Точечные и интервальные оценки случайной величины

Занятие 9

Определения

Приближенное значение случайной величины, вычисленное по ограниченному числу опытов, т. е. выборке, содержит элемент случайности и называется ***оценкой.***

Статистические оценки делятся на точечные и интервальные.

Оценка определяемая одним числом, называется ***точечной.***

Требования к оценке

Оценка $\tilde{a}$ для параметра a представляет собой функцию величин $x_1, x_2, ..., x_n$

$$\tilde{a} = a(x_1, x_2, ..., x_n)$$

1) **Несмещенность** – среднее значение средних величин равно среднему значению выборки

$$M(\tilde{a}) = a$$

Требования к оценке

2) **Состоятельность** – с увеличением числа опытов случайная величина $\tilde{a}$ приближается (сходится по вероятности) к параметру a

$$D(\tilde{a}) \to 0$$

$$Lim_{n\to\infty}\left(P(\tilde{a} - a) \geq \varepsilon\right) = 0$$

Требования к оценке

3) **Эффективность** – оценка обладает наименьшей дисперсией по сравнению с другими

$$D(\tilde{a}) = \min$$

Точечные оценки

Для математического ожидания: **среднее выборочное** наблюдаемых значений – это состоятельная и несмещенная оценка.

Эффективность оценки зависит от вида распределения случайной величины.

Для дисперсии : **выборочная дисперсия, среднее квадратичное отклонение и исправленная выборочная дисперсия**

$$\sigma_0^2 = \frac{n}{n-1} * \sigma^2$$

– несмещенные оценки.

Интервальная оценка

Для оценки точности и надежности вычисленного параметра $\widetilde{a}$ используют ***доверительные интервалы*** и ***доверительные вероятности***.

$$\left(\widetilde{a}-\varepsilon;\ \widetilde{a}+\varepsilon\right)$$

$\widetilde{a}$ - точечная оценка параметра;

ε- некоторая малая, положительная величина.

Определение

Доверительный интервал - интервал

$$I_\beta = \left(\widetilde{a} - \varepsilon; \widetilde{a} + \varepsilon\right)$$

который с вероятностью β **накрывает** истинное значение параметра.

β - достаточно большая вероятность, при которой событие можно считать практически достоверным.

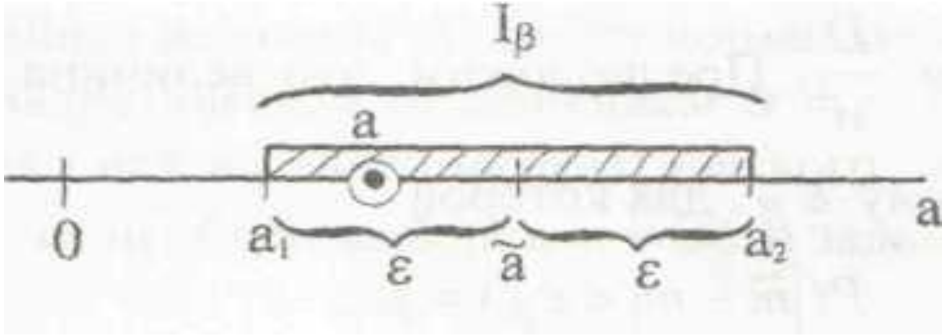
$\alpha = 1 - \beta$ - вероятность допустить ошибки при вычислении параметра (очень малая величина, уровень значимости).

Свойства интервальной оценки

1) ***Точность*** интервала определяется ε (чем меньше, тем точнее интервал).

2) **Надежность** интервала определяется доверительной вероятность β (надежностью).

Интервал может быть точным или надежным.



Доверительный интервал для математического ожидания

Пусть величина X имеет нормальное распределение с известным средним квадратическим отклонением σ, тогда доверительный интервал β надежностью имеет вид

$$\bar{x} - \sigma_m \cdot \arg \Phi\left(\frac{\beta}{2}\right), \bar{x} + \sigma_m \cdot \arg \Phi\left(\frac{\beta}{2}\right)$$

где $\Phi\left(\frac{\beta}{2}\right)$ - значение функции Лапласа, определяется по таблице, $\sigma_m = \frac{\sigma}{\sqrt{n}}$

Доверительный интервал для математического ожидания

В случае малой выборки $n < 30$ значение функции Лапласа заменяется на функцию Стьюдента (или t-распределения) с (n-1) степенями свободы

$$\bar{x} - \sigma_m \cdot t_\beta ,\ \bar{x} + \sigma_m \cdot t_\beta$$

Определяется по таблице.

Замечания

1. Минимальный объем выборки, который обеспечит требуемую оценку математического ожидания с достаточной точностью ε и надежностью β, можно найти по формуле

$$n = \frac{\left[\arg \Phi\left(\frac{\beta}{2}\right)\right]^2 \cdot \sigma^2}{\varepsilon^2}$$

2. При неограниченном возрастании объема выборки n распределение Стьюдента стремится к нормальному.

Пример

Количественный признак X генеральной совокупности распределен нормально.

По выборке объема n= 16 найдены выборочная средняя $\overline{x}$ =20,2 и σ= 0,8.

Оценить неизвестное математическое ожидание при помощи доверительного интервала с надежностью β = 0,95.

Решение

По таблице квантилей t-распределения находим значение $t_{0.95}(16-1)=2.13$.

Используя определение доверительного интервала находим границы:

$$\bar{x}-t_{\beta}\cdot\sigma=20.2-2.13\cdot\frac{0.8}{\sqrt{16}}=19.77$$

$$\bar{x}+t_{\beta}\cdot\sigma=20.2+2.13\cdot\frac{0.8}{\sqrt{16}}=20.63$$

$$(19.77;\ 20.63)$$

Доверительный интервал для среднего квадратичного отклонения

Пусть количественный признак X генеральной совокупности распределен нормально. Оценить неизвестное генеральное среднее квадратическое отклонение (найти доверительные интервалы, покрывающие параметр σ с надежностью β, т.е.

$$P\left(\sqrt{\sigma_0{}^2} - \varepsilon < \sigma < \sqrt{\sigma_0{}^2} + \varepsilon\right) = \beta$$

Доверительный интервал для среднего квадратичного отклонения

Для нахождения ε используют распределение χ^2

Его значение определяют по таблице квантилей χ^2 или q

$$\sqrt{{\sigma_0}^2}\cdot(1-q)<\sigma<\sqrt{{\sigma_0}^2}\cdot(1+q)$$

Где $\chi=\frac{\sqrt{{\sigma_0}^2}}{\sigma}\sqrt{n-1}$, величина распределенная по закону хи-квадрат,

σ_0 - исправленное среднее квадратичное отклонение

СПАСИБО ЗА ВНИМАНИЕ