

Точечные и интервальные оценки



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, \dots, x_n

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

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

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

$$\lim_{n \rightarrow \infty} (P(\tilde{a} - a) \geq \varepsilon) = 0$$

Требования к оценке

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

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

Точечные оценки

Среднее выборочное наблюдаемых значений – это состоятельная и несмещенная оценка.

Эффективность оценки зависит от вида распределения случайной величины.

Выборочная дисперсия, среднее квадратическое отклонение и исправленная выборочная дисперсия

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

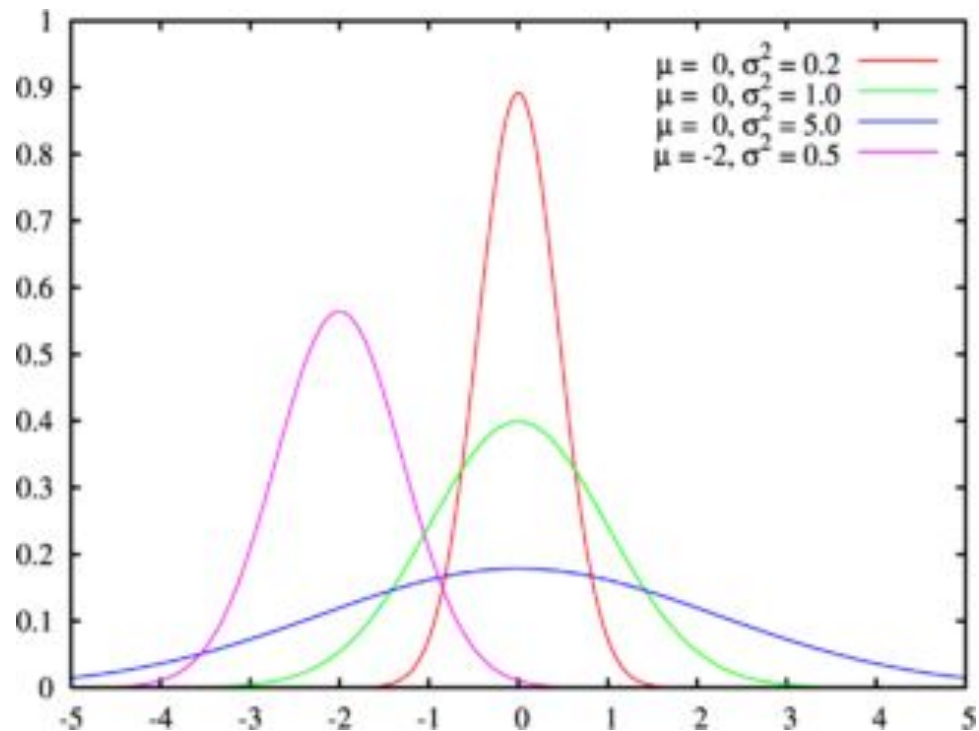
– несмещенные оценки.

• При обработке статистической информации широко используют распределение статистик, которые вычисляются по выборке из нормально распределенной генеральной совокупности.

Квантилью порядка p распределения случайной величины X называется действительное число X_p , удовлетворяющее уравнению $P(X < x_p) = p$. (определяются по специальным стат. таблицам)

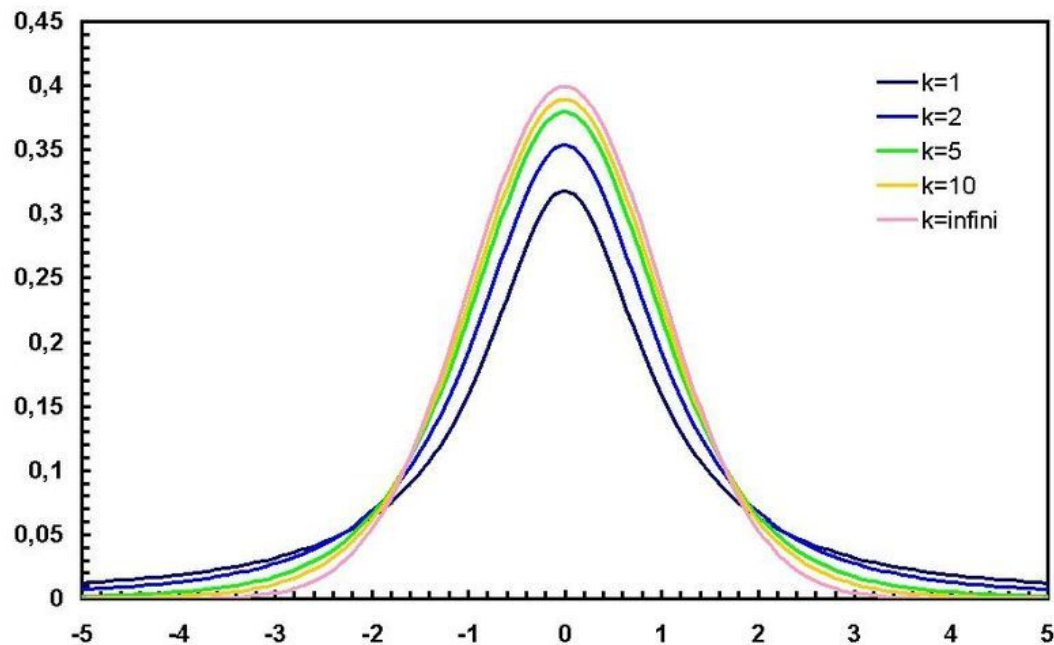
Виды распределений: Стандартное нормальное распределение

График плотности распределения



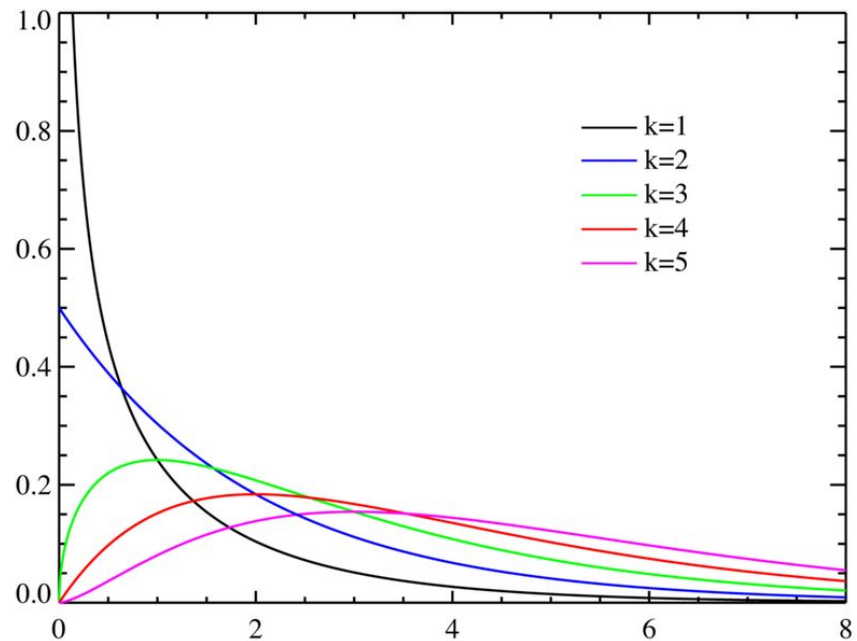
Виды распределений: t-распределение (Стьюдента)

Распределение Стьюдента с k степенями свободы, t-распределение симметрично. При увеличении объема выборки распределение стремится к нормальному.



Виды распределений: χ^2 распределение

Распределение с k – степенями свободы.



Интервальная оценка

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

$$(\tilde{a} - \varepsilon; \tilde{a} + \varepsilon)$$

\tilde{a} - точечная оценка параметра;

ε некоторая малая, положительная величина.

Определение

Доверительный интервал - интервал

$$I_{\beta} = (\tilde{a} - \varepsilon; \tilde{a} + \varepsilon)$$

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 вероятностью β (надежностью).

Интервал может быть точным или надежным.

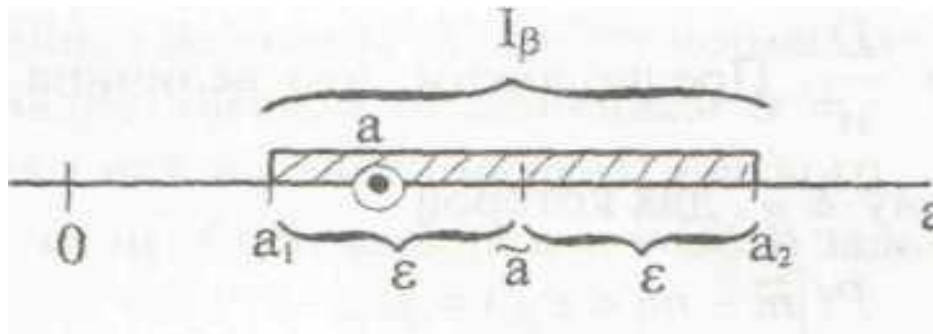


Таблица построения интервальных оценок

Параметр	Интервальная оценка	
Математическое ожидание	$\left(\bar{X}_s - A \frac{S_0}{\sqrt{n}}; \bar{X}_s + A \frac{S_0}{\sqrt{n}}\right)$	<p>Среднее выборочное $\bar{X}_s = \frac{\sum_{i=1}^k x_i n_i}{n}$;</p> <p>Распределение Стьюдента $A = t_{1-\frac{\alpha}{2}}(n-1)$ - если $n < 30$</p> <p>Стандартное нормальное распределение $A = U_{1-\frac{\alpha}{2}}$ - если $n \geq 30$;</p> <p>$S_0 = \sqrt{S_0^2}$ - среднеквадратическое отклонение</p>
Дисперсия	$\left(\frac{S_0^2(n-1)}{A_1}; \frac{S_0^2(n-1)}{A_2}\right)$	<p>Выборочная дисперсия $S_0^2 = \frac{\sum_{i=1}^k (x_i - \bar{x}_s)^2 n_i}{n-1}$;</p> <p>Распределение Хи-квадрат $A_2 = \chi_{1-\frac{\alpha}{2}}^2(n-1)$; $A_1 = \chi_{\frac{\alpha}{2}}^2(n-1)$</p>
Доля признака	$\left(\tilde{p} - A_{1-\frac{\alpha}{2}} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n}}; \tilde{p} + A_{1-\frac{\alpha}{2}} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n}}\right)$	<p>Стандартное нормальное распределение $A = U_{1-\frac{\alpha}{2}}$; $\tilde{p} = \frac{m}{n}$ - выборочная доля признака; m, n - число благоприятных и всевозможных исходов соответственно</p>

СПАСИБО ЗА ВНИМАНИЕ