

Martin Raic

5.02

računski: pisni
terorekion: ustni
projektna naloga

2 kolokvija

1. Centralni limitni izrek

$$S_n \sim B_n(n, \frac{1}{2}) \quad P(S_n = k) = \binom{n}{k} \cdot 2^{-n}$$

$$\frac{P(S_n = k+1)}{P(S_n = k)} = \frac{\binom{n}{k+1} 2^{-n}}{\binom{n}{k} 2^{-n}} = \frac{n!}{(k+1)!(n-k-1)!} \cdot \frac{k!}{n!} = \frac{n-k}{k+1}$$

Naj bo n sod $n=2m$

$P(S_n = k)$ je maksimalna pri $k=m$ DN
(glejamo razmerje $\frac{P(S_n = k+1)}{P(S_n = k)}$)

$d = 0, 1, \dots$

$$\frac{P(S_{2m} = m+d)}{P(S_{2m} = m)} = \frac{P(S_{2m} = m+1) P(S_{2m} = m+2) \dots P(S_{2m} = m+d)}{P(S_{2m} = m) P(S_{2m} = m+1) \dots}$$

$$= \frac{2m-m}{m+1} \cdot \frac{2m-m-1}{m+2} \dots \frac{2m-m-d+1}{m+d} =$$

$$= \frac{m \cdot (m-1) \cdot (m-2) \dots (m-d+1)}{(m+1) \cdot \dots \cdot (m+d)} =$$

$$= \left(\frac{1}{1 + \frac{1}{m}} \right) \left(\frac{1 - \frac{1}{m}}{1 + \frac{2}{m}} \right) \dots \left(\frac{1 - \frac{d-1}{m}}{1 + \frac{d}{m}} \right) \approx$$

$$\approx e^{-\frac{1}{m} - \frac{2}{m} - \dots - \frac{d-1}{m}} \cdot \frac{1}{m} - \frac{2}{m} \dots - \frac{d}{m} = e^{-\frac{d(d-1)}{2} - \frac{d(d+1)}{2}}$$

$$\ln(1+x) \approx x \quad 1+x \approx e^x \quad = e^{-\frac{d^2}{m}}$$

Torej tudi za $d = -1, -2, \dots$ zaradi sodosti
 $|d| \leq m: P(S_{2m} = m+d) \approx P(S_{2m} = m) \cdot e^{-\frac{d^2}{m}}$

$$\text{Izkaže se: } \lim_{m \rightarrow \infty} \sum_{d=-\infty}^{\infty} \left| \frac{P(S_{2m} = m+d)}{P(S_{2m} = m)} - e^{-\frac{d^2}{m}} \right| = 0$$

Izkaže se: $\lim_{m \rightarrow \infty} \sum_{d=-\infty}^{\infty} \left| \frac{P(S_{2m}=m+d)}{P(S_{2m}=m)} - e^{-\frac{d^2}{m}} \right| = 0$

$$\sum_{d=-\infty}^{\infty} \frac{P(S_{2m}=m+d)}{P(S_{2m}=m)} \overset{\text{skup vseh vrjetnosti}}{=} \frac{1}{P(S_{2m}=m)} \approx \sum_{d=-\infty}^{\infty} e^{-\frac{d^2}{m}}$$

$$\approx \int_{-\infty}^{\infty} e^{-\frac{x^2}{m}} dx = \sqrt{m} \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{m\pi}$$

$x = t\sqrt{m}$
 $dx = \sqrt{m} dt$

$$\Rightarrow P(S_{2m}=m) \approx \frac{1}{\sqrt{m\pi}}$$

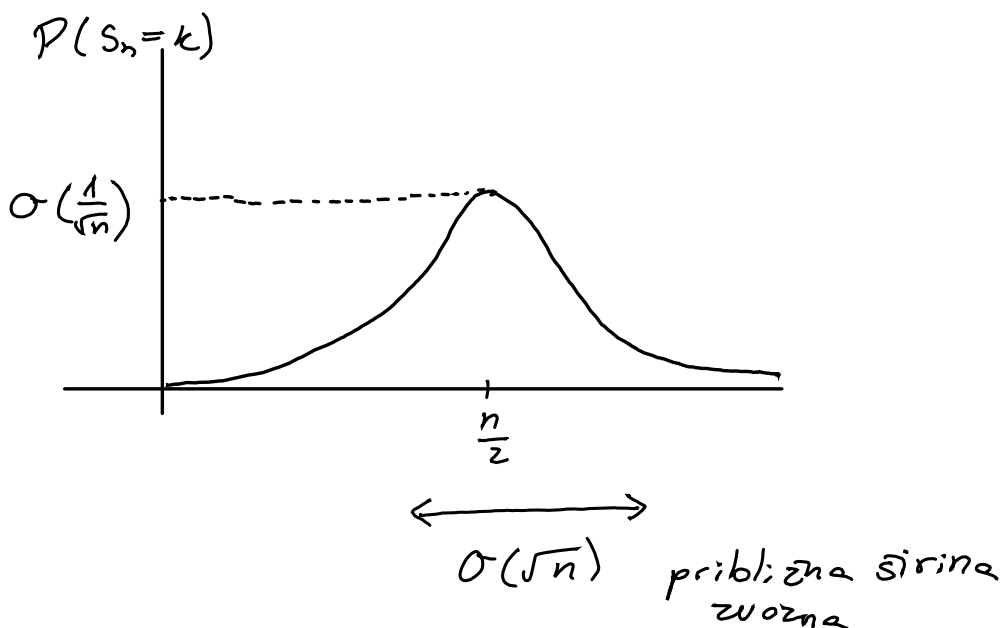
$$P(S_{2m}=m+d) \approx \frac{1}{\sqrt{m\pi}} e^{-\frac{d^2}{m}}$$

$d = k - m = k - \frac{n}{2}$

$$P(S_n=k) \approx \frac{1}{\sqrt{2\pi n}} e^{-\frac{2(k-\frac{n}{2})^2}{n}}$$

De Moivreova lokalna formula

Izkaže se, da to velja tudi za lihe n



$$P(a \leq S_n \leq b) = \sum_{k=a}^b P(S_n = k) \approx \sqrt{\frac{2}{n\pi}} \sum_{k=a}^b e^{-\frac{2(k - \frac{n}{2})^2}{n}} \approx$$

$$\approx \sqrt{\frac{2}{n\pi}} \sum_{k=a}^b \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} e^{-\frac{2(x - \frac{n}{2})^2}{n}} dx = \sqrt{\frac{2}{n\pi}} \int_{a-\frac{1}{2}}^{b+\frac{1}{2}} e^{-\frac{2(x - \frac{n}{2})^2}{n}} dx$$

$$a, b \in \mathbb{Z} + \frac{1}{2} \quad a \leq b$$

$$P(a \leq S_n \leq b) = P(a < S_n < b) = P(a + \frac{1}{2} \leq S_n \leq b - \frac{1}{2}) =$$

$$= \sqrt{\frac{2}{n\pi}} \int_a^b e^{-\frac{2(x - \frac{n}{2})^2}{n}} dx$$

To velja tudi za splošna $a, b \in \mathbb{R}$ $a \leq b$

Dobljeni približek za $P(S_n = k)$ je gostota porazdelitve $N(\frac{n}{2}, \frac{n}{4})$ v točki k .

↑ ← ←
pričakovana vrednost varianca

Podobno velja tudi za intervalske vrjetosti

Opazimo: $S \sim \text{Bin}(n, p)$ $E(S) = np$ $\text{var}(S) = np(1-p)$

Torej za $S \sim \text{Bin}(n, \frac{1}{2})$ velja $E(S) = \frac{n}{2}$ $\text{var}(S) = \frac{n}{4}$,

torej enako kot za normalno porazdelitev

Binomska porazdelitev ko gre n proti neskončno je take normalna porazdelitev, da ima isto pričakovano vrednost in varianco

Binomske porazdelitev $\text{Bin}(n, p)$

nastane kot porazdelitev vsote indikatorjev

$I_1 + \dots + I_n$ kjer so $I_1, \dots, I_n$ neodvisne

med sabo in $I_i \sim \text{Be}(p) = \begin{pmatrix} 0 & p \\ 1-p & 0 \end{pmatrix}$

Vsota veliko neodvisnih slučajnih spremenljivk z

Bernullijevo porazdelitvijo $\text{Be}(\frac{1}{2})$ je torej porazdeljena

približno normalno z ustrezno pričakovano

vrednostjo in varianco.

Primer na powerpointu

$X_1, \dots, X_n$ neodvisne. $\sim \begin{pmatrix} -1 & 0 & 9 \\ 0,5 & 0,45 & 0,05 \end{pmatrix}$

τ z verjetnostjo

5% zadene

10 € (glej priložnost)

$X_1, X_2, \dots, X_n$ neodvisne enakoporazdeljene

$$E(X_1) = \mu_1 \quad E(X_1^2) < \infty \quad \text{var}(X_1) = \sigma_1^2 \quad \sigma_1 > 0$$

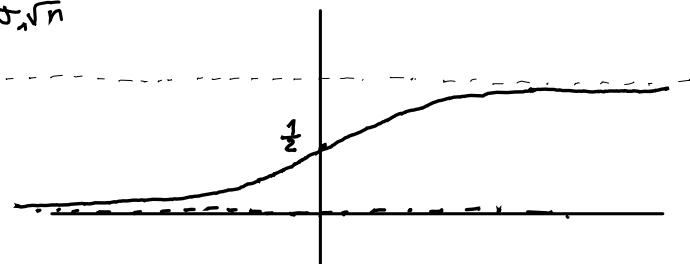
če definiramo $S = X_1 + \dots + X_n$ se izkaže

$$P(a \leq S_n \leq b) \approx \frac{1}{\sigma_1 \sqrt{2\pi n}} \int_a^b e^{-\frac{(x - n\mu_1)^2}{2n\sigma_1^2}} dx$$

Standardizacija: $E \rightarrow 0 \quad \text{var} \rightarrow 1$

$$t = \frac{x - n\mu_1}{\sigma_1 \sqrt{n}} \quad dt = \frac{dx}{\sigma_1 \sqrt{n}}$$

$$\frac{1}{\sqrt{2\pi}} \int_{\frac{a - n\mu_1}{\sigma_1 \sqrt{n}}}^{\frac{b - n\mu_1}{\sigma_1 \sqrt{n}}} e^{-\frac{t^2}{2}} dt = \Phi\left(\frac{b - n\mu_1}{\sigma_1 \sqrt{n}}\right) - \Phi\left(\frac{a - n\mu_1}{\sigma_1 \sqrt{n}}\right)$$



$$W_n := \frac{S_n - n\mu_1}{\sigma_1 \sqrt{n}} \quad E(W_n) = 0 \quad \text{var}(W_n) = 1$$

$\approx$ približno normalno porazdeljena

$$\Rightarrow P(a \leq S_n \leq b) = P\left(\frac{a - n\mu_1}{\sigma_1 \sqrt{n}} \leq W_n \leq \frac{b - n\mu_1}{\sigma_1 \sqrt{n}}\right) \approx$$

$$P(a \leq W_n \leq b) \approx \Phi(b) - \Phi(a) = P(a \leq Z \leq b)$$

$$Z \sim N(0, 1)$$

Dogovor: $Z \sim N(0, 1)$

Povsod kjer je $\leq$ je lahko tudi $<$

Izrek 1.1 (centralni limiti izrek; Paul Pierre Lévy)

Naj bo $X_1, X_2, \dots$ N&P, $S_1, S_2, \dots$ vsote X_i , $W_1, W_2, \dots$
 $W_n = \frac{S_n - n\mu_1}{\sigma_1 \sqrt{n}}$

$$\lim_{n \rightarrow \infty} P(W_n \leq w) = \Phi(w)$$

Jarl Voldemar Lindeberg:

$$\lim_{n \rightarrow \infty} \sup_{w \in \mathbb{R}} |P(W_n \leq w) - \Phi(w)| = \lim_{n \rightarrow \infty} |P(S_n \leq x) - \Phi(\frac{x - n\mu_1}{\sigma_1 \sqrt{n}})| = 0$$

Dokaz: opustimo (Poglej v zapisnik, ne učit na pamet)
velja za vse dokaze, ki jih [↑]opuščamo

← distribution

$$W_n \xrightarrow[n \rightarrow \infty]{d} N(0, 1)$$

Izrek 1.2

Če so $W_1, W_2, \dots$ kot prej, za vsako zvezno funkcijo $h: \mathbb{R} \rightarrow \mathbb{R}$, ki ima kvadratično omejeno rast ($\exists M \in \mathbb{R}, \forall x \in \mathbb{R}, |h(x)| \leq M(1+x^2)$)

velja:

$$\lim_{n \rightarrow \infty} \mathbb{E}(h(W_n)) = \mathbb{E}(h(Z))$$

Dokaz opuščamo

$Y_1, \dots, Y_n$ neodvisne z vsoto S

$Z_1, \dots, Z_n$ neodvisne z vsoto T

↖ z enakrat že n: trebade so normalne

Za $k=1, \dots, n$ privzamemo

- $\mathbb{E}(|Y_k|^3), \mathbb{E}(|Z_k|^3) < \infty$

- $\mathbb{E}(Y_k) = \mathbb{E}(Z_k) = 0$

- $\text{var}(Y_k) = \text{var}(Z_k) = \sigma_k^2 \quad \sigma_k \geq 0$

Tako je tudi $\mathbb{E}(S) = \mathbb{E}(T) = 0$ in

$$\text{var}(S) = \text{var}(T) = \sigma^2 := \sum_{k=1}^n \sigma_k^2; \quad \sigma > 0$$

Trditev 1.3

Za poljubno slučajno spremenljivko Y z vrednostmi na $[0, \infty)$ in poljuben $p \geq 1$

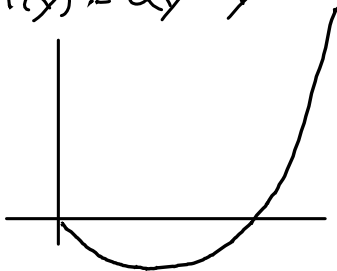
velja $(E(Y))^p \leq E(Y^p)$

$$(p=2, Y=|X| \quad \text{var}(X) = E(X^2) - E(X)^2 \geq 0)$$

Dokaz: BŠZS: $p > 1$, $P(Y > 0) > 0$, $E(Y^p) < \infty$

Podobna neenakost $E(Y) \leq a E(Y^p)$; $a > 0$

$$f(y) := ay^p - y$$



$$f'(y) = apy^{p-1} - 1 = 0$$

$$\Rightarrow y = (ap)^{-\frac{1}{p-1}} \text{ minimum}$$

$$f(y) \geq f((ap)^{-\frac{1}{p-1}})$$

$$ay^p - y \geq -\frac{p-1}{p} (ap)^{-\frac{1}{p-1}} \quad (\text{DN})$$

$$\Rightarrow E(Y) \leq a E(Y^p) + \frac{p-1}{p} (ap)^{-\frac{1}{p-1}} =: g(a)$$



$$g'(a) = E(Y^p) - ap^{-\frac{p}{p-1}} \quad (\text{DN})$$

$$ap = (E(Y^p))^{-\frac{p-1}{p}}$$

$$\Rightarrow E(Y) \leq g\left(\frac{1}{p} E(Y^p)\right)^{-\frac{p-1}{p}} = \underbrace{E(Y^p)^{\frac{1}{p}}}_{\text{DN}}$$

QED

$Y_1, \dots, Y_n$ neodv. z vsoto S $E(Y_k) = E(Z_k) = 0$
 $Z_1, \dots, Z_n$ neodv. z vsoto T $\text{var}(Y_k) = \text{var}(Z_k) = \sigma_k^2$
 $h: \mathbb{R} \rightarrow \mathbb{R}$

$$E(h(S)) - E(h(T))$$

Privzemimo, da je $h \in C^3(\mathbb{R})$ in $M_3 := \sup_{x \in \mathbb{R}} |h'''(x)| < \infty$

$$\left(h(x) = h(0) + h'(0)x + \frac{1}{2}h''(0)x^2 + \frac{1}{6}h'''(\xi)x^3 \right) \\ \Rightarrow h \in \mathcal{O}(x^3)$$

$$T = V_0 \longrightarrow V_1 \longrightarrow V_2 \longrightarrow \dots \longrightarrow V_n = S$$

$$V_k = Y_1 + Y_2 + \dots + Y_k + Z_{k+1} + \dots + Z_n$$

Privzemimo da so $Y_1, \dots, Y_n$ in $Z_1, \dots, Z_n$ vsi med seboj neodvisni.

$$E(h(S)) - E(h(T)) = \sum_{k=1}^n E(h(V_k)) - E(h(V_{k-1}))$$

$$U_k := Y_1 + \dots + Y_{k-1} + Z_{k+1} + \dots + Z_n$$

$$U_k + Y_k = V_k \quad U_k + Z_k = V_{k-1} \quad U_k, V_k \text{ in } Z_k \text{ so neodvisne}$$

Pozor! Tu nastopata U in V

$$h(V_k) = h(U_k) + h'(U_k)Y_k + \frac{1}{2}h''(U_k)Y_k^2 + R_k; |R_k| \leq \frac{1}{6}M_3|Y_k|^3$$

$$h(V_{k-1}) = h(U_k) + h'(U_k)Z_k + \frac{1}{2}h''(U_k)Z_k^2 + \tilde{R}_k \quad \text{Tist supremum} \\ |R_k| \leq \frac{1}{6}M_3|Y_k|^3 \quad |\tilde{R}_k| \leq \frac{1}{6}M_3|Z_k|^3$$

$$E(h'(U_k))E(Y_k) = 0$$

$$E(h(V_k)) = E(h(U_k)) + E(h'(U_k)Y_k) + \dots + E(R_k)$$

$$E(h(V_{k-1})) = E(h(U_k)) + E(h'(U_k)Z_k) + \dots + E(\tilde{R}_k)$$

$$E(h'(U_k))E(Z_k) = 0$$

$$E(\tilde{R}_k) \leq \frac{1}{6}M_3E(|Z_k|^3)$$

$$E(R_k) \leq \frac{1}{6}M_3E(|Y_k|^3)$$

$$E(h(V_k)) - E(h(V_{k-1})) \leq \frac{1}{6}M_3(E(|Y_k|^3) + E(|Z_k|^3)) \quad \nearrow \sum_{k=1}^n$$

$$E(h(S)) - E(h(T)) \leq \frac{1}{6}M_3 \sum_{k=1}^n (E(|Y_k|^3) + E(|Z_k|^3))$$

>>

((iz prejšnje strani))

$$\text{če } Z_k \sim N(0, \sigma_k^2) \Rightarrow T \sim N(0, \sigma^2)$$

27.2

$$Z_k \stackrel{d}{=} \sigma_k Z \quad (Z \sim N(0, 1)) \Rightarrow E(|Z_k|^3) = \sigma_k^3 E(|Z|^3)$$

↑
enako
porazdeljena

$$E(|Z|^3) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z|^3 e^{-\frac{z^2}{2}} dz \stackrel{DN}{=} \frac{4}{\sqrt{2\pi}}$$

$$N(\mu, \sigma^2)(h) := E(h(X)), \text{ kjer je } X \sim N(\mu, \sigma^2)$$

$$\text{sledi: } |E(h(S)) - N(0, 1)h| \leq \frac{1}{6} M_3 \sum (E|Y_k|^3 + \frac{4}{\sqrt{2\pi}} \sigma_k^3)$$

$$\begin{aligned} \text{Po trditvi 1.3 je } \sigma_k^3 &= E(|Z_k|^2)^{\frac{3}{2}} = \\ &= E(|Y_k|^2)^{\frac{3}{2}} \leq E(|Y_k|^3) \end{aligned}$$

Dokazeli smo

Izrek 1.4

Naj bodo $Y_1, Y_2, \dots, Y_n$ neodvisne slučajne sprem.

z $\mathbb{E}(|Y_k|^3) < \infty$ in $\mathbb{E}(Y_k) = 0, \text{var}(Y_k) = \sigma_k^2$

Naj bo $S = \sum_{k=1}^n Y_k$.

Torej za $\forall$ funkcijo $h \in \mathcal{C}^3(\mathbb{R})$ z $M_3 := \sup |h'''| < \infty$ velja:

$$|\mathbb{E}(h(S)) - N(0, \sigma^2)(h)| \leq \left(\frac{1}{6} + \frac{2}{3\sqrt{2\pi}}\right) M_3 \sum_{k=1}^n \mathbb{E}(|Y_k|^3),$$

kjer je $\sigma^2 = \sum_{k=1}^n \sigma_k^2$

Dokaz na prejšnji strani:

□

Izrek 1.5 (Berry-Essenov izrek)

Obstaja taka univerzalna konstanta C , da za poljubne neodvisne slučajne spremenljivke

$Y_1, \dots, Y_n$, ki izpolnjujejo pogoje prejšnjega izreka (z $E(|Y_k|^3) < \infty$ in $E(Y_k) = 0$, $\text{var}(Y_k) = \sigma_k^2$

$S = \sum_{k=1}^n Y_k$) velja

$$\sup_{y \in \mathbb{R}} |P(S \leq y) - \Phi(\frac{y}{\sigma})| = \sup_{w \in \mathbb{R}} |P(\frac{S}{\sigma} \leq w) - \Phi(w)| \\ \leq \frac{C}{\sigma^3} \sum_{k=1}^n E(|Y_k|^3), \text{ če je } \sigma > 0$$

Dokaz opustimo (ni ga v zapiskih)

2013: $C = 0,5583$

za enako porazdeljene: $C = 0,47$

1956: spodnje meje: $\frac{\sqrt{10}+3}{6\sqrt{2\pi}} = 0,4097$

Naj bodo $X_1, X_2, \dots$ neodvisne in enake porazdeljene
 $E(|X_1|^3) < \infty$, $E(X_1) = \mu_1$, $\text{var}(X_1) = \sigma_1^2$ ($\sigma_1 > 0$)

$$E(|X_1 - \mu_1|^3) = \gamma_1^3$$

$$S_n := \sum_{k=1}^n X_k \quad W_n := \frac{S_n - n\mu_1}{\sigma_1 \sqrt{n}}, \quad E(W_n) = 0, \quad \text{var}(W_n) = 1$$

$$Y_k := X_k - \mu_1, \quad E(Y_k) = 0$$

Fiksiramo n $S := \sum_{k=1}^n Y_k = S_n - n\mu_1 = W_n \sqrt{n} \sigma_1$

$$\sigma^2 := \text{var}(S) = n\sigma_1^2$$

$$\frac{1}{\sigma^3} \sum_{k=1}^n E(|Y_k|^3) = \frac{1}{n^{\frac{3}{2}} \sigma_1^3} n \gamma_1^3 = \frac{1}{\sqrt{n}} \left(\frac{\gamma_1}{\sigma_1} \right)^3 \xrightarrow{n \rightarrow \infty} 0$$

Hlrost konvergenca: $\frac{1}{\sqrt{n}}$

Posledica 1.6

Če so $X_1, X_2, \dots, S_n, W_n, \mu_1$ in σ_1 kot na prejšnji strani, velja:

$$\sup_{x \in \mathbb{R}} |P(S \leq x) - \Phi\left(\frac{x - n\mu_1}{\sigma_1 \sqrt{n}}\right)| = \sup_{w \in \mathbb{R}} |P(W_n \leq w) - \Phi(w)| \leq \frac{C}{\sqrt{n}} \left(\frac{\sigma_1}{\sigma_1}\right)^3$$

Dokaz na prejšnji strani

Posledica 1.7

$X_1, X_2, \dots, S_n, W_n, \mu_1, \sigma_1, \delta_1$ kot na dveh straneh nazaj,
za vsako funkcijo $h \in C^3(\mathbb{R})$ z $M_3 := \sup |h'''| < \infty$
velja $E(h(W_n)) - E(h(Z)) \leq \left(\frac{1}{6} + \frac{2}{3\sqrt{2\pi}}\right) \frac{M_3}{\sqrt{n}} \left(\frac{\delta_1}{\sigma_1}\right)^3$

$\bar{C}_2$ v izrek 1.4 vstavimo

$$\tilde{y}_k = \frac{X_k - \mu_1}{\sigma_1 \sqrt{n}} \quad \text{dobimo}$$

□

2. Konvergenca porazdelitev

Definicija I: Zaporedje realnih slučajnih spremenljivk

$X_1, X_2, \dots$ konvergira v porazdelitvi, k slučajni spremenljivki X , če za $\forall x \in \mathbb{R}$, za katerega je $\mathbb{P}(X=x)=0$ velja

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n < x) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X < x).$$

Če je $X_1 \sim \mu_1, X_2 \sim \mu_2, \dots$ pišemo

$$X_n \xrightarrow[n \rightarrow \infty]{d} X$$

$$\mu_n \xrightarrow[n \rightarrow \infty]{d} \mu$$

namesto d lahko pišemo tudi

W ali $\bar{S}$ (zaporedje porazdelitev **šibko** konvergira)

$\uparrow$
(konvergenca ki jo meriš z linearnimi funkcijami)

Opazka 2.1

Na ta način je, če so $X_n = C_n$ in $X = C$ ker konstantne, konvergenca v porazdelitvi ujema z običajno konvergenco zaporedij.

Naj bo $\lim_{n \rightarrow \infty} C_n = C$

$x < C$: za dovolj velike n je $C_n > x$

$$\Rightarrow \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x) = 0$$

$x > C$: za dovolj velike n $C_n < x$

$$\Rightarrow \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x) = 1$$

$x = C$: odvisno od tega na kakšen način konvergira C_n . kaotično obnašanje

DN: Dokazi še v obratno smer (če konvergira v porazdelitvi $\Rightarrow$ konvergira po točkah)

Opazka 2.2.

$$\text{Velja } \lim_{n \rightarrow \infty} P(X_n > x) = \lim_{n \rightarrow \infty} P(X_n \geq x) = P(X \geq x)$$

Primer:

Če je $a_1, a_2, \dots \in \mathbb{R}$ $\sigma_1, \sigma_2, \dots \in [0, \infty)$

in $a = \lim_{n \rightarrow \infty} a_n$ $\sigma = \lim_{n \rightarrow \infty} \sigma_n$ gre

$$N(a_n, \sigma_n) \xrightarrow[n \rightarrow \infty]{} N(a, \sigma^2)$$

Naj bo $\sigma > 0$. BŠZS: $\forall n \in \mathbb{N}, \sigma_n > 0$

$$P(X_n \leq x) = \Phi\left(\frac{x - a_n}{\sigma_n}\right)$$

$$\downarrow \text{p}$$

$$P(X < x) = \bar{\Phi}\left(\frac{x - a}{\sigma}\right)$$

Izkeže se, da je to res tudi, če je $\sigma = 0$ (DN)

Imenitnejša formulacija CLT, natančneje izreka 2.1

Če so $X_1, X_2, \dots$ neodvisne in enako porazdeljene
z $E(X_1) = \mu_1$ in $\text{var}(X_1) = \sigma_1^2$ ter $\sigma_1 > 0$ gre

$$\frac{X_1 + \dots + X_n - n\mu_1}{\sigma_1 \sqrt{n}} \xrightarrow[n \rightarrow \infty]{} N(0, 1)$$

□

Definicija: **Slučajna spremenljivka** na vrjetnostnem prostoru $(\Omega, \mathcal{F}, P)$ in z vrednostmi v metričnem prostoru (M, d) , je taka preslikava $X: \Omega \rightarrow M$, da če za vsako odprto množico $U \in M$ velja, da je $\{X \in U\}$ dogodek $(\{\omega \in \Omega; X(\omega) \in U\} \in \mathcal{F})$
 (M je lahko tudi samo topologija)

Opazka 2.3

če je X slučajna spremenljivka z vrednostmi v metričnem prostoru (M, d) in $h: M \rightarrow \mathbb{R}$ zvezna funkcija, je tudi $h(X)$ slučajna spremenljivka.

Za $\forall$ odprto množico $V \subseteq \mathbb{R}$ je namreč

$$\{h(X) \in V\} = \{X \in h^{-1}(V)\}$$

Definicija II: Zaporedje slučajnih spremenljivk $X_1, X_2, \dots$ z vrednostmi v metričnem prostoru (M, d) **v porazdelitvi konvergira** proti X , če za $\forall$ zvezno in omejeno funkcijo $h: M \rightarrow \mathbb{R}$ velja, da je $\lim_{n \rightarrow \infty} E(h(X_n)) = E(h(X))$

$$P(X_n=0) = 1 - \frac{1}{n}$$

$$// \quad P(X_n=n) = \frac{1}{n}$$

Primer: $X_n \sim \left(1 - \frac{1}{n}, \frac{1}{n}\right)$

$$E(h(X_n)) = h(0)\left(1 - \frac{1}{n}\right) + h(n) \frac{1}{n}$$

Če je h omejena gre to proti $h(0)$

Zato gre $X_n \xrightarrow[n \rightarrow \infty]{d} 0$. Vendar pa je

$$E(X_n) = 1, \text{ medtem ko je } E(0) = 0.$$

Izrek 2.4 (Eduard Helly, Hubbert Evelyn Gray)

Za $(M, d) = (\mathbb{R}, \text{običajna metrika})$ se definiciji

I in II ujemata

Dokaz: Opustimo

Opazka 2.5

Porazdelitev poljubne realne slučajne spremenljivke je natančno določena že z vrednostmi $\mathbb{P}(X \leq x)$, kjer x preteče neko gosto podmnožico realne osi (npr. $\{x; \mathbb{P}(X=x)=0\}$)

Podobno se za slučajno spremenljivko z vrednostmi v metričnem prostoru izkaže, da $\mathbb{E}(h(X))$, kjer h preteče omejene zvezne funkcije, natančno določa porazdelitev slučajne spremenljivke X ; tj. vse možne vrednosti $\mathbb{P}(X \in B)$, kjer B preteče Borellove množice.

Posledica 2.6

Če gre $X_n \xrightarrow[n \rightarrow \infty]{d} X$ kot tudi $X_n \xrightarrow[n \rightarrow \infty]{} \tilde{X}$,
mora biti $X \stackrel{d}{=} \tilde{X}$
 $\uparrow$ enako porazdeljeni

(To nam pove da je ta topologija Hausdorffova)

Izrek 2.7 (portmanteau)

(portmanteau ang. - dvodelen kovček za plašče
portemanteau fra. - stojala za plašče)

Za slučajne spremenljivke $X_1, X_2, \dots$ z vrednostmi v metričnem prostoru (M, d) so naslednje ekvivalentne:

$$1) \Leftrightarrow X_n \xrightarrow[n \rightarrow \infty]{d} X \text{ po definiciji II}$$

$$2) \Leftrightarrow \text{za } \forall \text{ odprto množico } U \subseteq M \text{ velja}$$

$$\liminf_{n \rightarrow \infty} P(X_n \in U) \geq P(X \in U)$$

$$3) \Leftrightarrow \text{za vsako zaprto množico } F \subseteq M \text{ velja}$$

$$\limsup_{n \rightarrow \infty} P(X_n \in F) \leq P(X \in F)$$

Dokaz opuščamo

Opazka 2.8

Naj bo (M, d) metrični prostor in $\tilde{M} \subseteq M$.

Torej je $(\tilde{M}, d)$ tudi metrični prostor in za dano preslikavo $\tilde{X}: \Omega \rightarrow \tilde{M}$ je ekvivalentno, da je to slučajna spremenljivka z vrednostmi v $(\tilde{M}, d)$ ali pa v (M, d) .

To je res zato, ker velja $\mathcal{T}(\tilde{M}) = \{U \cap \tilde{M}; U \in \mathcal{T}(M)\}$
 $\uparrow$ topologija od $\tilde{M}$

Opazka 2.9

Če so $M, \tilde{M}, d$ kot prej in $\tilde{X}_1, \tilde{X}_2, \dots$ in X slučajne sprem. z vrednostmi v $(\tilde{M}, d)$, gre $\tilde{X}_n \rightarrow X$ v okviru $(\tilde{M}, d) \Leftrightarrow$ če gre v okviru (M, d)

To sledi iz karakterizacije (2) v izreku portmanteau.

Opazka 2.10

naj bo (M, d) metrični prostor in $\tilde{M} \subseteq M$ odprta.

Če so $X_1, X_2, \dots$ slučajne spremenljivke z vrednostmi v M , $\tilde{X}$ slučajna sprem. z vrednostmi v $\tilde{M}$ in gre $X_n \rightarrow \tilde{X}$ velja

$$\lim_{n \rightarrow \infty} P(X_n \in \tilde{M}) = 1$$

Definicija: Naj bode $X_1, X_2, \dots$ in X slučajne spremenljivke definirane na istem vrjetnostnem prostoru in z vrednostmi v istem metričnem prostoru (M, d) . Pravimo, da zaporedje $X_1, X_2, \dots$ konvergira proti X **v vrjetnosti**, če za $\forall r > 0$.

$$\lim_{n \rightarrow \infty} \mathbb{P}(d(X_n, X) \geq r) = 0$$

Pišemo: $X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X$

Opazka 2.11

$$X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X \iff d(X_n, X) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0$$

$$d(X_n, X) = d_{\mathbb{R}}(d(X_n, X), 0)$$

Trditev 2.12.

Ko je $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = \tilde{X}_n) = 1$, gre

$$d(X_n, \tilde{X}_n) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0$$

Dokaz:

Predpostavka je ekvivalentna izjavi

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n \neq \tilde{X}_n) = 0. \text{ Tedyj pa tudi}$$

$$\text{za } \forall r > 0, \lim_{n \rightarrow \infty} \mathbb{P}(d(X_n, \tilde{X}_n) \geq r) = 0$$

qed ■

Trditev 2.13

Naj bodo $X_1, X_2, \dots, \tilde{X}_1, \tilde{X}_2, \dots$ in X slučajne spremenljivke z vrednostmi v metričnem prostoru (M, d) in naj gre $d(X_n, \tilde{X}_n) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0$.

Tedaj velja:

(1) Če gre $X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X$, gre tudi $\tilde{X}_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X$

(2) Če gre $X_n \xrightarrow[n \rightarrow \infty]{d} X$, gre tudi $\tilde{X}_n \xrightarrow[n \rightarrow \infty]{d} X$

Dokaz opustimo

Trditev 2.14

Če gre $X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X$, gre tudi $X_n \xrightarrow[n \rightarrow \infty]{d} X$

Dokaz:

Po opazki 2.11 gre $d(X_n, X) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0$.

Uporabimo trditev 2.13, kjer za X_n vzamemo X in za $\tilde{X}_n$ vzamemo X_n

	$\tilde{X}=1$	$\tilde{X}=2$
$X=1$	0	$\frac{1}{2}$
$X=2$	$\frac{1}{2}$	0

Torej obrat trditve ne velja. ■

Todilev 2.15

← konstanta

Če gre $X_n \xrightarrow{d} c$, gre tudi $X_n \xrightarrow{\mathbb{P}} c$

(za konvergenco proti konstanti velja torej ekvivalenca)

Dokaz: Za $r > 0$ je $K(c, r) = \{x \in M; d(x, c) < r\}$.

Po opazli 2.10 mora biti potem

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n \in K(c, r)) = 1$$

$\Leftrightarrow$

$$d(X_n, c) < r$$

qed $\blacksquare$

Posledica 2.16

Če je c konstanta, je $X_n \xrightarrow{d} c$

$$\Leftrightarrow d(X_n, c) \xrightarrow{\mathbb{P}} 0 \Leftrightarrow d(X_n, c) \xrightarrow{d} 0$$

Dokaz: $\square$

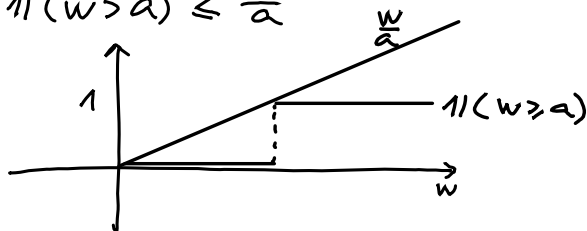
Trditev 2.17 (neenakba Markova)

Za nenegativno slučajno spremenljivko W
in poljuben $a > 0$ velja: $\mathbb{P}(W \geq a) \leq \frac{E(W)}{a}$

Dokaz:

Rezultat sledi iz neenakosti

$$\mathbb{1}(W > a) \leq \frac{W}{a}$$



qed ■

Trditev 2.18 za nek poljuben p

ko je $p > 0$ in so $W_1, W_2, \dots$ take spremenljivke,

da je $\lim_{n \rightarrow \infty} E(W_n^p) = 0$, gre tudi:

$$W_n \xrightarrow[n \rightarrow \infty]{d/P} 0$$

Dokaz:

$$\mathbb{P}(W_n \geq r) = \mathbb{P}(W_n^p \geq r^p) \leq \frac{E(W_n^p)}{r^p} \xrightarrow[n \rightarrow \infty]{} 0$$

qed ■

(Neki z razdeljco do 0)

Trditev 2.19 (Šibki zakon velikih števil)

Če so $X_1, X_2, \dots$ neodvisne in enako porazdeljene realne slučajne spr. z $E(X_1^2) < \infty$ in $E(X_1) = \mu_1$,
potem gre $\frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow \infty]{d/P} \mu_1$.

Dokaz:

Ekvivalentno je dokazati da gre $\frac{X_1 + \dots + X_n}{n} - \mu_1 \xrightarrow[n \rightarrow \infty]{P} 0$

$$E\left(\left(\frac{X_1 + \dots + X_n}{n} - \mu_1\right)^2\right) = \text{var}\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n^2} n \sigma_1^2 =$$

zamaknjeno & pričakovano vrednost, zato vemo $= \frac{\sigma_1^2}{n} \xrightarrow[n \rightarrow \infty]{} 0$

$E(\dots) = 0$ (ker $E(X_1) = \mu_1$)

rezultat sledi iz trditve 2.18

qed ■

Opazka 2.20

(1) predpostavka, da je $E(X_1^2) < \infty$ je premočna.
dovolj je, da je $E(|X_1|) < \infty$

(2) Tudi ob slednji predpostavki velja celo, da je

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \frac{X_1 + \dots + X_n}{n} = \mu_1\right) = \mathbb{P}\left(\{w; \lim_{n \rightarrow \infty} \frac{X_1(w) + \dots + X_n(w)}{n} = \mu_1\}\right) = 1$$

Dokaz opuščamo

Dokez:

če je $M' \rightarrow R$ vezna in omejena, velja tudi:

$$\lim_{n \rightarrow \infty} E((h \circ g)(x)) = E((h \circ g)(x))$$

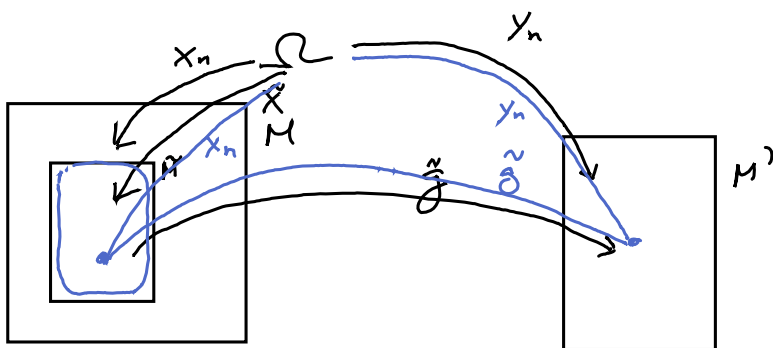
$$\lim_{n \rightarrow \infty} E(h(g(x))) = E(h(g(x)))$$

Izrek 2.22

- Naj bo (M, d) metrični prostor.
- $\tilde{M} \subseteq M$ podprostor,
- $X_1, X_2, \dots$ sluč. spr. z vrednostmi v M
- $\tilde{X}$ slučajna spremenljivka z vrednostmi v $\tilde{M}$
- in naj gre $X_n \xrightarrow[n \rightarrow \infty]{d} \tilde{X}$
- Naj bo (M', d) še en metrični prostor
- $\tilde{g}: \tilde{M} \rightarrow M'$ zvezna preslikava
- Naj bodo $Y_1, Y_2, \dots$ take sluč. sprem. z vrednostmi v M' , da je $\lim_{n \rightarrow \infty} P(X_n \in \tilde{M}, Y_n = \tilde{g}(X_n)) = 1$

Torej gre $Y_n \xrightarrow[n \rightarrow \infty]{d} \tilde{g}(\tilde{X})$

Dokaz: opustimo



$$\{X_n \in \tilde{M}, Y_n = \tilde{g}(X_n)\}$$

Modre puščice komutirajo

Če gre $X_n \xrightarrow[n \rightarrow \infty]{d} X$ in $Y_n \xrightarrow[n \rightarrow \infty]{d} Y$,
ali lahko rečemo, da gre $(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{d} (X, Y)$?

V splošnem ne.

Recimo, da gre $X_n \xrightarrow[n \rightarrow \infty]{d} X$ in $Y_n \xrightarrow[n \rightarrow \infty]{d} Y$ in
 $(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{d} (X, Y)$. Brez ko obstaja
tak slučajni vektor (X', Y') , da je $X' \stackrel{d}{=} X$
in $Y' \stackrel{d}{=} Y$, tade $(X, Y) \neq (X', Y')$, gre tudi
 $X_n \xrightarrow[n \rightarrow \infty]{d} X'$ in $Y_n \xrightarrow[n \rightarrow \infty]{d} Y'$, ne more pa biti
 $(X_n, Y_n) \rightarrow (X', Y')$
(po posledici 2.6)

	$Y=0$	$Y=1$	
$X=0$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$
$X=1$	$\frac{1}{6}$	$\frac{1}{3}$	
	$\frac{1}{3}$	$\frac{2}{3}$	

	$Y=0$	$Y=1$	
$X=0$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$
$X=1$	0	$\frac{1}{2}$	
	$\frac{1}{3}$	$\frac{2}{3}$	

Izrek 2.23

Naj bado:

- $X_1, X_2, \dots, X_n$ slučajne spremenljivke z vrednostmi v (M, d_M)
- $Y_1, Y_2, \dots$ sluč. spr. z vrednostmi v (N, d_N)
- Naj bo $C \in N$ konstanta
- Za posamezen $n \in \{1, 2, \dots\}$, naj bosta X_n in Y_n definirani na istem verjetnostnem prostoru.
- Opreмимо produkt $M \times N$ z metriko $d_{M \times N}$, ki izpolnjuje naslednje:
 - 1) $\forall x_1, x_2 \in M. d_{M \times N}((x_1, C), (x_2, C)) \leq d_M(x_1, x_2)$
 - 2) $\forall x \in M. \forall y \in N. d_{M \times N}((x, y), (x, C)) \leq d_N(y, C)$

Tedaj, če gre $X_n \xrightarrow[n \rightarrow \infty]{d} X$ in $Y_n \xrightarrow[n \rightarrow \infty]{d} C$, gre tudi: $(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{d} (X, C)$

Dokaz:

Ker je vložitev $L: M \rightarrow M \times N, X \mapsto (X, C)$ zvezna,

gre po trditvi 2.21 $(X_n, C) \rightarrow (X, C)$.

Ker gre $Y_n \xrightarrow[n \rightarrow \infty]{d} C$ gre po trditvi 2.15 tudi:

$Y_n \xrightarrow[n \rightarrow \infty]{P} C$, ker pomeni:

$$\forall r > 0, \lim_{n \rightarrow \infty} P(d_N(Y_n, C) \geq r) = 0$$

Po predpostavkah je torej tudi:

$$\lim_{n \rightarrow \infty} P(d_{M \times N}((X_n, Y_n), (X_n, C)) \geq r) = 0, \quad (\text{po 2})$$

ali ekvivalentno $d_{M \times N}((X_n, Y_n), (X_n, C)) \xrightarrow[n \rightarrow \infty]{P/d} 0$.

Rezultat sledi iz trditve 2.13 (2)

■
qed

Posledica 2.24 (šibki zakon velikih števil za slučajne vektorje)

če so

$$X_1 = \begin{bmatrix} x_1^{(n)} \\ \vdots \\ x_1^{(m)} \end{bmatrix}, \quad X_2 = \begin{bmatrix} x_2^{(n)} \\ \vdots \\ x_2^{(m)} \end{bmatrix}, \dots$$

nezavisni in enako porazdeljeni slučajni vektorji

z $E(\|X_1\|^2) < \infty$ in

$$\underline{\mu} = \begin{bmatrix} E(x_1^{(n)}) \\ \vdots \\ E(x_1^{(m)}) \end{bmatrix}, \text{ gre } \frac{\underline{X}_1 + \underline{X}_2 + \dots + \underline{X}_n}{n} \xrightarrow[n \rightarrow \infty]{} \underline{\mu}$$

Dokaz:

□

Izreki Slutskega (2.25, 2.26, 2.27, 2.28)

Posledica 2.25

Naj bodo $X_1, X_2, \dots$ in $Y_1, Y_2, \dots$ slučaj. vek. z vred. v $\mathbb{R}$
če gre $X_n \xrightarrow[n \rightarrow \infty]{d} X$ in $Y_n \xrightarrow[n \rightarrow \infty]{d} c$, kjer je c
konstanta, gre tudi $X_n + Y_n \xrightarrow{d} X + c$
in $X_n - Y_n \xrightarrow{d} X - c$

Dokaz: $\square$

Posledica 2.26

- Naj bodo $X_1, X_2, \dots$ slučaj. vek. z vrednostmi v $\mathbb{R}^n$
 - in $U_1, U_2, \dots$ realne slučaj. spr.
 - Naj gre $X_n \xrightarrow{d} X$ in $U_n \xrightarrow{d} U$
če je X ali U konstanta, gre tudi
 $\underline{U_n X_n} \xrightarrow{d} \underline{U X}$
-

Dokaz: $\square$

Trditve: 2.27

- Naj bodo $X_1, X_2, \dots$ in $Y_1, Y_2, \dots$ sluč. vekt. z vred. v $\mathbb{R}^m$
 - $U_1, U_2, \dots$ realne sluč. spr.
 - Za posamezen n naj bodo X_n, Y_n, U_n definirani na istem verjetnostnem prostoru.
 - Naj gre $X_n \xrightarrow{d} X$ in $U_n \xrightarrow{d} c$ $c \in \mathbb{R} - \{0\}$ konst.
 - Naj bo $\lim_{n \rightarrow \infty} P(X_n = U_n Y_n) = 1$
- Torej gre $Y_n \xrightarrow{d} \frac{1}{c} \cdot X$
-

Dokaz: Uporabimo 2.22

$$M := \mathbb{R}^m \times \mathbb{R} \quad M' := \mathbb{R}^m \quad \tilde{M} := \mathbb{R}^m \times (\mathbb{R} - \{0\})$$

Po izreku 2.23 gre $(X_n, U_n) \rightarrow (X, c)$

$$\tilde{g}: \tilde{M} \rightarrow \mathbb{R}^m \quad (x, u) \mapsto \frac{1}{u} x$$

$$\{(x_n, u_n) \in \tilde{M} \mid Y_n = g(X_n, U_n)\} = \{U_n \neq 0, X_n = U_n Y_n\}$$

ker je $\tilde{M}$ odprta in (X, c) zavzame vrednosti

v $\tilde{M}$ je po opazki 2.10 je

$$\lim_{n \rightarrow \infty} P((X_n, U_n) \in \tilde{M}) = \lim_{n \rightarrow \infty} P(U_n \neq 0) = 1$$

Sled: $P((X_n, U_n) \in \tilde{M}, Y_n = g(X_n, U_n)) \rightarrow 1$

in pogoji izreka 2.23 so izpolnjeni.

■ qed

Trditev 2.28

- Naj bodo $X_1, X_2, \dots$ in $U_1, U_2, \dots$ real. sluč. spr.
- Naj gre $X_n \xrightarrow{d} X$, $U_n \xrightarrow{d} c$, kjer je $c \in \mathbb{R} \setminus \{0\}$

Torej za

$$\forall a \in \mathbb{R}. P\left(\frac{X}{c} = a\right) = 0 \Rightarrow$$

$$\lim_{n \rightarrow \infty} P(U_n \neq 0, \frac{X_n}{U_n} < a) = \lim_{n \rightarrow \infty} P(U_n \neq 0, \frac{X_n}{U_n} \leq a) = P\left(\frac{X}{c} \leq a\right)$$

$$\lim_{n \rightarrow \infty} P(U_n \neq 0, \frac{X_n}{U_n} > a) = \lim_{n \rightarrow \infty} P(U_n \neq 0, \frac{X_n}{U_n} \geq a) = P\left(\frac{X}{c} \geq a\right)$$

Dokaz:

Uporabimo trditev 2.27 tako, da najprej

$$\text{postavimo } Y_n := \begin{cases} \frac{X_n}{U_n} & ; U_n \neq 0 \\ a+1 & ; U_n = 0 \end{cases}$$

↑ neko število, ki je večje od a

$$\text{Tako je } \{U_n \neq 0, \frac{X_n}{U_n} < a\} = \{Y_n < a\}$$

Po opazki 2.10 velja

$$\lim_{n \rightarrow \infty} P(U_n \neq 0) = \lim_{n \rightarrow \infty} P(U_n \neq 0, Y_n = \frac{X_n}{U_n}) = 1$$

Po trditvi 2.27 gre potem $Y_n \xrightarrow{d} \frac{X}{c}$ po

definiciji II. Po Hely-Brayevem izreku pa gre

$Y_n \xrightarrow{d} \frac{X}{c}$ tudi po definiciji I, torej velja

$$\lim_{n \rightarrow \infty} P(Y_n < a) = \lim_{n \rightarrow \infty} P(Y_n \leq a) = P\left(\frac{X}{c} \leq a\right)$$

Uporabimo

Drugo konvergenco dobimo, da X_1, X zamenjamo

$$\text{z } -X_n, -X$$

Mehak uvod v sklepno statistiko

Verjetnostno vprašanje: $P(A)$

Vprašanje sklepnje statistike:

Opazimo da se je zgodil dogodek A .

P_θ ; $\theta \in \Theta$

Kej lahko vsej približno povemo o θ

Thomas Bayes (ok 1701-1761)

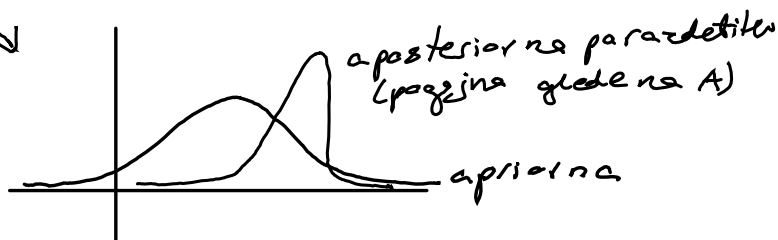
θ naj bo slučajna spremenljivka z neko

apriorna porazdelitev

$$P_+(A) = P(A|\theta=t)$$

Sir Ronald Aymler Fisher (1890-1962)

Sprejmemo šibkejše rezultate, ki so porazdelitev nečesa v zvezi z A in θ

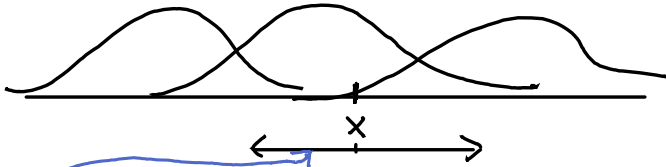


Torej porazdelitev boljše poznamo, ko vemo A

Zajed:

$$P = P_\mu \quad \mu \in \mathbb{R}$$

Recimo da pri P_μ velja $X \sim N(\mu, 1)$



cenilka / ocena za μ po metodi najvećjega
verjetja : $\hat{\mu} = x$:

pri čemer x je gostota največja za $\mu = x$

Sklepamo da je μ najbrže nekje na intervalu
zaupanja

Stopnja tveganja α : maksimalna dopustna
verjetnost zmot, tipično 5%

Naj bodo $x_1, x_2, \dots$ neodvisne in enako porazdeljene z $E(X_1) = \mu$ in $\text{var}(X_1) = \sigma^2$ $\sigma > 0$

Zanima nas μ , ki ni znan, želim pa so

$x_1, x_2, \dots, x_n$ - opažene vrednosti

Ena možnost je, da je znan tudi σ

$$\hat{\mu} = \bar{X} = \frac{x_1 + \dots + x_n}{n}$$

(izhaja iz zakona velikih števil $\bar{X} \xrightarrow[n \rightarrow \infty]{d} \mu$)

$\bar{X}$ je dosledna cenilka za μ (dobivemo vedno več informacij za μ (siloka doslednost))

Za koliko si upamo reči, da $\bar{X}$ odstopa od μ

želimo sklepati, da je $\bar{X} - \delta < \mu < \bar{X} + \delta$

Cilj: najti čim manjši δ za katerega bo $P(|\bar{X} - \mu| \geq \delta) < \alpha$
zmota

①

Če vemo, da je $X_1 \sim N(\mu, \sigma^2)$ in poznamo σ ,
vemo:

$$\bar{X} - \mu \sim N\left(0, \frac{\sigma^2}{n}\right)$$

$$\parallel$$

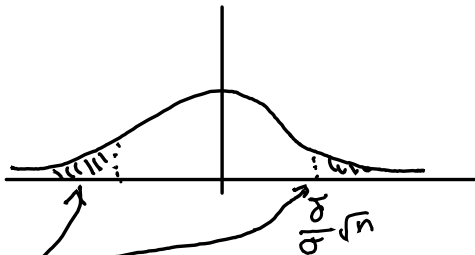
$$\frac{n\sigma^2}{n^2}$$

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

$$P(|\bar{X} - \mu| > \sigma) = P(|Z| > \frac{\sigma}{\sigma} \sqrt{n}) = 2P(Z > \frac{\sigma}{\sigma} \sqrt{n}) =$$

$$2(1 - \Phi(\frac{\sigma}{\sigma} \sqrt{n})) \leq \alpha$$

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



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

kritično območje

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

$$\alpha = 0,05 : \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 1,96$$

Interval zaupanja za μ je $\left[\bar{X} - \frac{\sigma}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right), \bar{X} + \frac{\sigma}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \right]$

Ta interval zaupanja je **eksakten**

② Če nismo prepričani, da so $x_1, \dots, x_n$ porazdeljene normalno, nam na pomoč priskoči centralni limitni izrek. Pri velikih n bo

$$P(|\bar{X} - \mu| \geq \frac{\sigma}{\sqrt{n}} \Phi^{-1}(1 - \frac{\alpha}{2})) \approx \alpha$$

α je zdaj le še **nominalna** stopnja tveganja, konstruirani interval zaupanja pa je **asimptotično eksakten**.

③ Kaj pa, če σ ne poznamo?
Ideja: Nadomestimo ga s prikol:škom

Ideja za oceno σ izhaja iz *empirične*
porazdelitve

$$\text{Emp}(x_1, x_2, \dots, x_n) := \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{pmatrix},$$

če so si $x_1, \dots, x_n$ različni, če pa

niso, pa ustrezne vrjetnosti seštejamo

$$\text{Npr: } \text{Emp}(2, 1, 2, 3, 7) = \begin{pmatrix} 1 & 2 & 3 & 7 \\ \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}$$

$$\bar{x} = \frac{x_1 + \dots + x_n}{n} = \mathbb{E}(W), \text{ kjer je } W \sim \text{Emp}(x_1, \dots, x_n)$$

Povprečje je empirična pričakovana vrednost
empirični standardni odklon pa je

$$\sigma_{\text{emp}} = \sqrt{\text{var}(W)} = \sqrt{\frac{(x_1 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}}$$

Cenilka za σ :

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Ali je cenilka na mestu?

$$\hat{\sigma} \xrightarrow[n \rightarrow \infty]{} \sigma$$

Iz zakona velikih števil sledi

$$\frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \dots + (x_n - \mu)^2}{n} \xrightarrow[n \rightarrow \infty]{} \sigma^2$$

$$\text{var}(W) = \text{var}(W - \mu)$$

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

1. zakon velikih števil
2. trditvi 2.21

Po izreku slutskega gre tudi

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

v okviru $\mathbb{R}$,
po opaz. 2.9 pa tudi v okviru $[0, \infty)$

Spet po trditvi 2.21 gre tudi $\hat{\sigma} \xrightarrow[n \rightarrow \infty]{} \sigma$

$$\frac{\hat{\sigma}}{\sigma} \rightarrow 1$$

zdi se:

$$\frac{\bar{X} - \mu}{\hat{\sigma}} = \frac{\frac{\bar{X} - \mu}{\sigma}}{\frac{\hat{\sigma}}{\sigma}} \xrightarrow[n \rightarrow \infty]{} N(0, 1)$$

slutski (2.27)
↓

Torej je $(\bar{X} - \frac{\hat{\sigma}}{\sqrt{n}} \Phi^{-1}(1 - \frac{\alpha}{2}), \bar{X} + \frac{\hat{\sigma}}{\sqrt{n}} \Phi^{-1}(1 - \frac{\alpha}{2}))$
asimptotično eksakten interval zaupanja
konvergira proti α ,
ko gre $n \rightarrow \infty$

Kaj če je $\hat{\sigma} = 0$? (ko so vse opazene vrednosti enake) Interval zaupanja ima še vedno smisel, ampak konvergenca pa nima pomena

μ je izven konstruiranega IZ v naslednjih primerih:

- * $\hat{\sigma} = 0$
- * $\hat{\sigma} \neq 0$ $\frac{\bar{X} - \mu}{\hat{\sigma}} \sqrt{n} \leq -\Phi^{-1}(1 - \frac{\alpha}{2}) = -\Phi^{-1}(\frac{\alpha}{2})$
- * $\hat{\sigma} \neq 0$ $\frac{\bar{X} - \mu}{\hat{\sigma}} \sqrt{n} \geq \Phi^{-1}(1 - \frac{\alpha}{2})$

Ker gre $\hat{\sigma} \xrightarrow[n \rightarrow \infty]{} \sigma > 0$, je $\lim P(\hat{\sigma} = 0) = 0$.

po trditvi 2.28 pa je

$$\lim P(\hat{\sigma} \neq 0, \frac{\bar{X} - \mu}{\hat{\sigma}} \sqrt{n} \leq \Phi^{-1}(\frac{\alpha}{2})) = P(Z \leq \Phi^{-1}(\frac{\alpha}{2}))$$

če to sklepanje uporabimo za $-X_1, -X_2, \dots$,
pa dobimo še:

$$\lim P(\hat{\sigma} \neq 0, \frac{\bar{X} - \mu}{\hat{\sigma}} \sqrt{n} \geq \Phi^{-1}(1 - \frac{\alpha}{2})) = \frac{\alpha}{2}$$

3. IZVLEČEK OPISNE STATISTIKE

statistika : $f(\overbrace{x_1, \dots, x_n}^{\text{opažene vrednosti}})$

n - število opažanj ali *numerus*

Primeri statistik :

1. aritmetična sredina ali povprečje

$$\bar{x} = \frac{x_1 + \dots + x_n}{n} = E(\text{Emp}(x_1, \dots, x_n))$$

2. empirična varianca

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \text{var}(\text{Emp}(x_1, \dots, x_n))$$

3. vrstilne statistike

$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ je ustrezna

permutacija zaporedja opaženih vrednosti

$x_{(k)}$ je k -ta vrstilna statistika

$x_{(1)}$ = minimum , $x_{(n)}$ = maksimum

Definicija: **Kvantil** slučajne spremenljivke X oz njene porazdelitve za vrjetnost $\alpha \in (0, 1)$ je poljubno število x_α , za katerega je

$$\mathbb{P}(X < x_\alpha) \leq \alpha \leq \mathbb{P}(X \leq x_\alpha)$$

če je X porazdeljena zvezno je $\mathbb{P}(X \leq x_\alpha) = \alpha$

Opazka 3.1.

Če je gostota na nekem odprtem intervalu $I \subseteq \mathbb{R}$ strogo pozitivna, drugje pa je enaka 0, je $F_X: I \rightarrow (0, 1)$ bijektivna in velja

$$x_\alpha = F_X^{-1}(\alpha)$$

$\Phi^{-1}(0,975) = 1,96$ je kvantil standardne porazdelitve za vrjetnost 0,975

Definicija: **Kvantil za delež** α , ki pripada danim podatkom, je kvantil njihove empirične porazdelitve za verjetnost α

Primer:

$$\text{Emp}(1,1,2,3,5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ \frac{2}{5} & \frac{1}{5} & \frac{1}{5} & \frac{0}{5} & \frac{1}{5} \end{pmatrix}$$

kvantil za delež $\frac{1}{2}$ je 2)

$$\mathcal{P}(X < 2) \leq \frac{1}{2} \leq \mathcal{P}(X \leq 2)$$

$$\text{Recimo da } x < 2 \Rightarrow \mathcal{P}(X \leq x) \leq \mathcal{P}(X < 2) = \frac{2}{5} < \frac{1}{2}$$

$$\text{Recimo da } x > 2 \Rightarrow \mathcal{P}(X < x) \geq \mathcal{P}(X \leq 2) = \frac{3}{5} > \frac{1}{2}$$

zato je $x=2$ edini možni kvantil za $\frac{1}{2}$

Kvantil za delež $\frac{3}{5}$ so $x \in [2, 3]$

$$\frac{2}{5} = \mathcal{P}(X < 2) \leq \mathcal{P}(X < x) \leq \mathcal{P}(X < 3) = \frac{3}{5}$$

$$\frac{3}{5} = \mathcal{P}(X \leq 2) \leq \mathcal{P}(X < x) \leq \mathcal{P}(X \leq 3) = \frac{4}{5}$$

Trditev 3.2: Vsi možni kvantili dane verjetnostne porazdelitve za dano verjetnost $\alpha \in (0, 1)$ (in z njimi vsi možni kvantili danih podatkov za delež za α) tvorijo zaprt interval $[x_{\alpha}^-, x_{\alpha}^+]$, kjer je x_{α}^- spodnji kvantil x_{α}^+ , ki je lahko izrajen ($\hat{x}_{\alpha}$), ne more pa biti prazen.

Dokaz: DN

$X_{\frac{1}{2}} = m$ imenujemo **mediana**



$X_{\frac{1}{3}}$ in $X_{\frac{2}{3}}$ sta prvi in drugi **tercil**

$X_{\frac{1}{4}}, X_{\frac{2}{4}} = m, X_{\frac{3}{4}}$ so **kvartili**

$X_{0.1}, X_{0.2}, \dots, X_{0.9}$ so **decili**

$X_{0.01}, \dots, X_{0.99}$ so **centili**; ali **percentili**

Kvantilom radi pripisujemo arbitrarne vrednosti.

Za mediano je to $\frac{X_{\frac{1}{2}}^- + X_{\frac{1}{2}}^+}{2}$.

Definicija: **Interkvartilni razmik** (interquartile range ali mid fifty) je razlika med 1. in 3. kvartilom

$$IQR := X_{\frac{3}{4}}^+ - X_{\frac{1}{4}}^-$$

Aritmetična sredina in mediana sta **meri centralne tendence**.

IQR in standardni odklon sta **meri razpršenosti**

Frekvence

Frekvenca je število enot z določeno lastnostjo, ki je navadno vezana na določeno statistično spremenljivko.

Relativna frekvenca ali **delež** je frekvenca, deljena s številom podatkov

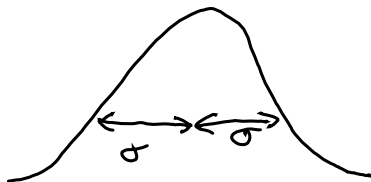
Podatke razvrstimo v razrede. Tako dobimo **histogram**

Kako široki naj bodo razredi?

Ideja: če podatki izhajajo iz neke zvezne porazdelitve, naj bo histogram čim bližje gostoti te porazdelitve.

$$\text{Širina} \approx \sigma \sqrt{\frac{24\sqrt{\pi}}{n}} \approx \frac{3,5\sigma}{\sqrt[3]{n}}$$

$$\text{širina} \propto \sigma \sqrt{\frac{24\sqrt{\pi}}{n}} \doteq \frac{3.5\sigma}{\sqrt[3]{n}}$$



$h \doteq$ širina stolpca
 $\nwarrow$ isti velikostni red

glavna paralelitve:

širina $\sim \sigma$

št. razredov $\sim \frac{\sigma}{h}$

enot $\sim n$

maksimalna pričakovana

frekvenca $\sim n : \frac{\sigma}{h} = \frac{nh}{\sigma}$

Tudi pričakovana

frekvenca za razred iz

glavnine

Posamezen podatek pade v razred iz

glavnine z verjetnostjo $\sim \text{Bin}(n, \frac{nh}{\sigma}) \sim \text{Pois}\left(\frac{nh}{\sigma}\right)$

Standardni odklon take frekven je $\sim \sqrt{\frac{nh}{\sigma}}$

= velikostni red
 suma

Pričakovana razlika med tipičnima sosednjima

razredoma $\sim \frac{nh}{\sigma} : \frac{\sigma}{h} = \frac{nh^2}{\sigma^2}$

$$\sqrt{\frac{nh}{\sigma}} \sim \frac{nh^2}{\sigma^2} \rightsquigarrow \frac{nh}{\sigma} \sim \frac{h^2 h^4}{\sigma^4} \rightsquigarrow h \sim \frac{\sigma}{\sqrt[3]{n}}$$

če izenacimo trend
 in sum do konstante

Pri: gaussovi porazdelitvi je

$$IQR = \sigma (\Phi^{-1}(\frac{3}{4}) - \Phi^{-1}(\frac{1}{4}))$$

Če v Scottovem pravilu σ nadomestimo

$$z \frac{IQR}{\Phi^{-1}(\frac{3}{4}) - \Phi^{-1}(\frac{1}{4})}, \text{ dobimo modificirano}$$

Freedman-Diaconisovo porazdelitev

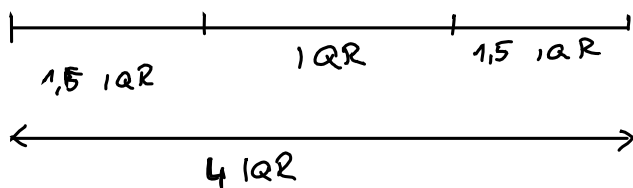
Vprašanje

Kaj naredi na nas vtis kolonine na histogramu? Ploščine

gostota frekvence $= \frac{\text{frekvenca}}{\text{širina razreda}}$

relativna gostota frekvence $= \frac{\text{frekvenca}}{\text{numerus} \cdot \text{širina razreda}}$

Splava se izločiti osamelce tj. vrednosti
vrednosti izven intervala $[x_{\frac{1}{4}} - \frac{3}{2}IQR, x_{\frac{3}{4}} + \frac{3}{2}IQR]$



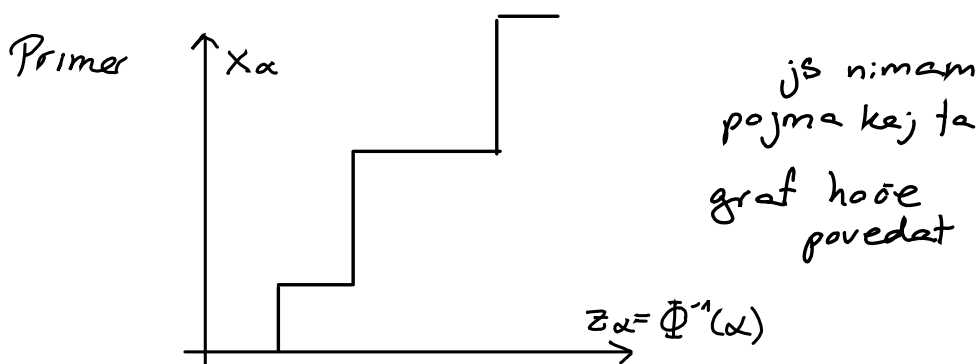
P (pri normalni porazdelitvi dobimo osamelce)

$$= 2 \cdot (1 - \Phi(4 \Phi^{-1}(\frac{3}{4}))) \approx 0,006977$$

Primerjalni kvantilni (Q-Q) grafikon

služi primerjavi dveh vrjetnostnih porazdelitev.

V najbolj "teoretični" različici ga sestavljajo vse možne točke (x, y) , kjer je x kvantil prve porazdelitve, y pa druge za $\alpha \in (0, 1)$



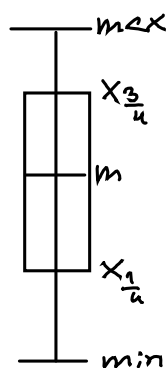
Dve zverni porazdelitvi se ujemata $\Leftrightarrow$

Njun primerjalni kvantilni grafikon je simetričen likih kvadrantov

Kadar pa empirično porazdelitev primerjamo z določeno teoretično porazdelitvijo navadno narišemo le kvantile za vrjetnosti $\frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}$. Kvantili za te vrjetnosti so vedno natančno določeni, enaki pa so ker $X_{(1)}, \dots, X_{(n)}$

Q-Q grafikon med dvema neizrojenima Gaussovima porazdelitvima je vedno premica. Poznamo Shapiro-Wilkov preizkus normalnosti

Škatla z brki



4. Ocenjevanje in napovedovanje

Opis problema:

Dobimo določene podatke za X , zanima nas Y

Primer: Problem trgovskega potnika

X = graf z uteženimi povezavami

Y = optimalni obhod

Matematični model $\rightarrow$ verjetnostni/stohastični model

$\downarrow$
statistični model

Stohastični model vsebuje slučaj: Vsaj ena od količin X in Y modeliramo kot slučajno.

Tako Y tipično ni natančno določen z X .

Statistični model običajno upošteva že ponavljanje

Era izmed glavnih nalog sklepane statistike je najti **opazljivo** količino $\hat{y}$, $\hat{y} = h(x)$, kjer je h merljiva, izračunljiva funkcija, da bo $\hat{y}$ čim bližje y .

V Bayesovem modelu sta X in Y slučajni spremenljivki na $(\Omega, \mathcal{F}, P)$

Če poznamo X in če Y živi v $\mathbb{R}$, je smiselna **napoved/prediktor** za Y slučajna spremenljivka $E(Y|X) = h(X)$

Diskreten primer: $\hat{y} = h(x)$ $h(x) = E(Y|X=x)$

Poznamo $P(X=x|Y=y)$. Privzamemo apriorno porazdelitev **napovedovalca/predikta** Y . Po Bayesovi formuli izračunamo $P(Y=y|X=x)$. Iz slednjih dobimo napoved $E(Y|X=x)$

Fisherjeva statistika na $(\Omega, \mathcal{F})$ pa gleda celo družino verjetnostnih mer $P_\vartheta, \vartheta \in \Theta$

Splošno: $X = F(\omega, \vartheta), Y = G(\omega, \vartheta)$

Dostikrat je Y karakteristika modula, torej

$Y = y = g(\vartheta)$. Tipično je $X = X(\omega)$

Primer:

Opazimo $X_1, \dots, X_n$, neodvisne, porazdeljene kot $N(\mu, \sigma^2)$

$$\Theta = \mathbb{R} \times (0, \infty) \quad \vartheta = (\mu, \sigma)$$

Ocenjujemo na primer:

$$\bullet g(\mu, \sigma) = \mu$$

$$\bullet g(\mu, \sigma) = \sigma^2$$

$$\bullet g(\mu, \sigma) = \mu^2$$

$$\bullet g(\mu, \sigma) = \mu + 2\sigma \quad (\text{Var})$$

Terminološko pojasnilo:

	deterministična	slučajna
poznamo	konstanta	opazanje
nepoznamo	karakteristika	predikat ali šum

če je predmet našega zanimanja karakteristika, potem namesto o predikatorju govorimo o **cenilki** preden opazimo podatke, in o **ceni** potem ko jih opazimo.

Primer: Vrnimo se k prejšnjemu primeru

$$X_k = \mu + \varepsilon_k \quad k \in \{1, \dots, n\}$$

$$\varepsilon_k \sim N(0, \sigma_k^2) \text{ so šum;}$$

smiselna cenilka za μ je

$$\hat{\mu} = \frac{x_1 + \dots + x_n}{n} = \bar{x}$$

Vrednote cenilk in prediktorjev

Naj bo $\hat{y} = h(X)$ cenilka za y
↖ slučajna spremenljivka

Definicija: Pričakovana ali srednja kvadratična napaka cenilke $\hat{y}$ karakteristike y je

$$MSE_v(\hat{y}|y) := E_v((\hat{y} - y)^2)$$

↖ mean squared error ↖ ne označuje pogojevanja ↖ pričakovana vrednost pri vrjetnostni meri P_θ

Definicija: Korenjena pričakovana kvadratična napaka cenilke $\hat{y}$ karakteristike y je:

$$RMSE_v(\hat{y}|y) := \sqrt{MSE_v(\hat{y}|y)}$$

Definicija: Pristranskost cenilke $\hat{y}$ za y je

$$Bias_v(\hat{y}|y) := E_v(\hat{y}) - y = E_v(\hat{y} - y)$$

Definicija: Cenilka je nepristranska, če za $\forall \theta \in \Theta$. $Bias(\hat{y}|y) = 0$

Opazka 4.1:

Če je cenilka nepristranska, je
 $MSE_{\hat{y}}(\hat{y}|y) = \text{var}_{\hat{y}}(\hat{y}) = \text{var}_{\hat{y}}(\hat{y} - y).$

V tem primeru označimo **standardno napako**

$$SE := \sqrt{\text{var}_{\hat{y}}(\hat{y})}$$

V splošnem pa velja:

$$\begin{aligned}\text{var}_{\hat{y}}(\hat{y}) &= \text{var}_{\hat{y}}(\hat{y} - y) = E_{\hat{y}}((\hat{y} - y)^2) - E(\hat{y} - y)^2 = \\ &= MSE_{\hat{y}}(\hat{y}|y) - (\text{Bias}_{\hat{y}}(\hat{y}|y))^2\end{aligned}$$

$$\Rightarrow MSE_{\hat{y}}(\hat{y}|y) = \text{var}_{\hat{y}}(\hat{y}) + (\text{Bias}(\hat{y}|y))^2$$

Definicija:

Zaporedje cenilk $\hat{y}_1, \dots$ je **šibko dosledno** za y , če v porazdelitvi konvergira proti y :

$$\forall \epsilon > 0. \hat{y}_n \xrightarrow[n \rightarrow \infty]{d_{1\theta}} y = y(v)$$

Trditev 4.2.

Br̄z ko je $\forall \vartheta \in \Theta. \lim_{n \rightarrow \infty} \text{MSE}_{\vartheta}(\hat{y}_n | y) = 0$, je
zaporedje cenilk $\hat{y}_1, \hat{y}_2, \dots$ šibko dosledno za y

Dokaz: Sledi iz trditve 2.18:

$$W_n \geq 0 \quad p > 0 \quad \lim_{n \rightarrow \infty} E(W_n^p) = 0 \Rightarrow W_n \xrightarrow[n \rightarrow \infty]{d} 0$$

Po posledici 2.16 velja:

$$X_n \xrightarrow[n \rightarrow \infty]{d} c \iff d(X_n, c) \xrightarrow[n \rightarrow \infty]{d} 0$$

Vstavimo:

$$X_n = \hat{y}_n, \quad c = y, \quad p = 2, \quad W_n = |\hat{y}_n - y|, \quad d(a, b) = |b - a|$$



Posledica: Če je zaporedje cenilk $\hat{y}_n$ za y
asimptotično nepristransko ($\forall \vartheta \in \Theta. \lim_{n \rightarrow \infty} E_{\vartheta}(\hat{y}) = y$)
in če tudi: $\forall \vartheta \in \Theta. \lim_{n \rightarrow \infty} \text{var}_{\vartheta}(\hat{y}_n) = 0$, je to
zaporedje šibko dosledno za y .

Primer:

Naj bodo $X_1, X_2, \dots$ enako porazdeljene s pričakovano vrednostjo μ in varianco σ^2 .

$$\mu = E(X_1) = g(\theta)$$

Tedaj je $\hat{\mu} = \bar{X}$ nepristranska cenilka za μ .

N: pa nujno šibko dosledna. Lahko je denimo

$$X_1 = X_2 = \dots \quad \text{in } \sigma > 0$$

Če pa prevzememo, da so $X_1, \dots$ nekorelirane pa je $\hat{\mu}$ tudi šibko dosledna.

$$MSE(\bar{X} | \mu) = \text{var}(\bar{X}) = \frac{\sigma^2}{n}$$

$$RMS(\bar{X} | \mu) = SE(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

Ke gre $n \rightarrow \infty$ gre $MSE \rightarrow 0$, torej je $\hat{\mu}$ šibko dosledna za μ

Če so $X_1, X_2, \dots$ neodvisne pa je po zakonu velikih števil (opazka 2.20 (1)) to velja tudi brez predpostavke o končnem drugem momentu. Zaporedje cenilk $X_1, X_2, \dots$ pa ni šibko dosledno za μ

V resnici je $\bar{X}$ najboljša nepristranska linearna cenilka (**NNLC**, bes linear unbiased estimator, **BLUE**) za μ .

Če je μ fiksna in so $c_1, \dots, c_n$ take konstante, da je $\hat{\mu} = c_1 X_1 + \dots + c_n X_n$ nepristranska cenilka za μ , velja:

$$\text{MSE}(\hat{\mu} | \mu) \geq \text{MSE}(\bar{X} | \mu)$$

$\text{var}(\hat{\mu})$

Za slednje je dovolj privzeti nekoreliranost

Če pa je model Gaussov ($X_1, X_2, \dots$ neodvisne in porazdeljene kot $N(\mu, \sigma^2)$), ter velja še eden izmed naslednjih pogojev:

- 1) σ je konstanten, torej znan
- 2) μ in σ se spreminjata neodvisno drug od drugega

Potem je $\bar{X}$ tudi najboljša nepristranska cenilka za μ nesploh. (Za tem stoji Rao-Cramérjeva ocena)

če dopustimo ker vse cenilke, pa tipično nobena ni enakomerno najboljša

če ocenjujemo μ^2 je tipično $\bar{X}^2$ pristranska, saj

$$E(\bar{X}^2) = \text{var}(\bar{X}) + E(\bar{X})^2 = \frac{\sigma^2}{n} + \mu^2$$

je pa asimptotično nepristranska:

$$\text{Bias}(\bar{X}^2 | \mu^2) = \frac{\sigma^2}{n} \xrightarrow{n \rightarrow \infty} 0 \text{ za } \forall \theta \in \Theta$$

je tudi šibko nedosledna (po teoremi 2.21:

če gre $X_n \xrightarrow{d} X$ in g zveza gre $g(X_n) \xrightarrow{d} g(X)$

Splošni primer, ko je $\hat{y} = h(x)$ prediktor za $y = G(w, \theta)$:

$$\text{Ostane } MSE_{\theta}(\hat{y} | y) = E_{\theta}((\hat{y} - y)^2).$$

Če za $\forall \theta \in \Theta$. $E_{\theta}(\hat{y}) = E_{\theta}(y)$, je nujno $MSE_{\theta}(\hat{y} | y) = \text{var}_{\theta}(\hat{y} - y)$, tipično pa to ni več enako $\text{var}_{\theta}(\hat{y})$

Primer:

Naj bodo $X_1, X_2, \dots$ nekorelirane in enako porazdeljene s pričakovano vrednostjo μ in varianco σ^2 .

$$X_i = \mu + \varepsilon_i$$

$\varepsilon_1, \varepsilon_2, \dots$ tvorijo **beli šum**, kar pomeni, da so nekorelirane s pričakovano vrednostjo 0 in enakimi variancami

Opazimo: $X = (X_1, \dots, X_n)$, zanimamo se $y = X_{n+1}$
 $y = \mu + \varepsilon_{n+1}$

Prediktor za y : $\hat{X}_{n+1} = \hat{y} = \frac{X_1 + \dots + X_n}{n} = \bar{X}$

Ta prediktor je nepristranski v smislu
 $E(\hat{X}_{n+1}) = E(X_{n+1})$

$$\begin{aligned} \text{MSE} &= \text{var}(\bar{X} - X_{n+1}) = \text{var}(\bar{X}) + \text{var}(X_{n+1}) = \\ &= \frac{\sigma^2}{n} + \sigma^2 = \sigma^2 \left(1 + \frac{1}{n}\right) \end{aligned}$$

5. Pričakovana vrednost in varianca večrazsežnih slučajnih objektov

Definicija: Za slučajni vektor $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$
definiramo pričakovano vrednost

$$E(X) = \begin{bmatrix} E(x_1) \\ \vdots \\ E(x_n) \end{bmatrix}$$

Trditev 5.1

Za deterministično matriko A in
slučajni vektor X kompatibilnih
dimenzij velja $E(AX) = AE(X)$

Dokaz: DV

Posledica 5.2

Če je u determinističen, X pa slučajni vektor
velja: $E(\langle X, u \rangle) = \langle E(X), u \rangle$
 $E(u^T X) = u^T E(X)$

Definicija: Za slučajno matriko $M = [M_{ij}]_{ij}$
definirajmo $E(M) := [E(M_{ij})]_{ij}$

Trditev 5.3:

Naj bo M slučajna matrika.

1) Velja $E(M^T) = E(M)^T$

2) Za poljubni deterministični matriki

A in B primernih velikosti, velja:

$$E(AM) = A E(M) \text{ in } E(MB) = E(M) B$$

Dokaz: DV

Trditev 5.4:

Za poljubne enako velike slučajne vektorje

$$E(X+Y) = E(X) + E(Y), \text{ ter za poljubni enako}$$

veliki slučajni matriki velja

$$E(M+N) = E(M) + E(N)$$

Dokaz: DV

Definicija: **Kovariančna matrica** vektorja X je

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

Kovariančna matrica med slučajnima vektorjema X, Y je

$$\text{cov}(X, Y) = \begin{bmatrix} \text{cov}(x_1, y_1) & \dots & \text{cov}(x_1, y_n) \\ \vdots & \ddots & \vdots \\ \text{cov}(x_n, y_1) & \dots & \text{cov}(x_n, y_n) \end{bmatrix}$$

Opazka 5.5: $\text{var}(X) = \text{cov}(X, X)$

Trditev 5.6:

$$\begin{aligned}\text{cov}(X, Y) &= \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))^T) = \\ &= \mathbb{E}(XY^T) - \mathbb{E}(X)\mathbb{E}(Y)^T\end{aligned}$$

Dokaz: $(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))^T =: [M_{ij}]_{ij}$

$$M_{ij} = (X_i - \mathbb{E}(X_i))(Y_j - \mathbb{E}(Y_j))$$

$$\Rightarrow \mathbb{E}(M_{ij}) = \text{cov}(X_i, Y_j)$$



Posledica 5.7:

Kovariančna metrika posamičnega slučajnega vektorja je simetrična, saj je slučajna metrika $(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))^T$ simetrična

Dokaz: $\square$

Opazka 5.8

- X, Y slučajna, vektora A, B deterministični matriki
velja: $\text{cov}(AX, BY) = E((AX - E(AX))(BY - E(BY))^T) =$
 $= E(A(X - E(X))(Y - E(Y))^T B^T) =$
 $A E((X - E(X))(Y - E(Y))^T) B^T =$
 $= A \text{cov}(X, Y) B^T$

Posledično je tudi $\text{var}(AX) = A \text{var}(X) A^T$

- $0 \leq \text{var}(\langle X, u \rangle) = \text{var}(u^T X) = u^T \text{var}(X) u =$
 $= \langle \text{var}(X) u, u \rangle$

Posledica 5.9: $\text{var}(X)$ je pozitivno semidefinitna

Toditev 5.10

Dana metrika je kovariančna metrika nekega slučajnega vektorja $\Leftrightarrow$ ko je pozitivno semidefinitna

Dokaz: ($\Rightarrow$) Posledica 5.9

($\Leftarrow$) Označimo dano metriko Σ in naj bo velikosti $n \times n$.

Vzemimo $Z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$, kjer so $z_1, \dots, z_n$ neodvisne in $z_i \sim N(0, 1)$. $\text{Var}(Z) = I_n$ seveda.

Oglejmo si slučajni vektor $X = AZ$.

Po opazki 5.8 je $\text{var}(X) = AA^T$.

Slučajni vektor X bo zadoščal zahtevi, če bo $AA^T = \Sigma$.

Ustrezna metrika pa obstaja, ker je Σ pozitivno semidefinitna (Cholesky

Posledica 5.11

Ze $\forall \mu \in \mathbb{R}^n$ in $\forall$ pozitivno semidefinitno metriko $\Sigma \in \mathbb{R}^{n \times n}$, $\exists$ slučaj. vekt. X z $E(X) = \mu$ in $\text{Var}(X) = \Sigma$

Opazka 5.12.

1) Za bločni slučajni vektor $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ je

$$E(X) = \begin{bmatrix} E(X_1) \\ E(X_2) \end{bmatrix}$$

2) Za bločno slučajno metriko

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}, \text{ je } E(M) = \begin{bmatrix} E(M_{11}) & E(M_{12}) \\ E(M_{21}) & E(M_{22}) \end{bmatrix}$$

3) Za bločne slučajne vektorja

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \text{ in } Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \text{ je } \text{cov}(X, Y) = \begin{bmatrix} \text{cov}(X_1, Y_1) & \text{cov}(X_1, Y_2) \\ \text{cov}(X_2, Y_1) & \text{cov}(X_2, Y_2) \end{bmatrix}$$

Če sta X in y enake velike: $\text{var}(x+y) = ?$

$$X+y = \begin{bmatrix} I_n & I_m \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leadsto \text{vsota lahko zapišemo z bloki}$$

Opazka 5.13.

$$\begin{aligned} \text{var}(X+y) &= \text{var} \left(\begin{bmatrix} I_n & I_m \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right) = \\ &= \begin{bmatrix} I_n & I_m \end{bmatrix} \text{var} \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) \begin{bmatrix} I_n \\ I_m \end{bmatrix} = \\ &= \begin{bmatrix} I_n & I_m \end{bmatrix} \begin{bmatrix} \text{var}(x) & \text{cov}(x,y) \\ \text{cov}(y,x) & \text{var}(y) \end{bmatrix} \begin{bmatrix} I_n \\ I_m \end{bmatrix} = \\ &= \text{var}(x) + \text{cov}(x,y) + \text{cov}(y,x) + \text{var}(y) \end{aligned}$$

Trditev 5.14

$$\text{cov}(y, x) = (\text{cov}(x, y))^T$$

Dokaz: DN

Definicija: Slučajna vektorja x in y sta **nekorelirana**, če je $\text{cov}(x, y) = 0$

Opazka 5.15

Poljubna neodvisna slučajna vektorja sta nekorelirana, če le obstajata njuni pričakovani vrednosti

Posledica 5.16

Če sta slučajna vektorja x in y enake velikosti in neodvisna, ter imata kovariančni matriki, velja: $\text{var}(x+y) = \text{var}(x) + \text{var}(y)$

6. Pridobivanje cenilk

(logično nadaljevanje 4. poglavja)

6.1 Metoda empirične porazdelitve

Opazimo vrednosti $x_1, x_2, \dots, x_n$, ki so enako porazdeljene: predstavljajo **vzorec** neke porazdelitve. Preučevana količina y nas se izraža kot karakteristika te porazdelitve

$$y = g(\theta) = \text{Char}_{\theta}(X_1) = \text{Char}(\text{porazdelitev za } X_1)$$

$\uparrow$ parameter porazdelitve

Tipična primera:

1) $\text{Char}_{\theta}(X_1) = E_{\theta}(X_1)$

2) $\text{Char}_{\theta}(X_1) = \text{var}_{\theta}(X_1)$

Cenilke za y je potem $\hat{y} = \text{Char}(\text{Emp}(x_1, \dots, x_n))$

1) $y = \mu = E_{\theta}(X_1)$: $\hat{y} = \hat{\mu} = \frac{x_1 + \dots + x_n}{n} = \bar{x}$

2) $y = \text{var}_{\theta}(X_1)$: $\hat{y} = \frac{(x_1 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$

kako dobri sta ti dve cenilki?

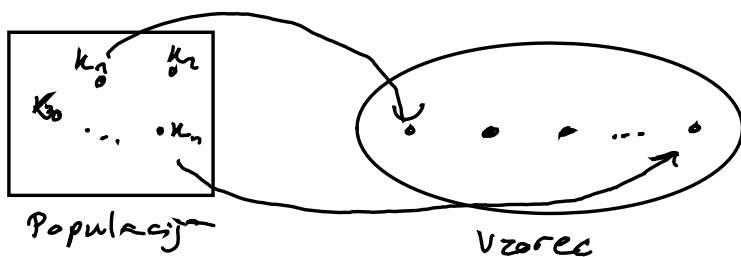
1) $\bar{x}$ je nepristranska. koliko je standardna napaka?

1.1) X_i nekorelirane, je $SE = \sqrt{\text{var}_e(\bar{x})} = \frac{\sigma}{\sqrt{n}}$

1.2) enostavno slučajno vzorčenje:

Na populaciji z N enotami: naj bo definirana statistična spremenljivka, brez funkcija $[N] \rightarrow \mathbb{R}: i \mapsto x_i$

Iz populacije vzamemo vzorec, ki naj ga sestavljajo enote z indeksi $k_1, \dots, k_n \in [N]$. Le ti so izbrani naključno.



O enostavnem slučajnem vzorčenju govorimo, če so vsi možni vzorci brez ponavljanja enako verjetni, tj. če je slučajni vektor $(k_1, \dots, k_n)$ porazdeljen enakomerno na množici $\{k_1, \dots, k_n; k_i \in [N], k_i \text{ vsi različni}\}$

$$X_i = x_{k_i} \quad i \in [n]$$

$$\text{var}(\bar{X}) = ?$$

$$X_1, \dots, X_N \sim \text{Emp}(x_1, \dots, x_N)$$

$$\mu := E(X_1) = \frac{x_1 + \dots + x_N}{N}$$

$$\sigma^2 := \text{var}(X_1) = \frac{(x_1 - \mu)^2 + \dots + (x_N - \mu)^2}{N}$$

$$SE^2 = \text{var}(\bar{X}) \quad SE^2 = \text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right)$$

$$= \frac{1}{n^2} \sum_{i,j} \text{cov}(X_i, X_j)$$

$$i=j \Rightarrow \text{cov}(X_i, X_j) = \sigma$$

$$i \neq j \Rightarrow$$

$$E(X_i X_j) = E(X_{K_i} X_{K_j}) =$$

$$= \sum_{k,l} x_k x_l \underbrace{P(K_i = k, K_j = l)}_{\begin{cases} \frac{1}{N(N-1)} & ; k \neq l \\ 0 & ; k = l \end{cases}}$$

$$= \sum_{k \neq l} x_k x_l \frac{1}{N(N-1)} =$$

$$= \frac{1}{N(N-1)} \sum_{k,l} x_k x_l - \frac{1}{N(N-1)} \sum_{k=1}^N x_k^2 =$$

$$= \frac{1}{N(N-1)} \left(\sum_{k=1}^N x_k \right)^2 - \frac{1}{N(N-1)} \sum_{k=1}^N x_k^2 =$$

$$= \frac{N}{N-1} \left(\frac{1}{N} \sum_{k=1}^N x_k \right)^2 - \frac{1}{N-1} \left(\frac{1}{N} \sum_{k=1}^N x_k^2 \right) =$$

$$= \frac{N}{N-1} E(X_1)^2 - \frac{1}{N-1} E(X_1^2) =$$

$$= \frac{1}{N-1} \sigma^2 + N\mu$$

$$\Rightarrow \text{cov}(X_i, X_j) = \frac{-\sigma^2}{N-1} \quad \text{za } i \neq j$$

Potem je $SE^2 = \frac{1}{n^2} (n\sigma^2 - n(n-1) \cdot \frac{1}{N-1} \sigma^2)$ za $i=j$

$$= \frac{\sigma^2}{n} \left(1 - \frac{n-1}{N-1} \right) = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)$$

korekcijski faktor
popravek za končne populacije $\uparrow$
(korekcijski člen) $\leftarrow$

Opomba: Tudi pri enostavnem slučajnem vzorcu se izkaže, da je $\bar{X}$ NNLC za μ .

② $\text{Cher} = \text{var}$

$$y = \text{var}(X_1) = \sigma^2$$
$$\hat{\sigma}^2 = \frac{(X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n}$$

Je ta cenika nepristranska?

koliko je torej $E(\hat{\sigma}^2)$?

$$\hat{\sigma}^2 = \frac{1}{n} \left\| \begin{bmatrix} x_1 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{bmatrix} \right\|^2 \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad 1_n = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$H = I_n - \frac{1}{n} 1_n 1_n^T$$

$$\bar{X} = \frac{1}{n} \cdot 1_n^T X$$

$$\hat{\sigma}^2 = \frac{1}{n} \|X - \frac{1}{n} 1_n 1_n^T X\|^2 = \frac{1}{n} \|HX\|^2 =$$
$$= \frac{1}{n} (HX)^T (HX) = \frac{1}{n} X^T H^T H X$$

Zaloga vrednosti matrice $1_n 1_n^T$?

$\frac{1}{n} 1_n (1_n^T X)$ ima vse komponente enake

$$X = \alpha 1_n: \frac{1}{n} 1_n 1_n^T X = \alpha \frac{1}{n} 1_n (\underbrace{1_n^T 1_n}_n) = \alpha 1_n = X$$

$1_n 1_n^T$ je torej projektor na podprstor
vektorjev, ki imajo vse komponente enake
 $\text{rang}(\frac{1}{n} 1_n 1_n^T) = 1$

$$(\frac{1}{n} 1_n 1_n^T)^T = \frac{1}{n} 1_n 1_n^T$$

Ore torej za ortogonalni projektor.

Torej je tudi H ortogonalni projektor
ranga $n-1$

$$\hat{\sigma}^2 = \frac{1}{n} X^T H X$$

$X_1, \dots, X_n$ enako porazdeljene

$$\mathbb{E}(X_1) = \mu, \quad \text{var}(X_1) = \sigma^2 \quad \sigma > 0$$

Cenilka za μ po metodi empirične porazdelitve:

$$\hat{\mu} = \bar{X} = \frac{1}{n} \mathbf{1}^T \underline{X} \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \mathbf{1}_n = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Cenilka za σ^2 po metodi empirične porazdelitve:

$$\underline{H} := I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \quad (\text{ortogonalna, range } n-1)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 = \frac{1}{n} X^T H X$$

$$H^2 = H^T = H$$

$$\text{Vemo: } \text{Var}(X) = \mathbb{E}(X X^T) - \mathbb{E}(X) \mathbb{E}(X)^T$$

$$\text{Točka 1: } \hat{\sigma}^2 = \frac{1}{n} \text{sl}(X^T H X) = \frac{1}{n} \text{sl}(H X X^T)$$

$$\text{Točka 2: } y_i := x_i - \mu, \quad y = X - \mu \mathbf{1}_n$$

$$\bar{y} = \frac{1}{n} \sum y_i = \bar{X} - \mu$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum (y_i - \bar{y})^2 = \frac{1}{n} y^T H y = \frac{1}{n} \text{sl}(H y y^T)$$

$$\mathbb{E}(\hat{\sigma}^2) = \frac{1}{n} \text{sl}(H \mathbb{E}(y y^T)) = \frac{1}{n} \text{sl}(H \text{var}(X))$$

i) $X_1, \dots, X_n$ nekorrelirane $\Rightarrow \text{var}(X) = \hat{\sigma}^2 I_n$ in

$$\mathbb{E}(\hat{\sigma}^2) = \frac{1}{n} \text{sl}(H) \sigma^2 = \frac{1}{n} \text{sl}(I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T) \sigma^2 = \frac{n-1}{n} \sigma^2$$

$\Rightarrow$ Popravljen vzorecna varianca je

$$\hat{\sigma}_+^2 = \frac{n}{n-1} \hat{\sigma}^2 = \frac{1}{n-1} \sum (x_i - \bar{X})^2 \quad \text{in je}$$

nepistranska za σ^2

Opomba: $\hat{\sigma}_+$ je tipično pristranska za σ

ii) Enostavno slučajno vzorčenje

Najbolj na vajah

Spomnimo se: standardna napaka nepristranske
cenilke $\bar{X}$ za μ : $SE = \frac{\sigma}{\sqrt{n}}$

$\hat{SE} = \frac{\hat{\sigma}}{\sqrt{n}}$ je cenilka za SE

$\hat{SE}^2 = \frac{\hat{\sigma}^2}{n} = \frac{1}{n} \sum (X_i - \bar{X})^2$ je cenilka za SE^2

$\hat{SE}_+^2 = \frac{\hat{\sigma}_+^2}{n} = \frac{1}{n(n-1)} \sum (X_i - \bar{X})^2$ je nepristranska
cenilka za SE^2

Metoda momentov je poseben primer metode
empirične razdelitve:

Teoretični moment $m_k = E(X_1^k)$ ocenimo
z empiričnim momentom

$$\hat{m}_k = \frac{1}{n} \sum X_i^k = m_k(\text{Emp}(X_1, \dots, X_n))$$

Proučevano količino Y izrazimo z momenti:

$$Y = g(m_1, \dots, m_k) \text{ in ocenimo z}$$

$$\hat{Y} = g(\hat{m}_1, \dots, \hat{m}_k)$$

Primer: $\hat{\sigma}^2 = m_2 - m_1^2$

$$\hat{\sigma}^2 = \hat{m}_2 - \hat{m}_1^2 = \frac{1}{n} \sum X_i^2 - \left(\frac{1}{n} \sum X_i\right)^2 = \frac{1}{n} \sum (X_i - \bar{X})^2$$

Metoda največjega verjetja

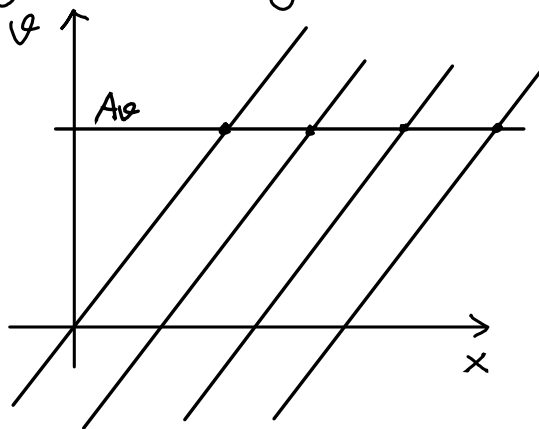
Definicija: **Verjetje** (likelihood) je verjetnostna funkcija ali gostota opažanja, ki jo gledamo kot funkcijo parametra modela.

Naj opažanje zavzame vrednosti v množici A , neodvisni od parametra ϑ in privzamemo še enega od naslednjih pogojev:

(D) A je števna in verjetje $L(\vartheta|x) = P_{\vartheta}(X=x) > 0$ za $\forall x \in A$

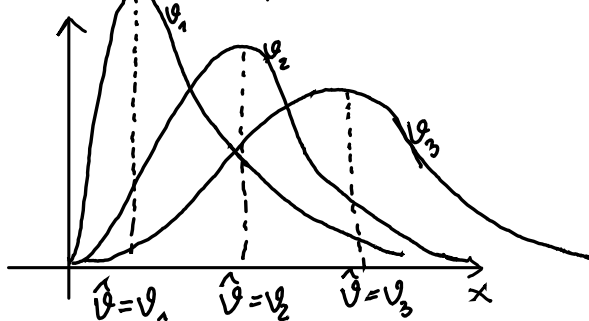
(Z) A je odprta podmnožica prostora $\mathbb{R}^n$, opažanje $\underline{X}$ je porazdeljeno zvezno in $L(\vartheta|x) = f_x(x|\vartheta) > 0$ za $\forall x \in A$

Kaj ne sme veljati:



Definicija: cenilka $\hat{\vartheta} = h(X)$ za ϑ po metodi največjega verjetja (CMNV) je tista cenilka, pri katerem $L(\vartheta|x)$ pri vsakem $x \in A$ doseže maksimum pri $\vartheta = h(x)$

Primer: $\Theta = \{\vartheta_1, \vartheta_2, \vartheta_3\}$



če je $\hat{\vartheta}$ cenilka za ϑ , je $g(\hat{\vartheta})$ cenilka za $g(\vartheta)$
Namesto L radi minimiziramo $l = \ln L$

Opažanje je čisto vektor iz neodvisnih in
enako porazdeljenih slučajnih spremenljivk
 $X_1, \dots, X_n$.

Verjetje za posamezno opažanje:

$$L_1(\vartheta|x) = P_{\vartheta}(X_1=x) \text{ ali } f_{X_1}(x|\vartheta)$$

$$L(\vartheta|x_1, \dots, x_n) = L_1(\vartheta|x_1) L_2(\vartheta|x_2) \dots L_1(\vartheta|x_n)$$

$$l_1(\vartheta|x) := \ln L(\vartheta|x)$$

$$l(\vartheta|x_1, x_2, \dots, x_n) = l_1(\vartheta|x_1) + l_2(\vartheta|x_2) + \dots + l_1(\vartheta|x_n)$$

Primer: Opazimo $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ neodvisne
 $\Theta = \mathbb{R} \times (0, \infty)$

$$\vartheta = (\mu, \sigma) \in \Theta$$

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

$$\ell_1(\mu, \sigma | x) = -\ln \sigma - \ln \sqrt{2\pi} - \frac{(x-\mu)^2}{2\sigma^2}$$

$$\ell(\mu, \sigma | x_1, \dots, x_n) = -n \ln \sigma - n \ln \sqrt{2\pi} - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

Fiksirajmo $x_1, \dots, x_n$. Pri katerih μ in σ je to maksimalno?

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{2\sigma^2} 2 \sum_{i=1}^n (x_i - \mu) = \frac{1}{\sigma^2} \left(\sum_{i=1}^n x_i - n\mu \right) = 0$$

$$\Rightarrow \mu = \frac{1}{n} \sum_{i=1}^n x_i \Rightarrow \hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\frac{\partial \ell}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2 = 0 \quad / \cdot \sigma^3$$

$$\Rightarrow n\sigma^2 = \sum_{i=1}^n (x_i - \mu)^2$$

$$\Rightarrow \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

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

V tem primeru (pri splešni porazdelitvi pa ne nujno) dobimo isti cenilki kot pri metodi empirične porazdelitve.

$\Theta \subseteq \mathbb{R}^n$ otvorna, $\vartheta = (\vartheta_1, \dots, \vartheta_p) \in \Theta$

Zbirna funkcija (score function):

$$S(\vartheta|x) := \nabla_{\vartheta} \ell(\vartheta|x) = \begin{bmatrix} \frac{\partial \ell}{\partial \vartheta_1} \\ \vdots \\ \frac{\partial \ell}{\partial \vartheta_p} \end{bmatrix} = \frac{\nabla_{\vartheta} L(\vartheta|x)}{L(\vartheta|x)}$$

Za $g: A \rightarrow \mathbb{R}$ definiramo:

$$\int_A g(x) m(dx) := \begin{cases} \sum_{x \in A} g(x) & ; (D) \\ \int_A g(x) dx & ; (Z) \end{cases}$$

$$\int_A L(\vartheta|x) m(dx) = 1 \quad / \frac{\partial}{\partial \vartheta_j}$$

(O1): $L(\vartheta|x)$ je parcijalno zvezno odvedljiva po ϑ in za $\forall \vartheta^* \in \Theta, \exists U^{o.d.r.} \subseteq \Theta, \vartheta^* \in U, \forall j$ velja:

$$\int \sup_{\vartheta \in U} \left| \frac{\partial L}{\partial \vartheta_j}(\vartheta|x) \right| m(dx) < \infty$$

$$\begin{aligned} \Rightarrow \int_A \frac{\partial L}{\partial \vartheta_j}(\vartheta|x) m(dx) &= 0 \\ &\parallel \\ \int_A \frac{\partial \ell}{\partial \vartheta_j}(\vartheta|x) L(\vartheta|x) m(dx) & \\ &\parallel \\ E\vartheta \left(\frac{\partial \ell}{\partial \vartheta_j}(\vartheta|x) \right) & \end{aligned}$$

(Opazka 6.1: Pri pogoju (O1) je $E\vartheta(S(\vartheta|x)) = 0$)

$$\begin{aligned} \frac{\partial^2 \ell}{\partial \vartheta_j \partial \vartheta_k}(\vartheta|x) &= \frac{\partial}{\partial \vartheta_k} \frac{\frac{\partial \ell}{\partial \vartheta_j}(\vartheta|x)}{L(\vartheta|x)} = \frac{\frac{\partial^2 \ell}{\partial \vartheta_j \partial \vartheta_k}(\vartheta|x) \cdot L(\vartheta|x) - \frac{\partial \ell}{\partial \vartheta_j}(\vartheta|x) \frac{\partial L}{\partial \vartheta_k}(\vartheta|x)}{(L(\vartheta|x))^2} \\ &= \frac{\frac{\partial^2 \ell}{\partial \vartheta_j \partial \vartheta_k}(\vartheta|x)}{L(\vartheta|x)} - \frac{\frac{\partial \ell}{\partial \vartheta_j}(\vartheta|x)}{\frac{\partial \ell}{\partial \vartheta_j}(\vartheta|x)} \cdot \frac{\partial \ell}{\partial \vartheta_k}(\vartheta|x) \end{aligned}$$

Imamo še močnejši pogoj:

(02): $L(\vartheta|x)$ je dvakrat parcialno zvezno odvedljiva po ϑ in za $\forall \vartheta^* \in \Theta \cap \mathcal{U}^{odp} \subseteq \Theta \ni \vartheta^*$. $\forall j, k$ velja

$$\int_A \sup_{\vartheta \in \mathcal{U}} \left| \frac{\partial^2 L}{\partial \vartheta_j \partial \vartheta_k}(\vartheta|x) \right| m(dx) < \infty$$

Če velja še (02) dobimo:

$$\frac{\partial^2}{\partial \vartheta_j \partial \vartheta_k} \int_A L(\vartheta|x) m(dx) = \int_A \frac{\partial^2 L}{\partial \vartheta_j \partial \vartheta_k}(\vartheta|x) m(dx) = E_{\vartheta} \left(\frac{\frac{\partial^2 L}{\partial \vartheta_j \partial \vartheta_k}(\vartheta|x)}{L(\vartheta|x)} \right) = 0$$

$$E_{\vartheta} \left(\frac{\partial^2 L}{\partial \vartheta_j \partial \vartheta_k}(\vartheta|x) \right) = -E_{\vartheta} \left(\frac{\partial L}{\partial \vartheta_j}(\vartheta|x) \cdot \frac{\partial L}{\partial \vartheta_k}(\vartheta|x) \right) =$$

$$\stackrel{6.1}{=} -\text{cov}_{\vartheta} \left(\frac{\partial L}{\partial \vartheta_j}(\vartheta|x), \frac{\partial L}{\partial \vartheta_k}(\vartheta|x) \right)$$

Hesseyeva matrika: $\mathcal{D}^2 g := \left[\frac{\partial^2 g}{\partial \theta_j \partial \theta_k} \right]_{j,k}$

Opazka 6.2: Pod pogoje (C1) in (C2) velja: $\text{var}_\nu(s(\nu|X)) = -E_\nu(\mathcal{D}_\theta^2 \ell(\nu|X))$

Definicija: Fisherjeva informacija je matrika:
 $Fl(\nu_1, \dots, \nu_n) := \text{var}_\nu(s(\nu|X)) = -E_\nu(\mathcal{D}_\theta^2 \ell(\nu|X))$

Naj bo $X = X = (X_1, \dots, X_n)$, kjer so X_i NEP

Definirajmo $Fl_1(\nu) := -E(\mathcal{D}_\theta^2 \ell_1(\nu|X_1))$

Opazka 6.3.

$$Fl(\nu) = n Fl_1(\nu) \quad (= -E(\mathcal{D}_\theta^2 \ell(\nu|X_1, \dots, X_n)))$$

Izrek 6.4.: Opazimo NEP $X_1, \dots, X_n$.

Naj bo $\Theta^{d \times p} \subseteq \mathbb{R}^r$ in $\vartheta^* \in \Theta$

Naj bo $Fl_n(\vartheta^*)$ obrnljiva.

Naj veljata (C1), (C2).

Naj bo $\hat{\vartheta}$ cenilka za ϑ po MNV kedar je slednja nedvoumno definirana, sicer pa poljubna, da je $E_{\vartheta^*}(\|\hat{\vartheta}\|^2) < \infty$

Tedaj velja:

$$1) \hat{\vartheta} \xrightarrow[n \rightarrow \infty]{P_{\vartheta^*}} \vartheta^*$$

$$2) \sqrt{n} B_{\vartheta^*}(\hat{\vartheta} | \vartheta^*) \xrightarrow[n \rightarrow \infty]{} 0$$

$$3) n \cdot \text{Var}_{\vartheta^*}(\hat{\vartheta}) \longrightarrow (Fl_n(\vartheta^*))^{-1}$$

$$\text{Oziramaj: } \text{Var}_{\vartheta^*}(\hat{\vartheta}) \sim (Fl(\vartheta^*))^{-1}$$

$$4) n E_{\vartheta^*}((\hat{\vartheta} - \vartheta^*)(\hat{\vartheta} - \vartheta^*)^T) \longrightarrow (Fl_n(\vartheta^*))^{-1}$$

Poleg tega za primerno funkcijo g in cenilko $g(\hat{\vartheta}) = \hat{y}$ za $y = g(\vartheta)$ velja:

$$1) \hat{y} \xrightarrow[n \rightarrow \infty]{P_{\vartheta^*}} y$$

$$2) \sqrt{n} B_{\vartheta^*}(\hat{y} | y) \xrightarrow[n \rightarrow \infty]{} 0$$

$$3) n \cdot \text{var}_{\vartheta^*}(\hat{y}) \xrightarrow[n \rightarrow \infty]{} (Dg(\vartheta^*))^T (Fl_n(\vartheta^*))^{-1} (Dg(\vartheta^*))$$

$$4) n \text{MSE}_{\vartheta^*}(\hat{y} | y) \xrightarrow[n \rightarrow \infty]{} \text{MSE}_{\vartheta^*}(\hat{y} | y)$$

Dokaz opustimo

7. O večrazsežni normalni porazdelitvi

Kaj bi si želeli?

- ① Če so $X_1, \dots, X_n$ porazdeljene normalno in neodvisne, naj bo $(X_1, \dots, X_n)$ porazdeljena večrazsežno normalno
- ② Če je X večrazsežni normalni slučajni vektor, B pa deterministična matrika, naj bo $a + BX$ spet večrazsežni normalni vektor.

$N(\mu, \sigma^2)$ imega skalo $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$$X \sim N(\mu, 0) \Leftrightarrow P(X=\mu) = 1$$

$$X \sim N(\mu, \sigma^2) \Rightarrow a + bX \sim N(a + b\mu, b^2\sigma^2)$$

$$Z \sim N(0, 1) \Rightarrow \sigma Z + \mu \sim N(\mu, \sigma^2)$$

Definicija: Standardna n -razsežna normalna porazdelitev je porazdelitev slučajnega vektora $\begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$, kjer so Z_i neodvisne standardne normalne slučajne spremenljivke

Opazka 7.1. To je porazdelitev z gostoto

$$\phi_n(z_1, \dots, z_n) = (2\pi)^{-\frac{n}{2}} e^{-\frac{z_1^2 + \dots + z_n^2}{2}}$$

Oziram:

$$\phi_n(z) = (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\|z\|^2} = (\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}z^T z}$$

Definicija. Večrazsežna normalna porazdelitev je taka, ki je enaka porazdelitvi slučajnega vektorja $AZ + \mu$, kjer je A deterministična metrika, μ deter. vektor, Z pa standardni večrazsežni normalni slučajni vektor. Vse kompatibilnih dimenzij

Opazka 7.2. Zgoraj definirana družina porazdelitev je najmanjša družina, ki zadošča

① in ②:

če je $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ bločni slučajni vektor z blokoma, ki sta neodvisna in porazdeljena večrazsežno normalno, je $X_1 \stackrel{d}{=} A_1 Z_1 + \mu_1$ in $X_2 \stackrel{d}{=} A_2 Z_2 + \mu_2$, kjer sta Z_1 in Z_2 neodvisna standardna normalna slučajna vektorja.

Tedaj je $\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$ spet standardni večrazsežni normalni vektor in velja

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} + \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \text{ torej}$$

je to večrazsežni normalni vektor

Opazka 7.3,

Če je $\underline{Z}$ n -vektor: standardni slučajni vektor in je $X = AZ + \mu$, potem je $E(X) = \mu$ in $\text{var}(X) = AA^T$

Opazka 7.4

Naj bo A neizrojena kvadratna matrika n in μ kot prej. Tedaj ima X gostoto:

$$f_X(x) = f_Z(A(x-\mu)) |\det A^{-1}| =$$

$$= \frac{1}{|\det A|} (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2} \|A^{-1}(x-\mu)\|^2}$$

$$\|A^{-1}(x-\mu)\|^2 = (A^{-1}(x-\mu))^T A^{-1}(x-\mu) =$$

$$= (x-\mu)^T A^T A^{-1} (x-\mu) =$$

$$= (x-\mu)^T (AA^T)^{-1} (x-\mu) = (x-\mu)^T \text{var}(X)^{-1} (x-\mu)$$

Če torej označimo $\Sigma = AA^T$, velja

$$f_X(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} e^{-\frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu)}$$

Poročilo takega slučajnega vektora X , je torej enolično določeno s pričakovano vrednostjo in kovariacijsko matriko

Izrek 7.5.

Porazdelitev poljubnega večrazsežnega normalnega slučajnega vektarja je enolično določena z njegovo pričakovano vrednostjo in kovariančno metriko

Dokaz: opustimo

Opazka 7.6.

Za $\forall \mu \in \mathbb{R}^n$ in $\forall \Sigma$ simetrične in pozitivno semidefinitne. $\exists$ večrazsežni normalni slučajni vektor X , za katerega je $E(X) = \mu$ in $\text{var}(X) = \Sigma$

Definicija. Za μ in Σ kot prej definiramo $N(\mu, \Sigma)$ kot večrazsežno normalno porazdelitev s pričakovano vrednostjo μ in kovariančno metriko Σ

Posledica 7.7

Na razsežnih slučajnih spremenljivkah se tu definirani koncept večrazsežne normalne porazdelitve ujema s klasičnim konceptom enorazsežne normalne porazdelitve

Opazka 7.8.

Če sta $X_1 \sim N(\mu_1, \Sigma_1)$ in $X_2 \sim N(\mu_2, \Sigma_2)$ neodvisne je $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}\right)$

Trditev 7.9.

Če je bločni slučajni vektor $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ porazdeljen večrazsežno normalno in sta njegovi bloke nekorrelirane, sta tudi neodvisne

Dokaz:

Če vzamemo neodvisne vektarje

$X_1' \stackrel{d}{=} X_1$ in $X_2' \stackrel{d}{=} X_2$, potem vemo,

da je $\begin{bmatrix} X_1' \\ X_2' \end{bmatrix}$ prav tako ra porazdeljen večrazsežno normalno in ima enako kovariančno matriko in pričakovano vrednost kot $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$.

Rezultat sledi iz izreka 7.5

Opazke 7, 10.

$$\text{če je } Z \sim N(0, I_n) \Rightarrow \|Z\|^2 \sim \chi^2(n) = \Gamma\left(\frac{n}{2}, \frac{n}{2}\right)$$

$$\text{Splošneje, če je } W \sim N\left(0, \begin{bmatrix} I_p & 0 \\ 0 & 0_{n-p} \end{bmatrix}\right)$$

$$\text{je } \|W\|^2 \sim \chi^2(p)$$

Trditav 7.11.

Če je H ortogonalni projektor range p in $X \sim N(0, H)$ je $\|H\|^2 \sim \chi^2(p)$

Dokaz: Obstaja ortogonalna preslikava Q , ki deluje na naslednji način:

$$\mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$\text{Im } H \rightarrow \mathbb{R}^p \times \{0\}^{n-p}$$

$$(\text{Im } H)^\perp \rightarrow \{0\}^p \times \mathbb{R}^{n-p}$$

$$\|X\|^2 = \|QX\|^2$$

$$QX \sim N(0, QHQ^T)$$

"
 $\begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix}$ je res če pogledamo lastnosti preslikave in ~~to~~ ker H glisa katere vektore

Zagled 7.12. $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ neovisne
 $\bar{X} = \frac{1}{n} \sum X_i$

$$\sum (X_i - \bar{X})^2 = ?$$

Spomnimo se: $y_i = X_i - \mu$ $\bar{y} = \bar{X} - \mu$

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (y_i - \bar{y})^2 = \|y - \bar{y} \cdot 1_n\|^2 =$$

$$= \|y - 1_n \cdot \frac{1}{n} 1_n^\top \bar{y}\|^2 = \|Hy\|^2$$

kjer je $H = I_n - \frac{1}{n} 1_n \cdot 1_n^\top$ ortogonalni projektor
 range $n-1$

$$\frac{1}{\sigma} y \sim N(0, I_n)$$

$$\frac{1}{\sigma} Hy \sim N(0, H) = N(0, HH^\top)$$

$$\Rightarrow \left\| \frac{1}{\sigma} Hy \right\|^2 = \frac{1}{\sigma^2} \sum (X_i - \bar{X})^2 \sim \chi^2_{(n-1)} = \Gamma\left(\frac{n-1}{2}, \frac{1}{2}\right)$$

$$\Rightarrow \sum (X_i - \bar{X})^2 \sim \Gamma\left(\frac{n-1}{2}, \frac{1}{2\sigma^2}\right)$$

Definicija: Studentova porazdelitev s p prostorskimi stopnjami je porazdelitev slučajne spremenljivke $\frac{Z}{\sqrt{G}}$, kjer sta $Z \sim N(0, 1)$ in $G \sim \Gamma(\frac{p}{2}, \frac{p}{2})$ neodvisni slučajni spremenljivki.

Opazka 7.13.: $p G \sim \Gamma(\frac{p}{2}, \frac{1}{2}) = \chi^2(p)$

Zgled 7.14. $X_1, \dots, X_n \sim N(\mu, \sigma^2)$
Vemo: $\frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0, 1)$

Kaj pa $\frac{\bar{X} - \mu}{\hat{\sigma}_+} \sqrt{n}$, kjer je $\hat{\sigma}_+^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

$$\frac{\frac{\bar{X} - \mu}{\sigma} \sqrt{n}}{\sqrt{\frac{\hat{\sigma}_+^2}{\sigma^2}}} \sim N(0, 1) \quad \sum (X_i - \bar{X})^2 \sim \Gamma\left(\frac{n-1}{2}, \frac{1}{2\sigma^2}\right)$$

$$\leadsto \frac{\sigma^2}{\hat{\sigma}_+^2} = \frac{1}{(n-1)\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \Gamma\left(\frac{n-1}{2}, \frac{n-1}{2}\right)$$

Manjke še neodvisnost

$\bar{X}$ in $\sum (X_i - \bar{X})^2$ sta neodvisni, ker sta

$\bar{X}$ in HX nekorrelirana. (Podrobnosti na vejah).

8. statistično sklepanje z nadzorovanim tveganjem

8.1 Množice zaupanja in nepovedne množice

Opazimo X , zanima nas Y .

Želimo sklepati: $Y_{\min} < Y < Y_{\max}$

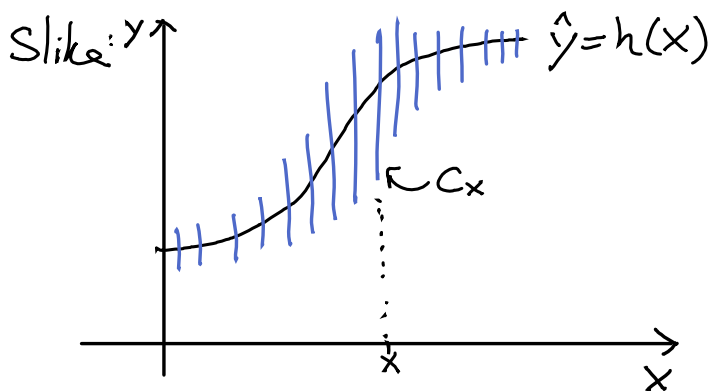
$$\begin{array}{cc} \parallel & \parallel \\ h_{\min}(X) & h_{\max}(X) \end{array}$$

Če je $Y = y = g(v)$ deterministična funkcija temu pravimo **interval zaupanja**, sicer pa temu pravimo **nepovedni interval**.

$$\forall \theta \in \Theta, P_{\theta}(Y_{\min} < Y < Y_{\max}) \geq 1 - \alpha$$

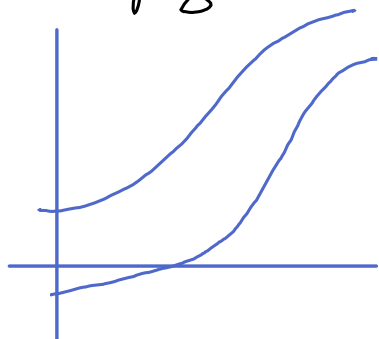
Splasnejše lahko sklepamo.

da je $Y \in C_X$, kjer je $C_X; X \in \dots$ nebor
nekih množic



Ideje za konstrukcijo:

① Celovit pogled



$$C = \{(x, y); y \in C_x\}$$

$$P_0((x, y) \in C) \geq 1 - \alpha$$

② Mera razhajanja $p(x, y)$

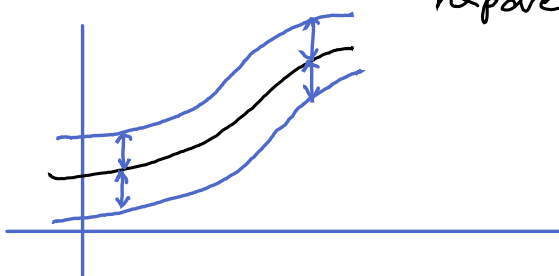
$$C = \{(x, y); p(x, y) < c_\alpha\}$$

Če imamo srečo, je p **pivotna funkcija** ker pomeni, da se porazdelitev slučajne spremenljivke $p(X, y)$ ne spreminja s θ . ← kar želimo opaziti
Torej lahko c_α dobimo kot minimalni c , za katerega je $P_{\theta \in \Theta}(p(X, y) \leq c) \geq 1 - \alpha$.
Torej je tudi $P_0(p(X, y) \leq c_\alpha) \geq 1 - \alpha$,
često (pogosto) pa tudi $P_0(p(x, y) < c_\alpha) \geq 1 - \alpha$

③ Uporaba točkovnih napovedi/cenilk

$$p(x, y) := |h(x) - y|, \text{ kjer je } \hat{y} = h(x)$$

napoved za y



$$\text{Splošneje: } p(x, y) = d(h(x), y)$$

Ponovimo standardni primer:

opazimo: $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ neodvisne

Recimo da poznamo σ in ocenjujemo μ

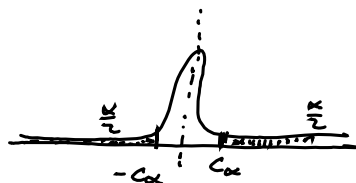
$$\hat{\mu} = \bar{X} \text{ nam da } g((x_1, \dots, x_n), \mu) = \left| \frac{x_1 + \dots + x_n}{n} - \mu \right|$$

To je pivotna funkcija, ker je

$$g(X_1, \dots, X_n, \mu) = |\bar{X} - \mu| \quad \text{in } \bar{X} - \mu \sim N\left(0, \frac{\sigma^2}{n}\right)$$

ne glede na μ

$$c_\alpha := \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}}$$



$$\bar{X} - c_\alpha < \mu < \bar{X} + c_\alpha$$

Napovedni interval za $X_{n+1} \sim N(\mu, \sigma^2)$ neodvisne
od $X_1, \dots, X_n$:

$$\left. \begin{array}{l} X_{n+1} \sim N(0, \sigma^2) \\ \bar{X} \sim N\left(0, \frac{\sigma^2}{n}\right) \end{array} \right\} \text{neodvisni} \Rightarrow X_{n+1} - \bar{X} \sim N\left(0, \left(1 + \frac{1}{n}\right)\sigma^2\right)$$

Napovedni interval $\bar{X} - c_\alpha^{\text{nap}} < X_{n+1} < \bar{X} + c_\alpha^{\text{nap}}$

$$c_\alpha^{\text{nap}} = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sigma \sqrt{1 + \frac{1}{n}}$$

Kaj pa če ne poznamo σ ?

Iz zgleda 7.14 vemo $\frac{\bar{X} - \mu}{\hat{\sigma}_+} \sqrt{n} \sim \text{Student}(n-1)$

$$\vartheta = (\mu, \sigma)$$

res, ne glede na μ in σ

Interval zaupanja bo zdaj spet $\bar{X} - c_\alpha < \mu < \bar{X} + c_\alpha$,

kjer bo $c_\alpha := F_{\text{student}(n-1)}^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\hat{\sigma}_+^2}{\sqrt{n}}$

Podobno dobimo tudi napovedni interval (DI):

$$c_\alpha^{\text{nap}} = F_{\text{student}(n-1)}^{-1}\left(1 - \frac{\alpha}{2}\right) \hat{\sigma}_+ \sqrt{1 + \frac{1}{n}}$$

pad ideje za konstrukcijo

④ Uporaba verjetja (le za $y = y = g(\theta)$)

Spomnimo se izreka 6.4: za cenilko $\hat{y}$
pa metodi največjega verjetja asimptotski:

$$MSE_{\theta}(\hat{y}|y) \sim \frac{(\nabla g(\theta))^T (F_{\theta}(\theta))^{-1} (\nabla g(\theta))}{n}$$

Izrek 8.5.

Naj bo $\hat{\theta}$ ocenilka za $\theta \in \Theta \subset \mathbb{R}^m$ po metodi največjega verjetja na podlagi neodvisnih in ensko porazdeljenih opazenj $X_1, \dots, X_n$ (kjer je redvomno definirana, sicer pa poljubna).

Naj bo $\theta^* \in \Theta$ in $FI_1(\theta^*)$ obrnljiva.

Tedaj pod podobnimi dodatnimi pogoji kot za izrek 6.4 velja

$$\sqrt{n}(\hat{\theta} - \theta^*) \xrightarrow[n \rightarrow \infty]{d, P_{\theta^*}} N(0, (FI_1(\theta^*))^{-1})$$

Nadalje naj bo $g: \Theta \rightarrow \mathbb{R}$ dovolj gladka funkcija ter $y^* = g(\theta^*)$ in $\hat{y} = g(\hat{\theta})$.

Če definiramo

$$RMSE^* = \sqrt{\frac{1}{n} ((\nabla g(\theta^*))^T \cdot (FI(\theta^*))^{-1} (\nabla g(\theta^*)))}$$

$$\hat{RMSE} = \sqrt{\frac{1}{n} ((\nabla g(\hat{\theta}))^T \cdot (FI(\hat{\theta}))^{-1} (\nabla g(\hat{\theta})))}$$

Velja še: $\frac{\hat{y} - y^*}{RMSE^*} \xrightarrow[n \rightarrow \infty]{d, P_{\theta^*}} N(0, 1)$

$$\frac{\hat{y} - y^*}{\hat{RMSE}} \xrightarrow[n \rightarrow \infty]{d, P_{\theta^*}} N(0, 1)$$

Če je torej $\hat{y} = g(\hat{\theta}) = h(X_1, \dots, X_n)$ in

$\hat{RMSE} = s(X_1, \dots, X_n)$ je

$$g(X_1, \dots, X_n, y) = \frac{h(X_1, \dots, X_n) - y}{s(X_1, \dots, X_n)}$$

asimptotično pivotna funkcija.

Dokaz: $\square$

Ideje:

- $l(\vartheta | x_1, \dots, x_n) = \sum_{i=1}^n l_1(\vartheta | x_i)$
- uporabimo večrazsežni CLI za $\sum_{i=1}^n \nabla l_1(\vartheta | x_i)$
- cenilke $\hat{y}$ je približno linearna funkcija vsote $\sum_{i=1}^n \nabla l_1(\vartheta^* | x_i)$
- izrek o konvergenci preslikanih slučajnih spremenljivk
- izrek Slutskega za deljenje

Dobimo asimptotično eksakten interval zaupanja:

$$\hat{y} - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \widehat{RMSE} < y < \hat{y} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \widehat{RMSE}$$

$\in o\left(\frac{1}{\sqrt{n}}\right)$

8.2 Preizkušanje domnev/ testiranje hipotez

Včasih imamo o ϑ kakšno idejo.

Temu pravimo **nizelne domneve** in jo
označimo s H_0 (splošno $\vartheta \in \Theta_0 \subseteq \Theta$)

Tipičen primer H_0 : „kovanec je pošten“

$$\vartheta = P(\text{pade grb}) \in [0, 1]$$

$$H_0: \vartheta = \frac{1}{2}$$

Negaciji: nizelne domneve pravimo **alternativna domneve** in jo ponavadi označimo s H_1

Če je opažanje neskladno s H_0 , to domnevo zavrnilo in sprejmemo H_1 . Sicer ne rečemo nič.

Da zmote torej po definiciji pride samo, če H_0 zavrnilo, čeprav velja.

- H_0 želimo zavrniti v oim več primerih ko ne drži. Moč preizkusa je funkcija:

$$\Theta - \Theta_0 \longrightarrow [0, 1]$$

$$v \mapsto P_v(H_0 \text{ zavrnemo})$$

- Za $v \in \Theta_0$ pa zahtevamo

$$P_v(H_0 \text{ zavrnemo}) = \alpha \begin{matrix} \text{eksakten} \\ \approx \text{aproximativen} \\ \rightarrow \text{eksakten} \\ \leq \text{konserватiven} \end{matrix}$$

stopnja tveganja

Zveza z množico zaupanja:

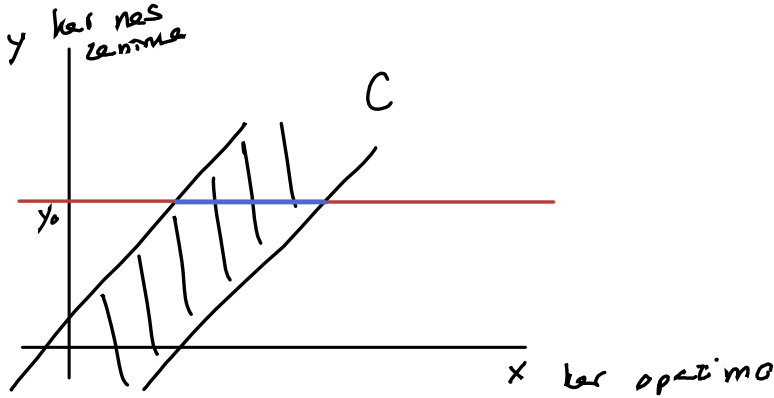
Vsaka domneva $v \in \Theta_0$ je za primerno karakteristiko

$y = g(v)$ divivalentna domnevi: $y = y_0$ npr.

$$g(v) = 1(v \in \Theta_0), y_0 = 1$$

Opazke 8.2

Nabor preizkusov domnev $H_0: y = y_0$ za vse možne y_0 je ekvivalenten množici zaupanja za y :



$$H_0: y = y_0 \Leftrightarrow (X, y_0) \notin C \Leftrightarrow y_0 \notin C_X \Leftrightarrow \underset{\substack{\uparrow \\ \text{mesa} \\ \text{razhajanja}}}{p(X, y_0)} \geq \alpha$$

Primer: Opezioma $S \sim \text{Bin}(n, \vartheta)$, kjer je n znan, ϑ pa nas zanima.
 $\hat{\vartheta} = \frac{S}{n}$

$$g(S, \vartheta) = \left| \frac{S}{n} - \vartheta \right|$$

$$\frac{S}{n} - \vartheta = \hat{\vartheta} - \vartheta \underset{\text{CLT}}{\sim} N\left(0, \frac{\vartheta(1-\vartheta)}{n}\right)$$

τ ϑ nastopa tukaj zato
 S ni niti približno
 pivotna funkcija

Lahko pa ϑ v varianci nadomestimo z
 oceno in dobimo Waldov interval zaupanja

$$\hat{\vartheta} - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{\hat{\vartheta}(1-\hat{\vartheta})}{n}} < \vartheta < \hat{\vartheta} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{\hat{\vartheta}(1-\hat{\vartheta})}{n}}$$

$H_0: \vartheta = \vartheta_0$ zavrneemo proti $H_1: \vartheta \neq \vartheta_0$ če je
 $|\hat{\vartheta} - \vartheta_0| \geq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{\vartheta_0(1-\vartheta_0)}{n}}$

Naravnnejša in tudi boljše različica preizkusa:
 $H_0: \vartheta = \vartheta_0$ zavrneemo, če je $|\hat{\vartheta} - \vartheta_0| \geq \Phi^{-1} \sqrt{\frac{\vartheta_0(1-\vartheta_0)}{n}}$

To vodi v izboljšeno množico zaupanja:
 $C = \{(s, \vartheta); |\frac{s}{n} - \vartheta| < \underbrace{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}_{z''} \sqrt{\frac{\vartheta(1-\vartheta)}{n}}\}$

Pr: danem S oziroma $\hat{\vartheta} = \frac{S}{n}$ dobimo

$$\frac{2n\hat{\vartheta} + z^2 - \sqrt{4nz^2\hat{\vartheta}(1-\hat{\vartheta}) + z^4}}{2(n+z^2)} < \vartheta < \frac{2n\hat{\vartheta} + z^2 + \sqrt{4nz^2\hat{\vartheta}(1-\hat{\vartheta}) + z^4}}{2(n+z^2)}$$

Willsonov interval zaupanja

V splošnem:

$$C = \{(x, y) : g(x, y) < c_\alpha(y)\}$$

τ ni nujno da je g pivotna, dovolj
 je da je porazdelitev $g(X, Y)$
 pri $\forall \vartheta$ natanko določena z $y = g(\vartheta)$

Včasih nas skrbijo odstopanja samo v eno smer

$H_0: y = y_0$ preizkušamo proti $H_1^+: y > y_0$ ali
 $H_1^-: y < y_0$

Primer: Loterija tedi, da vsaka druga srečka zadene

Če je n velik je pri H_0 $S \sim N(n\vartheta_0, \sqrt{n\vartheta_0(1-\vartheta_0)})$

$$\hat{\vartheta} = \frac{S}{n} \sim N\left(\vartheta_0, \sqrt{\frac{\vartheta_0(1-\vartheta_0)}{n}}\right)$$

$$\frac{\hat{\vartheta} - \vartheta_0}{\sqrt{\frac{\vartheta_0(1-\vartheta_0)}{n}}} \sim N(0, 1)$$

H_0 zavrnemo, če je $\frac{\vartheta - \vartheta_0}{\sqrt{\frac{\vartheta_0(1-\vartheta_0)}{n}}} \leq -\Phi^{-1}(1-\alpha) = \Phi^{-1}(\alpha)$



Dvostransko območje zavrnilne:



Preizkus z razmerjem verjetj:

$$LR := \frac{\sup_{\theta \in \Theta} L(\theta | X)}{\sup_{\theta \in \Theta_0} L(\theta | X)}$$

↑
likelihood
ratio

$\sim$ ničelna domneva $H_0: \theta \in \Theta_0$

$H_0: \theta \in \Theta_0$ zavržemo, če je $LR \geq c_\alpha$
Prag c_α določimo skladno s stopnjo tveganja α

Izrek 8.3 (Wilksov izrek)

Naj bode $X_1, X_2, \dots$ pri vsaki vrjetnosti meri:
 $\mathbb{P}_\theta$, $\theta \in \Theta$, neodvisne in enako porazdeljene
skalarne spremenljivke, Θ naj bo gladka mnogoterost
dimenzije p , $\Theta_0 \subseteq \Theta$ pa podmnogoterost dimenzije
 q . Obe naj bosta brez roba.

Pod določenimi dodatnimi pogoji gre tedaj

$$2 \ln(LR) \xrightarrow[n \rightarrow \infty]{d} \chi^2(p-q) \text{ pri vseh } \mathbb{P}_\theta; \theta \in \Theta_0.$$

Dokaz opuščamo

Na ta način dobimo Wilksov preizkus

$H_0: \theta \in \Theta_0$ zavrnemo, če je

$$2 \ln(LR) \geq F_{\chi^2(p-q)}^{-1}(1-\alpha)$$



Osnovni primer uporabe

$$P_\theta = \begin{pmatrix} a_0 & a_1 & \dots & a_p \\ \vartheta_0 & \vartheta_1 & \dots & \vartheta_p \end{pmatrix}$$

$$\vartheta = (\vartheta_0, \dots, \vartheta_p)$$

$$\Theta = \{ \vartheta \in \mathbb{R}^{p+1} : \sum_{i=0}^p \vartheta_i = 1 \wedge \vartheta_i > 0 \}$$

Za fiksne $\pi_0, \pi_1, \dots, \pi_p$ pričakujemo

$$H_0 : \vartheta_0 = \pi_0, \dots, \vartheta_p = \pi_p$$

$$\text{Opazimo } X_1, \dots, X_n \sim \begin{pmatrix} a_0 & a_1 & \dots & a_p \\ \vartheta_0 & \vartheta_1 & \dots & \vartheta_p \end{pmatrix}$$

$$\ln(LR) = \sup_{\vartheta \in \Theta} \sum_{i=1}^n \ell_n(\vartheta | X_i) - \sup_{\vartheta \in \Theta} \sum_{i=1}^n \ell_n(\vartheta | x_i)$$

$$\ell_n(\vartheta | x) = \begin{cases} \ln \vartheta_0 & ; x = a_0 \\ \ln \vartheta_1 & ; x = a_1 \\ \vdots & \\ \ln \vartheta_p & ; x = a_p \end{cases}$$

$$\ell(\vartheta_0, \vartheta_1, \dots, \vartheta_p | X_1, \dots, X_n) = \sum_{i=1}^n \sum_{j=0}^p \ln \vartheta_j \mathbb{1}(X_i = a_j) = \sum_{j=0}^p \ln \vartheta_j \underbrace{\sum \mathbb{1}(X_i = a_j)}_{N_j}$$

$\nwarrow$ frekvenca a_j

Lagrangevi multiplikatorji:

$$F = \sum N_j \ln \vartheta_j - \lambda \sum \vartheta_j$$

$$\frac{\partial F}{\partial \vartheta_j} = \frac{N_j}{\vartheta_j} - \lambda = 0 \quad \hat{\vartheta}_j = \frac{N_j}{\lambda}$$

$$\sum \hat{\vartheta}_j = 1 \Rightarrow \lambda = \sum_{j=0}^p N_j = n$$

$$\begin{aligned} \sup_{\vartheta \in \Theta} \ell(\vartheta_0, \dots, \vartheta_p | X_1, \dots, X_n) &= \sum_{j=0}^p N_j \ln \hat{\vartheta}_j \\ \sup_{\vartheta \in \Theta} \ell(\vartheta_0, \dots, \vartheta_p | X_1, \dots, X_n) &= \sum_{j=0}^p N_j \ln \pi_j \end{aligned}$$

$$2 \ln(LR) = 2 \sum N_j \ln \frac{\hat{\vartheta}_j}{\pi_j} = 2 \sum_{j=0}^p N_j \ln \frac{N_j}{n \pi_j}$$

$\tilde{N}_j = n \pi_j$ so pričakovane frekvence

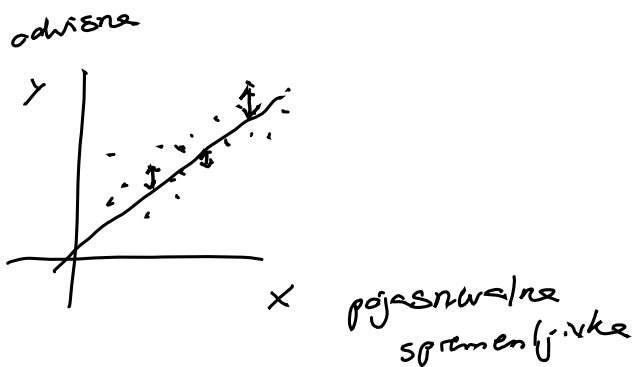
Minimalni pogoj za legitimnost: $\forall j, \tilde{N}_j \geq 5$

Tedaj bo porazdelitev $2 \ln(LR)$ vsaj približno blizu $\chi^2(p)$

Opazka 8.4

$$\text{Izkaže se, da je } 2 \ln(LR) \approx \sum_{j=0}^p \frac{(N_j - \tilde{N}_j)^2}{\tilde{N}_j}$$

3. Linearna regresija



To je **enostavna linearna regresija**, ki se ravna po predlogi: $y = a + bx$

Splošna linearna regresija

$$y = \beta_1 x_1 + \dots + \beta_p x_p$$

$x_1, \dots, x_p$.. pojasnevalne spremenljivke

y ... odvisna spremenljivka

$\beta_1, \dots, \beta_p$... koeficienti parametrov

Podatki: $x_{11}, \dots, x_{1p}; y_1$

$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad n > p$

$x_{n1}, \dots, x_{np}; y_n$

predložen sistem

$$Xp = y \quad p = ?$$

9.2 točkovno ocenjevanje koeficientov

Metoda najmanjših kvadratov
OLS

$$\begin{array}{lcl} y_1 & = & \beta_1 x_1 + \dots + \beta_p x_p + \varepsilon_1 \\ y_2 & = & \dots \dots \dots + \varepsilon_2 \\ \vdots & & \vdots \end{array} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

vektor
sumov

Poiščemo tisti β , pri katerem je
vsota $\|\varepsilon\|^2$ najmanjša

$$V = \text{Im } X \subseteq \mathbb{R}^n$$

Poiščemo $v \in V$ zatorega je $\|x - v\|^2$ minimalna

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

Statistični pogled

$$y = X\beta + \varepsilon \leftarrow \text{slučajni in neznan}$$

β parameter: determinističen, neznan
↑
karakteristične metrike

slučajni,
opazljiv

Milejšje minimalne predpostavke:

ε je beli šum

$$\mathbb{E}(\varepsilon) = 0$$

$$\text{var}(\varepsilon) = \sigma^2 I$$

homoskedastičnost in
nekoreliranost

Iščemo $\hat{\beta}$ parametra β

smiselne močes $X\beta = y$ po MNK

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

ke

kje dobra je ta cenilka?
 če je $c \in \mathbb{R}^T$ konstanten vektor, kako
 dobra je cenilka $c^T \hat{\beta}$ karakteristike
 $c^T \beta$?

Definicije:

Linearne oz. afine cenilke za količino
 g z vrednostmi v $\mathbb{R}$ na podlagi
 opazanja y z vrednostmi v $\mathbb{R}^n$
 cenilke oblike $\hat{g} = \psi^T y + \rho$, kjer je
 $\rho \in \mathbb{R}$ konstanta in $\psi \in \mathbb{R}^n$ konstantni vektor

Naj boljše nepristranske linearne cenilke
 NNLC oz BLUE za g je tista
 nepristranska linearne cenilka $\hat{g}$, da za
 poljubno alternativno nepristransko cenilko $\hat{g}$
 velja $\text{var}_g(Q) \geq \text{var}_g(\hat{g})$ za $\forall \hat{g}$ ^{linearno}

Izreke 9.1 (Gauss, Markov)

če je $c \in \mathbb{R}^T$ konstanten vektor, $\hat{\beta} := c^T \beta$ edina
NNLC za $z := c^T \beta$

Dalje v zapisih

Minimiziramo $\|y - X\beta\|^2$

$H = X(X^T X)^{-1}X^T$ je ortogonalni projektor
na $\text{Im} X$

$$\|(I_n - H)y\|^2 = \|\hat{\epsilon}\|^2 = \sum \epsilon_i^2 =: \text{RSS}$$

$y = X\hat{\beta} + \hat{\epsilon}$ $\hat{\epsilon}$ je vektor reziduirov ali ostankov

Trditel 9.2

$\sigma_+^2 = \frac{RSS}{n-p}$ je nepristranska cenilka za σ^2

Dakaz opuščamo

Možnejše predpostavke

dodamo $\varepsilon \sim N(0, \sigma^2 I_n)$

Oprezha 9.3. tedaj je $\hat{\beta}$ tudi cenilka po metodi največjega verjetja

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

$$SE = \sqrt{\text{var}(C^T \hat{\beta})}$$

Trditel 9.4.

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

Trditel 9.5.

Če definiramo $\hat{\beta}$ in $\hat{\sigma}_+^2$ sta neodvisni:

Posledica 9.6

$$SE = \sqrt{\hat{\sigma}_+^2 C^T (X^T X)^{-1} C}$$
$$\frac{C^T \hat{\beta} - C^T \beta}{\hat{SE}} \sim \text{Student}(n-p)$$

