Лекция 13.
Закон больших чисел. Неравенство Чебышева. Теоремы Чебышева и Бернулли.

Изучение статистических закономерностей позволило установить, что при некоторых условиях суммарное поведение большого количества случайных величин почти утрачивает случайный характер и становится закономерным (иначе говоря, случайные отклонения от некоторого среднего поведения взаимно погашаются). В частности, если влияние на сумму отдельных слагаемых является равномерно малым, закон распределения суммы приближается к нормальному. Математическая формулировка этого утверждения дается в группе теорем, называемой **законом больших чисел**.

Неравенство Чебышева.

Неравенство Чебышева, используемое для доказательства дальнейших теорем, справедливо как для непрерывных, так и для дискретных случайных величин. Докажем его для дискретных случайных величин.

Теорема 13.1(неравенство Чебышева). $p(|X - M(X)| < \varepsilon) \geq D(X) / \varepsilon^2$. (13.1)
Доказательство. Пусть X задается рядом распределения

X	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

Так как события $|X - M(X)| < \varepsilon$ и $|X - M(X)| \geq \varepsilon$ противоположны, то $p (|X - M(X)| < \varepsilon) + p (|X - M(X)| \geq \varepsilon) = 1$, следовательно, $p (|X - M(X)| < \varepsilon) = 1 - p (|X - M(X)| \geq \varepsilon)$. Найдем $p (|X - M(X)| \geq \varepsilon)$.
$D(X) = (x_1 - M(X))^2 p_1 + (x_2 - M(X))^2 p_2 + \ldots + (x_n - M(X))^2 p_n$. Исключим из этой суммы те слагаемые, для которых $|X - M(X)| < \varepsilon$. При этом сумма может только уменьшиться, так как все входящие в нее слагаемые неотрицательны. Для определенности будем считать, что отброшены первые k слагаемых. Тогда
$D(X) \geq (x_{k+1} - M(X))^2 p_{k+1} + (x_{k+2} - M(X))^2 p_{k+2} + \ldots + (x_n - M(X))^2 p_n \geq \varepsilon^2 (p_{k+1} + p_{k+2} + \ldots + p_n)$.
Отметим, что $p_{k+1} + p_{k+2} + \ldots + p_n$ есть вероятность того, что $|X - M(X)| \geq \varepsilon$, так как это сумма вероятностей всех возможных значений X, для которых это неравенство справедливо. Следовательно, $D(X) \geq \varepsilon^2 p(|X - M(X)| \geq \varepsilon)$, или $p (|X - M(X)| \geq \varepsilon) \leq D(X) / \varepsilon^2$. Тогда вероятность противоположного события $p(|X - M(X)| < \varepsilon) \geq D(X) / \varepsilon^2$, что и требовалось доказать.

Теоремы Чебышева и Бернулли.

Теорема 13.2 (теорема Чебышева). Если $X_1, X_2, \ldots, X_n$ – попарно независимые случайные величины, дисперсии которых равномерно ограничены ($D(X_i) \leq C$), то для сколь угодно малого числа ε вероятность неравенства

$$\left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{M(X_1) + M(X_2) + \ldots + M(X_n)}{n} \right| < e$$

будет сколь угодно близка к 1, если число случайных величин достаточно велико.
Замечание. Иначе говоря, при выполнении этих условий

$$\lim_{n \to \infty} p\left(\left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{M(X_1) + M(X_2) + \ldots + M(X_n)}{n} \right| < e \right) = 1.$$

Доказательство. Рассмотрим новую случайную величину $\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$ и найдем ее математическое ожидание. Используя свойства математического ожидания, получим,

что $M\left(\frac{X_1+X_2+...+X_n}{n}\right)=\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}$. Применим к $\overline{X}$ неравенство Чебышева:

$p(\left|\frac{X_1+X_2+...+X_n}{n}-\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}\right|<e)\geq 1-\frac{D\left(\frac{X_1+X_2+...+X_n}{n}\right)}{e^2}$. Так как рассматриваемые случайные величины независимы, то, учитывая условие теоремы, имеем: $D\left(\frac{X_1+X_2+...+X_n}{n}\right)=\frac{D(X_1)+D(X_2)+...+D(X_n)}{n^2}\leq\frac{Cn}{n^2}=\frac{C}{n}$. Используя этот результат, представим предыдущее неравенство в виде:

$p(\left|\frac{X_1+X_2+...+X_n}{n}-\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}\right|<e)\geq 1-\frac{C}{ne^2}$. Перейдем к пределу при $n\to\infty$: $\lim\limits_{n\to\infty}p(\left|\frac{X_1+X_2+...+X_n}{n}-\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}\right|<e)\geq 1$. Поскольку вероятность не может быть больше 1, можно утверждать, что

$\lim\limits_{n\to\infty}p(\left|\frac{X_1+X_2+...+X_n}{n}-\frac{M(X_1)+M(X_2)+...+M(X_n)}{n}\right|<e)=1$. Теорема доказана.

Следствие.
Если X_1, X_2, …, X_n – попарно независимые случайные величины с равномерно ограниченными дисперсиями, имеющие одинаковое математическое ожидание, равное a, то для любого сколь угодно малого ε > 0 вероятность неравенства $\left|\frac{X_1+X_2+...+X_n}{n}-a\right|<e$ будет как угодно близка к 1, если число случайных величин достаточно велико. Иначе говоря, $\lim\limits_{n\to\infty}p(\left|\frac{X_1+X_2+...+X_n}{n}-a\right|<e)=1$.

Вывод: среднее арифметическое достаточно большого числа случайных величин принимает значения, близкие к сумме их математических ожиданий, то есть утрачивает характер случайной величины. Например, если проводится серия измерений какой-либо физической величины, причем: а) результат каждого измерения не зависит от результатов остальных, то есть все результаты представляют собой попарно независимые случайные величины; б) измерения производятся без систематических ошибок (их математические ожидания равны между собой и равны истинному значению a измеряемой величины); в) обеспечена определенная точность измерений, следовательно, дисперсии рассматриваемых случайных величин равномерно ограничены; то при достаточно большом числе измерений их среднее арифметическое окажется сколь угодно близким к истинному значению измеряемой величины.

Теорема Бернулли.

Теорема 13.3 (теорема Бернулли). Если в каждом из n независимых опытов вероятность p появления события A постоянна, то при достаточно большом числе испытаний вероятность того, что модуль отклонения относительной частоты появлений A в n опытах от p будет сколь угодно малым, как угодно близка к 1:

$$\lim_{n\to\infty}p\left(\left|\frac{m}{n}-p\right|<e\right)=1. \qquad (13.2)$$

Доказательство. Введем случайные величины X_1, X_2, …, X_n, где X_i – число появлений A в i-м опыте. При этом X_i могут принимать только два значения: 1(с вероятностью p) и 0 (с вероятностью $q = 1 - p$). Кроме того, рассматриваемые случайные величины попарно независимы и их дисперсии равномерно ограничены (так как $D(X_i) = pq$, $p + q = 1$, откуда $pq \le ¼$). Следовательно, к ним можно применить теорему Чебышева при $M_i = p$:

$$\lim_{n\to\infty} p\left(\left|\frac{X_1 + X_2 + ... + X_n}{n} - p\right| < e\right) = 1.$$

Но $\frac{X_1 + X_2 + ... + X_n}{n} = \frac{m}{n}$, так как X_i принимает значение, равное 1, при появлении A в данном опыте, и значение, равное 0, если A не произошло. Таким образом,

$$\lim_{n\to\infty} p\left(\left|\frac{m}{n} - p\right| < e\right) = 1,$$

что и требовалось доказать.

Замечание. Из теоремы Бернулли *не следует*, что $\lim_{n\to\infty} \frac{m}{n} = p$. Речь идет лишь о *вероятности* того, что разность относительной частоты и вероятности по модулю может стать сколь угодно малой. Разница заключается в следующем: при обычной сходимости, рассматриваемой в математическом анализе, для всех n, начиная с некоторого значения, неравенство $\left|\frac{m}{n} - p\right| < e$ выполняется всегда; в нашем случае могут найтись такие значения n, при которых это неравенство неверно. Этот вид сходимости называют **сходимостью по вероятности**.

Лекция 14.
Центральная предельная теорема Ляпунова. Предельная теорема Муавра-Лапласа.

Закон больших чисел не исследует вид предельного закона распределения суммы случайных величин. Этот вопрос рассмотрен в группе теорем, называемых **центральной предельной теоремой.** Они утверждают, что закон распределения суммы случайных величин, каждая из которых может иметь различные распределения, приближается к нормальному при достаточно большом числе слагаемых. Этим объясняется важность нормального закона для практических приложений.

Характеристические функции.

Для доказательства центральной предельной теоремы используется метод характеристических функций.

Определение 14.1. **Характеристической функцией** случайной величины X называется функция

$$g(t) = M(e^{itX}) \qquad (14.1)$$

Таким образом, $g(t)$ представляет собой математическое ожидание некоторой комплексной случайной величины $U = e^{itX}$, связанной с величиной X. В частности, если X – дискретная случайная величина, заданная рядом распределения, то

$$g(t) = \sum_{k=1}^{n} e^{itx_k} p_k . \qquad (14.2)$$

Для непрерывной случайной величины с плотностью распределения $f(x)$

$$g(t) = \int_{-\infty}^{+\infty} e^{itx} f(x)dx. \qquad (14.3)$$

Пример 1. Пусть X – число выпадений 6 очков при одном броске игральной кости. Тогда по формуле (14.2) $g(t) = e^{it\cdot 0}\cdot\frac{5}{6} + e^{it\cdot 1}\cdot\frac{1}{6} = \frac{5+e^{it}}{6}$.

Пример 2. Найдем характеристическую функцию для нормированной непрерывной случайной величины, распределенной по нормальному закону $\left(f(x) = \frac{1}{\sqrt{2p}} e^{-\frac{x^2}{2}}\right)$. По формуле (14.3) $g(t) = \int_{-\infty}^{+\infty} e^{itx}\frac{1}{\sqrt{2p}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2p}}\int_{-\infty}^{+\infty} e^{itx-\frac{x^2}{2}} dx = e^{-\frac{t^2}{2}}$ (использовалась формула $\int_{-\infty}^{+\infty} e^{-Ax^2\pm 2Bx-C} dx = \sqrt{\frac{p}{A}} e^{-\frac{AC-B^2}{A}}$ и то, что $i^2 = -1$).

Свойства характеристических функций.

1. Функцию $f(x)$ можно найти по известной функции $g(t)$ по формуле

$$f(x) = \frac{1}{2p}\int_{-\infty}^{+\infty} e^{-itx} g(t)dt. \qquad (14.4)$$

(преобразование (14.3) называется *преобразованием Фурье*, а преобразование (14.4) – *обратным преобразованием Фурье*).

2. Если случайные величины X и Y связаны соотношением $Y = aX$, то их характеристические функции связаны соотношением

$$g_y(t) = g_x(at). \qquad (14.5)$$

3. Характеристическая функция суммы независимых случайных величин равна произведению характеристических функций слагаемых: для $Y = \sum_{k=1}^{n} X_k$

$$g_y(t) = g_{x_1}(t) \cdot g_{x_2}(t) \cdot \ldots \cdot g_{x_n}(t) \qquad (14.6)$$

Теорема 14.1 (центральная предельная теорема для одинаково распределенных слагаемых). Если $X_1, X_2, \ldots, X_n, \ldots$ - независимые случайные величины с одинаковым законом распределения, математическим ожиданием m и дисперсией σ^2, то при неограниченном увеличении n закон распределения суммы $Y_n = \sum_{k=1}^{n} X_k$ неограниченно приближается к нормальному.

Доказательство.

Докажем теорему для непрерывных случайных величин $X_1, X_2, \ldots, X_n$ (доказательство для дискретных величин аналогично). Согласно условию теоремы, характеристические функции слагаемых одинаковы: $g_x(t) = \int_{-\infty}^{+\infty} e^{itx} f(x)dx.$ Тогда по свойству 3 характеристическая функция суммы Y_n будет $g_{y_n}(t) = g_x^n(t).$ Разложим функцию $g_x(t)$ в ряд Маклорена:

$$g_x(t) = g_x(0) + g_x'(0)t + \left(\frac{g_x''(0)}{2} + a(t) \right) t^2, \text{ где } a(t) \to 0 \text{ при } t \to 0.$$

Найдем $g_x(0) = \int_{-\infty}^{+\infty} f(x)dx = 1,$ $g_x'(0) = \int_{-\infty}^{+\infty} ixe^{itx} f(x)dx \Big|_{t=0} = i\int_{-\infty}^{+\infty} xe^{itx} f(x)dx \Big|_{t=0} = i\int_{-\infty}^{+\infty} xf(x)dx = im.$

Если предположить, что $m = 0$ (то есть перенести начало отсчета в точку m), то $g_x'(0) = 0$.

$g_x''(0) = -\int_{-\infty}^{+\infty} x^2 e^{itx} f(x)dx \Big|_{t=0} = -\int_{-\infty}^{+\infty} x^2 f(x)dx = -s^2$ (так как $m = 0$). Подставив полученные результаты в формулу Маклорена, найдем, что

$$g_x(t) = 1 - \left(\frac{s^2}{2} - a(t) \right) t^2.$$

Рассмотрим новую случайную величину $Z_n = \frac{Y_n}{s\sqrt{n}}$, отличающуюся от Y_n тем, что ее дисперсия при любом n равна 0. Так как Y_n и Z_n связаны линейной зависимостью, достаточно доказать, что Z_n распределена по нормальному закону, или, что то же самое, что ее характеристическая функция приближается к характеристической функции нормального закона (см. пример 2). По свойству характеристических функций

$$g_{z_n}(t) = g_{y_n}\left(\frac{t}{s\sqrt{n}} \right) = \left(g_x\left(\frac{t}{s\sqrt{n}} \right) \right)^n = \left(1 - \left(\frac{s^2}{2} - a\left(\frac{t}{s\sqrt{n}} \right) \right) \frac{t^2}{ns^2} \right)^n.$$

Прологарифмируем полученное выражение:

$$\ln g_{z_n}(t) = n\ln(1-k), \text{ где } k = \left(\frac{s^2}{2} - a\left(\frac{t}{s\sqrt{n}} \right) \right) \frac{t^2}{ns^2}, \quad \lim_{n\to\infty} k = 0.$$

Разложим $\ln(1-k)$ в ряд при $n \to \infty$, ограничившись двумя членами разложения, тогда $\ln(1 - k) \approx - k$. Отсюда

$\lim_{n\to\infty} \ln g_{z_n}(t) = \lim_{n\to\infty} n\cdot(-k) = \lim_{n\to\infty}\left(-\frac{t^2}{2} + a\left(\frac{t}{s\sqrt{n}}\right)\frac{t^2}{s^2}\right) = -\frac{t^2}{2} + \lim_{n\to\infty}\frac{t^2}{s^2}a\left(\frac{t}{s\sqrt{n}}\right)$, где последний предел равен 0, так как $a(t)\to 0$ при $t\to 0$. Следовательно, $\lim_{n\to\infty}\ln g_{z_n}(t) = -\frac{t^2}{2}$, то есть $\lim_{n\to\infty} g_{z_n}(t) = e^{-\frac{t^2}{2}}$ - характеристическая функция нормального распределения. Итак, при неограниченном увеличении числа слагаемых характеристическая функция величины Z_n неограниченно приближается к характеристической функции нормального закона; следовательно, закон распределения Z_n (и Y_n) неограниченно приближается к нормальному. Теорема доказана.

А.М.Ляпунов доказал центральную предельную теорему для условий более общего вида:

Теорема 14.2 (теорема Ляпунова). Если случайная величина X представляет собой сумму очень большого числа взаимно независимых случайных величин, для которых выполнено условие:

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n} b_k}{\left(\sum_{k=1}^{n} D_k\right)^{\frac{3}{2}}}, \qquad (14.7)$$

где b_k – третий абсолютный центральный момент величины $X_к$, а D_k – ее дисперсия, то X имеет распределение, близкое к нормальному (условие Ляпунова означает, что влияние каждого слагаемого на сумму ничтожно мало).

Практически можно использовать центральную предельную теорему при достаточно небольшом количестве слагаемых, так как вероятностные расчеты требуют сравнительно малой точности. Опыт показывает, что для суммы даже десяти и менее слагаемых закон их распределения можно заменить нормальным.

Частным случаем центральной предельной теоремы для дискретных случайных величин является теорема Муавра-Лапласа.

Теорема 14.3 (теорема Муавра-Лапласа). Если производится n независимых опытов, в каждом из которых событие A появляется с вероятностью p, то справедливо соотношение:

$$p\left(a < \frac{Y-np}{\sqrt{npq}} < b\right) = \Phi(b) - \Phi(a), \qquad (14.8)$$

где Y – число появлений события A в n опытах, $q = 1 - p$.

Доказательство.

Будем считать, что $Y = \sum_{i=1}^{n} X_i$, где X_i – число появлений события A в i-м опыте. Тогда случайную величину $Z = \frac{Y - m_y}{s_y}$ (см. теорему 14.1) можно считать распределенной по нормальному закону и нормированной, следовательно, вероятность ее попадания в интервал (α, β) можно найти по формуле

$$p(a < Z < b) = \Phi(b) - \Phi(a).$$

Поскольку Y имеет биномиальное распределение, $m_y = np$, $D_y = npq$, $s_y = \sqrt{npq}$. Тогда $Z = \frac{Y-np}{\sqrt{npq}}$. Подставляя это выражение в предыдущую формулу, получим равенство (14.8).

Следствие.

В условиях теоремы Муавра-Лапласа вероятность $p_n(k)$ того, что событие A появится в n опытах ровно k раз, при большом количестве опытов можно найти по формуле:

$$p_n(k) \approx \frac{1}{\sqrt{npq}} \cdot j(x), \qquad (14.9)$$

где $x = \frac{k-np}{\sqrt{npq}}$, а $j(x) = \frac{1}{\sqrt{2p}} e^{-\frac{x^2}{2}}$ (значения этой функции приводятся в специальных таблицах).

Пример 3. Найти вероятность того, что при 100 бросках монеты число выпадений герба окажется в пределах от 40 до 60.

Применим формулу (14.8), учитывая, что $n = 0{,}5$. Тогда $np = 100 \cdot 0{,}5 = 50$, $\sqrt{npq} = \sqrt{100 \cdot 0{,}5 \cdot (1-0{,}5)} = 5$. Тогда, если $40 < Y < 60$, $-2 < \frac{Y-50}{5} < 2$. Следовательно,

$$p(40 < Y < 60) = p\left(-2 < \frac{Y-50}{5} < 2\right) = \Phi(2) - \Phi(-2) = 0{,}9772 - 0{,}0228 = 0{,}9544.$$

Пример 4. В условиях предыдущего примера найти вероятность того, что выпадет 45 гербов.

Найдем $x = \frac{45-50}{5} = -1$, тогда $p_{100}(45) \approx \frac{1}{5} \cdot j(-1) = \frac{1}{5} \cdot j(1) = \frac{1}{5} \cdot 0{,}2420 = 0{,}0484$.

Лекция 15.
Основные понятия математической статистики. Генеральная совокупность и выборка. Вариационный ряд, статистический ряд. Группированная выборка. Группированный статистический ряд. Полигон частот. Выборочная функция распределения и гистограмма.

Математическая статистика занимается установлением закономерностей, которым подчинены массовые случайные явления, на основе обработки статистических данных, полученных в результате наблюдений. Двумя основными задачами математической статистики являются:
- определение способов сбора и группировки этих статистических данных;
- разработка методов анализа полученных данных в зависимости от целей исследования, к которым относятся:
а) оценка неизвестной вероятности события; оценка неизвестной функции распределения; оценка параметров распределения, вид которого известен; оценка зависимости от других случайных величин и т.д.;
б) проверка статистических гипотез о виде неизвестного распределения или о значениях параметров известного распределения.
Для решения этих задач необходимо выбрать из большой совокупности однородных объектов ограниченное количество объектов, по результатам изучения которых можно сделать прогноз относительно исследуемого признака этих объектов.
Определим основные понятия математической статистики.
Генеральная совокупность – все множество имеющихся объектов.
Выборка – набор объектов, случайно отобранных из генеральной совокупности.
Объем генеральной совокупности *N* и объем выборки *n* – число объектов в рассматриваемой совокупности.
Виды выборки:
Повторная – каждый отобранный объект перед выбором следующего возвращается в генеральную совокупность;
Бесповторная – отобранный объект в генеральную совокупность не возвращается.

Замечание. Для того, чтобы по исследованию выборки можно было сделать выводы о поведении интересующего нас признака генеральной совокупности, нужно, чтобы выборка правильно представляла пропорции генеральной совокупности, то есть была **репрезентативной** (представительной). Учитывая закон больших чисел, можно утверждать, что это условие выполняется, если каждый объект выбран случайно, причем для любого объекта вероятность попасть в выборку одинакова.

Первичная обработка результатов.
Пусть интересующая нас случайная величина X принимает в выборке значение x_1 n_1 раз, x_2 – n_2 раз, …, $x_к$ – $n_к$ раз, причем $\sum_{i=1}^{k} n_k = n$, где n – объем выборки. Тогда наблюдаемые значения случайной величины $x_1, x_2, \ldots, x_к$ называют **вариантами**, а $n_1, n_2, \ldots, n_к$ – **частотами**. Если разделить каждую частоту на объем выборки, то получим **относительные частоты** $w_i = \frac{n_i}{n}$.
Последовательность вариант, записанных в порядке возрастания, называют **вариационным** рядом, а перечень вариант и соответствующих им частот или относительных частот – **статистическим рядом**:

x_i	x_1	x_2	…	x_k
n_i	n_1	n_2	…	n_k
w_i	w_1	w_2	…	w_k

Пример.
При проведении 20 серий из 10 бросков игральной кости число выпадений шести очков оказалось равным 1,1,4,0,1,2,1,2,2,0,5,3,3,1,0,2,2,3,4,1.Составим вариационный ряд: 0,1,2,3,4,5.
Статистический ряд для абсолютных и относительных частот имеет вид:

x_i	0	1	2	3	4	5
n_i	3	6	5	3	2	1
w_i	0,15	0,3	0,25	0,15	0,1	0,05

Если исследуется некоторый непрерывный признак, то вариационный ряд может состоять из очень большого количества чисел. В этом случае удобнее использовать **группированную выборку**. Для ее получения интервал, в котором заключены все наблюдаемые значения признака, разбивают на несколько равных частичных интервалов длиной h, а затем находят для каждого частичного интервала n_i – сумму частот вариант, попавших в i-й интервал. Составленная по этим результатам таблица называется **группированным статистическим рядом**:

Номера интервалов	1	2	…	k
Границы интервалов	$(a, a + h)$	$(a + h, a + 2h)$	…	$(b – h, b)$
Сумма частот вариант, попавших в интервал	n_1	n_2	…	n_k

Полигон частот. Выборочная функция распределения и гистограмма.

Для наглядного представления о поведении исследуемой случайной величины в выборке можно строить различные графики. Один из них – **полигон частот**: ломаная, отрезки которой соединяют точки с координатами (x_1, n_1), (x_2, n_2),…, (x_k, n_k), где x_i откладываются на оси абсцисс, а n_i – на оси ординат. Если на оси ординат откладывать не абсолютные (n_i), а относительные (w_i) частоты, то получим **полигон относительных частот** (рис.1)**.**

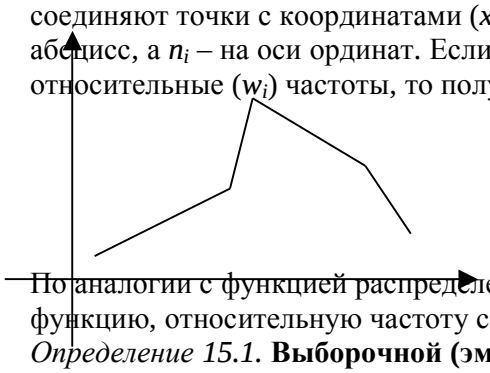

Рис. 1.

По аналогии с функцией распределения случайной величины можно задать некоторую функцию, относительную частоту события $X < x$.
Определение 15.1. **Выборочной (эмпирической) функцией распределения** называют функцию $F^*(x)$, определяющую для каждого значения x относительную частоту события $X < x$. Таким образом,

$$F^*(x) = \frac{n_x}{n}, \qquad (15.1)$$

где n_x – число вариант, меньших x, n – объем выборки.

Замечание. В отличие от эмпирической функции распределения, найденной опытным путем, функцию распределения $F(x)$ генеральной совокупности называют *теоретической функцией распределения*. $F(x)$ определяет вероятность события $X < x$, а $F^*(x)$ – его относительную частоту. При достаточно больших n, как следует из теоремы Бернулли, $F^*(x)$ стремится по вероятности к $F(x)$.
Из определения эмпирической функции распределения видно, что ее свойства совпадают со свойствами $F(x)$, а именно:

1) $0 \le F^*(x) \le 1$.
2) $F^*(x)$ – неубывающая функция.
3) Если x_1 – наименьшая варианта, то $F^*(x) = 0$ при $x \le x_1$; если $x_к$ – наибольшая варианта, то $F^*(x) = 1$ при $x > x_к$.

Для непрерывного признака графической иллюстрацией служит **гистограмма**, то есть ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные интервалы длиной h, а высотами – отрезки длиной n_i/h (гистограмма частот) или w_i/h (гистограмма относительных частот). В первом случае площадь гистограммы равна объему выборки, во втором – единице (рис.2).

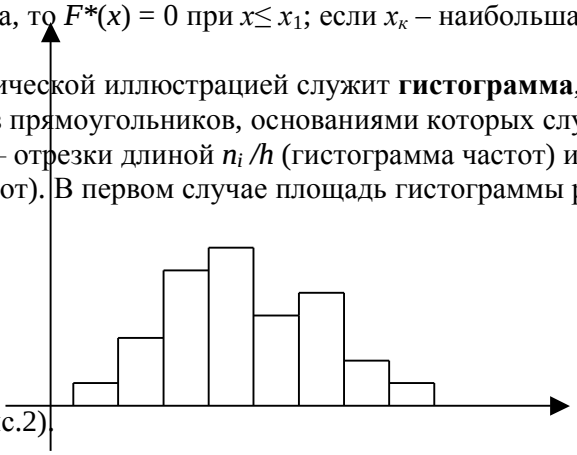

Рис.2.

Лекция 16.
Числовые характеристики статистического распределения: выборочное среднее, оценки дисперсии, оценки моды и медианы, оценки начальных и центральных моментов. Статистическое описание и вычисление оценок параметров двумерного случайного вектора.

Одна из задач математической статистики: по имеющейся выборке оценить значения числовых характеристик исследуемой случайной величины.

Определение 16.1. **Выборочным средним** называется среднее арифметическое значений случайной величины, принимаемых в выборке:

$$\bar{x}_B = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{n_1 x_1 + n_2 x_2 + ... + n_k x_k}{n} = \frac{\sum_{i=1}^{k} n_i x_i}{n}, \qquad (16.1)$$

где x_i – варианты, n_i - частоты.

Замечание. Выборочное среднее служит для оценки математического ожидания исследуемой случайной величины. В дальнейшем будет рассмотрен вопрос, насколько точной является такая оценка.

Определение 16.2. **Выборочной дисперсией** называется

$$D_B = \frac{\sum_{i=1}^{n} (x_i - \bar{x}_B)^2}{n} = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x}_B)^2}{n}, \qquad (16.2)$$

а **выборочным средним квадратическим отклонением** –

$$s_B = \sqrt{D_B}. \qquad (16.3)$$

Так же, как в теории случайных величин, можно доказать, что справедлива следующая формула для вычисления выборочной дисперсии:

$$D = \overline{x^2} - (\bar{x})^2. \qquad (16.4)$$

Пример 1. Найдем числовые характеристики выборки, заданной статистическим рядом

x_i	2	5	7	8
n_i	3	8	7	2

$$\bar{x}_B = \frac{2\cdot 3 + 5\cdot 8 + 7\cdot 7 + 8\cdot 2}{20} = 5{,}55; \quad D_B = \frac{4\cdot 3 + 25\cdot 8 + 49\cdot 7 + 64\cdot 2}{20} - 5{,}55^2 = 3{,}3475; \quad s_B = \sqrt{3{,}3475} = 1{,}83.$$

Другими характеристиками вариационного ряда являются:

- **мода M_0** – варианта, имеющая наибольшую частоту (в предыдущем примере $M_0 = 5$).
- **медиана m_e** - варианта, которая делит вариационный ряд на две части, равные по числу вариант. Если число вариант нечетно ($n = 2k + 1$), то $m_e = x_{k+1}$, а при четном $n = 2k$

$m_e = \frac{x_k + x_{k+1}}{2}$. В частности, в примере 1 $m_e = \frac{5+7}{2} = 6$.

Оценки начальных и центральных моментов (так называемые эмпирические моменты) определяются аналогично соответствующим теоретическим моментам:

- **начальным эмпирическим моментом порядка k** называется

$$M_k = \frac{\sum n_i x_i^k}{n}. \qquad (16.5)$$

В частности, $M_1 = \frac{\sum n_i x_i}{n} = \bar{x}_B$, то есть начальный эмпирический момент первого порядка равен выборочному среднему.

- **центральным эмпирическим моментом порядка k** называется

$$m_k = \frac{\sum n_i (x_i - \bar{x}_B)^k}{n}. \qquad (16.6)$$

В частности, $m_2 = \frac{\sum n_i (x_i - \bar{x}_B)^2}{n} = D_B$, то есть центральный эмпирический момент второго порядка равен выборочной дисперсии.

Статистическое описание и вычисление характеристик двумерного случайного вектора.

При статистическом исследовании двумерных случайных величин основной задачей является обычно выявление связи между составляющими.

Двумерная выборка представляет собой набор значений случайного вектора: (x_1, y_1), (x_2, y_2), …, (x_n, y_n). Для нее можно определить выборочные средние составляющих: $\bar{x}_B = \frac{\sum x_i}{n}$, $\bar{y}_B = \frac{\sum y_i}{n}$ и соответствующие выборочные дисперсии и средние квадратические отклонения.

Кроме того, можно вычислить **условные средние**: $\bar{y}_x$ - среднее арифметическое наблюдавшихся значений *Y*, соответствующих $X = x$, и $\bar{x}_y$ - среднее значение наблюдавшихся значений *X*, соответствующих $Y = y$.

Если существует зависимость между составляющими двумерной случайной величины, она может иметь разный вид: функциональная зависимость, если каждому возможному значению *X* соответствует одно значение *Y*, и статистическая, при которой изменение одной величины приводит к изменению распределения другой. Если при этом в результате изменения одной величины меняется среднее значение другой, то статистическую зависимость между ними называют корреляционной.

Лекция 17.

Основные свойства статистических характеристик параметров распределения: несмещенность, состоятельность, эффективность. Несмещенность и состоятельность выборочного среднего как оценки математического ожидания. Смещенность выборочной дисперсии. Пример несмещенной оценки дисперсии. Асимптотически несмещенные оценки. Способы построения оценок: метод наибольшего правдоподобия, метод моментов, метод квантили, метод наименьших квадратов, байесовский подход к получению оценок.

Получив статистические оценки параметров распределения (выборочное среднее, выборочную дисперсию и т.д.), нужно убедиться, что они в достаточной степени служат приближе-

нием соответствующих характеристик генеральной совокупности. Определим требования, которые должны при этом выполняться.
Пусть Θ* - статистическая оценка неизвестного параметра Θ теоретического распределения. Извлечем из генеральной совокупности несколько выборок одного и того же объема n и вычислим для каждой из них оценку параметра Θ: $\Theta_1^*, \Theta_2^*, ..., \Theta_k^*$. Тогда оценку Θ* можно рассматривать как случайную величину, принимающую возможные значения $\Theta_1^*, \Theta_2^*, ..., \Theta_k^*$.
Если математическое ожидание Θ* не равно оцениваемому параметру, мы будем получать при вычислении оценок систематические ошибки одного знака (с избытком, если $M(\Theta^*) > \Theta$, и с недостатком, если $M(\Theta^*) < \Theta$). Следовательно, необходимым условием отсутствия систематических ошибок является требование $M(\Theta^*) = \Theta$.

Определение 17.2. Статистическая оценка Θ* называется **несмещенной**, если ее математическое ожидание равно оцениваемому параметру Θ при любом объеме выборки:

$$M(\Theta^*) = \Theta. \qquad (17.1)$$

Смещенной называют оценку, математическое ожидание которой не равно оцениваемому параметру.
Однако несмещенность не является достаточным условием хорошего приближения к истинному значению оцениваемого параметра. Если при этом возможные значения Θ* могут значительно отклоняться от среднего значения, то есть дисперсия Θ* велика, то значение, найденное по данным одной выборки, может значительно отличаться от оцениваемого параметра. Следовательно, требуется наложить ограничения на дисперсию.

Определение 17.2. Статистическая оценка называется **эффективной**, если она при заданном объеме выборки n имеет наименьшую возможную дисперсию.

При рассмотрении выборок большого объема к статистическим оценкам предъявляется еще и требование состоятельности.

Определение 17.3. **Состоятельной** называется статистическая оценка, которая при $n \to \infty$ стремится по вероятности к оцениваемому параметру (если эта оценка несмещенная, то она будет состоятельной, если при $n \to \infty$ ее дисперсия стремится к 0).

Убедимся, что $\bar{x}_B$ представляет собой несмещенную оценку математического ожидания $M(X)$. Будем рассматривать $\bar{x}_B$ как случайную величину, а $x_1, x_2, \ldots, x_n$, то есть значения исследуемой случайной величины, составляющие выборку, – как независимые, одинаково распределенные случайные величины $X_1, X_2, \ldots, X_n$, имеющие математическое ожидание a. Из свойств математического ожидания следует, что

$$M(\bar{X}_B) = M\left(\frac{X_1 + X_2 + ... + X_n}{n}\right) = a.$$

Но, поскольку каждая из величин $X_1, X_2, \ldots, X_n$ имеет такое же распределение, что и генеральная совокупность, $a = M(X)$, то есть $M(\bar{X}_B) = M(X)$, что и требовалось доказать.
Выборочное среднее является не только несмещенной, но и состоятельной оценкой математического ожидания. Если предположить, что $X_1, X_2, \ldots, X_n$ имеют ограниченные дисперсии, то из теоремы Чебышева следует, что их среднее арифметическое, то есть $\bar{X}_B$, при увеличении n стремится по вероятности к математическому ожиданию a каждой их величин, то есть к $M(X)$. Следовательно, выборочное среднее есть состоятельная оценка математического ожидания.
В отличие от выборочного среднего, выборочная дисперсия является смещенной оценкой дисперсии генеральной совокупности. Можно доказать, что

$$M(D_B) = \frac{n-1}{n} D_\Gamma , \qquad (17.2)$$

где D_Γ – истинное значение дисперсии генеральной совокупности. Можно предложить другую оценку дисперсии – **исправленную дисперсию s^2**, вычисляемую по формуле

$$s^2 = \frac{n}{n-1} D_B = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x}_B)^2}{n-1} . \qquad (17.3)$$

Такая оценка будет являться несмещенной. Ей соответствует **исправленное среднее квадратическое отклонение**

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{k} n_i (x_i - \bar{x}_B)^2}{n-1}} . \qquad (17.4)$$

Определение 17.4. Оценка некоторого признака называется **асимптотически несмещенной**, если для выборки $x_1, x_2, \ldots, x_n$

$$\lim_{n\to\infty} \frac{x_1 + x_2 + \ldots + x_n}{n} = X , \qquad (17.5)$$

где X – истинное значение исследуемой величины.

Способы построения оценок.

1. Метод наибольшего правдоподобия.

Пусть X – дискретная случайная величина, которая в результате n испытаний приняла значения $x_1, x_2, \ldots, x_n$. Предположим, что нам известен закон распределения этой величины, определяемый параметром Θ, но неизвестно численное значение этого параметра. Найдем его точечную оценку.

Пусть $p(x_i, \Theta)$ – вероятность того, что в результате испытания величина X примет значение x_i. Назовем **функцией правдоподобия** дискретной случайной величины X функцию аргумента Θ, определяемую по формуле:

$$L(x_1, x_2, \ldots, x_n; \Theta) = p(x_1,\Theta)p(x_2,\Theta)\ldots p(x_n,\Theta).$$

Тогда в качестве точечной оценки параметра Θ принимают такое его значение $\Theta^* = \Theta(x_1, x_2, \ldots, x_n)$, при котором функция правдоподобия достигает максимума. Оценку Θ* называют **оценкой наибольшего правдоподобия**.

Поскольку функции L и $\ln L$ достигают максимума при одном и том же значении Θ, удобнее искать максимум $\ln L$ – **логарифмической функции правдоподобия**. Для этого нужно:

1) найти производную $\frac{d \ln L}{d\Theta}$;
2) приравнять ее нулю (получим так называемое *уравнение правдоподобия*) и найти критическую точку;
3) найти вторую производную $\frac{d^2 \ln L}{d\Theta^2}$; если она отрицательна в критической точке, то это – точка максимума.

Достоинства метода наибольшего правдоподобия: полученные оценки состоятельны (хотя могут быть смещенными), распределены асимптотически нормально при больших значениях n и имеют наименьшую дисперсию по сравнению с другими асимптотически нормальными оценками; если для оцениваемого параметра Θ существует эффективная оценка Θ*, то уравнение правдоподобия имеет единственное решение Θ*; метод наиболее полно использует данные выборки и поэтому особенно полезен в случае малых выборок.

Недостаток метода наибольшего правдоподобия: сложность вычислений.

Для непрерывной случайной величины с известным видом плотности распределения $f(x)$ и неизвестным параметром Θ функция правдоподобия имеет вид:

$$L(x_1, x_2, \ldots, x_n; \Theta) = f(x_1,\Theta)f(x_2,\Theta)\ldots f(x_n,\Theta).$$

Оценка наибольшего правдоподобия неизвестного параметра проводится так же, как для дискретной случайной величины.

2. Метод моментов.

Метод моментов основан на том, что начальные и центральные эмпирические моменты являются состоятельными оценками соответственно начальных и центральных теоретических моментов, поэтому можно приравнять теоретические моменты соответствующим эмпирическим моментам того же порядка.

Если задан вид плотности распределения $f(x, \Theta)$, определяемой одним неизвестным параметром Θ, то для оценки этого параметра достаточно иметь одно уравнение. Например, можно приравнять начальные моменты первого порядка:

$$\bar{x}_B = M(X) = \int_{-\infty}^{\infty} xf(x;\Theta)dx = j(\Theta),$$

получив тем самым уравнение для определения Θ. Его решение Θ^* будет точечной оценкой параметра, которая является функцией от выборочного среднего и, следовательно, и от вариант выборки:

$$\Theta = \psi(x_1, x_2, \ldots, x_n).$$

Если известный вид плотности распределения $f(x, \Theta_1, \Theta_2)$ определяется двумя неизвестными параметрами Θ_1 и Θ_2, то требуется составить два уравнения, например

$$\nu_1 = M_1, \quad \mu_2 = m_2.$$

Отсюда $\begin{cases} M(X) = \bar{x}_B \\ D(X) = D_B \end{cases}$ - система двух уравнений с двумя неизвестными Θ_1 и Θ_2. Ее решениями будут точечные оценки Θ_1^* и Θ_2^* - функции вариант выборки:

$$\Theta_1 = \psi_1(x_1, x_2, \ldots, x_n),$$
$$\Theta_2 = \psi_2(x_1, x_2, \ldots, x_n).$$

3. Метод наименьших квадратов.

Если требуется оценить зависимость величин y и x, причем известен вид связывающей их функции, но неизвестны значения входящих в нее коэффициентов, их величины можно оценить по имеющейся выборке с помощью метода наименьших квадратов. Для этого функция $y = \varphi(x)$ выбирается так, чтобы сумма квадратов отклонений наблюдаемых значений $y_1, y_2, \ldots, y_n$ от $\varphi(x_i)$ была минимальной:

$$\sum_{i=1}^{n} (y_i - j(x_i))^2 = \min.$$

При этом требуется найти стационарную точку функции $\varphi(x; a, b, c\ldots)$, то есть решить систему:

$$\begin{cases} \sum_{i=1}^{n} (y_i - j(x_i;a,b,c\ldots))\left(\frac{\partial j}{\partial a}\right)_i = 0 \\ \sum_{i=1}^{n} (y_i - j(x_i;a,b,c\ldots))\left(\frac{\partial j}{\partial b}\right)_i = 0 \\ \sum_{i=1}^{n} (y_i - j(x_i;a,b,c\ldots))\left(\frac{\partial j}{\partial c}\right)_i = 0 \\ \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

(решение, конечно, возможно только в случае, когда известен конкретный вид функции φ). Рассмотрим в качестве примера подбор параметров линейной функции методом наименьших квадратов.

Для того, чтобы оценить параметры a и b в функции $y = ax + b$, найдем $\left(\frac{\partial j}{\partial a}\right)_i = x_i; \left(\frac{\partial j}{\partial b}\right)_i = 1.$

Тогда $\begin{cases} \sum_{i=1}^{n} (y_i - (ax_i + b))x_i = 0 \\ \sum_{i=1}^{n} (y_i - (ax_i + b)) = 0 \end{cases}$. Отсюда $\begin{cases} \sum_{i=1}^{n} x_i y_i - a\sum_{i=1}^{n} x_i^2 - b\sum_{i=1}^{n} x_i = 0 \\ \sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i - bn = 0 \end{cases}$. Разделив оба полученных уравнения на n и вспомнив определения эмпирических моментов, можно получить выражения для a и b в виде:

$a = \frac{(K_{xy})_B}{(D_x)_B}$, $b = \bar{y}_B - \frac{(K_{xy})_B}{(D_x)_B}\bar{x}_B$. Следовательно, связь между x и y можно задать в виде:

$$y - \bar{y}_B = \frac{(K_{xy})_B}{(D_x)_B}(x - \bar{x}_B).$$

4. Байесовский подход к получению оценок.

Пусть (Y, X) – случайный вектор, для которого известна плотность $p(y|x)$ условного распределения Y при каждом значении $X = x$. Если в результате эксперимента получены лишь значения Y, а соответствующие значения X неизвестны, то для оценки некоторой заданной функции $\varphi(x)$ в качестве ее приближенного значения предлагается искать условное математическое ожидание $M(\varphi(x)\|Y)$, вычисляемое по формуле:

$y(Y) = \frac{\int j(x)p(Y \mid x)p(x)dm(x)}{q(Y)}$, где $q(y) = \int p(y \mid x)p(x)dm(x)$, $p(x)$ – плотность безусловного распределения X, $q(y)$ – плотность безусловного распределения Y. Задача может быть решена только тогда, когда известна $p(x)$. Иногда, однако, удается построить состоятельную оценку для $q(y)$, зависящую только от полученных в выборке значений Y.

Лекция 18.
Интервальное оценивание неизвестных параметров. Точность оценки, доверительная вероятность (надежность), доверительный интервал. Построение доверительных интервалов для оценки математического ожидания нормального распределения при известной и при неизвестной дисперсии. Доверительные интервалы для оценки среднего квадратического отклонения нормального распределения.

При выборке малого объема точечная оценка может значительно отличаться от оцениваемого параметра, что приводит к грубым ошибкам. Поэтому в таком случае лучше пользоваться *интервальными оценками*, то есть указывать интервал, в который с заданной вероятностью попадает истинное значение оцениваемого параметра. Разумеется, чем меньше длина этого интервала, тем точнее оценка параметра. Поэтому, если для оценки Θ* некоторого параметра Θ справедливо неравенство | Θ* - Θ | < δ, число δ > 0 характеризует **точность оценки** (чем

меньше δ, тем точнее оценка). Но статистические методы позволяют говорить только о том, что это неравенство выполняется с некоторой вероятностью.

Определение 18.1. **Надежностью (доверительной вероятностью)** оценки Θ* параметра Θ называется вероятность γ того, что выполняется неравенство | Θ* - Θ | < δ. Если заменить это неравенство двойным неравенством – δ < Θ* - Θ < δ, то получим:

$$p\,(\,\Theta^* - \delta < \Theta < \Theta^* + \delta\,) = \gamma.$$

Таким образом, γ есть вероятность того, что Θ попадает в интервал (Θ* - δ, Θ* + δ).

Определение 18.2. **Доверительным** называется интервал, в который попадает неизвестный параметр с заданной надежностью γ.

Построение доверительных интервалов.

1. Доверительный интервал для оценки математического ожидания нормального распределения при известной дисперсии.

Пусть исследуемая случайная величина X распределена по нормальному закону с известным средним квадратическим σ, и требуется по значению выборочного среднего $\bar{x}_B$ оценить ее математическое ожидание a. Будем рассматривать выборочное среднее $\bar{x}_B$ как случайную величину $\bar{X}$, а значения вариант выборки $x_1, x_2,\ldots, x_n$ как одинаково распределенные независимые случайные величины $X_1, X_2,\ldots, X_n$, каждая из которых имеет математическое ожидание a и среднее квадратическое отклонение σ. При этом $M(\bar{X}) = a$, $s(\bar{X}) = \frac{s}{\sqrt{n}}$ (используем свойства математического ожидания и дисперсии суммы независимых случайных величин). Оценим вероятность выполнения неравенства $|\bar{X} - a| < d$. Применим формулу для вероятности попадания нормально распределенной случайной величины в заданный интервал: $p\,(|X - a| < d) = 2\Phi\left(\frac{d}{s}\right)$. Тогда , с учетом того, что $s(\bar{X}) = \frac{s}{\sqrt{n}}$, $p\,(|\bar{X} - a| < d) = 2\Phi\left(\frac{d\sqrt{n}}{s}\right) =$

$=2\Phi(t)$, где $t = \frac{d\sqrt{n}}{s}$. Отсюда $d = \frac{ts}{\sqrt{n}}$, и предыдущее равенство можно переписать так:

$$p\left(\bar{x}_B - \frac{ts}{\sqrt{n}} < a < \bar{x}_B + \frac{ts}{\sqrt{n}}\right) = 2\Phi(t) = g\ . \qquad (18.1)$$

Итак, значение математического ожидания a с вероятностью (надежностью) γ попадает в интервал $\left(\bar{x}_B - \frac{ts}{\sqrt{n}}; \bar{x}_B + \frac{ts}{\sqrt{n}}\right)$, где значение t определяется из таблиц для функции Лапласа так, чтобы выполнялось равенство $2\Phi(t) = \gamma$.

Пример. Найдем доверительный интервал для математического ожидания нормально распределенной случайной величины, если объем выборки $n = 49$, $\bar{x}_B = 2{,}8$, $\sigma = 1{,}4$, а доверительная вероятность $\gamma = 0{,}9$.

Определим t, при котором $\Phi(t) = 0{,}9:2 = 0{,}45$: $t = 1{,}645$. Тогда

$2{,}8 - \frac{1{,}645 \cdot 1{,}4}{\sqrt{49}} < a < 2{,}8 + \frac{1{,}645 \cdot 1{,}4}{\sqrt{14}}$, или $2{,}471 < a < 3{,}129$. Найден доверительный интервал, в который попадает a с надежностью 0,9.

2. Доверительный интервал для оценки математического ожидания нормального распределения при неизвестной дисперсии.

Если известно, что исследуемая случайная величина X распределена по нормальному закону с неизвестным средним квадратическим отклонением, то для поиска доверительного интервала для ее математического ожидания построим новую случайную величину

$$T = \frac{\bar{x}_B - a}{\frac{s}{\sqrt{n}}}, \qquad (18.2)$$

где $\bar{x}_B$ - выборочное среднее, s – исправленная дисперсия, n – объем выборки. Эта случайная величина, возможные значения которой будем обозначать t, имеет распределение Стьюдента (см. лекцию 12) с $k = n - 1$ степенями свободы.

Поскольку плотность распределения Стьюдента $s(t,n) = B_n\left(1+\frac{t^2}{n-1}\right)^{-\frac{n}{2}}$, где

$B_n = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{p\,(n-1)}\,\Gamma\left(\frac{n-1}{2}\right)}$, явным образом не зависит от a и σ, можно задать вероятность ее попадания в некоторый интервал (- t_γ , t_γ), учитывая четность плотности распределения, следующим образом: $p\left(\left|\frac{\bar{x}_B - a}{\frac{s}{\sqrt{n}}}\right| < t_g\right) = 2\int_0^{t_g} s(t,n)dt = g$. Отсюда получаем:

$$p\left(\bar{x}_B - \frac{t_g s}{\sqrt{n}} < a < \bar{x}_B + \frac{t_g s}{\sqrt{n}}\right) = g. \qquad (18.3)$$

Таким образом, получен доверительный интервал для a, где t_γ можно найти по соответствующей таблице при заданных n и γ.

Пример. Пусть объем выборки n = 25, $\bar{x}_B$ = 3, s = 1,5. Найдем доверительный интервал для a при γ = 0,99. Из таблицы находим, что t_γ (n = 25, γ = 0,99) = 2,797. Тогда $3 - \frac{2{,}797 \cdot 1{,}5}{\sqrt{25}} < a < 3 + \frac{2{,}797 \cdot 1{,}5}{\sqrt{25}}$, или 2,161< a < 3,839 – доверительный интервал, в который попадает a с вероятностью 0,99.

3. Доверительные интервалы для оценки среднего квадратического отклонения нормального распределения.

Будем искать для среднего квадратического отклонения нормально распределенной случайной величины доверительный интервал вида (s – δ, s +δ), где s – исправленное выборочное среднее квадратическое отклонение, а для δ выполняется условие: p ($|\sigma - s| < \delta$) = γ.

Запишем это неравенство в виде: $s\left(1-\frac{d}{s}\right) < s < s\left(1+\frac{d}{s}\right)$ или, обозначив $q = \frac{d}{s}$,

$$s(1-q) < s < s(1+q). \qquad (18.4)$$

Рассмотрим случайную величину χ, определяемую по формуле

$$c = \frac{s}{s}\sqrt{n-1},$$

которая распределена по закону «хи-квадрат» с n-1 степенями свободы (см. лекцию 12). Плотность ее распределения

$$R(c,n)=\frac{c^{n-2}e^{-\frac{c^2}{2}}}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)}$$

не зависит от оцениваемого параметра σ, а зависит только от объема выборки n. Преобразуем неравенство (18.4) так, чтобы оно приняло вид $\chi_1 < \chi < \chi_2$. Вероятность выполнения этого неравенства равна доверительной вероятности γ, следовательно, $\int_{c_1}^{c_2} R(c,n)dc = g$. Предположим, что $q < 1$, тогда неравенство (18.4) можно записать так:

$$\frac{1}{s(1+q)} < \frac{1}{s} < \frac{1}{s(1-q)},$$

или, после умножения на $s\sqrt{n-1}$, $\frac{\sqrt{n-1}}{1+q} < \frac{s\sqrt{n-1}}{s} < \frac{\sqrt{n-1}}{1-q}$. Следовательно, $\frac{\sqrt{n-1}}{1+q} < c < \frac{\sqrt{n-1}}{1-q}$. Тогда $\int_{\frac{\sqrt{n-1}}{1+q}}^{\frac{\sqrt{n-1}}{1-q}} R(c,n)dc = g$. Существуют таблицы для распределения «хи-квадрат», из которых можно найти q по заданным n и γ, не решая этого уравнения. Таким образом, вычислив по выборке значение s и определив по таблице значение q, можно найти доверительный интервал (18.4), в который значение σ попадает с заданной вероятностью γ.

Замечание. Если $q > 1$, то с учетом условия $\sigma > 0$ доверительный интервал для σ будет иметь границы

$$0 < s < s(1+q). \qquad (18.5)$$

Пример.
Пусть $n = 20$, $s = 1{,}3$. Найдем доверительный интервал для σ при заданной надежности $\gamma = 0{,}95$. Из соответствующей таблицы находим q ($n = 20$, $\gamma = 0{,}95$) = 0,37. Следовательно, границы доверительного интервала: 1,3(1-0,37) = 0,819 и 1,3(1+0,37) = 1,781. Итак, $0{,}819 < \sigma < 1{,}781$ с вероятностью 0,95.