Точечные статистические оценки параметров распределения случайной величины

Статистической оценкой Θ^* неизвестного параметра теоретического распределения Θ называется функция от наблюдаемых значений случайной величины:

$$\Theta^* = f(x_1, x_2, \ldots, x_n)$$

Статистические оценки, определяемые одним числом, называются *точечными*.

Для того, чтобы оценки были надежными, к ним предъявляются требования *несмещенности, состоятельности и эффективности*.

Точечные статистические оценки

Оценка называется *несмещенной*, если её математическое ожидание равно оцениваемому параметру Θ при любом объеме выборки *n*, т.е.:

$$M\left(\Theta^*\right) = \Theta$$

Оценка Θ^* называется *эффективной*, если при заданном объеме выборки *n* она имеет наименьшую возможную дисперсию, т.е.: $D\left(\Theta^*\right) \to \min$

Оценка Θ^* называется *состоятельной* для параметра Θ, если она сходится по вероятности к оцениваемому параметру при неограниченном возрастании объема выборки, т.е.:

$$\lim_{N\to\infty} P\left(\left|\Theta - \Theta^*\right| \le \varepsilon\right) = 1$$

где $\varepsilon > 0$- скольугодно малое положительное число.

Точечные статистические оценки

Для того, чтобы несмещенная оценка была состоятельной, достаточно, чтобы было выполнено условие:

$$\lim_{N\to\infty} D\left(\Theta^*\right) = 0 \text{ .}$$

К наиболее часто используемым статистическим оценкам параметров распределения случайной величины можно отнести статистические оценки

- математического ожидания,
- дисперсии,
- асимметрии,
- эксцесса,
- моды и медианы.

Начальные и центральные моменты

Начальным эмпирическим моментом к-го порядка называется среднее значение к-х степеней элементов вариационного или интервального ряда:

$$\bar{M}_k = \frac{1}{N}\sum n_j x_j^k$$

В частности

$$\bar{x}_в = \bar{M}_1 = \frac{1}{N}\sum n_j x_j$$

Центральным эмпирическим моментом к-го порядка называется среднее значение к-х степеней разностей $x_j - \bar{x}_в$ для вариационного или интервального ряда:

$$\bar{m}_k = \frac{1}{N}\sum n_j \left(x_j - \bar{x}_в\right)^k$$

Начальные и центральные моменты

В частности:

$$D_в = \bar{m}_2 = \frac{1}{N}\sum n_j \left(x_j - \bar{x}_в\right)^2$$

Следует отметить что указанная статистическая оценка дисперсии является смещенной оценкой. Поэтому следует использовать так называемую исправленную оценку дисперсии:

$$s^2 = \frac{N}{N-1} D_в$$

Статистические оценки СКО вычисляются как корень квадратный из соответствующих оценок дисперсии:

$$\sigma_в = \sqrt{D_в} \quad ; \quad s = \sqrt{s^2}$$

Начальные и центральные моменты

Статистические оценки асимметрии и эксцесса вычисляются по формулам:

$$\bar{A}_s = \frac{\bar{m}_3}{s^3} \quad ; \quad \bar{E} = \frac{\bar{m}_4}{s^4} - 3$$

Приведенные формулы для вычисления статистических оценок параметров распределения случайной величины справедливы в рамках так называемого метода моментов, предложенного К. Пирсоном.

В соответствии с этим методом теоретические моменты распределения СВ приравниваются соответствующим эмпирическим моментам того же порядка. Основанием для этого является то, что эмпирические моменты являются несмещенными оценками соответствующих теоретических моментов.

Условные эмпирические моменты

Для упрощения вычислений эмпирических моментов вводят в рассмотрение так называемые условные варианты

$$u_j = \frac{x_j - C}{h}$$

где C – условный ноль, значение которого выбирается равным значению варианты интервального ряда, являющуюся средней или близкой к средней по значению в этом ряду. В результате все условные варианты оказываются целыми числами.

Вводятся в рассмотрение условные моменты к-го порядка:

$$\bar{M}_k^* = \frac{1}{N}\sum n_j \left(\frac{x_j - C}{h}\right)^k = \frac{1}{N}\sum n_j u_j^k$$

Связь условных эмпирических моментов с эмпирическими начальными и центральными моментами

Легко показать справедливость следующих соотношений:

$$\bar{x}_в = \bar{M}_1 = \bar{M}_1^* h + C$$

$$\bar{m}_2 = \left(\bar{M}_2^* - \left(\bar{M}_1^*\right)^2\right) h^2$$

$$\bar{m}_3 = \left(\bar{M}_3^* - 3\bar{M}_2^*\bar{M}_1^* + 2\left(\bar{M}_1^*\right)^3\right) h^3$$

$$\bar{m}_4 = \left(\bar{M}_4^* - 4\bar{M}_3^*\bar{M}_1^* + 6\bar{M}_2^*\left(\bar{M}_1^*\right)^2 - 3\left(\bar{M}_1^*\right)^4\right) h^4$$

Вычисление условных моментов

Рассматриваемый ниже метод вычислений носит название метода произведений. Вычисление условных моментов для интервального ряда удобно производить оформляя вычисления в виде следующей таблицы.

υ	n	u	n u	n u^2	n u^3	n u^4	n $(u+1)^4$
347.5	0.06034	-3	-0.18103	0.54310	-1.62931	4.88793	0.96552
382.5	0.11207	-2	-0.22414	0.44828	-0.89655	1.79310	0.11207
417.5	0.19828	-1	-0.19828	0.19828	-0.19828	0.19828	0.0
452.5	0.26724	0	0.0	0.0	0.0	0.0	0.26724
487.5	0.22414	1	0.22414	0.22414	0.22414	0.22414	3.58621
522.5	0.07759	2	0.15517	0.31034	0.62069	1.24138	6.28448
558.0	0.06034	3	0.18103	0.54310	1.62931	4.88793	15.44828
∑	1	–	-0.04310	2.26724	-0.25	13.23276	26.66379

Вычисление статистических оценок

Во 2-м столбце таблицы записываются относительные частоты. В 3-м столбце - условные варианты.
Суммы элементов 4,5.6 и 7 столбцов равны условным моментам соответствующего порядка.
Сумма элементов 8-го столбца является контрольной суммой. Для случая, когда во втором столбце таблицы записаны относительные частоты, должно быть выполнено равенство.

$$\sum n_j(u_j+1)^4 = \sum n_j u_j^4 + 4\sum n_j u_j^3 + 6\sum n_j u_j^2 + 4\sum n_j u_j + 1$$

После вычислений в приведенной таблице вычисляются нужные эмпирические начальные и центральные моменты, а затем, по указанным ранее формулам, статистические оценки математического ожидания, дисперсии, асимметрии и эксцесса.