

*ТОЧЕЧНАЯ И
ИНТЕРВАЛЬНАЯ
ОЦЕНКА СЛУЧАЙНОЙ
ВЕЛИЧИНЫ*

Одной из задач математической статистики, является определение параметров большого массива по исследованию его части.

Два вида оценки:

**Точечная оценка случайной величины;*

**Интервальная оценка случайной величины.*

ОЦЕНКИ ПАРАМЕТРОВ В ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ

Рассмотрим изучение некоторого количественного признака X . Его распределение в генеральной совокупности характеризуется параметрами, называемыми числовыми характеристиками генеральной совокупности. К ним относятся генеральная средняя, генеральная дисперсия и генеральное среднее квадратическое отклонение.

Определение. Генеральной средней \bar{x}_T называется среднее арифметическое всех значений изучаемого признака в генеральной совокупности:

$$\bar{x}_T = \frac{1}{N} \sum_{i=1}^N x_i,$$

где N - объем генеральной совокупности; x_i - значения признака для различных объектов генеральной совокупности с присвоенными им номерами $i = 1, 2, N$.

Можно показать, что генеральная средняя равна математическому ожиданию, но для определенности дальше мы будем пользоваться генеральной средней.

Как отмечалось ранее, рассчитать генеральную среднюю практически сложно или невозможно из-за большого объема генеральной совокупности, сложности измерений и проч. Поэтому для изучения генеральной совокупности из нее извлекают выборку относительно небольшого объема.

***Определение.** Выборочной средней \bar{x}_B называется среднее арифметическое всех значений изучаемого признака в выборке:*

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

где n - объем выборки; x_i - значения признака. Если некоторые значения x_i встречаются несколько раз (т. е. соответствующие частоты $m_i > 1$), то в сумму их следует подставлять столько же раз.

В математической статистике показывается, что в случае репрезентативной выборки математическое ожидание выборочной средней (для разных выборок) равно генеральной средней:

$$M(\bar{x}_B) = \bar{x}_T.$$

Для оценки генеральной средней, считается, что неизвестная генеральная средняя приблизительно равна рассчитанной выборочной средней:

$$\bar{x}_T \approx \bar{x}_B.$$

Это **точечная оценка** генеральной средней. Точечной она называется потому, что характеризуется одним числом - точкой на числовой оси.

Определение. Генеральной дисперсией σ_{Γ}^2 называется среднее арифметическое квадратов отклонений всех значений изучаемого признака X в генеральной совокупности от генеральной средней:

$$\sigma_{\Gamma}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}_{\Gamma})^2,$$

где x_i - значения признака; \bar{x}_{Γ} - генеральная средняя.

Генеральная дисперсия характеризует среднее отклонение значений признака в генеральной совокупности относительно генеральной средней.

Рассчитать генеральную дисперсию практически очень сложно. Поэтому для ее оценки вычисляют соответствующие характеристики в выборке.

Определение. Выборочной дисперсией σ_B^2 называется среднее арифметическое квадратов отклонений всех значений изучаемого признака в выборке относительно выборочной средней:

$$\sigma_B^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_B)^2,$$

где x_i - значение признака; \bar{x}_B - выборочная средняя. Если некоторые значения x_i встречаются несколько раз (т. е. соответствующие частоты $m_i > 1$), то в сумму их следует подставлять столько же раз.

Выборочная дисперсия характеризует среднее отклонение значений признака в выборке относительно выборочной средней.

Если данные представлены в виде интервального ряда распределения, то выборочную среднюю и выборочную дисперсию можно приближенно вычислить по следующим формулам:

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^k x_i^* m_i;$$

$$\sigma_B^2 = \frac{1}{n} \sum_{i=1}^k (x_i^* - \bar{x}_B)^2 m_i,$$

Определение. Генеральным средним квадратическим отклонением σ_Γ , называется квадратный корень из генеральной дисперсии:

$$\sigma_\Gamma = \sqrt{\sigma_\Gamma^2}.$$

Определение. Выборочным средним квадратическим отклонением σ_B называется квадратный корень из выборочной дисперсии:

$$\sigma_B = \sqrt{\sigma_B^2}.$$

ИНТЕРВАЛЬНАЯ ОЦЕНКА. ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ. ДОВЕРИТЕЛЬНАЯ ВЕРОЯТНОСТЬ.

При достаточно большом объеме выборки можно сделать вполне надежные заключения о параметрах генеральной совокупности. Однако на практике часто имеют дело с выборками небольшого объема ($n < 30$). При небольшом объеме выборки пользуются *интервальными оценками*. В этом случае указывается интервал (*доверительный интервал*).

Доверительный интервал – интервал, в котором с заданной *доверительной вероятностью* находится истинное значение случайной величины (среднее значение генеральной совокупности).

Доверительная вероятность – вероятность, с которой в заданном интервале (доверительном интервале) находится истинное значение случайной величины (среднее значение генеральной совокупности). Обычно в медико-биологических исследованиях доверительную вероятность принимают равной 0,95.

ТАБЛИЦА КОЭФФИЦИЕНТОВ СТЬЮДЕНТА

$f = n - 1$	Значение коэффициента $t_{P,f}$			
	$P = 0.9$	$P = 0.95$	$P = 0.99$	$P = 0.999$
1	6.31	12.71	63.66	636.6
2	2.92	4.30	9.93	31.60
3	2.35	3.18	5.84	12.90
4	2.13	2.78	4.60	8.60
5	2.02	2.57	4.03	6.90
6	1.94	2.45	3.71	5.96
7	1.90	2.37	3.50	5.40
8	1.86	2.31	3.36	5.04
9	1.83	2.26	3.25	4.78
10	1.81	2.23	3.17	4.60
20	1.73	2.09	2.85	3.85
120	1.66	1.98	2.62	3.37
∞	1.65	1.96	2.58	3.29

Доверительный интервал математически записывают так:

$$\bar{x}_e - mt_{\alpha;n} \leq x_z \leq \bar{x}_e + mt_{\alpha;n}, \text{ или}$$

$$\bar{x}_e \pm mt_{\alpha;n}, \text{ где}$$

$t_{\alpha;n}$ - коэффициент Стьюдента (величина табличная, размерности не имеет),

α - доверительная вероятность,

n - объем выборки;

m - ошибка среднего.

Чтобы определить доверительный интервал, необходимо:

1. Вычислить среднее значение выборки \bar{x}_e ;
2. Вычислить дисперсию для выборки σ_e^2 ;
3. Вычислить исправленную выборочную дисперсию:

$$S^2 = \frac{n}{n-1} \sigma_e^2$$

4. Вычислить ошибку среднего $m = \frac{S}{\sqrt{n}}$.

ЗАДАЧИ

Задача № 1 . При оценке жизненной емкости легких спринтеров был получен следующий вариационный ряд: $(X_i, л) - 5, 6, 7, 8, 9$; $m_i - 1, 2, 4, 3, 2$. Найдите дисперсию данной случайной величины, если ее выборочное среднее равно 7,25 л.

Решение:

1. Составим статистическое распределение:

X_i	5	6	7	8	9
m_i	1	2	4	3	2

Объем выборки составляет: $n=12$

Выборочное среднее равно: $X_{cp} = 7,25$

Выборочная дисперсия рассчитывается по формуле:

$$\overline{D_B} = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x}_B)^2 \cdot m_i$$

$$\overline{D_B} = \frac{1}{12} [(5-7,25)^2 + (6-7,25)^2 \cdot 2 + (7-7,25)^2 \cdot 4 + (8-7,25)^2 \cdot 3 + (9-7,25)^2 \cdot 2] =$$

$$= \frac{1}{12} [5,0625 + 1,5625 \cdot 2 + 0,0625 \cdot 4 + 0,5625 \cdot 3 + 3,0625 \cdot 2] = \frac{1}{12} [5,0625 + 3,125 + 0,25 + 1,6875 + 6,125] = \frac{16,25}{12} = 1,35.$$

Задача № 2. При исследовании частоты дыхания гимнастов было установлено, что среднее квадратическое отклонение равно 1,1. Найдите выборочную дисперсию данного показателя.

Решение:

$$\sigma_B = \sqrt{D_B}, \text{ следовательно } (\sigma_B)^2 = D_B.$$

$$\sigma_B = 1,1;$$

$$D_B = (1,1)^2 = 1,21$$

Ответ: $D_B = 1,21.$

Задача № 3. При исследовании двигательных функций каратистов было установлено, что выборочная дисперсия показателя теппинг-теста равна 1,69. Найдите среднее квадратическое отклонение данного показателя.

Решение:

$$\sigma_B = \sqrt{D_B}$$

$$D_B = 1,69;$$

$$\sigma_B = \sqrt{1,69} = 1,3 \text{ уд/с}$$

Ответ: $\sigma_B = 1,3 \text{ уд/с}$

Задача № 4. При исследовании скорости распространения механической волны на поражённых участках кожи у больных псориазом в регрессирующей стадии были получены следующие результаты, м/с: 38, 39, 41, 41, 38, 43, 40, 40, 42, 38, 38, 39, 38, 41, 42, 41, 42, 41, 39, 43, 42, 43, 40, 39, 40, 38, 43, 42, 39, 42.

Дать точечную оценку распределения (рассчитать математическое ожидание, дисперсию, среднее квадратическое отклонение);

Дать интервальную оценку данного показателя на уровне доверительной вероятности $\alpha = 0,95$.

Решение:

1. Построение вариационного ряда.

Объем выборки $n=30$ элементов.

X_i	38	39	40	41	42	43
m_i	6	5	4	5	6	4

Точечная оценка распределения

Выборочная средняя - $\bar{X}_B = \frac{1}{n} \sum_{i=1}^k x_i \cdot m_i$

$$\bar{X}_B = \frac{1}{30}(38 \cdot 6 + 39 \cdot 5 + 40 \cdot 4 + 41 \cdot 5 + 42 \cdot 6 + 43 \cdot 4) = 40,4 \text{ м/с.}$$

Выборочная дисперсия- $\bar{D}_B = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x}_B)^2 \cdot m_i$

$$\bar{D}_B = \frac{1}{30}[(38-40,4)^2 \cdot 6 + (39-40,4)^2 \cdot 5 + (40-40,4)^2 \cdot 4 + (41-40,4)^2 \cdot 5 + (42-40,4)^2 \cdot 6 + (43-40,4)^2 \cdot 4] = 2,97$$

Выборочное среднее квадратическое отклонение

$$\sigma_B = \sqrt{\bar{D}_B}; \quad \sigma_B = \sqrt{2,97} = 1,72$$

Интервальная оценка: $P=0,95$, $t_{\alpha,n}=2,04$.

Доверительный интервал:

$$\bar{x}_B - mt_{\alpha,n} \leq x_T \leq \bar{x}_B + mt_{\alpha,n}$$

m - ошибка среднего:

$$m = \frac{\sigma_B}{\sqrt{n}}; \quad m = \frac{1,72}{\sqrt{30}} = 0,31$$

Доверительный интервал:

$$40,4 - 0,6324 \leq x_T \leq 40,4 + 0,6324$$

$$39,77 \leq x_T \leq 41,03$$

Ответ: $39,77 \leq x_T \leq 41,03$

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. Измерения вязкости крови у 10 пациентов с полицитемией дали следующие результаты: (X_i , мПа*с) – 14,9; 15,1; 16,2; 17,3; 18,2; 18,5, m_i – 1, 2, 3, 2, 1, 1. Найдите выборочную дисперсию данного показателя, если его выборочное среднее равно 16,5 мПа*с. (1,49).
2. При исследовании частоты дыхания гимнастов было установлено, что среднее квадратическое отклонение равно 1,5. Найдите выборочную дисперсию данного показателя.
3. При оценке жизненной емкости легких ($X, л$) спринтеров было установлено, что выборочная дисперсия данного показателя составляет 1,35. Найдите выборочное среднее квадратическое отклонение.
4. При исследовании скорости распространения механической волны в коже щеки после процедуры криомассажа у пациенток с сухим типом кожи были получены следующие результаты, м/с: 38, 58, 46, 39, 49, 62, 62, 49, 43, 44, 68, 41, 54, 64, 64, 38, 58, 46, 39, 49, 62, 62, 49, 43, 44, 68, 41, 54, 64, 64. Дать точечную оценку распределения (рассчитать математическое ожидание, дисперсию, среднее квадратическое отклонение); дать интервальную оценку данного показателя на уровне доверительной вероятности $\alpha = 0,95$.