ЛЕКЦИЯ 13

Теория вероятностей и математическая статистика

Интервальное оценивание

Интервальное оценивание

- Пусть, как обычно, имеется выборка из распределения с неизвестным параметром θ. До сих пор мы занимались «точечным оцениванием» неизвестного параметра — находили приближенное значение параметра (его оценку).

  Интервальное оценивание
- Существует другой подход к оцениванию, при котором мы указываем интервал, накрывающий параметр с заданной наперед вероятностью. Такой подход называется «интервальным оцениванием».

<u>Определение.</u>

Доверительным интервалом **уровня значимости α ($0< \alpha <1$) для параметра θ называется интервал $I=[I_1, I_2]$, для которого выполняется условие:**

$$P(I_1(X) \leq \theta \leq I_2(X)) = 1 - \alpha.$$

- **Число $1 - \alpha$ называется *доверительной вероятностью*, а $I_1(X)$, $I_2(X)$**

– нижней и верхней доверительными границами.

Смысл доверительного интервала

- Таким образом, $(1 - \alpha)$ –доверительный интервал – случайный интервал, который с вероятностью $1 - \alpha$ накрывает истинное значение параметра θ.
- (Параметр – неслучайная величина, а границы интервала случайны, поэтому читают формулу как «интервал накрывает параметр», а не как « параметр попадает в интервал»).

Уровень значимости α

- Его обычно берут равным одному из чисел 0.001, 0.005, 0.01, 0.05, 0.1. Уровень значимости выражает ошибку доверительного интервала. Чем меньше α , тем больше доверительная вероятность и тем надежнее доверительный интервал, но более надежный интервал является более широким и менее информативным. Стандартный уровень значимости α =0.05. Соответствующий доверительный интервал называется 95% –м.

Доверительный интервал для параметра a нормального распределения N(a,σ).

Известно, что $\dfrac{a-\bar{x}}{\dfrac{\sigma}{\sqrt{n}}} \in N(0,1).$

Запишем уравнение:

$$P\left(\gamma_1 \le \frac{a-\bar{x}}{\dfrac{\sigma}{\sqrt{n}}} \le \gamma_2\right) = 1-\alpha.$$

$\gamma_1 = u_{\alpha/2},\ \ \gamma_2 = u_{1-\alpha/2}.$

Продолжение

Разрешим неравенство

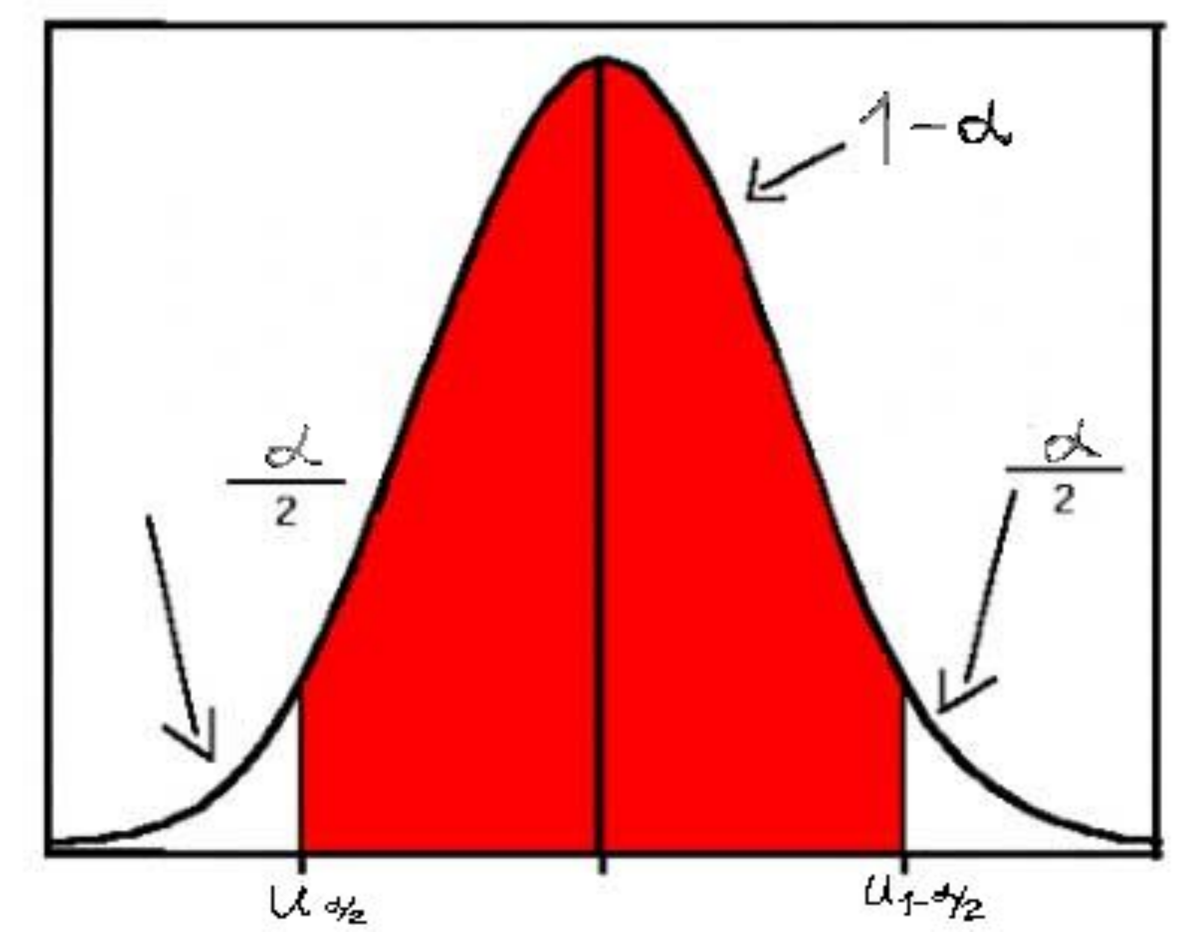
относительно а:

$$P(\bar{x}+\gamma_1\cdot\frac{\sigma}{\sqrt{n}}\le a\le\bar{x}+\gamma_2\cdot\frac{\sigma}{\sqrt{n}})=1-\alpha.$$

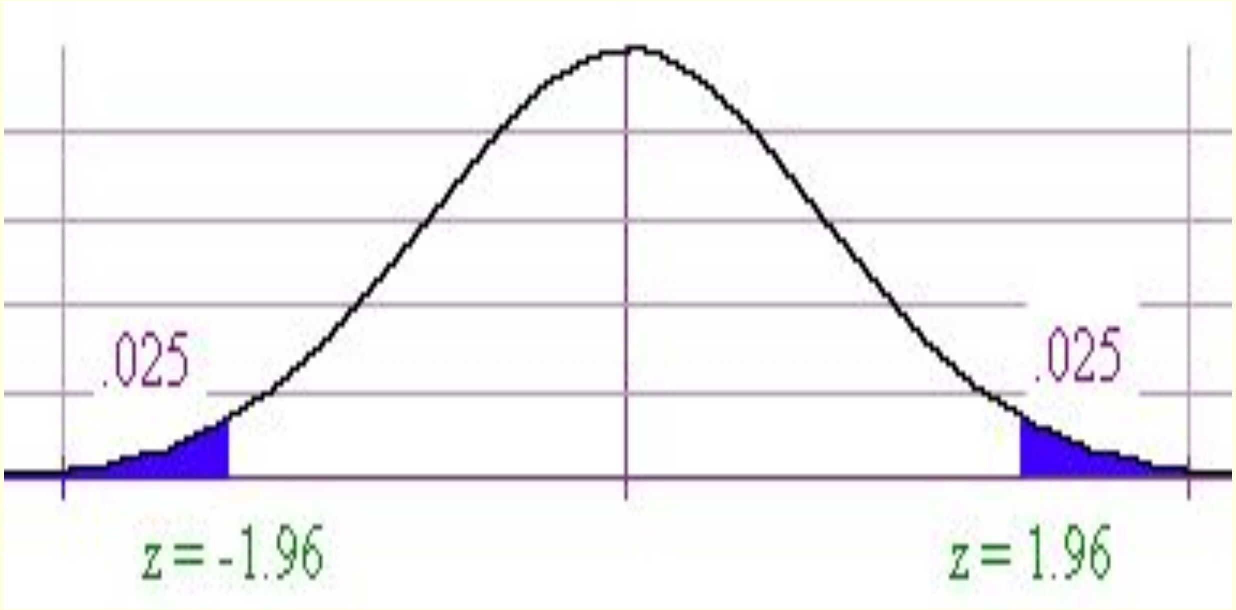
$$\gamma_1=u_{\alpha/2},\ \ \gamma_2=u_{1-\alpha/2},\qquad u_{\alpha/2}=-u_{1-\alpha/2}.$$

$$P(\bar{x}-u_{1-\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}\le a\le\bar{x}+u_{1-\alpha/2}\cdot\frac{\sigma}{\sqrt{n}})=1-\alpha.$$

Квантили нормального распределения



Квантили нормального распределения



Окончательный ответ:

Доверительный интервал для а :

$$I_a = \left[\bar{x} - u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \quad \bar{x} + u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right].$$

Замечание

- Значения квантилей находят в таблицах. Приведем наиболее часто используемые квантили стандартного нормального распределения N(0,1):
  - $u_{0.95} = 1.64$,
  - $u_{0.975} = 1.96$,
  - $u_{0.995} = 2.58$.

Пример

- Найти доверительный интервал уровня значимости $\alpha = 0.05$ для неизвестного математического ожидания a нормально распределенного признака X генеральной совокупности, если $\sigma = 5$, выборочное среднее $=14$ и объем выборки $n = 25$.

Доверительный интервал для a:

$$I_a = \left[\bar{x} - u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \quad \bar{x} + u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right].$$

Все величины известны. Подставив $u_{0.975} = 1.96$, $\bar{X} = 14$, $\sigma = 5$, $n = 25$, окончательно получим искомый доверительный интервал

$$12{,}04 \le a \le 15{,}96.$$

Схема построения доверительного интервала

- Т.о., надо взять статистику $G(x, \theta)$, такую, что она сама зависит от параметра θ, а ее распределение от θ не зависит, записать уравнение

$$P(\gamma_1 \le G(x, \theta) \le \gamma_2) = 1 - \alpha,$$

и разрешить неравенство под знаком вероятности относительно параметра θ.

Как найти γ_1 и γ_2

- В качестве γ_1 и γ_2 будем использовать квантили распределения статистики $G(x, \theta)$:
- $\gamma_1 = G_{\alpha/2}$, $\gamma_2 = G_{1-\alpha/2}$

- **<u>Напоминание.</u>** ***Квантиль порядка q отсекает слева 100·q% значений случайной величины.***

Доверительный интервал для параметра а нормального распределения N(а,σ) (при *неизвестном* σ)

Мы не можем использовать статистику $\dfrac{a-\bar{x}}{\frac{\sigma}{\sqrt{n}}}$, т.к. σ неизвестно.

Но мы знаем похожую статистику, не использующую σ :

$$\frac{a-\bar{x}}{\frac{\bar{s}}{\sqrt{n}}} \in T_{n-1}.$$

Запишем уравнение:

$$P\left(\gamma_1 \le \frac{a-\bar{x}}{\frac{\bar{s}}{\sqrt{n}}} \le \gamma_2\right) = 1-\alpha. \qquad \gamma_1 = {}_{n-1,\alpha/2}, \;\; \gamma_2 = {}_{n-1,1-\alpha/2}.$$

Здесь t.

Окончательный ответ (при *неизвестном* σ) :

Доверительный интервал для а:

$$I_a = \left[\bar{x} - t_{n-1,1-\alpha/2} \cdot \frac{\bar{s}}{\sqrt{n}}, \quad \bar{x} + t_{n-1,1-\alpha/2} \cdot \frac{\bar{s}}{\sqrt{n}} \right].$$

Доверительный интервал для параметра σ нормального распределения N(a,σ).

Известно, что $\frac{nS^2}{\sigma^2} = \chi^2_{n-1}$.

Запишем уравнение:

$$P(\gamma_1 \leq \frac{nS^2}{\sigma^2} \leq \gamma_2) = 1 - \alpha.$$

$$\gamma_1 = \chi^2_{n-1, \alpha/2}, \quad \gamma_2 = \chi^2_{n-1, 1-\alpha/2}.$$

Продолжение

Разрешим неравенство

относительно

$$P(\gamma_1 \le \frac{nS^2}{\sigma^2} \le \gamma_2) = P(\frac{\gamma_1}{nS^2} \le \frac{1}{\sigma^2} \le \frac{\gamma_2}{nS^2}) =$$

$$= P(\frac{nS^2}{\gamma_2} \le \sigma^2 \le \frac{nS^2}{\gamma_1}) = 1 - \alpha.$$

$$\gamma_1 = \chi^2_{n-1, \alpha/2}, \quad \gamma_2 = \chi^2_{n-1, 1-\alpha/2}.$$

Окончательный ответ:

Доверительный интервал для σ:

$$I_\sigma = \left[\frac{S\sqrt{n}}{\sqrt{\chi^2_{n-1,1-\alpha/2}}}, \quad \frac{S\sqrt{n}}{\sqrt{\chi^2_{n-1,\alpha/2}}} \right].$$

Асимптотический доверительный интервал

- Если оценка параметра асимптотически нормальна и несмещена, (например,

является о.м.п.), то

$$\frac{\theta - \hat{\theta}}{\sqrt{D\hat{\theta}}} \to u \in N(0,1).$$

Асимптотический доверительный интервал

$$\text{Если } \frac{\theta - \hat{\theta}}{\sqrt{D\hat{\theta}}} \to u \in N(0,1), \text{ то}$$

можно записать уравнение:

$$P\left(u_{\alpha/2} \le \frac{\theta - \hat{\theta}}{\sqrt{D\hat{\theta}}} \le u_{1-\alpha/2}\right) = 1 - \alpha.$$

- **Разрешив неравенство относительно θ, получим доверительный интервал для параметра θ значимости α.**