

Centralni limitni izrek

$X_1, X_2, \dots$ neodvisne enako porazdeljene (NEP)
slučajne spremenljivke z $E(X_1^2) < \infty$.

Označimo $S_n = X_1 + \dots + X_n$

Za n dovolj velik je

$$P(a \leq S_n \leq b) \approx \Phi\left(\frac{b - nE(X_1)}{\sqrt{n \operatorname{var}(X_1)}}\right) - \Phi\left(\frac{a - nE(X_1)}{\sqrt{n \operatorname{var}(X_1)}}\right)$$

$$\Phi(x) = P(Z \leq x) ; Z \sim N(0, 1)$$

Kocke vzrjemo 105-krat

a) Oцени vrjetnost da pade < 350 pik.

$X_i, \dots$ pike na i -ti kocki

X_i -ji so enako porazdeljeni: $E(X_i) = 3,5$

$$P(\infty \leq S_{105} \leq 350) \approx$$

$$E(X_i^2) = \sum_{i=0}^6 \frac{1}{6} i^2 = \frac{1}{6} \cdot \frac{6 \cdot 7 \cdot 13}{6} = \frac{7 \cdot 13}{6}$$

$$\text{var}(X_i) = -\frac{7^2}{4} + \frac{7 \cdot 13}{6} = \frac{35}{12}$$

$$P(S < 350) \approx \Phi\left(\frac{350 - 105 \cdot \frac{7}{2}}{\sqrt{105 \cdot \frac{35}{12}}}\right) = \Phi(-1) \approx 0,159 \dots$$

?? pri aproksimaciji je to enako

$$P(S \leq 349) = P(S < 350) = 0,145$$

Dejanska vrjetnost: 0,152

$$P(S < 349,5) = \dots \approx 0,152$$

V povprečju je srednja vrednost najboljša aproksimacija

b) Oцени vjerojatnost da 6 pada najviše 10-krat
 $I_1, \dots$ na item mogu pade 6

$$P(I_i = 1) = \frac{1}{6} \quad I_i \sim \text{Ber}\left(\frac{1}{6}\right)$$

$$\text{var}(I_i) = \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{6^2}$$

$$P(\bar{S}_{105} \leq 10) = P(\bar{S} < 10,5) \approx \Phi\left(\frac{10 - \frac{105}{6}}{\sqrt{105 \cdot \frac{5}{36}}}\right) \approx 0,03$$

dejaansko: 0,027

visoka relativna nagiba

Problem: nesimetrična vjerojatnost: $\frac{1}{6}$ pada $\frac{5}{6}$

ali pa $\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \dots$

V sloveniji živi 51% žensk. koliko klicov mora opraviti u Mediani, da bo v anketi sodelovalo vsaj pol žensk z verjetnostjo vsaj 0,9?

$X_i \dots$ poklicane je ženska indikator
 $x_i \sim \text{Ber}(p) \quad p = 0,51$

$$X_1 + \dots + X_n = S_n$$

$$P(S_n \geq \frac{n}{2}) \geq 0,9$$

$$\text{"}$$

$$P(S_n \leq \frac{n}{2}) < 0,1 \quad \text{"}^2$$

$$P(S_n \leq \frac{n}{2}) \approx \Phi\left(\frac{\frac{n}{2} - np}{\sqrt{np(1-p)}}\right) < 0,1$$

$$\Phi(z) < 0,1$$

$$z < -1,29$$

$$\sqrt{n}\left(\frac{1}{2} - p\right) < -1,29 \sqrt{p(1-p)}$$

$$\sqrt{n} > \frac{1,29 \sqrt{p(1-p)}}{-(\frac{1}{2} - p)} = \frac{1,29 \sqrt{0,51 \cdot 0,49}}{-(-0,01)} = \frac{1,29 \cdot 0,5}{0,01} = 32,25$$

n mora biti vsaj 33
 mora biti vsaj 33

3. $X_i \sim \text{Gamma}(a, \lambda)$ neod

Izračunaj $P(X_1 + \dots + X_n \leq \frac{cna}{\lambda})$

Momentna rodovna funkcija $M_X(t) = E(e^{tX})$

X, Y neodvisni:

$$M_{X+Y}(t) = E(e^{t(X+Y)}) = E(e^{tX} \cdot e^{tY}) \stackrel{\text{neodvisnost}}{=} E(e^{tX} e^{tY}) = M_X(t) \cdot M_Y(t)$$

$$M_{X_1}(t) = E(e^{tX_1}) = \int_0^{\infty} e^{tx} \frac{\lambda^a x^{a-1}}{\Gamma(a)} e^{-\lambda x} dx =$$

$$\frac{\lambda^a}{\Gamma(a)} \int_0^{\infty} e^{x(t-\lambda)} x^{a-1} dx \dots = \frac{1}{(1 - \frac{t}{\lambda})^a}$$

Momentna rodovna funkcija enolično določa porazdelitev

$$M_{X_1+X_2}(t) = \frac{1}{(1 - \frac{t}{\lambda})^{2a}} = M_{\Gamma(2a, \lambda)}(t)$$

$$\Rightarrow \sim \Gamma(2a, \lambda) \Rightarrow S_n \sim \Gamma(na, \lambda)$$

$$\begin{aligned} P(S_n \leq \frac{cna}{\lambda}) &= \int_0^{\frac{cna}{\lambda}} \frac{\lambda^{na} x^{na-1} e^{-\lambda x}}{\Gamma(na)} dx = \\ &= \int_0^{\frac{cna}{\lambda}} \frac{(\lambda x)^{na-1} \lambda e^{-\lambda x}}{\Gamma(na)} dx = \int_0^{cna} \frac{1}{\Gamma(na)} u^{na-1} e^{-u} du \end{aligned}$$

DN: uporabi centralni limitni izrek

Ocene za napake pri CLT

če imajo X_i tudi tretji moment
velja ocena

$$\sup_{x \in \mathbb{R}} |P(S_n \leq x) - \Phi(\frac{x - nE(X_1)}{\sqrt{n \text{var}(X_1)}})| \leq \frac{0,5583}{(n \text{var}(X_1))^{\frac{3}{2}}} n E(|X_1 - E(X_1)|^3)$$

da se posplošiti za primer
ko $X_1, \dots, X_n$ niso enako porazdeljeni:

① $X_1, X_2, \dots$ neodv. in $X_k \sim \exp(\frac{1}{k})$

Ali v tem primeru velja CLT?

$$S_n = \sum_{i=1}^n X_i \quad E(X_k) = k \quad \text{var}(X_k) = k^2$$

$$\text{var}(S_n) = \sum_{i=1}^n \text{var}(X_i) = \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

↑
neodvisnost

$$E(|X_k - E(X_k)|^3)$$

$$\gamma_1^3 = E(|X_1 - E(X_1)|^3) = \frac{12}{e} - 2 \quad \leftarrow \text{namig}$$

$$X_k \stackrel{d}{=} k X_1$$

$$\mathcal{P}(X_k > x) = e^{-\frac{1}{k}x} \quad \mathcal{P}(kX_1 > x) = \mathcal{P}(X_1 > \frac{x}{k}) = e^{-\frac{1}{k}x}$$

↑
preživetvena
funkcija

$$\Rightarrow E(|X_k - E(X_k)|^3) = E(|kX_1 - kE(X_1)|^3) = k \cdot \gamma_1^3$$

Oglejmo si

$$\frac{\sum_{k=1}^n E(|X_k - E(X_k)|^3)}{\text{var}(S_n)^{\frac{3}{2}}} = \frac{\gamma_1^3 (1 + \dots + n^3)}{\left(\frac{n(n+1)(2n+1)}{6}\right)^{\frac{3}{2}}} = \frac{\gamma_1^3 \frac{n^2(n+1)^2}{4}}{\left(\frac{n(n+1)(2n+1)}{6}\right)^{\frac{3}{2}}}$$

$$= O\left(\frac{n^4}{n^{\frac{3}{2}}}\right) = O\left(\frac{1}{\sqrt{n}}\right) \rightarrow 0$$

$\Rightarrow$ CLT velja

②

$$p \in (0, 1) \quad S \sim \text{Bin}(n, p)$$

$$\text{ocen: } \left| P(S \leq x) - \Phi\left(\frac{x - np}{\sqrt{np(1-p)}}\right) \right| = *$$

Obstajajo $X_1, \dots, X_n \sim \text{Bernuli}(p)$ neodvisne
 $S = \sum X_i$

Poglejmo si

$$\begin{aligned} E(|X_1 - \underbrace{E(X_1)}_p|)^3 &= E(|X_1 - p|^3) = \\ &= |1-p|^3 \cdot p + (1-p) |p|^3 = p(1-p)((1-p)^2 + p^2) \leq p(1-p) \end{aligned}$$

$$* \leq \frac{c \cdot np(1-p)}{(np(1-p))^{\frac{3}{2}}} = \frac{c}{\sqrt{np(1-p)}}$$

③

$X_1, \dots, X_{100}$ neadv. enako porazdeljene

$$X_1 \sim \begin{pmatrix} -1 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

$$S = 10X_1 + X_2 + \dots + X_{100}$$

Ocenj. vrijednost

$$P(S < 41)$$

$\frac{1}{3}$
" "

$\frac{2}{3}$
" "

$$\star = P(S < 41 | X_1 = -1)P(X_1 = -1) + P(S < 41 | X_1 = 0)P(X_1 = 0)$$

$$P(S < 41 | X_1 = -1) = P(\tilde{S} < 51)$$

$$P(S < 41 | X_1 = 1) = P(\tilde{S} < 31)$$

$$\star \approx \frac{1}{3} \Phi\left(\frac{50 - 95 \cdot \frac{1}{3}}{\sqrt{88 \cdot \frac{8}{9}}}\right) + \frac{2}{3} \Phi\left(\frac{30 - 33}{\sqrt{88}}\right) \approx 0,57 \dots$$

Točan rezultat $\approx 0,569$

če bi uprabil ali direktno na S bi
dobili $\approx 0,6$

$$X_1, X_2, \dots \text{ neodv} \quad X_n \sim \begin{pmatrix} -n & 0 & n \\ \frac{1}{2n^2} & 1 - \frac{1}{n^2} & \frac{1}{2n^2} \end{pmatrix}$$

Pokaži, da CLT ne velja

Recimo da velja CLT

$$\text{Potem } \lim_{n \rightarrow \infty} P(X_1 + \dots + X_n)$$

$$E(S_n) = \sum E(X_k) = 0$$

$$\text{var}(X_n) = E(X_n^2) = 1 \Rightarrow \text{var}(S_n) = n$$

$$W_n = \frac{S_n - 0}{\sqrt{n}} \quad P(W_n \in [-\frac{1}{2}, \frac{1}{2}]) = \Phi(\frac{1}{2}) - \Phi(-\frac{1}{2}) = 1 - 2\Phi(-\frac{1}{2}) \approx 0,38$$

$$n \geq 4$$

$$\begin{aligned} P(-\frac{1}{2} \leq W_n \leq \frac{1}{2}) &= P(-\frac{1}{2} \leq \frac{S_n}{\sqrt{n}} \leq \frac{1}{2}) = \\ &= P(\frac{\sqrt{n}}{2} \leq S_n \leq \frac{\sqrt{n}}{2}) \geq P(X_2 = X_3 = \dots = X_n = 0) \end{aligned}$$

⋮

$$\frac{n+1}{2^n} > \frac{1}{2} \quad \text{za } n \geq 4 \quad *$$

Naj bodo $X_1, X_2, \dots$ slučajne spremenljivke z vrednostmi sl. spr. $(X_n)_{n \in \mathbb{N}}$ konvergirajo proti sl. spr. X v paradoksu, če

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x) \quad \text{za } \forall x \in \mathbb{R}, \text{ kjer}$$

je F_X zvezna

$$X_n = \frac{1}{n}$$

Očitno bi želeli, da konvergira proti 0

$$F_{X_n}(x) = \begin{cases} 1 & x \geq \frac{1}{n} \\ 0 & \text{sicer} \end{cases} \quad \text{Oglej}$$

$$F_X(x) = \begin{cases} 1 & x \geq 0 \\ 0 & \text{sicer} \end{cases}$$

Oglejmo si, kakšno je točno neravnost

$$F_X(0) = 1 \quad F_{X_n}(0) = 0 \quad \forall n \in \mathbb{N} \quad *$$

Oglejmo si, da po definiciji, vsa pravda
kjer je uveljavljena

May bo $x \neq 0$

$$F_X(x) = 1 \quad \text{če } x > 0$$

$$F_X(x) = 0 \quad \text{če } x < 0$$

$$F_{X_n}(x) = 1 \quad \text{če } x > \frac{1}{n}$$

$$F_{X_n}(x) = 0 \quad \text{če } x \leq \frac{1}{n}$$



$$X_n \xrightarrow{d} X, \text{ kjer } X \sim \text{Pois}(\lambda)$$

$$\lim F_{X_n}(x) = F_X(x), \text{ kjer je } F_X \text{ zwezna}$$

$$P(X=k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

$$\lim_{n \rightarrow \infty} \frac{P(X_n=k)}{P(X_n=k-1)}$$

$$\frac{P(X_n=k)}{P(X_n=k-1)} = \frac{\binom{n}{k} p_n^k (1-p_n)^{n-k}}{\binom{n}{k-1} p_n^{k-1} (1-p_n)^{n-k+1}} =$$

$$= \frac{n-k+1}{k} \frac{p_n}{1-p_n}$$

$$n p_n \rightarrow \lambda \Rightarrow p_n \rightarrow 0$$

$$\lim k = \frac{n p_n}{k(1-p_n)} + \frac{p_n(1-k)}{k(1-p_n)} \xrightarrow{n \rightarrow \infty} \frac{\lambda}{k}$$

$$P(X_n=k) = \frac{P(X_n=k)}{P(X_n=k-1)} \cdot \frac{P(X_n=k-1)}{P(X_n=k-2)} \cdot \dots \cdot \frac{P(X_n=1)}{P(X_n=0)} P(X_n=0)$$

$$= \prod_{i=0}^k \frac{\lambda}{i} (1-p_n)^n = \lambda (1-p_n)^k \prod_{i=0}^k \frac{1}{i} =$$

$$P(X_n=0) = (1-p_n)^n = (1-p_n)^{\frac{n p_n}{p_n}} \Rightarrow e^{-\lambda}$$

$$= \frac{\lambda e^{-\lambda}}{k!}$$

$$\lim F_{X_n}(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^{\lfloor x \rfloor} P(X_n=k) = \sum \frac{\lambda e^{-\lambda}}{k!} = F_X(x)$$

(M, d) metrični prostor

$$\tilde{M} \subseteq M$$

$\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}$ si spr. z vred. v $\tilde{M}$

1) $x_n \xrightarrow[n \rightarrow \infty]{d} \tilde{x}$, $(\tilde{M}, d)$ zaplen $(M, d) \rightarrow$
vse $\tilde{M}$ zaplen v X

$$X_1, X_2, \dots \quad N \in \mathcal{P} \quad \geq E(X_1) = 0$$

$$\text{var}(X_1) = \sigma^2 \quad \Rightarrow \quad E(X_1^2) = \sigma^2$$

$$\text{var}(X_1^2) = \tau^2$$

a) kein $\frac{1}{\sqrt{n}}$ konvergenz

$$\begin{array}{c} \downarrow \\ S \\ \text{" } E(X^2) \\ \hline \frac{\sum x_i^2 - n\sigma^2}{\sqrt{n}} \end{array}$$

$y_1, y_2, \dots$ NEP $E(y_1^2) < \infty$
 Potem $S = \sum y_i$ $\frac{s - E(y_1)n}{\sqrt{\text{Var}(y_1)n}} \xrightarrow{d} N(0, 1)$

$$Y_i = X_i^2$$

$$\text{Var}(X_i^2) = \tau_2$$

$$\tau \frac{S - n\sigma^2}{\sqrt{\tau^2 n}} \rightarrow \tau \cdot N(0, 1) = N(0, \tau^2)$$

b) prüfen: $\sqrt{n} \bar{X}^2 \xrightarrow{d} 0$

$$\bar{X} = \frac{X_1 + \dots + X_n}{n} =$$

$$\sqrt{n} \bar{X} \cdot \bar{X} \quad \bar{X} \xrightarrow{d} 0 \leftarrow E(x_1)$$

To ist ein ~~stark~~ starker ~~zu~~^{zu} sehr sein
 Part ist ein starkes

$$\vec{x} \cdot (\sqrt{n} \vec{x})$$

$$\sqrt{n} \bar{X} = \frac{\sigma X_1 + \dots + X_n - \sigma \cdot n}{\sqrt{n} \sigma} \xrightarrow{d} N(0, 1)(x)$$

$$\Rightarrow \sqrt{n}X^2 \xrightarrow{d} 0$$

c) $\mu_n^2 = \frac{\sum (x_j - \bar{x})^2}{n}$

konvergenz $\sqrt{n} (M_n^2 - \sigma^2)$

$$= \frac{\sum x_j^2 - 2 \sum x_j \bar{x} + \sum \bar{x}^2}{\sqrt{n}} - n\sigma^2$$

$$\rightarrow N(0, \tau^2)(x) + 0 + 0$$

Naj bodo $X_1, X_2, \dots$ slučaj. sprem. definirane na $(\Omega, \mathcal{F}, P)$. Pravimo da X_n konvergira k slučaj. spr. X

a) v vrjetnosti, če $\forall \varepsilon > 0. \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0$

b) skoraj gotovo če $P(\lim X_n = X) = 1$

$$P(\{\omega \in \Omega \mid \lim X_n(\omega) = X(\omega)\}) = 1$$

c) $E(P(X_n)) \xrightarrow{d} E(P(X))$

$$P \rightarrow d$$

$$d \not\rightarrow P$$

$$X(0, 1)$$

$$X_n = -X$$

$$X_n \xrightarrow{d} X$$

$$\lim P(|X_n - X| > \varepsilon) =$$

$$\lim (P(2|X| > \varepsilon) =$$

$$\lim P(|X| > \frac{\varepsilon}{2}) = P(|X| > \frac{\varepsilon}{2})$$

$$\neq 0 \quad \forall \varepsilon > 0$$

$$P \Rightarrow d$$

$$\forall \varepsilon > 0 \lim P(|X_n - X| > \varepsilon) = 0$$

$E(f(X_n))$ f zweigeteilt, ungerade polynom

$$E(f(X_n)) \quad I_n := \{ |X_n - X| > \varepsilon \}$$

$$E(f(X_n)) = E(f(X_n)(I_n + I_n^c)) =$$

$$= E(f(X_n)I_n) + E(f(X_n)I_n^c)$$

$$E(|f(X_n) - f(X)|) = E(|f(X_n)| 1_{I_n}) + E(|f(X_n)| 1_{I_n^c}) \leq C\varepsilon$$

$$\parallel \int_0^\infty P(|X_n - X| > \varepsilon) d\varepsilon = 0$$

und

$$X_n \xrightarrow{s.o} X \text{ o.e. } \mathcal{P}(\lim_{n \rightarrow \infty} X_n = X) = 1$$

$$\text{oziroma } \mathcal{P}(\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1$$

$$s.g. \Rightarrow \mathcal{P}$$

$$\mathcal{P} \neq s.g.$$

$$s.g. \Rightarrow \mathcal{P}$$

← zraven so lahke omega

$$\{\lim_{n \rightarrow \infty} X_n = X\} = \{\forall \varepsilon > 0. \exists N > 0. n \geq N \Rightarrow |X_n - X| < \varepsilon\}$$

$$\bigcap_{\varepsilon > 0} \bigcup_{N \geq 0} \bigcap_{n \geq N} \{|X_n - X| < \varepsilon\} =$$

$$= \bigcap_{m \in \mathbb{N}} \underbrace{\bigcup_{N \geq 0} \bigcap_{n \geq N} \{|X_n - X| < \frac{1}{m}\}}_{A_m \text{ padajoče v } m}$$

$$= \bigcup_{M \geq 0} \underbrace{\bigcap_{m \in \mathbb{N}} \{|X_m - X| < \frac{1}{m}\}}_{B_M \text{ naraščajoče v } N}$$

Vemo:

$$\mathcal{P}(\bigcap_{m \in \mathbb{N}} A_m) = 1$$

$$\text{Vemo } \lim_{m \rightarrow \infty} \mathcal{P}(A_m) = 1$$

$$\lim_{m \rightarrow \infty} \mathcal{P}(\bigcup_{N \geq 0} B_N) = 1 \quad A_m \supseteq \bigcap_{N \geq 0} A_m \xrightarrow{1=1} \mathcal{P}(\bigcap_{N \geq 0} A_m) \leq \mathcal{P}(A_m) = 1$$

$$\lim_{m \rightarrow \infty} \lim_{N \rightarrow \infty} \mathcal{P}(B_N) = 1$$

$$B_N \subseteq \{|X_N - X| < \frac{1}{N}\}$$

$$\lim_{N \rightarrow \infty} \mathcal{P}(\{|X_N - X| < \frac{1}{N}\}) = 1$$

$P \neq S.g.$

$$x_1(\omega) = 1_{[0,1]}(\omega)$$

$$x_2(\omega) = 1_{[0, \frac{1}{2}]} \quad x_3 = 1_{[\frac{1}{2}, 1]}$$

$$x_4 = 1_{[0, \frac{1}{4}]} \quad x_5 = 1_{[\frac{1}{4}, \frac{1}{2}]} \quad x_6 = 1_{[\frac{1}{2}, \frac{3}{4}]} \quad x_7 = 1_{[\frac{3}{4}, 1]}$$

$$X = 0$$

$$x_n: \text{---} \square \text{---}$$

Naj bo M in M' metrični prostora in
 $g: M \rightarrow M'$ preslikava s lastnostmi: (sternovskega)
 nezveznosti $D_g = \{x \in M \mid g \text{ ni zvezna v } x\}$

Naj bodo $x_1, x_2, \dots, x$ slučaj. spren. z vrednostmi
 v M in naj bo $P(X \in D_g) = 0$.

$$\text{če } X_n \xrightarrow[n \rightarrow \infty]{d.p./s.p.} X \Rightarrow g(X_n) \xrightarrow[n \rightarrow \infty]{d.p./s.p.} g(X)$$

27.3

$$X_n \xrightarrow{P} 0 \quad \sqrt{n} X_n \xrightarrow{d} N(0, \sigma^2)$$

Pričemo da je $P(X_n = 0) = 0$

Po izreku gre $\sin(X_n) \xrightarrow{d} 0$
 $\sin(0)$

$$\sqrt{n} \sin(X_n) = \frac{X_n}{X_n} \sqrt{n} \sin(X_n) \sim 1 \cdot N(0, \sigma^2) = N(0, \sigma^2)$$



gre prot 1 je konstant

SKLEPNA STATISTIKA

Pri verjetnosti pri neznan verjetnostni meri P in znanem dogodku A računamo $P(A)$

Pri statistiki opazimo, da se je zgodil dogodek A , ne vemo pod katero verjetnostno mero P_ϑ ; $\vartheta \in \Theta$.

Radi bi sklepali pri katerem $\vartheta \in \Theta$ se dogodek A zgodil: z največjo verjetnostjo

Zgled:

Kovanci z verjetnostjo grba $\vartheta \in [0, 1]$

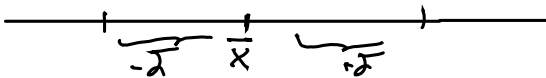
Pri verjetnosti fiksiramo in odgovarjamo na vprašanje, da pri prvih petih metih opazimo 1 grb
 $X \sim \text{Bin}(5, \vartheta)$

Iščemo ϑ

① $x_1, x_2, \dots, x_n$ neadv. enako porazdeljene
 $x_1 \sim N(\mu, \sigma^2)$

Recimo da je σ^2 poznam. Poišoi interval
 zaupanja za stopnjo napake α

$$\bar{x} \xrightarrow[n \rightarrow \infty]{\text{zakonski}} \mu = E(x_1)$$



$$P(\mu \in (\bar{x} - \sigma, \bar{x} + \sigma)) \geq 1 - \alpha$$

$$P(\bar{x} - \sigma < \mu < \bar{x} + \sigma) \geq 1 - \alpha$$

$$P(|\bar{x} - \mu| > \sigma) \leq \alpha$$

$$\bar{x} = \frac{x_1 + \dots + x_n}{n} \sim N(\mu, n\sigma^2)$$

$$\text{var}(\bar{x}) = \frac{n\sigma^2}{n^2}$$

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \sqrt{n} \frac{\bar{x} - \mu}{\sigma} \sim N(0, 1)$$

$$1 - \Phi\left(\frac{\sqrt{n}}{\sigma} d\right) = P\left(Z \geq \frac{\sqrt{n}}{\sigma} d\right) \leq \frac{\alpha}{2}$$

$$d \geq \frac{1.96}{\sqrt{n}}$$

② Kot prej μ σ nista poznane
 Interval zaupanja za σ^2 ?

$$\sigma^2 = E((X - \mu)^2)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 \xrightarrow{n \rightarrow \infty} \sigma^2$$

Sicer $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

izkaže se, da je $\frac{(n-1)\hat{\sigma}^2}{\sigma^2} = Q \sim \chi^2_{n-1}$

Q vzame vrednosti $[0, \infty)$



se odločimo da na doli strani vzamemo $\frac{\alpha}{2}$

$$P(J_- < Q < J_+) \geq 1 - \alpha$$

$$J_+ = F_{\chi^2_{n-1}}^{-1}\left(1 - \frac{\alpha}{2}\right)$$

$$J_- = F_{\chi^2_{n-1}}^{-1}\left(\frac{\alpha}{2}\right)$$

$$P(J_- < Q < J_+) = 1 - \alpha$$

$$P\left(\underbrace{\frac{(n-1)\hat{\sigma}^2}{\sigma} < \sigma < \frac{(n-1)\hat{\sigma}^2}{\sigma}}\right) = 1 - \alpha$$

iz

Definiramo kvantilno funkcijo $\tilde{F}^{-1}$ kot
 $\tilde{F}^{-1}: (0,1) \rightarrow \mathbb{R}$
 $\alpha \mapsto \min \{x \in \mathbb{R} \mid F(x) \geq \alpha\}$

Pokaži da je $\tilde{F}^{-1}(\alpha)$ kvantil reda α

$$y = \tilde{F}^{-1}(\alpha) \quad F(y) \geq \alpha$$

Vzemimo $x < y \Rightarrow F(x) < \alpha$

$$\lim_{x \rightarrow y} F(x) < \alpha \Rightarrow F(y) \leq \alpha$$

Trditev: Naj bo $U \sim U(0, 1)$. Definiramo
 $X = \tilde{F}^{-1}(U)$ potem je X porazdeljen z F

lema: $\tilde{F}^{-1}(\alpha) \leq x \Leftrightarrow \alpha \leq F(x)$

$\Leftarrow$) Naj bo $\alpha \leq F(x)$

$$\tilde{F}^{-1}(\alpha) = \min \{ y \in \mathbb{R} : F(y) \geq \alpha \}$$

$$\Rightarrow F(\tilde{F}^{-1}(\alpha)) \leq F(x) \leadsto \tilde{F}^{-1}(\alpha) \leq x$$

ker je F naraščajoča

$(\Rightarrow)$

$$\tilde{F}^{-1}(\alpha) \leq x$$

$$\forall y : F(y) \geq \alpha \Rightarrow F(x) \geq \alpha$$

$$\alpha \leq F(y) \leq F(x)$$

Dalje zbiraj

127

$$X \sim F$$

$$P(X \leq x) = F(x)$$

$$P(\tilde{F}^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$$

Točkasto ocenjevanje

Imamo vzorec X (vektor) in s porazdelitvijo povezane količine ϑ
količine ϑ ocenimo s cenilko $\hat{\vartheta} = f(X)$

$X_1, \dots, X_n$ i.i.i.d.

$X_i \sim U(0, a)$ $a > 0$

oceniti a

$$\text{Vemo: } E(X_i) = \frac{a}{2} = \mu$$

$$\bar{x} = \hat{\mu}$$

$$\hat{a} = 2\bar{x}$$

$$MSE(\hat{a}) = ? \quad MSE = \text{Var} + \text{Bias}^2$$

$$\text{Bias}(\hat{a}) = E(2\bar{x}) - a = 2 \cdot \frac{a}{2} - a = 0$$

$$\text{Var}(\hat{a}) = \text{Var}\left(2 \sum_{i=1}^n X_i\right) = \sum \text{Var}(X_i) = 4 \cdot \frac{a^2}{12} \cdot n = \frac{a^2 n}{3n^2}$$

$$MSE = \frac{a^2}{3n} \xrightarrow{n \rightarrow \infty} 0 \Rightarrow \hat{a} \text{ je dosledna}$$

X

17. u.

$X_i \sim N(\mu, \sigma^2)$ P_i : ketegori c bo
 cenilke najboljša

$$MSE(c\hat{\sigma}^2) = c^2 \left(\frac{(n-1)^2}{n^3} \sigma^4 + \frac{(n-2n+3)(n-1)}{n^3} \sigma^4 \right) -$$

$$- 2c\sigma^4 \frac{n-1}{n} + \sigma^4$$

$$= \left(c^2 \frac{n^2-1}{n^2} - c \cdot \frac{2(n-1)}{n} + 1 \right) \sigma^4$$

$$\frac{\partial}{\partial c} (\dots) = \left(c \frac{2(n^2-1)}{n^2} - \frac{2(n-1)}{n} \right) \sigma^4 = 0$$

$$\leadsto c = \frac{n^2}{2(n^2-1)} \cdot \frac{2(n-1)}{n} = \frac{n}{n+1}$$

X, Y slučajne vektore ~~de~~ dimenzije

Poznamo $\text{var}(X) = \Sigma_{11}$ $\text{var}(Y) = \Sigma_{22}$

$$\text{cov}(X, Y) = \Sigma_{12}$$

$$\text{var}(X+Y) = \text{cov}(X+Y, X+Y) =$$

$$= \text{cov}\left(\begin{bmatrix} I_n & I_n \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}, \begin{bmatrix} I_n & I_n \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}\right) =$$

$$\begin{bmatrix} I_n & I_n \end{bmatrix} \text{cov}\left(\begin{bmatrix} X \\ Y \end{bmatrix}, \begin{bmatrix} X \\ Y \end{bmatrix}\right) \begin{bmatrix} I_n & I_n \end{bmatrix}^T =$$

$$\text{cov}(X)$$

...
vektor

$X \sim QX$ $\Leftrightarrow \forall Q$ orthogonale

$$Q^T Q = I$$

$$\text{var}(X) = \begin{bmatrix} \text{cov}(x_1, x_1), \dots \\ \vdots \end{bmatrix}$$

$$\text{var}(X) = \text{var}(QX) = Q \text{var}(X) Q^T$$

$$\Rightarrow Q \text{var}(X) = \text{var}(X) Q$$

~~#~~
 A symmetrisch $\Rightarrow \exists Q, D$ $A = Q D Q^T$ D diag

$$\Rightarrow \text{var}(X) \text{ diagonal}$$

$$E(X) = 1 \cdot a \quad \text{var}(X) = I_n + b 1_n \cdot 1_n^T$$

$$E(X^T (I_n + \frac{1}{n} 1_n \cdot 1_n^T) X) \neq$$

$$Y = X^T (I_n + \frac{1}{n} 1_n \cdot 1_n^T) X$$

$$E(Y) = E(\text{sl}(Y)) \text{ , ker je } Y \text{ skalar}$$

$$\text{sl}(AB) = \text{sl}(BA)$$

$$\Rightarrow \text{sl}(Y) = \text{sl}((I_n + \frac{1}{n} 1_n \cdot 1_n^T) X X^T)$$

$$E(\text{sl}(Y)) = \text{sl}(I_n + \frac{1}{n} 1_n \cdot 1_n^T \underbrace{E(X X^T)}_{\text{var}(X) + E(X)E(X)^T}) =$$

$$\Rightarrow E(Y) = \text{sl}((I_n + \frac{1}{n} 1_n \cdot 1_n^T)(I_n + b \cdot 1_n \cdot 1_n^T + a^2 1_n \cdot 1_n^T)) =$$

$$= \text{sl}(I_n + 2(\frac{1}{n} + b + a^2) 1_n \cdot 1_n^T + \frac{1}{n}(b + a^2) n 1_n^T) =$$

$$= 1 + 2(\frac{1}{n} + b + a^2) + b + a^2 = \underline{\underline{1 + \frac{2}{n} + 3b + 3a^2}}$$

Membela sem
entrat grej.

$X_1, \dots, X_n$ MEP

$$f_X(x, \vartheta, \lambda) = \lambda e^{-\lambda(x-\vartheta)} \cdot \mathbb{1}_{x \geq \vartheta}$$

a) Določi cenilke po metodi momentov ϑ, λ

$$E(X) = \int_{\mathcal{R}_+} x f_X(x, \vartheta, \lambda) dx$$

Opazimo $X \stackrel{d}{=} Y + \vartheta$ $Y \sim \text{Exp}(\lambda)$

$$\Rightarrow E(X) = \frac{1}{\lambda} + \vartheta$$

$$\text{var}(X) = \frac{1}{\lambda^2}$$

$$E(X^2) = \text{var}(X) + (E(X))^2 = \frac{2}{\lambda^2} + 2\frac{\vartheta}{\lambda} + \vartheta^2$$

$$\vartheta = E(X) - \frac{1}{\lambda}$$

$$E(X^2) = \frac{1}{\lambda^2} + E(X^2)$$

$$\leadsto \lambda = \frac{1}{\sqrt{E(X^2) - E^2(X)}}$$

$$\vartheta = E(X) - \sqrt{E(X^2) - E^2(X)}$$

$$\lambda = \frac{1}{\sqrt{\overline{X^2} - \bar{X}^2}} \quad \vartheta = \bar{X} - \sqrt{\overline{X^2} - \bar{X}^2}$$

b) Cenilke po metodi največjega vjerjetja

$$L(\lambda, \vartheta | x_1, \dots, x_n) = \prod_{i=1}^n f_X(x_i)$$

$$= \lambda^n e^{-\lambda(\sum x_i - n\vartheta)} \mathbb{1}_{x_1 \geq \vartheta} \dots \mathbb{1}_{x_n \geq \vartheta}$$

$$\ell = \ln L = \ln \lambda^n - \lambda(\sum x_i - n\vartheta) + \sum_{i=1}^n \ln(\mathbb{1}_{x_i \geq \vartheta})$$

$$\frac{\partial}{\partial \lambda} \ell = n \cdot \frac{1}{\lambda} - \sum x_i + n\vartheta + \sum_{i=1}^n \ln(\mathbb{1}_{x_i \geq \vartheta})$$

L hočemo maksimizirati $\Rightarrow$ hočemo ℓ maksimizirati

$$\vartheta: x_i \geq \vartheta \quad \forall i$$

zaradi člene $+n\vartheta$ želimo ϑ čim večjo

$$\Rightarrow \hat{\vartheta} = \min(x_i) = x_{(1)}$$

$$\frac{n}{\lambda} = \sum x_i - n \underbrace{x_{(1)}}_{\text{oznaka da je največji}}$$

$$\hat{\lambda} = \frac{n}{\sum x_i - n x_{(1)}} = \frac{1}{(\bar{x} - x_{(1)})}$$

$x_1, \dots, x_n$ NET

$$f_X(x, \vartheta) = e^{\vartheta - x} \mathbb{1}_{x > \vartheta} \quad ; \quad \vartheta > 0 \text{ nepoznat}$$

Dalo bi ϑ MNV

$$L = \prod_{i=1}^n f_X(x_i, \vartheta) = e^{-\sum x_i + n\vartheta} \mathbb{1}_{x > \vartheta} = L(\vartheta | x_1, \dots, x_n)$$

$$\ln L = \ell = n\vartheta - \sum x_i + \ln \mathbb{1}_{x > \vartheta}$$

maksimiziramo: $n\vartheta + \ln \mathbb{1}_{x > \vartheta}$

Predpostavljamo, da je ϑ lahko vzamemo
Rabimo $\min(x_i) - \epsilon$

Ample naenkrat ϵ ni najboljši
 $\Rightarrow$ cenilke po MNV ne dostajajo

V praksi je isto če vzamemo $\leq$ tvoj
je $\min(x_i)$ najboljša cenilka

Fisherjeva informacija

Pri modelih $X = (X_1, \dots, X_n)$ in z neznanim parametrom $\vartheta \in \mathbb{R}$, je dana z

$$F_1(\vartheta) = E \left(\left(\frac{\partial \ell(\vartheta | X)}{\partial \vartheta} \right)^2 \right) = - E \left(\frac{\partial^2 \ell(\vartheta | X)}{\partial \vartheta^2} \right)$$

Velja Rao-Cramerjeva ocena

$$SE^2 = \text{var}(\hat{\vartheta}) \geq \frac{1}{F_1(\vartheta)}$$

za nepristransko cenilko $\hat{\vartheta}$

Če je model dovolj regularen za cenilko $\hat{\vartheta}$

po MMV velja $\hat{\vartheta} \rightarrow \vartheta \quad E(\hat{\vartheta}) \rightarrow \vartheta$

$$SE^2 \sim E((\hat{\vartheta} - \vartheta)^2) \sim \frac{1}{F_1(\vartheta)}$$

Če so $X_1, \dots, X_n$ NEP je $F_1(\vartheta) = n F_{1,1}(\vartheta)$

$$F_{1,1}(\vartheta) = E \left(\frac{\partial \ell_1(\vartheta | X_1)}{\partial \vartheta} \right)^2$$

$$X_1, \dots, X_n \sim \text{Exp}\left(\frac{1}{\alpha}\right) \quad N \in \mathcal{P}$$

Določimo Fisherjevo transformacijo za α

$$f_X(x) = \frac{1}{\alpha} e^{-\frac{1}{\alpha}x}$$

~~$$l = -n \frac{1}{\alpha} + \sum \frac{1}{\alpha} x_i$$~~

$$l_1 = \ln \frac{1}{\alpha} - \frac{1}{\alpha} x_1 = -\ln \alpha - \alpha^{-1} x_1$$

$$\frac{\partial l_1}{\partial \alpha} = -\frac{1}{\alpha} + \frac{1}{\alpha^2} x_1$$

$$\frac{\partial^2 l_1}{\partial \alpha^2} = \frac{1}{\alpha^2} - 2 \frac{1}{\alpha^3} x_1$$

$$F_1(\alpha) = n E \left(2 \frac{1}{\alpha^3} x_1 - \frac{1}{\alpha^2} \right) =$$

$$= 2n \frac{1}{\alpha^3} \underbrace{E(X_1)}_{\alpha} - \frac{n}{\alpha^2} = \frac{n}{\alpha^2}$$

od zadnjice vemo: $\hat{\alpha} = \bar{x}$ sledi:

$$SE^2 \geq \frac{1}{F_1(\alpha)} = \frac{\alpha^2}{n} \quad \square$$

Ocena F. inf. je $F(\hat{\alpha})$.

Nadaljevanje yapa od prej:
Če je $\theta \in \mathbb{R}^d$ in imamo

karakteristiko $z = h(\theta)$ po MNV in ker

$\hat{z} = h(\hat{\theta})$ za θ po MNV.

$$\text{Velja: } F(\theta) = \left[E \left(\frac{\partial \ell(\theta|x)}{\partial \theta_i} \frac{\partial \ell(\theta|x)}{\partial \theta_j} \right) \right]_{i,j=1}^d = \\ = - \left[E \left(\frac{\partial^2 \ell(\theta|x)}{\partial \theta_i \partial \theta_j} \right) \right]_{i,j=1}^d$$

Če je $\hat{z}$ nepristranske je

$$SE^2 = \text{var}(\hat{z}) \geq \nabla h(\theta)^T F^{-1}(\theta) \nabla h(\theta)$$

Predstela isto kot prej.

Od zadnjih $X_1, \dots, X_n \sim P$

$$X_1 \sim \begin{pmatrix} -1 & 0 & 1 & 2 \\ 1-2a-b & a & b & a \end{pmatrix}$$

$$\hat{a} = \frac{1}{4}$$

$$\hat{b} = \frac{3}{10}$$

ocenjujemo $a^2 \cdot b = h(a, b) = \xi$

a) oceni ξ po MNV

b) poišči: asimptotič $SE(\xi)$

a) $\hat{\xi} = h(\hat{a}, \hat{b}) = \frac{3}{160}$

b) $\nabla h(a, b)^T FI^{-1}(a, b) \nabla h(a, b)$

$$\nabla h(a, b)^T = \begin{bmatrix} 2ab \\ a^2 \end{bmatrix}$$

$$FI(a, b) = ?$$

$$\frac{\partial l_1}{\partial a} = \begin{pmatrix} \frac{-2}{1-2a-b} & \frac{1}{a} & 0 \\ 0 & 0 & 1 \\ \frac{1}{a} & 0 & 0 \end{pmatrix}$$

$$\frac{\partial l_1}{\partial b} = \begin{pmatrix} \frac{-1}{1-2a-b} & -1 \\ 0 & 0 \\ \frac{1}{b} & 1 \\ 0 & 2 \end{pmatrix}$$

$$E \left(\frac{\partial l_1(a, b | X)}{\partial a} \right)^2 = \sum_{i=1}^2 (\dots)^2 P(X=i)$$

$$= \frac{4}{(1-2a-b)^2} (1-2a-b) + 2 \frac{1}{a^2} \cdot a + 0 =$$

$$= \frac{4}{1-2a-b} + \frac{2}{a^2} = \frac{2-2b}{a(1-2a-b)}$$

$$E \left(\frac{\partial l_1}{\partial b} \right) = \frac{2}{1-2a-b}$$

$$E \left(\left(\frac{\partial l_1}{\partial b} \right)^2 \right) = \frac{1-2a}{b(1-2a-b)}$$

$$FI_1(a, b) =$$

$$FI_1(a, b) = \frac{1}{1-2a-b} \begin{bmatrix} \frac{2-2b}{a} & 2 \\ 2 & \frac{1-2a}{b} \end{bmatrix}$$

$$\Rightarrow FI_1^{-1} = \frac{1}{2} \begin{pmatrix} a(1-2a) & -2ab \\ -2ab & b(2-b) \end{pmatrix}$$

$$FI^{-1} = \frac{1}{n} FI_1^{-1}$$

b) $\nabla h^T FI^{-1} \nabla h = \left(\frac{a^3 b}{n} (a - 5ab + 2b) \right)^{\frac{1}{2}}$

Standardna p -razsežna normalna porazdelitev je porazdelitev vektora
 $z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$ kjer so z_i N(0,1).

Splošna n -razsežna normalna porazdelitev je
vsaka porazdelitev slučajnega vektora
 $Az + u$

Naledge

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim N(0, \begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix})$$

Doaci $P(x > 0, y > 0)$

$$\exists A \quad \Sigma = AA^T$$

$$AZ \sim N(0, I)$$

$$a_{ij} = \sqrt{\Sigma_{ij} - \sum_{k=1}^{j-1} a_{jk}^2}$$

$$i \geq j: \frac{\Sigma_{ij} - \sum_{k=1}^{j-1} a_{ik} a_{jk}}{a_{jj}}$$

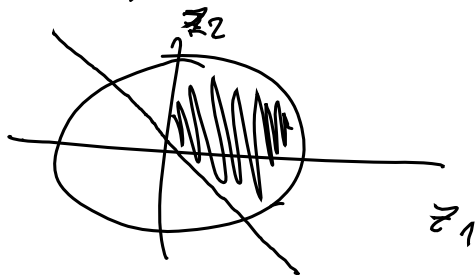
$$\begin{bmatrix} \sqrt{8} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{5 - \frac{1}{2}} \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{3\sqrt{2}}{2} \end{bmatrix} = A^T$$

$$V = AZ = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$P(x > 0, y > 0)$$

$$P(2\sqrt{2}z_1 > 0, \frac{\sqrt{2}}{2}z_1 + \frac{3\sqrt{2}}{2}z_2 > 0) =$$

$$= P(z_1 > 0, z_2 > -\frac{1}{3}z_1)$$



$$P = \text{area} = \frac{\frac{\pi}{2} + \frac{1}{3} \arctan \frac{1}{3}}{\pi} =$$

$$= \frac{1}{2} + \frac{1}{3\pi} \arctan \frac{1}{3}$$

Naj bo $Z \sim N(\vec{0}, I_n)$. Potem je $\|Z\| \sim \chi^2(n) = T(\frac{n}{2}, \frac{1}{2})$
 studentova porazdelitev.

$$T = \frac{\bar{Z}}{\sqrt{\frac{1}{p} K}} \quad Z \sim N(0, 1) \quad Z, K \text{ neodvisni}$$

$$K \sim T(\frac{p}{2}, \frac{1}{2}) \approx \chi^2(p)$$

$$T \sim \text{student}(p)$$

$$\frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0, 1)$$

$$\hat{\sigma}_+^2 = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\frac{\bar{X} - \mu}{\hat{\sigma}_+} \sqrt{n} = \frac{\frac{\bar{X} - \mu}{\sigma} \sqrt{n}}{\sqrt{\frac{\hat{\sigma}_+^2}{\sigma^2}}} \sim N(0, 1)$$

$$\frac{\hat{\sigma}_+}{\sigma} = \sqrt{\frac{1}{\sigma^2} \frac{1}{n-1} \sum (X_i - \bar{X})^2} = \|HX\|^2 \sim \chi^2(n-1)$$

Neodvisnost?

Povzgi je pokazati, da sta neodvisna $\bar{X}$ in HX

Ker sta $\bar{X}$ in HX normalne slučajne vektore,
 neodvisnost sledi iz nekoreliranosti

$$\begin{aligned} \text{Cov}(\bar{X}, HX) &= \text{Cov}\left(\frac{1}{n} \mathbf{1}_n^T X, HX\right) = \\ &= \frac{1}{n} \mathbf{1}_n^T \text{Cov}(X, X) H^T = \frac{1}{n} \mathbf{1}_n^T \sigma^2 H = \end{aligned}$$

$$= \frac{1}{n} \mathbf{1}_n^T \sigma^2 \left(I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) = 0$$

Y relage

Naj bodo $X_1, \dots, X_n \text{ i.i.d. } N(\mu, \sigma^2)$

Skonstruiraj interval zaupanja za stopnjo $1-\alpha$ za:

a) μ , če poznamo σ ?

$$\hat{\mu} = \bar{X}$$

$$C_\alpha = ?$$

$$\mathbb{P}(\bar{X} - C_\alpha < \mu < \bar{X} + C_\alpha) \geq 1 - \alpha$$

$$\mathbb{P}(\mu - C_\alpha \leq \bar{X} \leq \mu + C_\alpha)$$

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

$$Y = \sqrt{n} \frac{\bar{X} - \mu}{\sigma} \sim N(0, 1)$$

$$\Phi(-x) \approx 1 - \Phi(x)$$

$$= \mathbb{P}\left(-\frac{\sqrt{n}}{\sigma} C_\alpha \leq Y \leq \frac{\sqrt{n}}{\sigma} C_\alpha\right) =$$

$$= \Phi\left(\frac{\sqrt{n}}{\sigma} C_\alpha\right) - \Phi\left(-\frac{\sqrt{n}}{\sigma} C_\alpha\right) =$$

$$= 2\Phi\left(\frac{\sqrt{n}}{\sigma} C_\alpha\right) - 1 \geq 1 - \alpha$$

$$\Rightarrow \Phi\left(\frac{\sqrt{n}}{\sigma} C_\alpha\right) \geq 1 - \frac{\alpha}{2}$$

$$C_\alpha \geq \frac{\sigma}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

b) stopnja zaupanja za μ , če α ni poznan

$$\hat{\sigma}_+^2 = \sqrt{\frac{1}{n} \sum (X_i - \bar{X})^2}$$

$$\mathbb{P}(\bar{X} - c_\alpha < \mu < \bar{X} + c_\alpha) \leq 1 - \alpha$$

$$T = \frac{\bar{X} - \mu}{\hat{\sigma}_+} \sqrt{n} \sim \text{Student}(n-1)$$

$$\mathbb{P}(\mu - c_\alpha < \bar{X} < \mu + c_\alpha) =$$

$$= \mathbb{P}\left(-\frac{c_\alpha}{\hat{\sigma}_+} \sqrt{n} < T < \frac{c_\alpha}{\hat{\sigma}_+} \sqrt{n}\right) = 1 - \alpha$$

$$F\left(+\frac{c_\alpha}{\hat{\sigma}_+} \sqrt{n}\right) - F\left(-\frac{c_\alpha}{\hat{\sigma}_+} \sqrt{n}\right) = 1 - \alpha$$

$$2F\left(c_\alpha \frac{\sqrt{n}}{\hat{\sigma}_+}\right) - 1 = 1 - \alpha$$

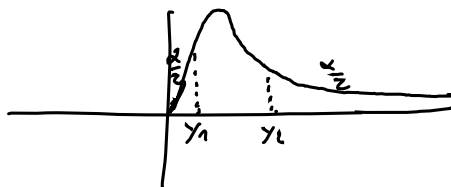
$$\Rightarrow c_\alpha = \frac{\hat{\sigma}_+}{\sqrt{n}} F^{-1}\left(1 - \frac{\alpha}{2}\right)$$

c) za σ , če μ ni poznan

~~$$[\sigma_+ - c_\alpha, \sigma_+ + c_\alpha]$$~~

$$\frac{\hat{\sigma}_+^2}{\sigma^2} = \frac{1}{\sigma^2} \frac{1}{n-1} \sum (x_i - \bar{x})^2 \sim \chi^2(n-1)$$

$$\mathcal{P} \left(y_1 \leq \frac{\hat{\sigma}_+^2}{\sigma^2} \leq y_2 \right) = 1 - \alpha$$



$$\mathcal{P} \left(F^{-1}(\alpha/2) \leq \frac{\hat{\sigma}_+^2}{\sigma^2} \leq F^{-1}(1 - \alpha/2) \right) =$$

$$= \mathcal{P} \left(\frac{1}{F(1 - \alpha/2)} \hat{\sigma}_+^2 \leq \sigma^2 \leq \frac{\hat{\sigma}_+^2}{F^{-1}(\alpha/2)} \right)$$

interval zaupanja: $\left[\sqrt{\frac{\hat{\sigma}_+^2}{F(1 - \alpha/2)}}, \sqrt{\frac{\hat{\sigma}_+^2}{F^{-1}(\alpha/2)}} \right]$

ϑ ... vrijednost razpoka
 β ... stopnja zaupanja

Konfidenčni interval zaupanja za ϑ glede na n poskusov

Če se poskus ponese, vzamemo $[0, 1]$,
sicer vzamemo $[0, \vartheta^*]$

Iščemo ϑ^* najmanjši možni

$$P(\vartheta \in [0, \vartheta^*]) \geq \beta$$

$$\underline{P(\vartheta > \vartheta^*)} \leq 1 - \beta = P(\vartheta \text{ vsej en se ponese}) \geq \beta$$

$$1 - P(\text{noben se ne ponese}) = 1 - (1 - \vartheta)^n \geq \beta$$

$$(1 - \vartheta)^n \leq 1 - \beta$$

$$1 - \vartheta \leq \sqrt[n]{1 - \beta}$$

$$\vartheta \geq 1 - \sqrt[n]{1 - \beta}$$

$$\vartheta^* = 1 - \sqrt[n]{1 - \beta}$$

Naj bo $\hat{\theta}$ cenilka po MNU po devoli
lepem modelu

Pri devolji velja, da je $\hat{\theta}$ porazdeljena približno
 $N(0, \frac{1}{FI(\theta)})$

Ta nam da asimptotični interval zaupanja $1-\alpha$ dolike
$$\left[\hat{\theta} - \frac{F^{-1}\left(1 - \frac{\alpha}{2}\right)}{\sqrt{FI(\hat{\theta})}}, \hat{\theta} + \dots \right]$$

$$X_1, \dots, X_n \quad X_i \sim N(\mu, \sigma^2) \\ 12 \quad \& \quad \sigma$$

$$\left[\hat{\sigma}_+^2 \sqrt{\frac{n-1}{F_{\chi^2(n-1)}^{-1}(1-\frac{\alpha}{2})}}, \hat{\sigma}_+^2 \sqrt{\frac{n-1}{F_{\chi^2(n-1)}^{-1}(\frac{\alpha}{2})}} \right]$$

$$\hat{\sigma}_+^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

asimptotiki?

$$L(\mu, \sigma) \\ L_1(\mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$l(\mu, \sigma) = \sum_{i=1}^n \ln\left(\frac{1}{\sigma \sqrt{2\pi}}\right) - \frac{(x-\mu)^2}{2\sigma^2}$$

$$\text{where se } \hat{\mu} = \bar{x} \\ \hat{\sigma}^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$$

$$\frac{\partial^2}{\partial \mu^2} l_1 = -\frac{1}{\sigma^2}$$

$$\frac{\partial^2}{\partial \mu \partial \sigma} l_1 = -\frac{2(x_1 - \mu)}{\sigma^3}$$

$$\frac{\partial^2}{\partial \sigma^2} l_1 = \frac{1}{\sigma^2} - \frac{3(x_1 - \mu)^2}{\sigma^4}$$

$$F_1(\mu, \sigma) = E \begin{bmatrix} -\frac{1}{\sigma^2} & -\frac{2(x_1 - \mu)}{\sigma^3} \\ -\frac{2(x_1 - \mu)}{\sigma^3} & \frac{1}{\sigma^2} - \frac{3(x_1 - \mu)^2}{\sigma^4} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & -\frac{1}{\sigma^2} + \frac{3}{\sigma^2} \end{bmatrix}$$

$$\sigma^2 = h(\theta) = \rho r_2(\theta) \quad \theta = \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}$$

$$\nabla h^T F^{-1} \nabla h$$

$$\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \ddots & \\ & \ddots \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\sigma^2}{2n} \sim \text{var}(\hat{\sigma}^2)$$

$$\hat{\sigma}^2 = \frac{F_{N(\mu, \sigma^2)}^{-1}(1-\frac{\alpha}{2})}{\sqrt{\frac{2n}{\lambda \hat{\sigma}}}}, \quad \hat{\sigma}^2 = \frac{F^{-1}(1-\frac{\alpha}{2})}{\sqrt{\frac{\sigma^2-1}{2n}}}$$

Domneva H_0 in H_1

če je model parametriziran s $\vartheta \in \Theta$
lahko H_0 določimo z $\vartheta \in \Theta_0$.

Alternativno domeno lahko določimo
kar z $\Theta_1 = \Theta - \Theta_0$

Pri stopnji tveganja α skonstruiramo
preizkus, da velja

$$P_0(H_0 \text{ zavrnemo} | H_0 \text{ drži}) \leq \alpha$$

Moč preizkusa

$$P(H_0 \text{ zavrnemo} | H_1 \text{ drži})$$

Odbijemo o zavrnitvi H_0 sprejmemo, če
preizkusna statistika pade v kritično domeno K_α

	H_0 drži	H_1 drži
H_0 zavrnemo	napaka tipa 1 $P \leq \alpha$	✓ $P(\dots) = 1 - \beta$ ← Moč testa
H_1 ne zavrnemo	✓	napaka tipa 2 $P(\dots) = \beta$

Opazimo $X_1, \dots, X_n \text{ i.i.d. } X_i \sim N(\mu, \sigma^2)$

Vemo $\sigma = 5$

Preizkusimo $H_0: \mu = 100$

a) $H_1: \mu < 100$

Preizkus stopnje tveganja α

$$\bar{X} = \frac{1}{n} (X_1 + \dots + X_n)$$

Zavrtnemo če: $\bar{X} < k_\alpha$

$$P(\bar{X} < k_\alpha \mid \mu < 100) =$$

$P_0(\bar{X} < k_\alpha) \leq \alpha$
 ← išemo najmanjši možni k_α , da je moč testa največja

$$X_1 \sim N(100, 5^2)$$

$$\bar{X} \sim N(100, \frac{5^2}{n})$$

$$\text{var}(\frac{1}{n} \bar{X}) = \frac{1}{n^2} \sum \text{var}(X_i) = \frac{1}{n} \text{var}(X_1)$$

$$P(\frac{\bar{X} - 100}{5} \sqrt{n} < \frac{k_\alpha - 100}{5} \sqrt{n}) \leq \alpha$$

$$\Phi(\frac{k_\alpha - 100}{5} \sqrt{n}) \leq \alpha$$

$$k_\alpha = \frac{5}{\sqrt{n}} \Phi^{-1}(\alpha) + 100$$

b) $H_1: \mu \neq 100$

$$\begin{array}{c} \text{-----} \\ \mu_0 - c \quad \mu_0 \quad \mu_0 + c \end{array}$$

$$P_0(\bar{X} - c \leq \mu \leq \bar{X} + c) = P(100 - c \leq \bar{X} \leq 100 + c) =$$

$$= P(-\frac{c\sqrt{n}}{5} \leq \frac{\bar{X} - 100}{5} \sqrt{n} \leq \frac{c\sqrt{n}}{5}) =$$

$$1 - (\Phi(\frac{c\sqrt{n}}{5}) - \Phi(-\frac{c\sqrt{n}}{5})) = \alpha$$

$$1 - \alpha = 1 - 2\Phi(-\frac{c\sqrt{n}}{5})$$

$$\frac{c\sqrt{n}}{5} = \Phi^{-1}(1 - \frac{\alpha}{2})$$

$$c = \frac{5}{\sqrt{n}} \Phi^{-1}(1 - \frac{\alpha}{2})$$

$$X_1 \dots X_5 \text{ i.i.d. } \sim N(1, \sigma^2)$$

Z razporeda naj oziroma
 iz vzorca izračunamo varianco σ^2 po
 slednji formuli

$$\hat{\sigma}^2 = \frac{1}{5} \sum (X_i - 1)^2$$

$$T = \frac{\hat{\sigma}^2 \cdot 5}{\sigma^2} \sim \chi^2(5) \quad \text{ker je } \text{var}(X_i) = \sigma^2$$

Investiramo, če $T < c$

$P_{\sigma^2}(T < c)$ je funkcija
 od σ^2

↑
 ker je pri $\sigma^2 = \sigma_0^2$
 odločitev najbolj
 korektna

$\Rightarrow 1 - P_{\sigma_0^2}(T < c)$ verjetnost, da σ^2 torej
 je večji, kot bi bilo odločeno pri σ_0^2

$$P_{\sigma_0^2}(T_{\sigma_0^2} < c) = \alpha$$

$$c = F_{\chi^2(5)}^{-1}(\alpha)$$

$$\text{Investiramo, če } \frac{\hat{\sigma}^2 \cdot 5}{\sigma_0^2} < F_{\chi^2(5)}^{-1}(\alpha)$$

Preizkus na podlagi razmerja
verjetij

$$\text{Imejmo: } H_0 \text{ s } \Theta_0 \\ H_1 \text{ s } \Theta_1 = \Theta - \Theta_0$$

Preizkus na podlagi raz. verjetij temelji na
statistiki

$$\lambda = \frac{\sup_{\theta \in \Theta} L(\theta|x)}{\sup_{\theta \in \Theta} L(\theta|x)}$$

oz

$$\lambda = \frac{\sup_{\theta \in \Theta_1} L(\theta|x)}{\sup_{\theta \in \Theta_0} L(\theta|x)}$$

Wilksov izrek

$$2 \ln \lambda \xrightarrow[n \rightarrow \infty]{d} \chi^2(\dim \Theta - \dim(\Theta_0))$$

$X \dots$ št. licen v kpru (dnevno)
 $Y \dots$ št. licen v cepe

$$\bar{x} = 20 \quad \bar{y} = 22 \quad \text{št. licen} \sim \text{Po}(\lambda_x) \\ \sim \text{Po}(\lambda_y)$$

$$H_0: \lambda_x = \lambda_y$$

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$L(\lambda_x, \lambda_y | x, y) = \prod_{i=1}^n L(\lambda_{x_i}, \lambda_{y_i} | x_i, y_i)$$

$$\Theta = [0, \infty)^2$$

$$\Theta_0 = [0, \infty)$$

$$l(\lambda_x, \lambda_y | x, y) = -n\lambda_x - n\lambda_y + \sum x_i \ln(\lambda_x) - \ln(x_i!) + y_i \ln(\lambda_y) - \ln(y_i!)$$

$$a) \quad l = -2n\lambda + \sum (x_i + y_i) \ln \lambda$$

$$\frac{\partial l}{\partial \lambda} = 2n + \sum \frac{x_i + y_i}{\lambda} = 0$$

$$\Rightarrow \lambda = \frac{\sum x_i + y_i}{2n} = \frac{1}{2} (\bar{x} + \bar{y})$$

$$b) \quad \frac{\partial l}{\partial \lambda_x}, \frac{\partial l}{\partial \lambda_y}, \dots, \lambda_x = \bar{x}, \lambda_y = \bar{y}$$

$$\Lambda = \frac{\prod \frac{\bar{x}^{x_i}}{x_i!} e^{-\bar{x}} \frac{\bar{y}^{y_i}}{y_i!} e^{-\bar{y}}}{\prod \frac{(\frac{1}{2}(\bar{x} + \bar{y}))^{x_i}}{x_i!} e^{-\frac{1}{2}(\bar{x} + \bar{y})}} =$$

$$= \prod_{i=1}^n \left(2 \frac{\bar{x} \bar{y}}{\bar{x} + \bar{y}} \right)^{x_i} = \left(2 \frac{\bar{x} \bar{y}}{\bar{x} + \bar{y}} \right)^{n\bar{x} + n\bar{y}}$$

Uporabimo Wilksov izsek

$$2 \ln(\Lambda) \xrightarrow[n \rightarrow \infty]{} \chi^2(1)$$

$$P(2n(\bar{x} + \bar{y}) \ln\left(\frac{\bar{x} \bar{y}}{\bar{x} + \bar{y}}\right) > d) = \alpha$$

$$F_{\chi^2(1)}^{-1}(1 - \alpha) = d$$

$$\text{zapravi se zamenja za } 2 \ln \Lambda > d$$

Linearna regresija

Model $Y = X\beta + \epsilon$

Y slučajen, opazan

cilj: računamo β in variance ϵ

NNLC za β je

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

Na realn; osi imamo $\beta_1, \beta_2, \beta_3$

Izmerimo $\beta_1, \beta_2, \beta_3, \beta_1 - \beta_2, \beta_2 - \beta_3, \beta_1 - \beta_3$

$$y_1 = \beta_1 + \varepsilon_1$$

$$y_2 = \beta_2 + \varepsilon_2$$

$$y_3 = \beta_3 + \varepsilon_3$$

$$y_4 = \beta_1 - \beta_2 + \varepsilon_4$$

$$y_5 = \beta_2 - \beta_3 + \varepsilon_5$$

$$y_6 = \beta_1 - \beta_3 + \varepsilon_6$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ 1 & -1 & & & & \\ & & 1 & -1 & & \\ 1 & & & & -1 & \end{bmatrix}$$

$$\beta = (X^T X)^{-1} X^T y$$

Velj.:

$$\text{var } \hat{\beta} = \sigma^2 (X^T X)^{-1}$$

$$\hat{\sigma}^2 = \frac{\|Y - X\hat{\beta}\|^2}{n-m}$$

nepriistranske cenilke za σ^2

če je model gaussov ($\epsilon \sim N(0, \sigma^2 I)$) je

$$\frac{\hat{\beta} - \beta}{\sigma} \sim N(0, (X^T X)^{-1})$$

$$\text{za } \frac{\hat{\sigma}^2 (n-m)}{\hat{\sigma}^2} \sim \chi^2(n-m)$$

kako točke

↑
skup podatkov

↖ praznjenje ↗

↖ neopredeljeno ↗