8. Понятие интервальной оценки числовой характеристики случайной величины. Доверительная вероятность (надежность). Доверительный интервал. Метод построения интервальной оценки с помощью центральных статистик.

Точечная оценка даёт нам лишь приближенное значение, т.е вероятность того, что оценка параметра совпадает с его истинным значением никогда не равна 1

Если для больших выборок это ещё ок, вероятность стремиться к одному (свойство состоятельности), то для маленьких не всё так радужно

поэтому делают следующим образом

1. По выборке определяют точечную оценку $\widehat{\theta}$ неизвестного параметра θ;
2. Затем задается вероятность γ (надежность), и по определенным правилам находят такие значения статистик $T_1(x_1, \ldots, x_n)$ и $T_2(x_1, \ldots, x_n)$, $T_1 < T_2$, чтобы выполнялось соотношение:

$$P(T_1 < \theta < T_2) = \gamma.$$

ОПР (интервальной оценки)

Интервальная оценка параметра θ - это числовой интервал (T_1, T_2), в котором с заданной вероятность γ находится параметр θ

Границы T_1 и T_2 определяются по выборочным данным(случайным точкам), то они тоже будут случайными величинами. Кроме того границы также зависят от γ

- с **ростом γ длина интервала увеличивается**, с **уменьшением уменьшается**

ОПР (довертильного интервала и вероятности)

(T_1, T_2) - это доверительный интервал, а γ - доверительная вероятность(надежность)

Интервальная оценка должна удовлетворять следующим требованиям:

1. Должна быть точной, т.е ширина интервала $\Delta = \frac{(T_1 - T_2)}{2}$ должна быть как можно меньше
2. γ должна принимать высокие значения, обычно берут всё что больше 0.9

Также вводят понятие уровня значимости α. Математически он равен 1 - γ, по сути обозначает вероятность ошибки оценки, т.е **вероятность, с которой оцениваемый параметр в интервале не находится**

Как искать интервалы

интервалы ищут с помощью центральной статистики и магии

ОПР(центральной статистики)

$W(x, \theta)$- центральная статистика, если:

1. $W(x, \theta)$ строго монотонная функция и она непрерывна по θ
2. $F_W(x)$ - функция распределения W не зависит от параметра θ

ОПР(Квантиль)

число z_γ - называют γ-квантилем распределения F(x),если выполняется условие F(z_γ) = γ

по сути z_γ - обозначает правый конец интервала (-∞, z_γ), в который попадёт точка с вероятность γ

вычислять квантили можно с помощью обратной функции распределения(ищи её в экселе, юпитере, статистике)

Зная всё это добро сможем найти границы следующим образом

$$P(T_1 < W(x, \theta)) = \gamma = 1 - \alpha$$

$$1 - P(T_1 < W(x, \theta)) = 1 - \gamma = \alpha = P(W(x, \theta) < T_1)$$

ну и всё зная функцию распределения $W(x, \theta)$ - $F_W(x)$ находим теперь α-квантиль он и будет левой границей T_1

Поступаем аналогично и находим другой конец интервала. STONKS

Вероятность не попасть в интервал (T_1, T_2) равна α. Предполагают, что вероятность оказаться левее или правее одинакова и составляет $\frac{\alpha}{2}$.
Т.к $1 - \alpha = \gamma$, то :

- $\frac{\alpha}{2} = \frac{1-\gamma}{2}$
- $1 - \frac{\alpha}{2} = \frac{1+\gamma}{2}$

поэтому можно записать так:

$$P(z_{\frac{1-\gamma}{2}} < W(x, \theta) < z_{\frac{1+\gamma}{2}}) = \gamma$$

Замечание

Если у распределения плотность распределния симетрична относительно оси ординат, то квантили тоже будут симетричными $z_\gamma = z_{1-\gamma}$

Иногда интервал выбирают симметрично относительно какой-то оценки параметра θ, т.е $(\theta - \Delta, \theta + \Delta)$

здесь Δ - это абсолютная погрешность оценки