Классические довер-ые интервалы.

- для $p = P(X_i = 1)$
- для $\mu = E(X_i)$
- для $\sigma^2 = Var(X_i)$
- для $p_1 - p_2$
- для $\mu_1 - \mu_2$
- для σ_1^2 / σ_2^2

Задача 1: $X_i \in \{1, 0\}$ $X_i \sim Bern(p)$ p – неизв. величина

выборка: $X_1, \dots, X_n$ – незав

наивная точечная оценка $\hat{p} = \dfrac{X_1 + \dots + X_n}{n}$ ← и метод м-в и мет. максим-правд

$\text{plim}\, \hat{p} = p$ (по ЗБЧ)

CI для p? (n велико)

a) $Var(\hat{p})$?

б) как сост-но оценить $Var(\hat{p})$?

$$\text{plim} \frac{\widehat{Var}(\hat{p})}{Var(\hat{p})} = 1$$

в) асу. CI для p?

---

sol-n:

a) $Var(\hat{p}) = \frac{1}{n^2} \cdot Var(X_1 + \dots + X_n) = \frac{n}{n^2} Var(X_i) =$ $\quad X_i \in \{0, 1\}$, $E(X_i^2) = E(X_i)$

$= \frac{1}{n} \cdot (E(X_i^2) - (E(X_i))^2) = \frac{1}{n}(p - p^2) = \frac{p(1-p)}{n}$

содержит неизв p не знаем.

б) !! наивная plug-in estimator $\widehat{Var}(\hat{p}) = \dfrac{\hat{p}(1-\hat{p})}{n}$

$$\text{plim} \frac{\widehat{Var}(\hat{p})}{Var(\hat{p})} = \text{plim} \frac{\hat{p}(1-\hat{p})}{p(1-p)} = 1 \; !!$$

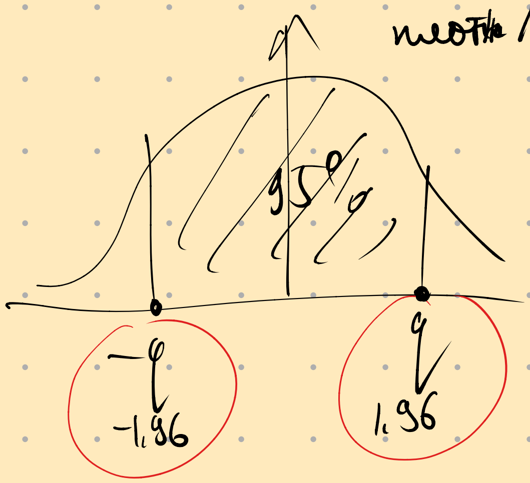
в) ЦПТ: $\dfrac{\hat p - E(\hat p)}{\sqrt{Var(\hat p)}} \xrightarrow{n\to\infty} N(0;1)$

новое

n - велико

$$P\left(-q \le \frac{\hat p - p}{\sqrt{\frac{p(1-p)}{n}}} \le q\right) \approx 0{,}95$$

плотн $N(0;1)$



$$-q \le \frac{\hat p - p}{\sqrt{\frac{p(1-p)}{n}}} \le q$$

$$\begin{cases} \hat p - p \le q \cdot \sqrt{\frac{p(1-p)}{n}} \\ \hat p - p \ge -q\sqrt{\frac{p(1-p)}{n}} \end{cases}$$

↳ p?

решим ур-ие

упрощ. (уже делали)

подменим знам-тель на $\hat{Var}(\hat p)$

$$-q \le \frac{\hat p - p}{\sqrt{\frac{\hat p(1-\hat p)}{n}}} \le q$$

... [решаем нер-во]

$$p \in \left[\hat p - q\sqrt{\frac{\hat p(1-\hat p)}{n}}; \hat p + q\cdot\sqrt{\frac{\hat p(1-\hat p)}{n}}\right]$$

[более грубый, но более широкий]

$$(\hat p - p)^2 = q^2 \cdot \frac{p(1-p)}{n}$$

p – неизв. вер

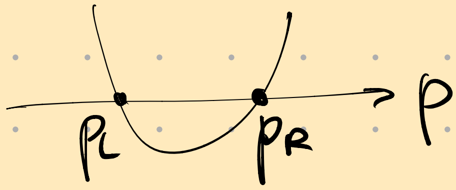
$\hat p$ – оценка (наив.)

q – [не $1-p$] - квантиль

$$n(p^2 + \hat p^2 - 2p\hat p) = q^2 p - q^2 p^2$$

$$(n+q^2)p^2 - (q^2 + 2\hat p n)\cdot p + n\hat p^2 = 0.$$

ветви вверх



$$D = (q^2 + 2\hat p n)^2 - 4(n+q^2)\cdot n\hat p^2$$

$$p \in \left[\frac{q^2+2\hat p n - \sqrt{D}}{2(n+q^2)};\ \frac{q^2+2\hat p n+\sqrt{D}}{2(n+q^2)}\right]$$

для ур-ня дов 95%

$q = 1{,}96$

$\hat p = \frac{X_1+\ldots+X_n}{n}$

и знаем

[хорошо работает даже для небольших выборок]

! Вариант с корнями обеспечивает более близкую фактич. вер-сть покрытия к заказанной (номин-ой).

Задача 2

X_i ~ незав $X_i \sim N(\mu;\sigma^2)$ μ, σ^2 неизв.

$X_1, \ldots, X_n$

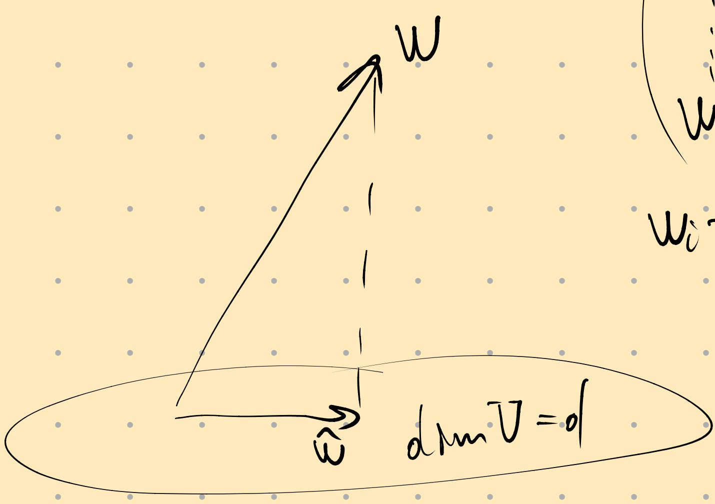
[n может быть мало!]

очень сильная

- постройте точный CI для σ^2.

а) как распр-на СВ $\frac{\sum_{i=1}^n (X_i-\bar X)^2}{\sigma^2}$?

б) постройте CI для σ^2



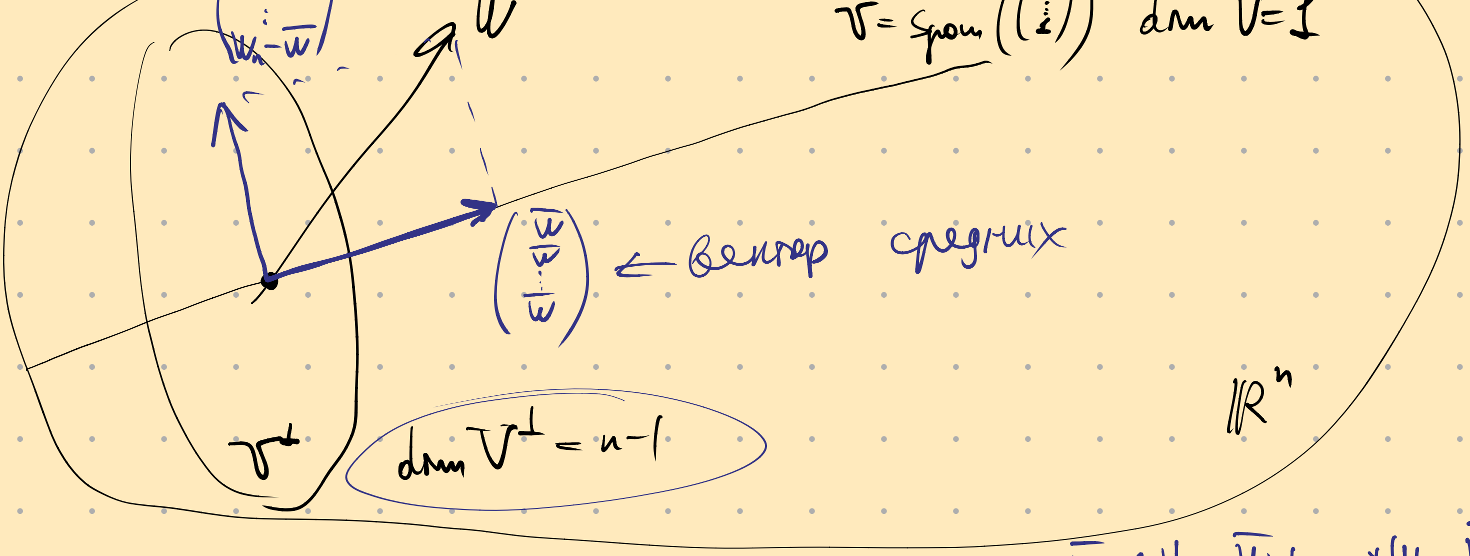
$$\begin{pmatrix} W_1 \\ \vdots \\ W_n \end{pmatrix} \sim N\left(\begin{pmatrix} 0\\0\\0\\0\\ \vdots \\ 0\end{pmatrix}; \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix}\right)$$

W_i ~ незав $N(0;1)$

$\|\hat W\|^2 \sim \chi^2_d$ ← кв. длина проекции

d – разм-сть пр-ва, куда проецирую

$\begin{pmatrix} W_1-\bar W \\ W_2-\bar W \end{pmatrix}$ ← вектор отклонений от средних



$$\sum_{i=1}^{n} (W_i - \bar W)^2 \sim \chi^2_{n-1}$$

$$\underbrace{W_1 - \bar W + W_2 - \bar W + \ldots + W_n - \bar W}_{= 0}$$

$$\sum W_i - n\cdot \bar W = \sum W_i - \sum W_i = 0$$

$X_i \sim N(\mu; \sigma^2)$ μ неизв.! σ^2 неизв.

$$W_i \sim N(0;1) \iff X_i \sim N(\mu;\sigma^2)$$

$$X_i = \mu + \sigma \cdot W_i$$

$$X_i - \bar X = \mu + \sigma \cdot W_i - \frac{\sum_{j=1}^{n}(\mu + \sigma W_j)}{n} = \mu + \sigma W_i - \frac{n\mu + \sigma \cdot \sum W_j}{n} =$$

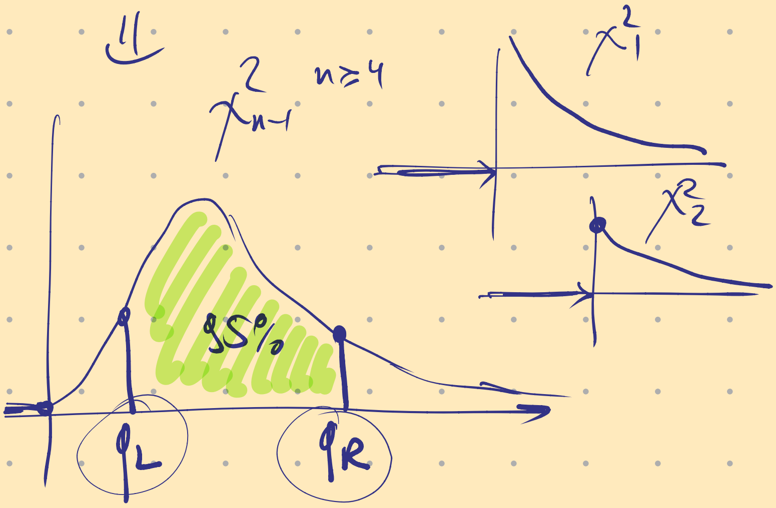
$$= \mu + \sigma W_i - \mu - \sigma \cdot \bar W =$$

$$= \sigma (W_i - \bar W)$$

a) $$R = \frac{\sum (X_i - \bar X)^2}{\sigma^2} = \frac{\sum(\sigma(W_i - \bar W))^2}{\sigma^2} = \underbrace{\sum (W_i - \bar W)^2}_{} = \|\text{проекция } W \text{ на } V^{\perp}\|^2$$

$$R \sim \chi^2_{n-1}$$

б) CI



$n \in \{2, 3, \ldots\}$

$$P(q_L \leq \quad \leq q_R) = 0.95$$

$$q_L \le \frac{\sum (X_i - \bar{X})^2}{\sigma^2} \le q_R$$

CI (точный!) $\sigma^2 \in \left[\frac{\sum (X_i-\bar{X})^2}{q_R}; \frac{\sum (X_i-\bar{X})^2}{q_L}\right]$

[при жёсткой $X_i \sim N(\mu;\sigma^2)$]

[асимпт. ничего не надо или
$n\to\infty$ $X_i \sim$ DISTR
конечн. 4-ый момент]

	n-мало	n-велико ($n\to\infty$)
$[X_i \sim N(\mu;\sigma^2)]$ жёсткие предп.	☺	☺
$[X_i \sim$ одинак. распр.$]$ мягкие предп. $X_i \sim$ одинак. распр.	Нет стат.	ЦПТ asy ☺

$P(X_i \le t) = F(t)$

не зав. от i

Бабушка Акиная

Задача 3

$X_1, X_2, \dots, X_n$ незав. одинак. распр.

$\boxed{\text{сцен. 1}}$ n-велико ($n\to\infty$) $E(X_i)$, $Var(X_i)$ конечн

$\boxed{\text{сцен 2}}$ n-любое, но жёсткая предп. $X_i \sim N(\mu;\sigma^2)$

в) $\boxed{\text{сцен 1}}$: asy 95% CI для μ

г) $\boxed{\text{сцен 2}}$: точный 95% CI для μ.

а) наивная точечная оценка $\hat{\mu}$?

б) состоят. точечная оценка для $\hat{\sigma}^2$?

а) $\hat{\mu} = \frac{X_1 + X_2 + \dots + X_n}{n}$ ☺ $E(\hat{\mu}) = \mu$ несмещ.

$\text{plim}\, \hat{\mu} = \mu$ состоят.

б) $\hat{\sigma}^2_B = \frac{\sum(X_i-\bar{X})^2}{n}$ — число слаг-х квадратов

$\hat{\sigma}^2_{UN} = \frac{\sum(X_i-\bar{X})^2}{n-1}$ — размер-ть пр-ва, куда проец.

$E(\hat{\sigma}^2_B) = \frac{n-1}{n}\cdot\sigma^2$ $\qquad$ $E(\hat{\sigma}^2) = \sigma^2$

обе — сост.

"смещённая" $\qquad$ "несмещённая"

"нескорректир-ная" $\qquad$ "скорректир-ая"

$\operatorname{plim} \hat{\sigma}^2_B = \operatorname{plim} \hat{\sigma}^2_{UN} = \sigma^2$

sample variance = выборочная дисперсия.

в) $\boxed{\text{asy CI}}$ $\qquad n \to \infty$

ЦПТ. $\qquad \frac{\hat{\mu} - E(\hat{\mu})}{\sqrt{Var(\hat{\mu})}} \xrightarrow{n\to\infty} N(0;1)$

ЦПТ + лемма Слуцкого $\qquad \frac{\hat{\mu} - E(\hat{\mu})}{\sqrt{\widehat{Var}(\hat{\mu})}} \xrightarrow{n\to\infty} N(0;1)$

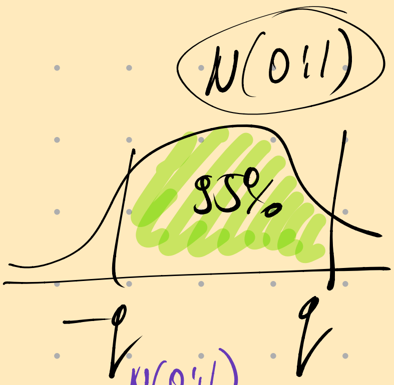
можно исп-ть: $\qquad \hat{\sigma}^2_B$ и $\hat{\sigma}^2_{UN}$

$$Var(\hat{\mu}) = Var\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

$$\widehat{Var}(\hat{\mu}) \to \hat{\sigma}^2_{UN}/n, \quad \hat{\sigma}^2_B/n$$

$\boxed{\text{готовый рецепт.}}$ $\qquad n \to \infty$

$$-q \le \frac{\hat{\mu} - \mu}{\sqrt{\hat{\sigma}^2/n}} \le q$$



$\Downarrow$ X_i ~ один.

классич. асимпт. интервал для μ:

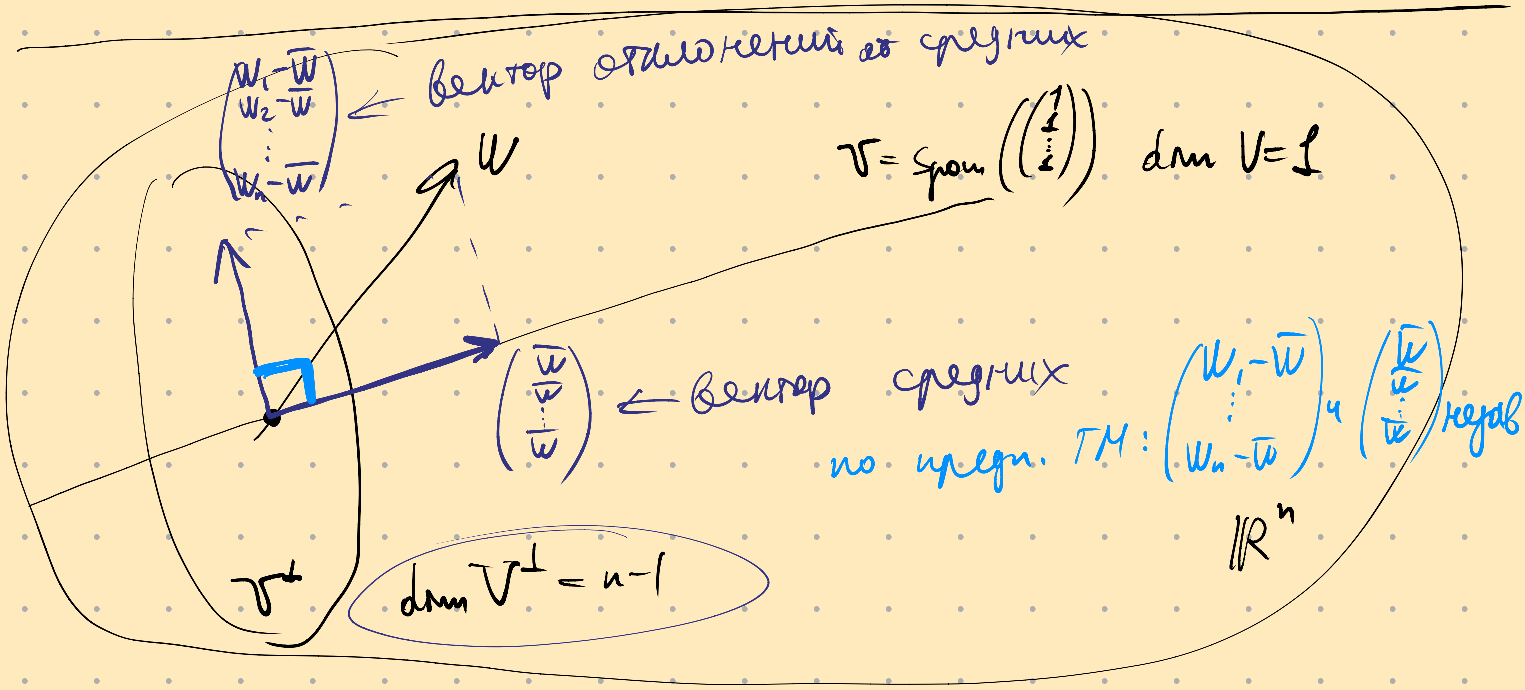
$$\left[\hat{\mu} - \underset{N(0;1)}{q}\cdot\sqrt{\frac{\hat{\sigma}^2}{n}};\ \hat{\mu} + \underset{N(0;1)}{q}\cdot\sqrt{\frac{\hat{\sigma}^2}{n}}\right]$$

2) n – любое, $X_i \sim N(\mu; \sigma^2)$ незав

$\bar{X} \sim N\left(\mu; \frac{\sigma^2}{n}\right)$ — точное

$\frac{\sum (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$ — точное

$\boxed{\text{незав!}}$



точное распр-ие.

незав $\frac{N(0;1)}{\sqrt{\chi^2_d / d}} \sim t_d$

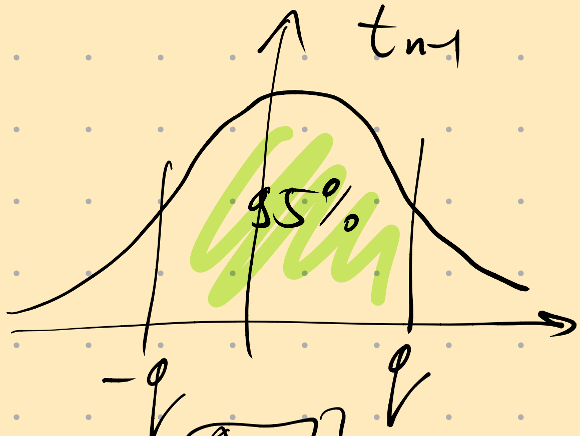
$$\frac{\overbrace{(\bar{X} - \mu) \Big/ \sqrt{\frac{\sigma^2}{n}}}^{N(0;1)}}{\underbrace{\sqrt{\frac{\sum (X_i - \bar{X})^2}{\sigma^2} \Big/ (n-1)}}_{\sqrt{\chi^2_{n-1} / n-1}}} \sim t_{n-1}$$

незав

$$\frac{\bar{X} - \mu}{\sqrt{\frac{\sum (X_i - \bar{X})^2}{n \cdot (n-1)}}} \sim t_{n-1}$$

$$\frac{\bar X-\mu}{\sqrt{\hat\sigma^2_{UN}/n}} = \frac{\bar X-\mu}{\sqrt{\hat\sigma^2_B/(n-1)}} \underset{\text{точно}}{\sim} t_{n-1}$$

$$-q \le \frac{\bar X-\mu}{\sqrt{\hat\sigma^2_{UN}/n}} \le q$$



точный при жёстких предп. $X_i \sim N(\mu;\sigma^2)$

$$\mu \in \left[\bar X - \underset{t_{n-1}}{q}\cdot\sqrt{\frac{\hat\sigma^2_{UN}}{n}};\ \bar X + \underset{t_{n-1}}{q}\cdot\sqrt{\frac{\hat\sigma^2_{UN}}{n}}\right]$$

ересь! ошибочно (иногда) пишут, что треб-ся только $\bar X \underset{\text{точн.}}{\sim} N$