Занятие 3

ТОЧЕЧНЫЕ ОЦЕНКИ НЕИЗВЕСТНЫХ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ ИССЛЕДУЕМОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Краткие теоретические сведения

Пусть известен закон распределения для СВ X (нормальный, Пуассона, показательный и т. д.), зависящий от одного или нескольких параметров. Требуется по выборке, полученной в результате n экспериментов, оценить (найти приближенно) неизвестный параметр распределения. Оценка, определяемая одним числом, называется ***точечной***.

Оценкой неизвестного параметра распределения называют функцию от наблюдаемых значений случайной величины.

Оцениваемый параметр обозначим через θ, а его оценку – $\overline{\theta}$: $\overline{\theta} = f(x_1, x_2, ..., x_n)$.

$\overline{\theta}$ является случайной величиной.

Точечная оценка должна удовлетворять определенным требованиям:

– быть несмещенной, т. е. $M(\overline{\theta}) = \theta$;

– быть эффективной, т. е. если неизвестный параметр имеет несколько оценок, вычисленных по выборкам одного и того же объема n, то нужно выбрать ту, которая имеет наименьшую дисперсию;

– быть состоятельной, т. е. при $n \to \infty$ стремится по вероятности к оцениваемому параметру: $\lim\limits_{n\to\infty} P(\left|\overline{\theta} - \theta\right| < \varepsilon) = 1 \;\; \forall \varepsilon > 0$.

Выборочная средняя $\overline{x}_{\text{в}}$ является несмещенной, состоятельной и эффективной для математического ожидания генеральной совокупности.

Несмещенной и состоятельной оценкой для дисперсии генеральной совокупности является исправленная выборочная дисперсия:

$$D_{\text{и}} = S^2 = \frac{n}{n-1} D_{\text{в}}; \qquad D_{\text{и}} = \frac{\sum\limits_{i=1}^{k} \left(x_i - \overline{x}_{\text{в}} \right)^2 \cdot n_i}{n-1}.$$

Исправленным средним квадратическим отклонением является

$$S = \sqrt{D_{\text{и}}}; \qquad S = \sqrt{S^2} = \sqrt{\frac{n}{n-1} D_{\text{в}}}.$$

Для вычисления $\overline{x}_{\text{в}}$, $D_{\text{в}}$ используют наиболее распространенный метод – метод произведений.

Целесообразно составить таблицу:

1) в первый столбец записывают x_i, располагая в порядке возрастания;

2) во второй столбец записывают число экспериментов, в которых получено соответствующее значение x_i и сумму $\sum\limits_{i=1}^{k} n_i$;

3) в третий столбец записывают условные варианты $U_i = \dfrac{x_i - b}{h}$, где h – шаг разбиения, b – «ложный нуль».

В качестве b берется варианта, стоящая посередине вариационного ряда, или варианта, имеющая максимальную частоту. В клетках над нулем пишут последовательно –1, –2, –3, …, а под нулем пишут 1, 2, 3, …;

4) находят $n_i \cdot U_i$ и сумму $\sum\limits_{i=1}^{k} n_i \cdot U_i$ и записывают в четвертый столбец;

5) в пятый столбец записывают $n_i \cdot U_i^2$ и сумму $\sum_{i=1}^{k} n_i \cdot U_i^2$;

6) в шестой столбец записывают $n_i(U_i+1)^2$ и сумму $\sum_{i=1}^{k} n_i(U_i+1)^2$.

Если сумма $\sum_{i=1}^{k} n_i(U_i+1)^2$ будет равна сумме $\sum_{i=1}^{k} n_i + 2\sum_{i=1}^{k} n_i \cdot U_i + \sum_{i=1}^{k} n_i \cdot U_i^2$, то вычисления произведены правильно.

Определяем условные начальные моменты 1-го и 2-го порядков:

$$\overline{\nu}_1 = \frac{1}{n}\sum_{i=1}^{k} U_i \cdot n_i; \qquad \overline{\nu}_2 = \frac{1}{n}\sum_{i=1}^{k} U_i^2 \cdot n_i.$$

Находим выборочное среднее $\overline{x}_{\text{в}}$ и выборочную дисперсию $D_{\text{в}}$:

$$\overline{x}_{\text{в}} = \overline{\nu}_1 \cdot h + b; \qquad D_{\text{в}} = \left(\overline{\nu}_2 - \left(\overline{\nu}_1\right)^2\right) \cdot h^2.$$

Пример 3.1. Методом произведений вычислить выборочную среднюю и выборочную дисперсию по данным выборки (табл. 3.1).

Таблица 3.1

x_i	12	14	16	18	20	22
n_i	5	15	50	16	10	4

Решение. В качестве «ложного нуля» возьмем варианту 16. Следовательно, $b = 16$, $U_i = \dfrac{x_i - 16}{2}$.

Результаты вычислений сведем в табл. 3.2.

Таблица 3.2

x_i	n_i	U_i	$n_i \cdot U_i$	$n_i \cdot U_i^2$	$n_i(U_i+1)^2$
12	5	–2	–10	20	5
14	15	–1	–15	15	0
$\boxed{16}$	50	0	0	0	50
18	16	1	16	16	64
20	10	2	20	40	90
22	4	3	12	36	64
$\sum$	100	–	23	127	273

Контроль: 273 = 100 + 46 + 127.

Равенство выполнено, следовательно, таблица заполнена верно.

Вычислим условные начальные моменты:

$$\overline{\nu}_1 = \frac{1}{100} \cdot 23 = 0,23; \qquad \overline{\nu}_2 = \frac{127}{100} = 1,27.$$

Вычислим выборочную среднюю и выборочную дисперсию: $\overline{x}_{\text{в}} = 0,23 \cdot 2 + 16 = 16,46$; $D_{\text{в}} = (1,27 - (0,23)^2) \cdot 4 = (1,27 - 0,0529) \cdot 4 =$ $= 1,2171 \cdot 4 = 4,8684 \approx 4,87$.

Определим исправленную выборочную дисперсию: $D_{\text{и}} = S^2 = \dfrac{100}{99} \cdot 4,87 = 4,92$, и исправленное среднее квадратическое отклонение: $S = \sqrt{4,92} = 2,22$.

Получим несмещенные оценки для математического ожидания, дисперсии и среднего квадратического отклонения.

Задачи для аудиторной работы

Задача 3.1.

По данным выборки (табл. 3.3) найти несмещенные оценки для математического ожидания и дисперсии.

Таблица 3.3

x_i	48	52	56	60	64	68	72	76	80	84
n_i	2	4	6	8	12	30	18	8	7	5

Задача 3.2.

Найти несмещенные оценки для математического ожидания и дисперсии по данным выборки (табл. 3.4).

Таблица 3.4

x_i	7,9	8,1	8,3	8,5	8,7	8,9
n_i	5	20	80	95	40	10

Задача 3.3.

По данным выборки (табл. 3.5) найти:

а) несмещенные оценки для математического ожидания и дисперсии;

б) коэффициент асимметрии и эксцесс.

Таблица 3.5

$x_i - x_{i+1}$	[0–5)	[5–10)	[10–15)	[15–20)	[20–25)	[25–30)
n_i	49	41	26	14	17	3