27. Интервальные оценки параметров распределения

Определение 27.1. *Доверительным интервалом для несмещённого параметра a называют интервал $(a_1; a_2)$ со случайными границами, зависящими от наблюдений: $a_1 = a_1(x_1, \ldots, x_n)$, $a_2 = a_2(x_1, \ldots, x_n)$, накрывающий неизвестный параметр с заданной вероятностью γ: $P\big(a \in (a_1; a_2)\big) = \gamma$.*

Вероятность γ называется доверительной вероятностью или надежностью доверительного интервала.

Обычно γ задают равным 0,9; 0,95; 0,99 и более.

Доверительный интервал $\mathbf{I}_\gamma$ *для неизвестного математического ожидания нормального распределения при известной дисперсии имеет вид:*

$$\mathbf{I}_\gamma \left(\overline{x} - \tau_{\frac{\gamma}{2}} \frac{\sigma}{\sqrt{n}};\ \overline{x} + \tau_{\frac{\gamma}{2}} \frac{\sigma}{\sqrt{n}}\right), \tag{27.1}$$

где величина $\tau_{\frac{\gamma}{2}}$ определяется из уравнения:

$$\Phi(\tau_{\frac{\gamma}{2}}) = \frac{\gamma}{2} \tag{27.2}$$

Доверительный интервал $\mathbf{I}_\gamma$ *для неизвестного математического ожидания нормального распределения при неизвестной дисперсии имеет вид*:

$$\mathbf{I}_\gamma \left(\overline{x} - t_\gamma \frac{S^*}{\sqrt{n}};\ \overline{x} + t_\gamma \frac{S^*}{\sqrt{n}}\right), \tag{27.3}$$

где величина t_γ определяется по таблице приложения 3 критических точек распределения Стьюдента для $\alpha = 1 - \gamma$ и $k = n - 1$ или с помощью компьютера из уравнения для функции распределения Стьюдента $F_{st}(x)$ с $n-1$ степенью свободы:

$$F_{st}(t_\gamma) = \frac{1+\gamma}{2}, \tag{27.4}$$

где $\overline{x}$ и S^* — соответственно выборочное среднее и исправленное СКО.

Если независимые случайные величины $\xi_i \sim N(a; \sigma)$, $i = 1, \ldots, n$, то случайная величина

$$t = \frac{\bar{\xi} - a}{S^*/\sqrt{n}} \tag{27.5}$$

имеет распределение Стьюдента с $n-1$ степенью свободы.

Пример 27.1. *Найти доверительный интервал для оценки с надежностью 0,99 неизвестного математического ожидания a нормального распределения, если среднее квадратическое отклонение* $\sigma = 3$, *выборочное среднее* $\overline{x} = 32$ *и объём выборки* $n = 36$.

◀ В данном примере воспользуемся выражением (27.1). Поскольку здесь $\gamma = 0{,}99$, то параметр $\tau_{\frac{\gamma}{2}}$ найдем с помощью равенства:

$$\Phi(\tau_{\frac{\gamma}{2}}) = \frac{\gamma}{2} = \frac{0{,}99}{2} = 0{,}495.$$

Отсюда по таблицам нормированная функции Лапласа, приложение 2, определим $\tau_{\frac{\gamma}{2}} = 2{,}57$. Здесь левая граница интервала

$$\overline{x} - \tau_{\frac{\gamma}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 32 - 2{,}57 \cdot \frac{3}{6} = 32 - 1{,}285 = 30{,}715.$$

Правая граница определится как $32 + 1{,}285 = 33{,}285$. Таким образом, искомый доверительный интервал для математического ожидания a будет

$$30{,}715 < a < 33{,}285.$$

Полученный результат означает, что с вероятностью 0,99 математическое ожидание генеральной совокупности находится в интервале
$\mathbf{I}_\gamma = \Big(30{,}715;\ 33{,}285\Big)$. ▶

Ответ: $\boxed{\mathbf{I}_\gamma = \Big(30{,}715;\ 33{,}285\Big).}$

Пример 27.2. *Из генеральной совокупности извлечена выборка объёма* $n = 16$:

x_i	3,5	4,1	4,7	5,4	5,6	6,2
m_i	2	3	2	4	3	2

Оценить с надежностью 0,95 *математическое ожидание a нормально распределённой случайной величины по выборочному среднему с помощью доверительного интервала.*

◀ В данном случае дисперсия неизвестна и доверительный интервал определяется по формуле (27.3). Выборочное среднее вычислим по формуле

(26.2):

$$\overline{x} = \frac{1}{16}(2 \cdot 3,5 + 3 \cdot 4{,}1 + 2 \cdot 4{,}7 + 4 \cdot 5{,}4 + 3 \cdot 5{,}6 + 2 \cdot 6{,}2) \approx 4{,}9698.$$

Выборочную дисперсию удобнее искать с помощью выражения (26.7):

$S^2 = \frac{1}{16}\sum_{i=1}^{6} m_i x_i^2 - (\overline{x})^2 =$

$= \frac{1}{16}(2 \cdot 3,5^2 + 3 \cdot 4{,}1^2 + 2 \cdot 4{,}7^2 + 4 \cdot 5{,}4^2 + 3 \cdot 5{,}6^2 + 2 \cdot 6{,}2^2) - 4{,}9698^2 \approx 0{,}7309.$

Согласно (26.8), исправленная выборочная дисперсия

$$S^{*2} = \frac{16}{15} \cdot 0{,}7309 = 0{,}7796.$$

Отсюда находим исправленное СКО $S^* \approx 0{,}883$. При $\alpha = 1 - \gamma = 0{,}05$ и числе степеней свободы $n - 1 = 15$ из таблицы приложения 3 определим $t_\gamma = 2{,}13$.

Для вычисления значений критических точек распределения Стьюдента в Excel встроена функция СТЬЮДЕНТ.ОБР.2Х. Набираем команду =СТЬЮДЕНТ.ОБР.2Х(0,05;15) и получает значение 2,13145.

Находим радиус доверительного интервала

$$\varepsilon = t_\gamma \cdot \frac{S^*}{\sqrt{n}} = 2{,}13\frac{0{,}883}{\sqrt{16}} \approx 0{,}47.$$

Находим доверительный интервал

$$I_\gamma = (\overline{x} - \varepsilon;\ \overline{x} + \varepsilon) = \mathbf{(4{,}499;\ 5{,}439)}. \blacktriangleright$$

Ответ: $\boxed{I_\gamma = (4{,}499;\ 5{,}439).}$

27.1. Гистограмма относительных частот

Пример 27.3. *Из генеральной совокупности, являющейся нормальной случайной величиной ξ, извлечена выборка объёма $n = 50$, результаты которой сгруппированы с постоянным размахом интервала h=4 и помещены в таблицу.*

i	1	2	3	4	5	6
$(x_i; x_{i+1}]$	(10; 14]	(14; 18]	(18; 22]	(22; 26]	(26; 30]	(30; 34]
m_i	4	11	15	12	6	2

1) Построить гистограмму относительных частот по сгруппированным данным, где m_i — частота попадания вариант в промежуток $(x_i; x_{i+1}]$.

2) Найти эмпирическую функцию распределения $F^*(x)$.

3) На основании данного распределения, найти выборочную среднюю $\overline{x}$, несмещённую выборочную дисперсию S^{*2}, моду M_o и медиану M_e.

4) Найти доверительные интервалы для оценки, с надежностью $\gamma_1 = 0{,}9$, $\gamma_2 = 0{,}95$ и $\gamma_3 = 0{,}99$, неизвестного математического ожидания $M(\xi) = a$ генеральной совокупности в предположении, что она распределена нормально.

1) ◀Находим вектор относительных частот наблюдений, попавших в i-тый интервал $P_i^* = \dfrac{m_i}{n}$.
$P^* = [4/50, 11/50, 15/50, 12/50, 6/50, 2/50] = [0.08, 0.22, 0.3, 0.24, 0.12, 0.04]$.

Делим полученный вектор относительных частот на размах интервалов $h = 4$, получаем вектор плотности относительной частоты.
$\dfrac{P^*}{h} = [4/200, 11/200, 15/200, 12/200, 6/200, 2/200] =$
$= [0.02, 0.055, 0.075, 0.06, 0.03, 0.01]$.

Строим график, рис. 58, состоящий из шести прямоугольников ширина каждого из них равна 4, а высота $\dfrac{P_i^*}{h}$. Площадь полученной фигуры равна 1. Если увеличивать объём выборки и количество интервалов, то верхняя линия гистограммы приближается к функции плотности генеральной совокупности непрерывной случайной величины. ▶

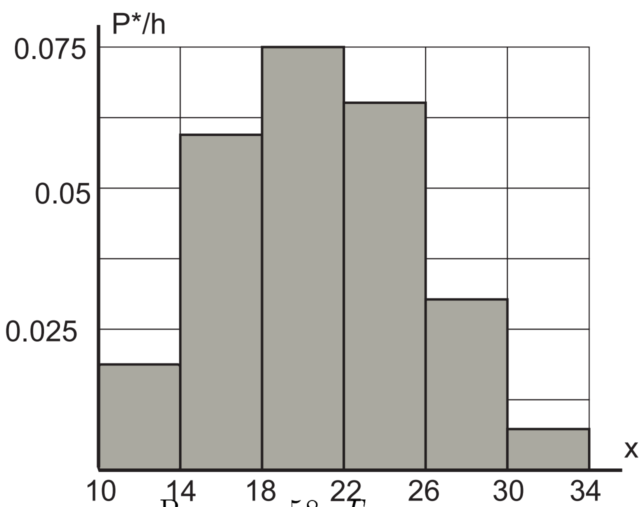

Рисунок 58. *Гистограмма*

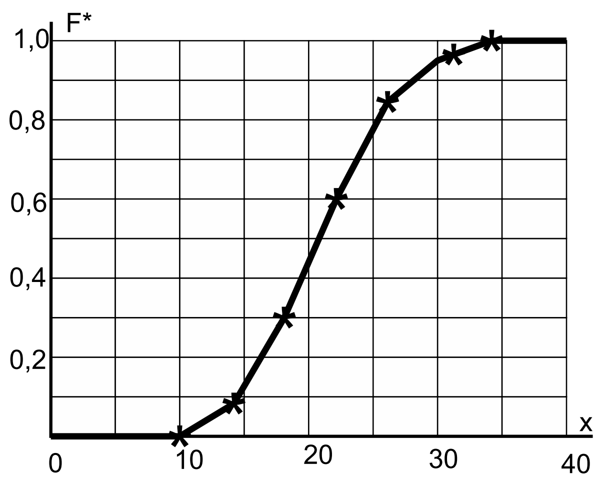

Рисунок 59. *Функция распределения $F^*(x)$*

2) ◀ Находим эмпирическую функцию распределения по формуле $F^*(x) = P^*(\xi < x) = \dfrac{n_x}{n}$, где n_x — число наблюдений меньших x. При задании выборки в виде группового распределения, где значения частот относятся

к центру интервалов (массив координат $D = [12, 16, 20, 24, 28, 32]$). Предполагаем, что на каждом отрезке исследуемая случайная величина распределена равномерно, следовательно функция распределения линейная. Поведение эмпирической функции распределения $F^*(x)$ на каждом i-том отрезке аппроксимируем непрерывной линейно возрастающей функцией. Тогда функцию распределения строим в виде кусочно-линейной кривой. Уравнение прямой, проходящей через две точки с координатами (x_i, y_i) и $(x_i + \Delta x, y_i + \Delta y_i)$, записывается в виде $y = y_i + \dfrac{\Delta y_i}{\Delta x}(x - x_i)$. Для нашего примера $x_i = 10 + 4(i-1)$, $i = 1, 2, 3, 4, 5, 6, 7$, а $y_1 = 0$, $y_i = y_{i-1} + P_i$, $i = 2, 3, 4, 5, 6, 7$, т.е, массивом $Y = [0, 0.08, 0.3, 0.6, 0.84, 0.96, 1]$.
$\Delta y_i = [0.08, 0.22, 0.3, 0.24, 0.12, 0.04]$.

$$F^*(x) = \begin{cases} 0 & \text{при} \quad x \leqslant 10, \\ 0{,}02(x-10) & \text{при} \quad x \in (10, 14], \\ 0{,}08 + 0,055(x-14) & \text{при} \quad x \in (14, 18], \\ 0{,}3 + 0,075(x-18) & \text{при} \quad x \in (18, 22], \\ 0{,}6 + 0,06(x-22) & \text{при} \quad x \in (22, 26], \\ 0{,}84 + 0,03(x-26) & \text{при} \quad x \in (26, 30], \\ 0{,}96 + 0,01(x-30) & \text{при} \quad x \in (30, 34], \\ 1 & \text{при} \quad x > 36. \end{cases}$$

Строим график эмпирической функции распределения $F^*(x)$, рис. 59. ►

3) ◄ Находим выборочную среднюю $\overline{x}$, несмещённую выборочную дисперсию S^{*2}, моду M_o и медиану M_e.

По формуле (26.2) находим выборочную среднюю $\overline{x} = \dfrac{1}{n}\sum\limits_{i=1}^{k} m_i d_i =$

$= \dfrac{1}{50}(4 \cdot 12 + 11 \cdot 16 + 15 \cdot 20 + 12 \cdot 24 + 6 \cdot 28 + 2 \cdot 32) = 20{,}88.$

По формуле (26.6) находим выборочную дисперсию

$\overline{x^2} = \dfrac{1}{n}\sum\limits_{i=1}^{k} m_i \cdot d_i^2 =$

$= \dfrac{1}{50}(4 \cdot 12^2 + 11 \cdot 16^2 + 15 \cdot 20^2 + 12 \cdot 24^2 + 6 \cdot 28^2 + 2 \cdot 32^2) = 461{,}12.$

$S^2 = \overline{x^2} - (\overline{x})^2 = 461{,}12 - 20{,}88^2 = 25{,}146.$

По формуле (26.8) находим несмещённую выборочную дисперсию:
$S^{*^2} = \dfrac{n}{n-1}S^2 = \dfrac{50}{49} \cdot 25{,}146 = 25{,}659.$

Несмещённое выборочное средне квадратическое отклонение $S^* = \sqrt{S^{*^2}} =$

$= 5{,}066$. Для выборки, представленной в виде сгруппированного распределения, значение моды аппроксимируются в некоторую точку модального интервала (внутри которого находится максимальное значение):

$$M_0 = X_0 - h\frac{f_{mo} - f_{mo-1}}{f_{m0-1} - 2f_{mo} + f_{mo+1}},$$

где X_O — нижнее значение модального интервала; f_{mo}, f_{mo-1}, f_{mo+1} — значение частот в модальном, предыдущем и следующем интервалах, соответственно; h — размах модального интервала.

Для решаемой задачи: $mo = 3$, $X_0 = 18$, $f_{mo} = 15, f_{mo-1} = 11$, $f_{mo+1} = 12$, $h = 4$.

$$M_0 = 18 - 4\frac{15 - 11}{11 - 30 + 12} \approx 20{,}286.$$

Для выборки, представленной в виде сгруппированного распределения, значение медианы аппроксимируются в некоторую точку M_e медианного интервала по формуле:

$$M_e = X_0 + h\frac{0{,}5\sum\limits_{k=1}^{s} f_k - \sum\limits_{k=1}^{me-1} f_k}{f_{me}},$$

где X_O — нижняя граница, в котором находится медиана (медианный интервал me); f_{me} — значение частоты в медианном интервале; h — размах медианного интервала.

$$M_e = 18 + 4\frac{25 - (4 + 11)}{15} \approx 20{,}667.$$

►

4) ◄Найти доверительные интервалы для оценки с надежностью $\gamma_1 = 0{,}9$, $\gamma_2 = 0{,}95$ и $\gamma_3 = 0{,}99$, неизвестного математического ожидания $M(\xi) = a$ генеральной совокупности в предположении, что она распределена нормально.

Доверительный интервал для неизвестного математического ожидания нормального распределения при неизвестной дисперсии вычисляем по формуле (27.3) $\left(\overline{x} - t_\gamma\frac{S^*}{\sqrt{n}};\ \overline{x} + t_\gamma\frac{S^*}{\sqrt{n}}\right)$, где величина t_γ определяется по таблице приложения 3 критических точек распределения Стьюдента для $\alpha = 1 - \gamma$ и

$k = n-1$ или с помощью компьютера из уравнения для функции распределения Стьюдента $F_{st}(x)$ с $k = n-1 = 49$ степенью свободы. $\alpha_1 = 0{,}1$, $\alpha_2 = 0{,}05$, $\alpha_3 = 0{,}01$.

$t_{\gamma_1} = 1{,}675; \quad t_{\gamma_2} = 2{,}01; \quad t_{\gamma_3} = 2{,}68.$

Находим радиусы доверительных интегралов $\varepsilon_i = t_{\gamma_i}\dfrac{S^*}{\sqrt{n}}, i = 1, 2, 3.$

$$\varepsilon_1 = 1{,}675\frac{5{,}065}{\sqrt{50}} = 1{,}22;\ \ \varepsilon_2 = 2{,}01\frac{5{,}065}{\sqrt{50}} = 1{,}44;\ \ \varepsilon_3 = 2{,}68\frac{5{,}065}{\sqrt{50}} = 1{,}92.$$

Отметим, что с увеличением надёжности радиус доверительного интервала увеличивается.

Находим доверительные интервалы.

$I_{\gamma_1} = (19{,}68; 22{,}08); \quad I_{\gamma_2} = (19{,}44; 22{,}32); \quad I_{\gamma_3} = (18{,}96; 22{,}8).$

Это означает, что выполняются вероятностные равенства

$P(19{,}68 < M(\xi) < 22{,}08) = 0{,}9;$

$P(19{,}44 < M(\xi) < 22{,}32) = 0{,}95;$

$P(18{,}96 < M(\xi) < 22{,}8) = 0{,}99.$ ►

Пример 27.4. *Случайная величина ξ генеральной совокупности распределена нормально, при этом известно средне квадратическое отклонение $\sigma = 2$.*

1) С надежностью $\gamma_1 = 0{,}9$, $\gamma_1 = 0{,}95$ и $\gamma_2 = 0{,}99$ найти доверительные интервалы для оценки математического ожидания генеральной совокупности, если объем выборки $n = 40$ и среднее выборочное $\overline{x} = -5$.

2) Как изменятся радиусы доверительных интервалов при увеличении объёма выборки в четыре раза при тех же значениях среднего квадратического отклонения и надёжности.

Доверительный интервал для неизвестного математического ожидания нормального распределения при известной дисперсии имеет вид (27.1): $\left(\overline{x} - \tau_{\gamma/2}\dfrac{\sigma}{\sqrt{n}};\ \overline{x} + \tau_{\gamma/2}\dfrac{\sigma}{\sqrt{n}}\right)$, где величина $\tau_{\gamma/2}$ определяется из уравнения (27.2): $\Phi(\tau_{\gamma/2}) = \frac{\gamma}{2}$.

1) ◄ Найдём значения величин $\tau_{\gamma_1/2}, \tau_{\gamma_2/2}, \tau_{\gamma_3/2}$.

$$\Phi(\tau_{\gamma_1/2}) = \frac{\gamma_1}{2} = \frac{0{,}9}{2} = 0{,}45.$$

Отсюда, по таблицам функции Лапласа определим $\tau_{\gamma_1/2} = 1{,}65$.

Аналогично,

$$\Phi(\tau_{\gamma_2/2}) = 0{,}475 \Rightarrow \tau_{\gamma_2/2} = 1{,}96.$$

$$\Phi(\tau_{\gamma_3/2}) = 0{,}495 \Rightarrow \tau_{\gamma_3/2} = 2{,}58.$$

Находим радиусы доверительных интегралов $\varepsilon_i, i = 1, 2, 3$,

$$\varepsilon_i = \tau_{\gamma_i/2}\frac{\sigma}{\sqrt{n}}.$$

$\varepsilon_1 = 0{,}522; \quad \varepsilon_2 = 0{,}620; \quad \varepsilon_3 = 0{,}816.$

Мы видим, что радиусы доверительных интегралов с увеличением надёжности расширяются.

Находим доверительные интервалы.

$I_{\gamma_1} = (-5{,}522; -4{,}478);$

$I_{\gamma_2} = (-5{,}62; -4{,}38);$

$I_{\gamma_3} = (-5{,}816; -4{,}184).$ ►

2) ◄ При увеличении объёма выборки в четыре раза, радиус доверительного интервала уменьшиться в два раза

$\varepsilon_1 = 0{,}261; \quad \varepsilon_2 = 0{,}310; \quad \varepsilon_3 = 0{,}409$

и доверительные интервалы будут равны

$I_{\gamma_1} = (-5{,}261; -4{,}439);$

$I_{\gamma_2} = (-5{,}31; -4{,}69);$

$I_{\gamma_3} = (-5{,}408; -4{,}592).$ ►

27.2. Выборочный коэффициент коррел

Рассмотрим выборку объёма n из генеральной совокупности значений двумерной случайной величины $(\xi; \eta)$, т.е. n пар наблюдений $(x_i; y_i)$. Поскольку многие значения в этой выборке могут повторяться, их заносят в так называемую корреляционную таблицу (табл. 27.2). В первом столбце этой таблицы перечислены значения x_i, во втором — y_i в виде вариационных рядов. На пересечении i-й строки

Таблица 27.2

Корреляционная таблица					
$\xi\backslash\eta$	y_1	y_2	$\dots$	y_s	n_{i*}
x_1	n_{11}	n_{12}	$\dots$	n_{1s}	n_{1*}
x_2	n_{21}	n_{22}	$\dots$	n_{2s}	n_{2*}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_k	n_{k1}	n_{k2}	$\dots$	n_{ks}	n_{k*t}
n_{*j}	n_{*t1}	n_{*2}	$\dots$	n_{*s}	n

и j-го столбца — соответствующая частота n_{ij}, т.е. количество раз, которое наблюдение $(x_i; y_j)$ встретилось в выборке. При обработке корреляционной таблицы в последнем столбце указывают сумму частот по строкам $n_{i*} = \sum\limits_{j=1}^{s} n_{ij}$, а в последней строке — сумму частот по столбцам $n_{*j} = \sum\limits_{i=1}^{k} n_{ij}$. Сумма всех элементов последнего столбца или строки даст объём выборки

$$n = \sum_{i=1}^{k}\sum_{j=1}^{s} n_{ij} = \sum_{i=1}^{k} n_{i\cdot} = \sum_{j=1}^{s} n_{\cdot j}.$$

Первый и последний столбцы корреляционной таблицы образуют статистическое распределение выборки случайной величины ξ, а первая и последняя строки образуют выборку случайной величины η.

$$\overline{x} = \frac{\sum\limits_{i=1}^{k} n_{i\cdot} x_i}{n}, \quad \overline{x^2} = \frac{\sum\limits_{i=1}^{k} n_{i\cdot} x_i^2}{n}, \quad S*_x^2 = \overline{x^2} - \overline{x}^2,$$

$$\overline{y} = \frac{\sum\limits_{j=1}^{s} n_{\cdot j} y_j}{n}, \quad \overline{y^2} = \frac{\sum\limits_{j=1}^{s} n_{\cdot j} y_j^2}{n}, \quad S*_y^2 = \overline{y^2} - \overline{y}^2.$$

$$\overline{xy} = \frac{\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{s} n_{ij} x_i y_j}{n}. \qquad (27.6)$$

Определение 27.2. ***Выборочным коэффициентом корреляции*** $r^*_{\xi\eta}$ *называется:*

$$r^*_{\xi\eta} = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{S_x \cdot S_y}. \qquad (27.7)$$

Выборочный коэффициент корреляции является статистической оценкой коэффициента корреляции и он обладает следующими свойствами:

1) $r^*_{\xi\eta} = r^*_{\eta\xi}$;

2) Выборочный коэффициент корреляции находится в пределах от -1 до 1: $-1 \leqslant r^*_{\xi\eta} \leqslant 1$;

3) $|r^*_{\xi\eta}| = 1$ тогда и только тогда, когда между значениями x_i и y_i имеется линейная зависимость. Чем ближе $r^*_{\xi\eta}$ к нулю, тем хуже эта зависимость аппроксимируется линейной.

Пример 27.5. *Дана корреляционная таблица случайного вектора* (ξ, η). *Найти выборочный коэффициент корреляции.*

$\xi\backslash\eta$	*1*	*2*	*4*
5	*22*	*2*	*3*
10	*0*	*20*	*6*
15	*1*	*12*	*12*
20	*4*	*8*	*10*

◀ Выпишем расширенную корреляционная таблицу

$\xi\backslash\eta$	1	2	4	n_{i*}
5	22	2	3	27
10	0	20	6	26
15	1	12	12	25
20	4	8	10	22
n_{*j}	27	42	31	100

Объём выборки $n = 100$. По формулам (27.6), определяем характеристики выборки:

$$\overline{x} = \frac{135 + 260 + 375 + 440}{100} = \frac{1210}{100} = 12{,}1; \qquad \overline{y} = \frac{27 + 84 + 124}{100} = \frac{235}{100} = 2{,}35;$$

$$\overline{x^2} = \frac{17700}{100} = 177; \qquad \overline{y^2} = \frac{691}{100} = 6{,}91;$$

$$S_x^2 = 177 - 12{,}^2 = 30{,}59; \qquad S_y^2 = 6{,}91 - 2{,}35^2 = 1{,}3875;$$

$$\overline{xy} = \frac{3125}{100} = 31{,}25;$$

$r^*_{\xi\eta} = \frac{\overline{xy} - \overline{x}\ \overline{y}}{S_x\ S_y} \approx$ **0,432**.▶

Ответ: $\boxed{r^*_{\xi\eta} \approx 0{,}432.}$

Пример 27.6. *Среднемесячная заработная плата (тыс. руб.) в Москвы в 2021-2022 годах составила по отраслям:*

Отрасль	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
2021 год	*112*	*71*	*41*	*136*	*91*	*41*	*78*	*71*	*107*	*103*
2022 год	*126*	*75*	*59*	*138*	*107*	*43*	*95*	*80*	*117*	*110*

Найти коэффициент корреляции

Взято из реальных источников. Ссылка активная на март 2023г.

https://www.audit-it.ru/inform/zarplata/index.php?id_region=184

1. Производство напитков.
2. Производство пищевых продуктов.
3. Производство одежды.
4. Производство лекарственных средств и материалов.
5. Производство мебели.
6. Ремонт и монтаж машин и оборудования.
7. Производство компьютеров.
8. Строительство зданий.
9. Научные исследования и разработки.
10. Образование.

◀

$$\overline{x} = \frac{112 + 71 + \ldots + 100}{10} = 85{,}1; \qquad \overline{y} = \frac{126 + 75 + \ldots + 110}{10} = 95;$$

$$\overline{x^2} = \frac{112^2 + 71^2 + \ldots + 100^2}{10} = 8090{,}7; \qquad \overline{y^2} = \frac{126^2 + 75^2 + \ldots + 110^2}{10} = 9853{,}8;$$

$$S_x^2 = 8090{,}7 - 85{,}1^2 = 848{,}69; \qquad S_y^2 = 9853{,}8 - 95^2 = 828{,}8;$$

$$\overline{xy} = \frac{112 \cdot 126 + \ldots + 100 \cdot 110}{10}\frac{89063}{10} = 8906{,}3;$$

$r^*_{\xi\eta} = \frac{\overline{xy} - \overline{x}\ \overline{y}}{S_x\ S_y} \approx$ **0,98**.

Что говорит о тесной взаимосвязи с переменными. ▶

Ответ: $\boxed{r^*_{\xi\eta}0{,}98.}$

Задания для самостоятельной работы

27.1. Известно, что случайная величина ξ генеральной совокупности распределена нормально, и при объёме выборки $n = 20$, при этом известно среднее квадратическое отклонение $\sigma = 5$, и радиус доверительного интервала равен 1,2. Каким будет радиус доверительного интервала при увеличении объёма выборки в девять раз, при тех же значениях среднего квадратического отклонения и надежности?

27.2. Известно, что случайная величина ξ генеральной совокупности распределена нормально, при этом известно среднее квадратическое отклонение $\sigma = 3$. С надежностью $\gamma = 0{,}975$ найти доверительный интервал для оценки математического ожидания генеральной совокупности, если объём выборки $n = 100$ и выборочное среднее 15.

27.3. В таблице 27.3 даны 30 вариантов заданий выборок объёма $n = 10$. По данным в таблице результатам измерений найти доверительные интервалы для оценки неизвестного математического ожидания a нормального распределения с заданной надежностью γ.

Указание. Выборочное среднее и исправленное выборочное СКО находить соответственно по формулам (26.1) и (26.8), доверительный интервал по формуле (27.3).

Таблица 27.3

	1	2	3	4	5	6	7	8	9	10	γ
1	4,50	4,51	4,52	4,53	4,54	4,49	4,54	4,47	4,49	4,46	0,95
2	4,43	4,41	4,39	4,45	4,40	4,35	4,42	4,40	4,37	4,38	0,99
3	30,10	30,40	30,30	30,00	29,45	29,65	30,05	30,15	29,90	30,00	0,999
4	80,20	80,10	80,30	79,70	79,80	79,80	80,10	80,00	79,70	80,30	0,95
5	5,40	5,41	5,40	5,42	5,39	5,38	5,38	5,37	5,35	5,40	0,99
6	14,28	14,26	14,27	14,30	14,31	14,32	14,31	14,29	14,30	14,26	0,999
7	20,12	20,11	20,10	20,10	19,98	19,97	20,02	20,03	20,02	20,10	0,999
8	36,41	36,42	36,44	36,45	36,48	36,49	36,46	36,45	36,42	36,38	0,99
9	14,46	14,45	14,41	14,40	14,42	14,48	14,50	14,42	14,45	14,44	0,95
10	80,30	80,30	80,20	80,30	80,20	79,80	79,80	79,70	79,90	80,50	0,95
11	15,38	15,40	15,41	15,42	15,43	15,37	15,36	15,36	15,34	15,33	0,99
12	16,06	16,20	16,16	16,00	16,15	16,05	16,01	16,03	16,02	16,02	0,999
13	5,90	5,88	5,97	5,95	5,93	5,85	5,88	5,87	5,87	5,90	0,95
14	7,71	7,73	7,70	7,72	7,68	7,67	7,65	7,63	7,70	7,71	0,99
15	15,21	15,28	15,25	15,24	15,17	15,18	15,20	15,19	15,26	15,22	0,999
16	2,58	2,50	2,57	2,49	2,49	2,51	2,55	2,53	2,53	2,55	0,95
17	4,38	4,40	4,37	4,42	4,35	4,40	4,39	4,45	4,43	4,41	0,99
18	14,41	14,45	14,46	14,40	14,48	14,42	14,50	14,45	14,42	14,44	0,999
19	10,25	10,23	10,24	10,18	10,17	10,18	10,19	10,20	10,19	10,17	0,999
20	5,97	5,88	5,90	5,93	5,95	5,88	5,85	5,87	5,90	5,87	0,99
21	16,02	16,03	16,02	16,05	16,01	16,15	16,00	16,16	16,20	16,06	0,95
22	42,80	42,72	42,75	42,90	42,98	42,85	42,07	42,93	42,77	42,83	0,95
23	3,44	3,47	3,38	3,39	3,46	3,49	3,39	3,47	3,46	3,45	0,99
24	36,40	36,70	36,90	36,80	36,30	36,70	36,90	36,90	36,40	36,00	0,999
25	8,35	8,40	8,38	8,44	8,45	8,44	8,37	8,39	8,36	8,42	0,95
26	15,28	15,25	15,21	15,24	15,18	15,17	15,20	15,19	15,26	15,22	0,99
27	27,90	27,30	26,90	27,30	27,40	27,50	27,00	27,60	27,70	27,40	0,999
28	7,71	7,72	7,70	7,73	7,67	7,68	7,65	7,63	7,71	7,70	0,95
29	5,41	5,40	5,42	5,40	5,39	5,38	5,37	5,38	5,35	5,40	0,99
30	28,70	28,30	28,80	28,80	28,00	28,10	27,90	28,70	28,10	28,20	0,999

27.4. Лаборатория электролампового завода провела испытания 100 ламп на продолжительность горения и получила следующие результаты (в часах): 812, 817, 828, 833, 841, 820, 822, 825, 826, 824, 826, 825, 817, 826, 834, 818, 842, 826, 837, 827, 821, 835, 823, 824, 815, 833, 830, 824, 816, 828, 822, 826, 827, 822, 837, 816, 825, 810, 823, 831, 826, 814, 838, 831, 824, 812, 827, 839, 828, 836, 815, 836, 817, 828, 823, 832, 819, 826, 818, 820, 811, 828, 810, 822, 836, 816, 829, 821, 833, 821, 829, 823, 832, 823, 831, 826, 832, 827, 829, 826, 836, 821, 838, 818 , 822, 819, 823, 828, 826, 820, 825, 828, 822, 835, 824, 825, 820, 829, 825, 824];

В предположении, что случайная величина ξ — продолжительность горения лампы является нормальной случайной величиной, провести исследования по схеме примера 27.3. Диапазон наблюдений разбить на 6 интервалов.

27.5. Дана корреляционная таблица случайного вектора (ξ, η). Найти выборочный коэффициент корреляции.

$\xi \backslash \eta$	2	4	8
-1	18	2	3
0	6	20	6
2	1	12	15
4	4	8	5

Ответ: $r^*_{\xi\eta} \approx 0{,}317$.

27.6. Найти коэффициент корреляции между производительностью труда (выпуск продукции в тыс. руб. на одного рабочего) ξ и затратами на покупку дополнительного оборудования (тыс. руб.) η для 10 предприятий города по следующим данным:

ξ	4,6	4	4,8	5,3	5	5,8	7	7,2	7,8	10,8
η	8,5	8,7	9,1	9,3	10,2	11	12,5	12,2	14	13,2

Ответ: $r^*_{\xi\eta} \approx 0{,}863$.