частоту события $X < x$. Обозначим функцию распределения через $F^*(x)$; если n_x – число вариант, меньших x, n – объем выборки, то по определению

$$F^*(x) = n_x/n.$$

Из определения эмпирической функции следует, что $F^*(x)$ обладает следующими свойствами:

1) значения функции $F^*(x)$ принадлежат отрезку [0; 1];
2) $F^*(x)$ – неубывающая функция;
3) если X_{min} – наименьшая, X_{max} – наибольшая варианта, то $F^*(x) = 0$, при $X < X_{min}$; $F^*(x) = 1$, при $X \geq X_{max}$.

Выборочную функцию распределения можно задать таблично или графически.

Пример. Найти эмпирическую функцию по данному распределению выборки:

варианты x_i	5	7	10	12
частоты n_i	3	5	10	2

Решение. Объем выборки $n = \sum_{i=1}^{4} n_i = 3 + 5 + 10 + 2 = 20$.

Наименьшая варианта $X_{min} = 5$, поэтому $F^*(x) = 0$, при $X \leq 5$.

Результаты вычислений $F^*(x)$ приведены в таблице

Таким образом, искомая эмпирическая функция определяется формулами

$$F^*(x) = \begin{cases} 0, \text{ если } x \leq 5 \\ 0{,}15, \text{ если } 5 < x \leq 7 \\ 0{,}4 \text{ если } 7 < x \leq 10 \\ 0{,}9, \text{ если } 10 < x \leq 12 \\ 1, \text{ если } x > 12 \end{cases}$$

x	$F^*(x)$
$x \le 5$	0
$5 < x \le 7$	$P_i^* = \dfrac{3}{20}$
$7 < x \le 10$	$P_i^* + P_2^* = \dfrac{3}{20} + \dfrac{5}{20} = \dfrac{8}{20}$
$10 < x \le 12$	$P_i^* + P_2^* + P_3^* = \dfrac{18}{20}$
$12 < x$	$P_i^* + P_2^* + P_3^* + P_4^* = 1$

График этой функции изображен на рис. 30

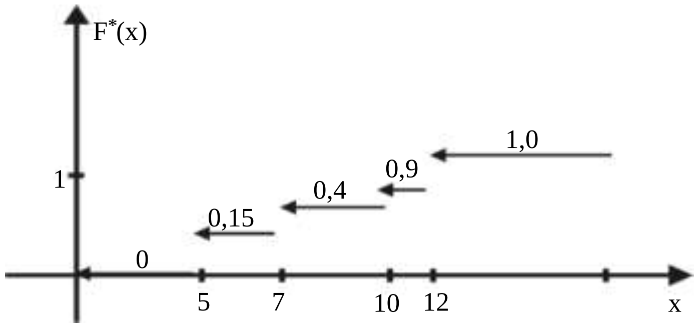

Рис.30.

Глава 8. Статистические оценки параметров распределения

8.1. Выборочная средняя и выборочная дисперсия

Оценки параметров генеральной совокупности, полученные на основании выборки, называются **статистическими**. Пусть случайная величина X имеет распределение $F(X, \theta)$, где θ - неизвестный параметр. Нужно определить приближенное значение этого параметра (оценить параметр θ) по некоторой выборке x_1, x_2, … , x_n. Оценку параметра θ обозначим через $\tilde{\theta}$, статистическая оценка является случайной величиной и зависит от x_1, x_2, … , x_n, т. е.

$$\tilde{\theta} = \tilde{\theta}\,(x_1, x_2, \dots, x_n)$$

В каждой из i – той серии n испытаний $\tilde{\theta}$ принимает некоторое значение $\tilde{\theta}_i$. Поэтому можно говорить о распределении этой величины и о числовых характеристиках распределения. Оценка параметра $\tilde{\theta}$ называется **несмещенной**, если математическое ожидание $\tilde{\theta}$ равно θ, т.е.

$$M(\tilde{\theta}) = \theta$$

и **смещенной**, если

$$M(\tilde{\theta}) \neq \theta.$$

Оценка $\tilde{\theta}$ параметра θ называется **состоятельной**, если при любом ε > 0

$$\lim_{n\to\infty} P\left(\left|\tilde{\theta} - \theta\right| < \varepsilon\right) = 1.$$

Оценка $\tilde{\theta}$ называется **эффективной**, если при заданном n она имеет наименьшую дисперсию, т.е. $D(\tilde{\theta}) = D_{\min}$.

Оценка $\tilde{\theta}$ имеет практическую ценность, если она является несмещенной, состоятельной и эффективной.

Если статистическая оценка характеризуется одним числом, она называется точечной. К числу таких оценок относятся выборочная средняя и выборочная дисперсия.

Выборочной средней $x_в$ называется среднее арифметическое значений признака выборочной совокупности. Если все значения $x_1, x_2, \ldots, x_n$ различны, то

$$X_В = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i .$$

Если значения $x_1, x_2, \ldots, x_k$, имеют соответственно частоты $n_1, n_2, \ldots, n_k$, причем $n_1 + n_2 + \ldots + n_k = n$, то

$$x_В = \frac{n_1 x_1 + n_2 x_2 + \ldots + n_k x_k}{n} .$$

<u>Замечание.</u> Выборочную среднюю обозначают также $\overline{X}$. Выборочную среднюю принимают в качестве оценки генеральной средней. Можно показать, что эта оценка является несмещенной и состоятельной.

Выборочной дисперсией $D_{в}$ называют среднее арифметическое квадратов отклонения наблюдаемых значений выборки от их среднего значения $x_{в}$. Если все значения x_1, x_2, … , x_n признака выборки объема n различны, то

$$D_{в} = \frac{1}{n}\sum_{i=1}^{k}(x_i - x_{в})^2$$

Если значения $x_1, x_2, \ldots, x_k$ имеют соответственно частоты $n_1, n_2, \ldots, n_k$, причем

$$n_1 + n_2 + \ldots + n_k = n,$$

то

$$D_{в} = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - x_{в})^2 .$$

Выборочное среднее квадратическое отклонение $\sigma_{в}$ определяется формулой

$$\sigma_{в} = \sqrt{D_{в}} .$$

Для вычисления выборочной дисперсии можно пользоваться формулой

$$D_{в} = x_{в}^2 - (x_{в})^2 ,$$

где

$$x_{в} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i \ , \quad x_{в}^2 = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 \ .$$

Выборочная дисперсия оценивает дисперсию генеральной совокупности и является смещенной оценкой. Чтобы получить несмещенную оценку генеральной дисперсии выборочную дисперсию умножают на величину $\frac{n}{n-1}$ и получают так называемую **эмпирическую** (или «**исправленную**») дисперсию

$$S^2 = \frac{n}{n-1} \cdot D_{в}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{k} n_i \left(x_i - x_в\right)^2$$

При $n \ge 50$ практически нет разницы между оценками $D_в$ и S^2. Эти оценки являются состоятельными оценками генеральной дисперсии. Для оценки среднего квадратического отклонения генеральной совокупности служит **«исправленное» среднее квадратическое отклонение**, или эмпирический стандарт

$$S = \sqrt{S^2} = \sqrt{\frac{1}{n-1}\cdot\sum_{i=1}^{k} \mathrm{n_i} \left(x_i - x_в\right)^2}\,.$$

В практических вычислениях для удобства расчетов при определении статистических оценок переходят к условным вариантам. Например, если варианты x_i – большие числа, то используют разности

$$u_i = x_i - c$$

где c – произвольно выбранное число (ложный нуль), такое, при котором условные варианты принимают небольшие численные значения.
В этом случае

$$x_в = c + \frac{\left(\sum_{i=1}^{k} n_i u_i\right)}{\mathrm{n}},$$

$$D_в = u_в^2 - \left(u_в\right)^2,$$

где

$$u_в = \frac{\left(\sum_{i=1}^{k} n_i u_i\right)}{\mathrm{n}},\quad u_в^2 = \frac{\left(\sum_{i=1}^{k} n_i u_i^2\right)}{\mathrm{n}}.$$

Пример. Найти выборочную среднюю $x_в$ и несмещенные оценки дисперсии S^2 и стандартного отклонения S на основании данного распределения выборки (57, 71, 66, 76, 70, 68, 74, 68, 69).

Решение. Выборка содержит мало повторяющихся элементов, поэтому вариационный ряд не составляем. Выберем в качестве ложного нуля число $c = 65$. Вычисления оформим в виде таблицы

x_i	$x_i - c$	$(x_i - c)^2$
57	-8	64
71	6	36
66	1	1
76	11	121
70	5	25
68	3	9
74	9	81
68	3	9
69	4	16
$\sum$	34	362

Имеем,

$$n = 9;$$

$$x_в = 34/9 + 65 = 3{,}78 + 65 = 68{,}78;$$

$$D_В = \frac{362}{9} - \left(\frac{34}{9}\right)^2 = 25{,}95 .$$

$$S^2 = \frac{9}{8} \cdot 25{,}95 = 29{,}19$$

$$S = \sqrt{29{,}19} = 5{,}40 .$$

Ответ. $x_в = 68{,}78$; $S^2 = 29{,}19$; $S = 5{,}40$.

8.2. Интервальные оценки параметров распределений

При выборках небольшого объема статистическая оценка $\tilde{\theta}$ может значительно отличаться от истинного значения параметра θ. Поэтому в случае малых значений объемов выборки n часто используют интервальные оценки.

Интервальной оценкой называют числовой интервал $\left(\bar{\theta}^{(1)}; \bar{\theta}^{(2)}\right)$, определяемый по результатам выборки, в котором с вероятностью γ находится неизвестное значение параметра генеральной совокупности θ, т.е.

$$P\left(\bar{\theta}^{(1)} \le \theta \le \bar{\theta}^{(2)}\right) = \gamma$$

Вероятность γ называют **доверительной вероятностью** или **надежностью**; обычно выбирают значения для вероятности γ близкими к единице (0,95; 0,98; 0,99 и т.д.). Интервал $\left(\bar{\theta}^{(1)}; \bar{\theta}^{(2)}\right)$ называется **доверительным интервалом**. Найдем доверительный интервал из условия

$$P\left(\left|\theta - \tilde{\theta}\right| < \delta\right) = \gamma,$$

где δ > 0 – некоторое число.
Отсюда следует, что

$$P\left(\tilde{\theta} - \delta < \theta < \tilde{\theta} + \delta\right) = \gamma$$

Эта формула означает, что доверительным интервалом является промежуток $\left(\tilde{\theta} - \delta, \tilde{\theta} + \delta\right)$. Концы доверительного интервала называют доверительными границами. Доверительные границы являются случайными величинами (они зависят от результатов выборки).
Доверительная вероятность γ, точность оценки δ и объем выборки n связаны между собой определенным соотношением. При этом если известны две какие-то величины из этих трех, то можно найти третью величину.

8.3. Интервальные оценки для генеральной средней

Рассмотрим правила построения доверительного интервала для оценки математического ожидания *a* нормального распределения при известном значении среднего квадратического отклонения σ. Пусть случайная величина *X* генеральной совокупности имеет нормальное распределение с известным значением среднего квадратического отклонения σ и неизвестным значением математического ожидания *a*.

Пусть из генеральной совокупности взята случайная выборка $x_1, x_2, \ldots, x_n$ и вычислена выборочная средняя $x_в$. Найдем доверительный интервал, в котором содержится *a*, с надежностью γ. Заметим, что элементы выборки $x_1, x_2, \ldots, x_n$ являются случайными величинами имеющими то же распределение, что и случайная величина X генеральной совокупности (нормальное). Поэтому, величина $x_в$ также имеет нормальное распределение, причем,

$$M(x_в) = a .$$

Найдем среднее квадратическое отклонение

$$\sigma(x_В) = \sqrt{D(X_В)} = \sqrt{\frac{D(x_1) + D(x_2) + \ldots + D(x_n)}{n^2}} = \sqrt{\frac{n\sigma^2}{n^2}} = \frac{\sigma}{\sqrt{n}}$$

Т.о.,

$$M(X_в) = a ,$$

$$\sigma(X_В) = \frac{\sigma}{\sqrt{n}} .$$

Потребуем, чтобы

$$P(|x_В - a| < \delta) = \gamma .$$

Поскольку случайная величина $X_в$ имеет нормальное распределение, то

$$P(|x_В - a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma(X_в)}\right) .$$

Отсюда

$$P(|x_В - a| < \delta) = 2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right) .$$

Обозначим,

$$t = \frac{\delta\sqrt{n}}{\sigma}$$

Тогда,

$$\delta = \frac{t\sigma}{\sqrt{n}}$$

Получим,

$$P\left(|x_в - a| < \frac{t\sigma}{\sqrt{n}}\right) = 2\Phi(t)$$

или

$$P\left(x_в - \frac{t\sigma}{\sqrt{n}} < a < x_в + \frac{t\sigma}{\sqrt{n}} \right) = 2\Phi(t)$$

где $x_в$ – выборочное среднее.

Поскольку величина P известна и равна γ, т.е.

$$\gamma = 2\Phi(t),$$

то доверительный интервал

$$\left(x_в - \frac{t\sigma}{\sqrt{n}},\ x_в + \frac{t\sigma}{\sqrt{n}} \right)$$

покрывает неизвестный параметр a с надежностью γ и точностью оценки

$$\delta = \frac{t\sigma}{\sqrt{n}}.$$

Число t определяется равенством $\gamma = 2\Phi(t)$ и находится с помощью таблицы для функции Лапласа (Приложение 2).

Пример. Случайная величина X имеет нормальное распределение с известным средним квадратическим отклонением $\sigma = 3$.

$$P\left(x_в - \frac{t\sigma}{\sqrt{n}} < a < x_в + \frac{t\sigma}{\sqrt{n}} \right) = 2\Phi(t) = 3.$$

Найти доверительный интервал для оценки неизвестного математического ожидания, если объем выборки $n = 36$ и задана надежность оценки $\gamma = 0{,}95$.

Решение. По таблице интегральной функции Лапласа $\Phi(t)$ (Приложение 2) и из условия $\gamma = 0{,}95$ найдем

$$2\Phi(t) = 0{,}95;$$

$$t = 1{,}96$$

Тогда точность оценки

$$\delta = \frac{t\sigma}{\sqrt{n}} = 1{,}96 \cdot \frac{3}{\sqrt{36}} = 0{,}98$$

Отсюда, доверительный интервал имеет вид

$$x_в\text{-}0{,}98 \le a \le x_в + 0{,}98 .$$

Ответ: $(x_в\text{-}0{,}98,\ x_в + 0{,}98)$.

Рассмотрим теперь правила построения доверительного интервала для оценки математического ожидания a нормального распределение при **неизвестном значении среднего квадратического отклонения** σ.

В этом случае для построения интервальной оценки генеральной средней a применяется случайная величина

$$T = \frac{x_в\text{-}a}{S} \cdot \sqrt{n} ,$$

где $x_в$ – выборочная средняя; S – «исправленное» среднее квадратическое отклонение; n – объем выборки.

Известно, что случайная величина T имеет распределение Стьюдента с $\nu = n - 1$ степенями свободы. Распределение Стьюдента определяется параметром n и не зависит от неизвестных параметров a и σ. В таблице Приложения 4 для различных ν приводятся числа вероятность превышения которых случайной величиной T равна заданному значению уровня значимости $\alpha = 1 - \gamma$.

Найдем вероятность

$$P\left(\left| \frac{X_в\text{-}a}{\frac{S}{\sqrt{n}}} \right| < t_\gamma \right) = \gamma .$$

Отсюда следует, что

$$P\left(x_в\text{-}t_\gamma \cdot \frac{S}{\sqrt{n}} < a < x_в + t_\gamma \cdot \frac{S}{\sqrt{n}} \right) = \gamma .$$

Здесь число t_γ должно быть найдено по таблице Приложения 4. Таким образом, доверительный интервал

$$\left(x_в\text{-}t_\gamma \cdot \frac{S}{\sqrt{n}} < a < x_в + t_\gamma \cdot \frac{S}{\sqrt{n}} \right)$$

покрывает неизвестный параметр a с надежностью γ.

Пример. Случайная величина X имеет нормальное распределение с неизвестным математическим ожиданием a и неизвестным средним квадратическим отклонением σ. По выборке объема n = 16 получены значения выборочной средней $x_в$ = 20,2 и «исправленное» среднее квадратическое отклонение S = 0,8. Оценить неизвестное математическое ожидание (найти доверительный интервал) с надежностью γ = 0,95.

Решение. По таблице распределения Стьюдента (Приложение 4) найдем значение t_γ.

Имеем n = 16; γ = 0,95; n - 1 = 15; α = 1 - γ = 0,05; t_γ = 2,13.

Тогда имеем

$$x_в\text{-}t_\gamma \cdot \frac{S}{\sqrt{n}} = 20{,}2 - 2{,}13 \cdot \frac{0{,}8}{\sqrt{16}} = 19{,}77$$

$$x_в + t_\gamma \cdot \frac{S}{\sqrt{n}} = 20{,}2 + 2{,}13 \cdot \frac{0{,}8}{\sqrt{16}} = 20{,}63$$

Значит, с надежностью γ = 0,95 математическое ожидание a заключено в интервале (19,77; 20,63).

Ответ. (19,77; 20,63).

8.4. Интервальные оценки для генеральной дисперсии

Пусть из генеральной совокупности X, распределенной по нормальному закону, взята случайная выборка объемом n и вычислена исправленная дисперсия S^2. Найдем доверительный интервал, покрывающий δ с надежностью γ.

Для доверительного интервала должно выполняться условие

$$P\left(\left|\sigma - S\right| < \delta\right) = \gamma$$

$$P\left(S - \delta < \sigma < S + \delta\right) = \gamma.$$

Введем параметр

$$q = \frac{\sigma}{S}.$$

Тогда получим условие

$$P\left(S(1-q) < \sigma < S(1+q)\right) = \gamma.$$

Будем рассматривать случайную величину (хи - квадрат)

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

Известно, что случайная величина χ имеет распределение Пирсона плотность распределения которого зависит только от n и не зависит от оцениваемого параметра δ.
Выполним некоторые преобразования в неравенствах для доверительного интервала (будем предполагать, что **0 < q < 1**).

$$S(1-q) < \sigma < S(1+q).$$

Отсюда,

$$\frac{1}{S(1+q)} < \frac{1}{\sigma} < \frac{1}{S(1-q)}$$

$$\frac{\sqrt{n-1}}{(1+q)} < \frac{S\sqrt{n-1}}{\sigma} < \frac{\sqrt{n-1}}{(1-q)}.$$

Тогда вероятность равна

$$P\left(\frac{\sqrt{n-1}}{(1+q)} < \frac{S\sqrt{n-1}}{\sigma} < \frac{\sqrt{n-1}}{(1-q)}\right) = \gamma.$$

Из этого уравнения по таблице Приложения 3 для распределения Пирсона (χ^2 - распределения) можно найти параметр q. Получив значение q, найдем доверительный интервал $S(1-q) < \sigma < S(1+q)$.

<u>Пример</u>. Генеральная совокупность имеет нормальное распределение. По выборке объема n = 25 получено значение «исправленного» среднего квадратического отклонения S = 0,8. Найти доверительный интервал, покрывающий генеральное среднее квадратическое отклонение δ с надежностью γ = 0,95.

<u>Решение</u>. По таблице (Приложение 3) находим q = 0,32. Найдем доверительный интервал

$$0{,}8(1-0{,}32) < \sigma < 0{,}8(1+0{,}32)$$

$$0{,}54 < \sigma < 1{,}06$$

<u>Ответ</u>. $0{,}54 < \sigma < 1{,}06$.

Рассмотрим теперь случай **q > 1**.
Т.к. σ > 0 всегда, то неравенство

$$S(1-q) < \sigma < S(1+q)$$

можно записать в виде

$$0 < \sigma < S(1+q).$$

Отсюда

$$\frac{1}{\sigma} > \frac{1}{S(1+q)}, \quad \frac{S\sqrt{n-1}}{\sigma} = \frac{\sqrt{n-1}}{1+q}.$$

Значит, случайная величина

$$\chi = \frac{S\sqrt{n-1}}{\sigma}$$

должна удовлетворять условию

$$\frac{\sqrt{n-1}}{1+q} < \chi < \infty .$$

Значение q можно найти по таблице Приложения 3.

<u>Пример</u>. Генеральная совокупность имеет нормальное распределение. По выборке объема n = 10 получено значение «исправленного» среднего

квадратического отклонения S = 0,16. Найти доверительный интервал, покрывающий генеральное среднее квадратическое отклонение σ с надежностью γ = 0,999.

Решение. По таблице Приложения 3 при условии, что n = 10, γ = 0,999 находим q = 1,8.

Найдем доверительный интервал $0 < \sigma < 0{,}16 \cdot (1 + 1{,}8)$.

Ответ. $0 < \sigma < 0{,}45$.