VERJETNOST IN STATISTIKA

2 LETO

1 SLUČAJNI VEKTORJI

↳ **DEFINICIJA**

Naj bo $(\Omega, \mathcal{F}, P)$ verjetnostni prostor. Preslikavi $X: \Omega \to \mathbb{R}$ rečemo SLUČAJNA SPREMENLJIVKA (SS), če so vse množice $\{X \leq x\} = \{\omega ; X(\omega) \in (-\infty, x]\} = X^{-1}((-\infty, x])$ dogodki. [$\mathcal{F}$ je tu σ-algebra]

↳ **ZGLED**

Kako bi na ta način formalirali met kovanca?

Intuitivno: $X=1$ če vržemo cifro

$X=0$ če vržemo grb

$\Omega = \{$prostor dopustnih metov$\}$. $X(\text{met}) = \begin{cases} 1 & ; \text{ cifra} \\ 0 & ; \text{ grb} \end{cases}$

↳ Če je X s.s., lahko definiramo KOMULATIVNO PORAZDELITVENO FUNKCIJO (K.P.F) $F_X: \mathbb{R} \to [0,1]$ s predpisom

$$\boxed{F_X(x) = P(X \leq x) = P(X^{-1}((-\infty, x]))}$$

↳ **DEFINICIJA**

Preslikava $X = (X_1, X_2, \dots, X_n): \Omega \to \mathbb{R}^n$ je SLUČAJNI VEKTOR (SV), če so $\{X_1 \leq x_1, \dots, X_n \leq x_n\} = X_1^{-1}((-\infty, x_1]) \times \dots \times X_n^{-1}((-\infty, x_n])$ dogodki za vse realne n-terice $(x_1, \dots, x_n) \in \mathbb{R}^n$

K.P.F. SV X je $F_X: \mathbb{R}^n \to [0,1]$, definirana s predpisom $F_X(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n)$

↳ **LASTNOSTI K.P.F.**

(1) • LIMITA V $-\infty$

$$\lim_{x_n \to -\infty} F(x_1, \dots, x_{n-1}, x_n) = \lim_{x_n \to -\infty} P(X \in (-\infty, x_1] \times \dots \times (-\infty, x_{n-1}] \times (-\infty, x_n]) = \lim_{k \to \infty} P(X \in K \times (-\infty, -k]) =$$
$$= P\left(\bigcap_{k=1}^{\infty} \{X_1 \leq x, \dots, X_{n-1} \leq x_{n-1}, X_n \leq -k\}\right) = P\left(\bigcap_{k=1}^{\infty} \{X \in K \times (-\infty, -k]\}\right) = P(\emptyset) = 0$$

zveznost od spodaj

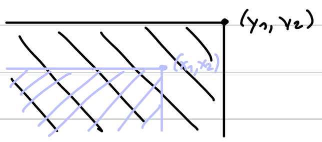
Enako velja: $\lim_{x_i \to -\infty} F_X(x_1, \dots, x_n) = 0$ za $\forall i$

• LIMITA V ∞

$$\lim_{(x_1, \dots, x_n) \to (\infty, \infty, \dots, \infty)} F_X(x_1, \dots, x_n) = P\left(\bigcup_{k=1}^{\infty} \{X_1 \leq k, X_2 \leq k, \dots X_n \leq k\}\right) = P(X \leq \mathbb{R}^n) = P(\Omega) = 1$$

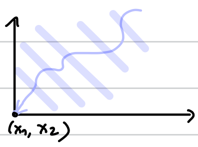
zveznost od spodaj

(2) MONOTONOST: Če $x_i \leq y_i$ za vse i, potem $F_X(x_1, \dots, x_n) \leq F_X(y_1, \dots, y_n)$. Ker je $(-\infty, x_1) \times \dots \times (-\infty, x_n) \subset (-\infty, y_1] \times \dots \times (-\infty, y_n]$ je $P(\dots_x) \leq P(\dots_y)$ zaradi monotonosti verjetnosti



(3) ZVEZNOST Z DESNE: če $(y_1, \dots, y_n) \downarrow (x_1, \dots, x_n)$ velja $F_X(x_1, \dots, x_n) \downarrow F_Y(y_1, \dots, y_n)$

$$\lim_{\forall i : y_i \downarrow x_i} F_X(y_1, \dots, y_n) = F_X(x_1, \dots, x_n)$$

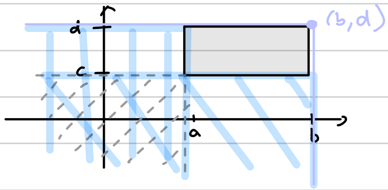


OPOMBA: Lastnosti (1),(2),(3) karakterizijo abstraktne k.p.f. v primeru $n=1$, torej v primeru s.s.

(4) Oglejmo si še lastnosti za s.v. Vzemimo $n=2$. Oglejmo si primer $(X,Y):\Omega\to\mathbb{R}^2$

$$P((X,Y)\in(a,b]\times(c,d]) = F_X(b,d)-F_X(a,d)-F_X(b,c)+F_X(a,c)\geq 0$$

za $a<b$, $c<d$

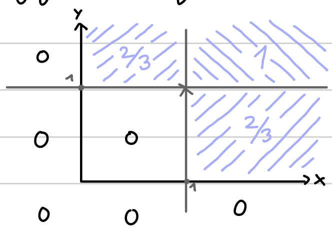


OPOMBA: Lastnosti (1),(2),(3),(4) karakterizirajo k.p.f. dvorazsežnega SV

Očitna posplošitev velja tudi za $n\geq 3$ (kvadri)

ZGLED

Oglejmo si funkcijo $F:\mathbb{R}^2\to[0,1]$



Dokazali smo da 1,2,3 niso dovolj za karakterizacijo $1-\frac{2}{3}-\frac{2}{3}+0 \cancel{\geq} 0$

Ker je $\{X_1\leq x_1,\dots,X_{n-1}\leq x_{n-1}\}=\bigcup_{k=1}^{\infty}\{X_1\leq x_1,\dots,X_{n-1}\leq x_{n-1}, X_n\leq k\}$,

so posledično $(X_1,\dots,X_{n-1})$ podvektorji SV in tudi SV.

DEFINICIJA

Vsi PODVEKTORJI SV $(X_1,\dots,X_n)$ so slučajni vektorji.

Npr. $(X_1,\dots,X_{n-1})$: $\{X_1\leq x,\dots,X_{n-1}\leq x_{n-1}\}=\bigcup_{k=1}^{\infty}\{X_1\leq x_1,\dots,X_{n-1}\leq x_{n-1}, X_n\leq k\}$ $(*)$

Posebej, če je $(X_1,\dots,X_n)$ SV, so X_i SS (obratna implikacija je očitna). Iz $(*)$ je tudi razvidno

$$\boxed{F_{(X_1,\dots,X_{n-1})}(x_1,\dots,x_{n-1})=\lim_{x_n\to\infty}F_{(X_1,\dots,X_n)}(x_1,\dots,x_n)}$$

POSLEDICA ? SHADY

$(X_1,\dots,X_n)$ je S.V. natanko tedaj, ko so njegovi komplementi SS.

ZGLED

$P(X\in[110,120]$ in $Y\in[80,90])\implies B=[110,120]\times[80,90]$

$P(X\geq 6$ in $Y=0)=P(X\geq 6)\cdot P(Y=0)\implies B=[6,\infty)\times\{0\}$

OPOMBA

i) $\{X\in(-\infty,b]\}$ dogodek $\implies\{X\in(b,\infty)\}$ dogodek

ii) $\{X\in(a,b]\}$ dogodek

iii) Iz $[a,b]=\bigcap_{n=1}^{\infty}(a-\frac{1}{n},b]$ sledi $\{X\in[a,b]\}$ je dogodek. Posebej $\{X=a\}$ je dogodek

iv) Iz $(a,b)=\bigcup_{k=1}^{\infty}(a,b-\frac{1}{n}]$ je $\{X\in(a,b)\}$ dogodek

Če je $X:\Omega\to\mathbb{R}^n$ slučajni vektor, so množice $\{X\in B\}$ dogodki za vse Borelove množice B.

↳ LANSKA OPOMBA → lepše napisano

Naj bo $X: \Omega \to \mathbb{R}$ SS. Potem velja:

→ $\{X \in (a,b]\} = \{X \in (-\infty,b]\} \cap \{X \in (-\infty,a]\}^c$

→ $\{X \in (a,b)\} = \bigcup_{c_n \uparrow b} \{X \in (a,c_n]\}$

→ $\{X \in [a,b]\} = \bigcap_{n=1}^{\infty} \{X \in (a-\frac{1}{n}, b+\frac{1}{n})\}$

→ $\{X = x\}$

Zgornje lahko "kombiniramo" (števne unije, preseki, komplementi ; ŠTEVNO). Izkaže se, da obstajajo verjetnosti $P(X \in B)$, kjer je B poljubna "Borelova" množica $\subset \mathbb{R}$.

↳ BORELOVE MNOŽICE

- Borelove množice $\subset \mathbb{R}$
  - → vsi intervali
  - → vse elementarne množice

  \+ števne unije
  \+ števni preseki
- Borelove množice $\subset \mathbb{R}^2$
  - → vsi krogi
  - → vse elementarne množice
  - → vse metrično odprte množice
  - → vse zaprte množice

  \+ števne unije
  \+ števni preseki
- Borelove množice $\subset \mathbb{R}^n$
  - → podobno kot v $\mathbb{R}^2$ in posplošeno

ZVEZNI SLUČAJNI VEKTORJI

↳ **DEFINICIJA**

Slučajni vektor $X: \Omega \to \mathbb{R}^n$ ima „zvezno" gostoto $f_X: \mathbb{R}^n \to [0,\infty)$, če za vsako Borelovo množico $B \subset \mathbb{R}^n$ velja

$$\boxed{P(X \in B) = \int_B f_X(x)\,dx}$$

, kjer si lahko mislimo, da gre za izlimitirani Riemannov integral.

↳ $x = (x_1, \dots, x_n)$, $dx = dx_1, \dots, dx_n$

Funkcija $f: \mathbb{R}^n \to \mathbb{R}$ je „zvezna", če je množica točk nezveznosti zanemarljiva za n-terni integral.

↳ **ZGLED**

$X \sim U((a,b))$, če ima X gostoto $f_X(x) = \frac{1}{b-a} \mathbb{1}_{(a,b)}(x)$. Lahko bi vzeli $f_X(x) = \begin{cases} \frac{1}{b-a} \cdot \mathbb{1}_{[a,b]}(x) & ; x \neq \frac{a+b}{2} \\ 42 & ; x = \frac{a+b}{2} \end{cases}$

↳ Osredotočimo se na dvorazsežne slučajne vektorje $(X,Y): \Omega \to \mathbb{R}^2$ z "zvezno" gostoto. Tedaj je

$$\boxed{F_{(X,Y)}(x,y) = P((X,Y) \times (-\infty, x] \times (-\infty, y]) = \int_{(-\infty,x]\times(-\infty,y]} f_{(X,Y)}(u,v)\,du\,dv = \int_{-\infty}^{x}\left(\int_{-\infty}^{y} f_{(X,Y)}(u,v)\,dv\right)du = \int_{-\infty}^{y}\left(\int_{-\infty}^{x} f_{(X,Y)}(u,v)\,du\right)dv}$$

Za martingalni/robni k.p.f velja $F_X(x) = \lim_{y\to\infty} F_{(X,Y)}(x,y) = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{x} f_{(X,Y)}(u,v)\,du\right)dv = \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f_{(X,Y)}(u,v)\,dv\right)du$ [in podobno za F_Y]

Od tod sledi $f_X(x) = F_X'(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,v)\,dv$ za skoraj vse x (*)

osnovni izrek analize

Podobno $f_Y(y) = F_Y'(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(u,y)\,du$ (**)

(formalneje: če je $f_{(X,Y)}: \mathbb{R}^2 \to [0,\infty)$ gostota vektorja (X,Y) sta z izrazoma (*), (**) definirani (možni) marginalni gostoti)

↳ Najpomembnejša dvorazsežna zvezna porazdelitev je **DVORAZSEŽNA NORMALNA PORAZDELITEV**

<u>DEF</u>: Naj bosta μ_x, μ_y realni števili in $\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$ (simetrična) pozitivno definitna matrika, kjer je σ_{xy} realno število.

[OP: Σ je pd $\iff \sigma_x^2\sigma_y^2 - \sigma_{xy}^2 > 0$]. Pravimo, da je s.v. (X,Y) porazdeljen normalno po zakonu $N\left(\begin{bmatrix}\mu_x\\ \mu_y\end{bmatrix}, \begin{bmatrix}\sigma_x^2 & \sigma_{xy}\\ \sigma_{xy} & \sigma_y^2\end{bmatrix}\right)$, če ima gostoto

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi}\frac{1}{\sqrt{\det\Sigma}}\, e^{-\frac{1}{2}\left\langle \Sigma^{-1}\left(\begin{bmatrix}x\\y\end{bmatrix} - \begin{bmatrix}\mu_x\\ \mu_y\end{bmatrix}\right), \begin{bmatrix}x\\y\end{bmatrix} - \begin{bmatrix}\mu_x\\ \mu_y\end{bmatrix}\right\rangle}$$

$$= \frac{1}{2\pi}\frac{1}{\sqrt{\det\Sigma}}\, e^{-\frac{1}{2}\frac{1}{\sigma_y^2\sigma_x^2 - \sigma_{xy}^2}\left(\sigma_x^2(x-\mu_x)^2 - 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_x^2(y-\mu_y)^2\right)}$$

Vpeljemo $\rho = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2}\sqrt{\sigma_y^2}}$. Vidimo, da lahko σ_{xy} (kot parameter) nadomestimo z parametrom $\rho \in (-1,1)$.

$$\Sigma^{-1} = \frac{1}{\det\Sigma}\begin{bmatrix}\sigma_y^2 & -\sigma_{xy}\\ -\sigma_{xy} & \sigma_x^2\end{bmatrix}$$

Gostoto $f_{(X,Y)}$ lahko izrazimo kot: $f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right)}$

$$= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left\langle \begin{bmatrix}1 & -\rho\\ -\rho & 1\end{bmatrix}\begin{bmatrix}\frac{x-\mu_x}{\sigma_x}\\ \frac{y-\mu_y}{\sigma_y}\end{bmatrix}, \begin{bmatrix}\frac{x-\mu_x}{\sigma_x}\\ \frac{y-\mu_y}{\sigma_y}\end{bmatrix}\right\rangle}$$

ta kvadratna forma je konstanta

Nivojnice funkcijo so elipse → glej lanski zvezek

Izkaže se, da velja $\mu_x = E(X)$, $\mu_y = E(Y)$, $\sigma_x^2 = Var(X)$, $\sigma_y^2 = Var(Y)$, $\sigma_{xy} = Cov(X,Y)$

Posplošimo: s.v. $X: \Omega \to \mathbb{R}^n$ je porazdeljen po zakonu $N(\mu, \Sigma)$, kjer je $\mu \in \mathbb{R}^n$ in $\Sigma \in \mathbb{R}^{n\times n}$ s.p.d., če ima n-razsežno gostoto

$$f_X(x) = (2\pi)^{-\frac{n}{2}} (\det\Sigma)^{-\frac{1}{2}}\, e^{-\frac{1}{2}\langle \Sigma^{-1}(x-\mu), x-\mu\rangle}$$

2 NEODVISNOST SLUČAJNIH SPREMENLJIVK

DEFINICIJA – NEODVISNOST

komponente vektorja $(X_1,\dots,X_n)$ so neodvisne, če velja $F_{(X_1,\dots,X_n)}(x_1,\dots,x_n) = F_{X_1}(x_1)\cdot F_{X_2}(x_2)\cdot\ldots\cdot F_{X_n}(x_n)$ za vse realne n-terice $x_1,\dots,x_n$.

Zgornjo enakost lahko prepišemo v: $P(X_1\in(-\infty,x_1] \text{ in } \dots \text{ in } X_n\in(-\infty,x_n]) = P(X_1\in(-\infty,x_1])\cdot\ldots\cdot P(X_n\in(-\infty,x_n])$

Izkaže se, da so s.s. $X_1,\dots,X_n$ neodvisne $\iff$ $\boxed{P\left(\bigcap_{i=1}^{n}\{X_i\in B_i\}\right) = \prod_{i=1}^{n} P(X_i\in B_i)}$ za vse n-terice Borelovih množic $B_1,\dots,B_n\subset\mathbb{R}$

- Posebej s.s. X in Y sta neodvisni, če velja $P(X\in A \text{ in } Y\in B) = P(X\in A)\cdot P(Y\in B)$ za vse Borelove pare $A,B\subset\mathbb{R}$.
- Če so $X_1,\dots,X_n$ neodvisne, sta vsaki X_i, X_j $(i\neq j)$ neodvisni, obratno pa ne velja (kot pri neodvisnosti n-terice dogodkov)

TRDITEV

Naj ima slučajni vektor (X,Y) gostoto $f_{(X,Y)}$. Tedaj sta X in Y neodvisni $\iff$ $\boxed{f_{(X,Y)}(x,y) = f_X(x)\cdot f_Y(y)}$

za skoraj vse realne $(x,y)\in\mathbb{R}^2$

To dodamo, saj je za tiste podmnožice, ki imajo mero 0, jih integral na katerega je vezana gostota „ne vidi"

DOKAZ $(\Rightarrow)$: Vemo $\int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(s,t)\,ds\,dt = \int_{-\infty}^{x} f_X(u)\,du\cdot\int_{-\infty}^{y} f_Y(v)\,dv$ za $\forall x,y$

Odvajamo: $\int_{-\infty}^{y} f_{(X,Y)}(x,v)\,dv = f_X(x)\cdot\int_{-\infty}^{y} f_Y(v)\,dv$ (kjer sta leva in desna stran odvedljivi)

uporabimo osnovni izrek analize

Odvajamo še po y in ustrezen rezultat sledi $\left[f_{(X,Y)}(x,y) = f_X(x)\cdot f_Y(y) \quad \checkmark\right]$

$(\Leftarrow)$ Privzamemo $f_{(X,Y)}(u,v) = f_X(u)\cdot f_Y(v)$ za skoraj vse pare $(u,v)\in\mathbb{R}^2$

Sledi
$$\int_{(-\infty,x]\times(-\infty,y]} f_{(X,Y)}(s,y)\,du\,dv = \int_{(-\infty,x]\times(-\infty,y]} f_X(u)\cdot f_Y(v)\,du\,dv$$

Lahko bi vzeli $A\times B$ za kateri koli Borelovi množici A in B

(Fubinijev izrek da je dvojni integral enak „dvakratnim")

kar po Fubinijevem izreku pomeni ravno $F_{(X,Y)}(x,y) = F_X(x)\cdot F_Y(y)$

ZGLED: Naj bo $(X,Y)\sim N\left(\begin{bmatrix}\mu_X\\ \mu_Y\end{bmatrix}, \begin{bmatrix}\sigma_X^2 & \sigma_{X,Y}\\ \sigma_{XY} & \sigma_Y^2\end{bmatrix}\right)$. Kdaj sta X,Y neodvisni?

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi}\frac{1}{\sigma_X\sigma_Y\sqrt{1-\rho^2}}\cdot e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)}$$

na obeh straneh so gostote gladke

$$f_X(x) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_X}e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} \qquad f_Y(y) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_Y}e^{-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2}$$

KAKO SMO IZRAČUNALI ZGORNJI INTEGRAL, DA SMO DOBILI ROBNE PORAZDELITVE

PO X, da dobiš $f_X(x)$
- ignorirali bi konstante na začetku
- $\int e^{-(ax^2+bx)}$
- nova spremenljivka
- dobiš gama funkcijo...

· če velja $f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y)$ za skoraj vse pare x,y v tem primeru (gladkost) sledi enakost za vse pare x,y

Pri $(x,y) = (\mu_X, \mu_Y)$ dobimo enakost $\frac{1}{2\pi} \cdot \frac{1}{\sigma_X \sigma_Y \sqrt{1-\rho^2}} = \frac{1}{\sigma_X \sqrt{2\pi}} \cdot \frac{1}{\sigma_Y \sqrt{2\pi}}$

od tod sledi $\rho = 0 \iff \sigma_{X,Y} = 0$

· če je $\overset{ro}{\rho} = 0$, je očitno $f_{(X,Y)}(x,y) = f_X(x) f_Y(y)$ povsod

Vidimo: komponenti dvorazsežnega normalnega vektorja sta neodvisni $\iff$ nekorelirani

↳ PODOBNO. Če je $(X_1, \dots, X_n)$ n-razsežno normalno porazdeljen, so $X_1, X_2, \dots, X_n$ neodvisne $\iff X_1, \dots, X_n$ so paroma nekorelirane $\iff X_1, \dots, X_n$ so paroma neodvisne

neodvisno $\iff$ nekoreliranost
samo ko so porazdeljene normalno, drugače ni $\iff$

↳ POSLEDICA TRDITVE

Naj ima (X,Y) gostoto $f_{(X,Y)}$. Tedaj sta X in Y neodvisni $\iff f_{(X,Y)}(x,y) = \phi(x) \cdot \psi(y)$ za skoraj vse x in y za neki integrabilni funkciji ϕ in ψ.

IDEJA DOKAZA: za neki $a>0$ je $\int_{-\infty}^{\infty} a\, \phi(x)\, dx = \int_{-\infty}^{\infty} \frac{1}{a} \psi(y)\, dy = 1$. Sledi, da sta $\frac{1}{a} \cdot \psi$ in $a \cdot \phi$ različici gostot f_X in f_Y.

3 TRANSFORMACIJE SLUČAJNIH SPREMENLJIVK IN VEKTORJEV

↳ **MOTIVACIJA**: Če poznamo porazdelitev s.v. X in je g primerna funkcija kako lahko "izrazimo" porazdelitev transformacije $g(X)$

↳ **V DISKRETNEM PRIMERU** je to preprosto: če je X diskreten s.v. z vrednostmi $\{x_i \mid i \in I\}$ in je g "katerakoli" funkcija, je $g(X)$ slučajni vektor (ali s.s.) z vrednostmi $\{y_j \mid j \in J\}$ za $\{y_j\}_{j\in J}$: $y_j = g(x_i)$

Očitno je $P(g(X)=y_j) = \sum_i \mathbb{1}_{\{y_j\}}(g(x_i)) \cdot P(X=x_i)$

ZGLED: $X \sim \begin{pmatrix} -1 & 0 & 1 \\ 1/3 & 1/2 & 1/6 \end{pmatrix} \implies X^2 \sim \text{"}\begin{pmatrix} 1 & 0 & 1 \\ 1/3 & 1/2 & 1/6 \end{pmatrix}\text{"} = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$

TRANSFORMACIJA SS. $X: \Omega \to (a,b)$

↳ Privzamemo, da je X "zvezna" s.s., za katero je $P(X \in (a,b)) = 1$ in naj bo $g:(a,b) \to (c,d)$ odvedljiva bijekcija.

Zanima nas porazdelitev transformacije $g(X)$ [odv → zveznost (nujno more biti zvezna)]

<u>OPOMBA</u>: dopuščamo tudi neomejene intervale

↳ Privzemimo, da je g strogo naraščajoča. Naj bo $z \in \mathbb{R}$.

bijekcija mora biti monotona, torej je lahko samo strogo naraščajoča ali strogo padajoča

Računamo
$$F_{g(X)}(z) = P(g(X) \leq z) = \begin{cases} P(X \leq g^{-1}(z)) = F_X(g^{-1}(z)) & ; z \in (c,d) \\ 1 & ; z \geq d \\ 0 & ; z \leq c \end{cases}$$

Če ima X gostoto f_X, sledi
$$f_{g(X)}(z) = F'_{g(X)}(z) = \begin{cases} f_X(g^{-1}(z)) \cdot (g^{-1}(z))' & ; z \in (c,d) \\ 0 & ; \text{sicer} \end{cases} = \mathbb{1}_{(c,d)}(z) \cdot \frac{f_X(g^{-1}(z))}{g'(g^{-1}(z))}$$

OP: Formula sledi iz verižnega pravila!

Za DN "izračunajte" $F_{g(X)}(z)$ v primeru padajoče funkcije g.

pazi, kaj se zgodi v ulomku

↳ V primeru gostote za odvedljivo bijekcijo g v splošnem velja
$$\boxed{f_{g(X)}(z) = \mathbb{1}_{(c,d)}(z) \cdot \frac{f_X(g^{-1}(z))}{|g'(g^{-1}(z))|}}$$
TRANSFORMACIJSKA FORMULA

OP: absolutno vrednost moramo dodati, saj je v primeru padajoče g odvod negativen, kar pa nebi bilo vredu, saj mora biti gostota vedno pozitivna

↳ **ZGLED**

$X \sim N(0,1) \implies f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

Naj bo $a \neq 0, b \in \mathbb{R}$ in $g(x) = ax+b$ (linearna bijekcija)

Po transformacijski formuli je
$$\boxed{f_{aX+b}(z) = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{z-b}{a}\right)^2/2} \cdot \frac{1}{|a|}}$$

Prepoznamo gostoto porazdelitve $N(b, a^2)$

↳ **ZGLED**: Nemonotone transformacije

Zanima nas "$N(0,1)^2$" oziroma X^2, če je $X \sim N(0,1)$
[Ta je "problem", saj kvadratična funkcija ni injektivna] ⟶

OP: Ideja tega zgleda: Torej v primeru da nimamo injektivne g, lahko domnevamo razdelimo na injektivne dele in potem računamo naprej

Izračunajmo kar k.p.f.: Vzemimo $z \geq 0$, saj je za $z < 0$ itak verjetnost 0

$$F_{X^2}(z) = P(X^2 \leq z) = P(|X| \leq \sqrt{z}) = P(-\sqrt{z} \leq X \leq \sqrt{z}) \overset{\text{zveznost}}{=} P(X \leq \sqrt{z}) - P(X \leq -\sqrt{z}) =$$
$$= F_X(\sqrt{z}) - F_X(-\sqrt{z})$$

↳ Sledi $f_{X^2}(z) = \left[F'_{X^2}(z) = (F_X(\sqrt{z}) - F_X(-\sqrt{z}))'\right] = f_X(\sqrt{z}) \cdot \frac{1}{2\sqrt{z}} + f_X(-\sqrt{z}) \cdot \frac{1}{2\sqrt{z}}$

V primeru $X \sim N(0,1)$ dobimo $f_{X^2}(z) = \frac{1}{\sqrt{z}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z}$

To je gostota porazdelitve X^2 (hi-kvadrat z eno prostorsko stopnjo) oziroma $\Gamma(\frac{1}{2}, \frac{1}{2})$

↳ **KOMENTAR**: V literaturi uporabljamo $N(\mu, \overset{\text{disperzija}}{\sigma^2})$ ali $N(\mu, \overset{\text{odklon}}{\sigma})$. Menda nihče ne uporablja $N(\mu, \sqrt{\Sigma})$ za večrazsežno normalno. Mi bomo uporabljali kar $N(\mu, \Sigma)$

TRANSFORMACIJE SLUČAJNIH VEKTORJEV

↳ **TRDITEV**: Naj bo X s.v. z vrednostmi v $\mathbb{R}^n$ in naj bo $h: \mathbb{R}^n \to \mathbb{R}$ ZVEZNA funkcija. Tedaj je $h(X)$ slučajna spremenljivka

DOKAZ: $\{h(X) \leq z\} = \{h(X) \in (-\infty, z]\} = \{X \in \underbrace{h^{-1}(\overbrace{(-\infty, z]}^{\text{zaprta}})}_{\text{praslika}}\}$

Praslika zaprte množice je zaprta, torej Borelova, torej gre za dogodek

↳ **ZGLED**: Obravnavajmo vprašanje porazdelitve vsote $X+Y$, tu je $h: \mathbb{R}^2 \to \mathbb{R}$
$h(x,y) = x+y$

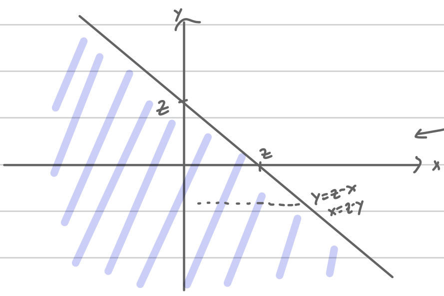
Privzamemo seveda, da ima (X,Y) gostoto $f_{(X,Y)}$

$$F_{(X,Y)}(z) = P(X+Y \leq z) = P((X,Y) \in \{(x,y) \mid x+y \leq z\}) = \int_{\{(x,y) \mid x+y \leq z\}} f_{(X,Y)}(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_{(X,Y)}(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{z} f_{(X,Y)}(u-y,y)\,du\,dy$$

iščemo ravninski integral

integral s parametrom

$u = x+y$
$du = dx$



$$= \int_{-\infty}^{z}\left(\int_{-\infty}^{\infty} f_{(X,Y)}(u-y,y)\,dy\right)du$$

Po osnovnem izreku analize sledi $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(z-y,y)\,dy$

Če sta X in Y neodvisni, se gostota f_{X+Y} poenostavi v $f_{X+Y} = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\,dy$

Tej konstrukciji (kako iz dveh gostot naredimo "tretjo") rečemo KONVOLUCIJA GOSTOT f_X in f_Y

↳ V splošnem porazdelitev funkcij $h(x_1, \dots x_n)$ najenostavneje "izračunamo" tako, da $h: \mathbb{R}^n \to \mathbb{R}$ dopolnimo do diferenciabilne bijekcije $g: D \to E$ med odprtima podmnožicama D in E prostora $\mathbb{R}^n$

Torej: $g(x_1, \dots, x_n) = \underbrace{(h(x_1, \dots, x_n), g_2(x_1, \dots, x_n), \dots, g_n(x_1, \dots, x_n))}_{\text{da bo bijekcija}}$

Tedaj je $h(x_1, \dots, x_n)$ marginalna/robna porazdelitev porazdelitve $g(x_1, \dots, x_n)$

↳ **ZGLED**: $h(X, Y) = X + Y$ lahko dopolnimo do $g(X, Y) = (X + Y, Y)$

↳ **TRDITEV**: Naj bo $f_X: \mathbb{R}^n \to [0, \infty)$ gostota "zveznega" vektorja X, za katerega je $P(X \in D) = 1$ in naj bo $g: D \to E$ diferenciabilna bijekcija med odprtima množicama D in E v $\mathbb{R}^n$. Tedaj ima transformacija $g(X)$ gostoto

$$\boxed{f_{g(X)}(z) = f_X(g^{-1}(z)) \, |\det J g^{-1}(z)| \cdot \mathbb{1}_E(z) = \mathbb{1}_E(z) \cdot \frac{f_X(g^{-1}(z))}{|\det J g(g^{-1}(z))|}}$$

OP: Ponovi Jacobija iz analize

DOKAZ: Naj bo C Borelova $(\subset E)$. $P(g(X)) = C = P(\underbrace{g^{-1}(C)}_{\text{praslika Borelove množice}}) = \int_{g^{-1}(C)} f_X(x) \, dx = \int_C f_X(g^{-1}(z)) \cdot |\det J g^{-1}(z)| \, dz$

izrek o zamenjavi spremenljivk

$z = g(x)$
$x = g^{-1}(z)$

(zamenjava spremenljivk v n-ternem integralu)

7.3

↳ Poglejmo si tradicionalen zapis za dvorazsežnih s.v.:

Standardno pišemo $g(X, Y) = (U, V)$ oziroma $U = U(X, Y)$, $V = V(X, Y)$, ter $(X, Y) = g^{-1}(U, V)$ oziroma $X = X(U, V)$, $Y = Y(U, V)$

Transformacijska formula se glasi:

$$f_{(U,V)}(u, v) = f_{(X,Y)}(x(u,v), y(u,v)) \left| \det \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}_{(u,v)} \right|$$

det; pove, po katerih sprem. računamo; abs

$$\boxed{f_{(U,V)}(u, v) = \frac{f_{(X,Y)}(x(u,v), y(u,v))}{\left| \det \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}_{(x(u,v), y(u,v))} \right|}}$$

↳ **ZGLED**

① Izračunajmo gostoto vsote $X+Y$ za "zvezen" s.v. (X, Y):

Funkcijo $U = X + Y$ dopolnimo do diferenciabilne bijekcije $\mathbb{R}^2 \to \mathbb{R}^2$ s funkcijo $V = Y$ $(g(X, Y) = (X + Y, Y))$

Inverz je podan z $X = U - V$, $Y = V$. Sledi:

$$f_{(X+Y, Y)}(u, v) = \frac{f_{(X,Y)}(u - v, v)}{\left| \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}_{\dots} \right|} = f_{(X,Y)}(u - v, v)$$

Gostota s.s. $U = X + Y$ je robna gostota, torej $f_{(X+Y)}(u) = \int_{-\infty}^{\infty} f_{(X,Y)}(u - v, v) \, dv$

② Izračunajmo še gostoto produkta $U = X \cdot Y$

Funkcijo U dopolnimo do diferenciabilne bijekcije $\overbrace{\underbrace{\mathbb{R} \times (\mathbb{R} \setminus \{0\})}_{D} \to \mathbb{R} \times (\mathbb{R}\setminus\{0\})}^{\text{v def sta to } \epsilon}$ s funkcijo $V = Y$. Inverz je podan z $X = \frac{U}{V}$, $Y = V$

Sledi $f_{(X\cdot Y, Y)}(u,v) = \dfrac{f_{(X,Y)}(\frac{u}{v}, v)}{\left| \left\| \begin{matrix} y & x \\ 0 & 1 \end{matrix} \right\|_{(\frac{u}{v}, v)} \right|} = \frac{1}{|v|} \cdot f_{(X,Y)}(\frac{u}{v}, v)$ za $v \neq 0$

in $f_{XY}(u) = \int_{-\infty}^{\infty} \frac{1}{|v|} f_{(X,Y)}(\frac{u}{v}, v)\, dv$

↳ TRDITEV

Naj bo $X \sim N(\mu, \Sigma)$, kjer je $\mu \in \mathbb{R}^n$ in $\Sigma \in \mathbb{R}^{n\times n}$ s.p.d., in naj bosta $A \in \mathbb{R}^{n\times n}$ obrnljiva matrika ter $\nu \in \mathbb{R}^n$ poljuben vektor. Tedaj velja $\boxed{A \cdot X + \nu \sim N(A \cdot \mu + \nu, A \Sigma A^T)}$

simetrična, pozitivno definitna

saj je pričakovana vrednost linearen operator

var je kvadratičen operator, ne moremo dati $A^2\Sigma$, saj potem var ni s.p.d.

$Var(aX + b) = a^2 \cdot Var(X)$

↳ varianca je kvadratičen operator, ki ne prepozna konstant

DOKAZ

Gre za transformacijo $g: \mathbb{R}^n \to \mathbb{R}^n$, $g(x) = Ax + \nu$

↳ ker je A tudi obrnljiva in je funkcija linearna, bo tudi funkcija obrnljiva, torej imamo inverz

[odvod funkcije je tista funkcija, ki jo v točki najbolj aproksimira]

Inverz: $g^{-1}(z) = A^{-1}(z - \nu)$

Seveda je $Jg(x) = A$

$Jg(x): \mathbb{R}^n \to \mathbb{R}^n$ bomo interpretirali kot preslikavo

Po transformacijski formuli je $f_{AX+\nu}(x) = \dfrac{f_X(A^{-1}(z-\nu))}{|\det A|}$

$$= (2\pi)^{-\frac{n}{2}} \cdot (\det \Sigma)^{-\frac{1}{2}} \cdot \frac{1}{|\det A|}\, e^{-\frac{1}{2}\langle \Sigma^{-1}(A^{-1}(z-\nu)-\mu),\, A^{-1}(z-\nu)-\mu \rangle}$$

multiplikativnost A

$$= (2\pi)^{-\frac{n}{2}} \cdot (\det \Sigma)^{-\frac{1}{2}} \cdot \frac{1}{\sqrt{(\det A)^2}}\, e^{-\frac{1}{2}\langle \Sigma^{-1}(\underbrace{A^{-1}(z-\nu)-\mu}_{A^{-1}(z-\nu-A\mu)}),\, \underbrace{A^{-1}(z-\nu)-\mu}_{A^{-1}(z-\nu-A\mu)} \rangle}$$

$$= (2\pi)^{-\frac{n}{2}} \cdot (\det \Sigma)^{-\frac{1}{2}} \cdot \frac{1}{\sqrt{\det(AA^T)}}\, e^{-\frac{1}{2}\langle \Sigma^{-1}(A^{-1}(z-\nu)-\mu),\, A^{-1}(z-\nu)-\mu \rangle}$$

$$= (2\pi)^{-\frac{n}{2}} \cdot \det(A\Sigma A^T)^{-1/2}\, e^{-\frac{1}{2}\langle (A\Sigma A^T)^{-1}(z-(A\mu+\nu)),\, z-(A\mu+\nu) \rangle}$$

4 PRIČAKOVANA VREDNOST SLUČAJNIH SPREMENLJIVK IN VEKTORJEV

↳ **DEFINICIJA**

Naj bo X "zvezna" s.s. z gostoto $f_X : \mathbb{R} \to [0,\infty)$. Pričakovana vrednost s.s. X je

$$\boxed{E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx}$$, če obstaja $\left(\iff \int_{-\infty}^{\infty} |x| \cdot f_X(x)\,dx < \infty\right)$

Za s.v. $(X_1, \dots, X_n)$ definiramo $E((X_1,\dots,X_n)) = (E(X_1), \dots, E(X_n))$

↳ **KOMENTAR**: Smisel ima tudi definicija $E(X) = \int_{\mathbb{R}^n} x \cdot f_X(x)\,dx$ za s.v. X z gostoto $f_X : \mathbb{R}^n \to [0,\infty)$.

Res: $\int_{\mathbb{R}^n} x_i\, f_X(x_1,\dots,x_n)\, dx_1, \dots, dx_n = \int_{x_i} x_i \left(\int_{\mathbb{R}^{n-1}} f_X(x_1,\dots,x_n)\, dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n \right) dx_i = \int_{x_i=-\infty}^{\infty} x_i \cdot f_{X_i}(x_i)\, dx_i = E(X_i)$

Podobno definiramo pričakovano vrednost slučajnih matrik

↳ **ZGLED**

- Če je $X \sim \text{Cauchy}(0,1)$, $E(X)$ ne obstaja. Pišemo: $E(\text{Cauchy}(0,1))$ ne obstaja
- $E(U(a,b)) = \frac{a+b}{2}$
- $E(N(\mu,\sigma^2)) = \mu$

↳ **TRDITEV**

Naj bo $g : \mathbb{R} \to \mathbb{R}$ zvezna (ali $g : (a,b) \to \mathbb{R}$ zvezna, ki jo razširimo z 0 zunaj (a,b)) in naj bo X „zvezna" s.s. z gostoto f_X. Tedaj je: $\boxed{E(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\,dx}$, če obstaja.

DOKAZ: Za primer, ko je $P(X \in (a,b)) = 1$ in je $g : (a,b) \to (c,d)$ odvedljiva bijekcija. Tedaj vemo, da ima $g(X)$ gostoto:

$$f_{g(X)}(z) = f_X(g^{-1}(z)) \cdot |(g^{-1})'(z)| \cdot \mathbb{1}_{(c,d)}(z)$$

Zato $E(g(X)) = \int_c^d f_X(g^{-1}(z)) \cdot \left| \frac{d}{dz} \cdot (g^{-1}(z)) \right| dz \overset{z=g(x)}{=} \int_a^b g(x) \cdot f_X(x)\,dx$

Zaradi urejenih mej imamo abs. ???

↳ **ZGLED**

Posebej izrazimo $E(|X|)$ za "zvezno" s.s. z gostoto f_X. Najprej $F_{|X|}(z) = P(|X| \le z) = \mathbb{1}_{(0,\infty)}(z) \cdot P(X \in [-z,z]) = \mathbb{1}_{(0,\infty)}(z) \cdot (F_X(z) - F_X(-z))$

Sledi $f_{|X|}(z) = \mathbb{1}_{(0,\infty)}(z) \cdot (f_X(z) + f_X(-z))$ in zato $\boxed{E(|X|) = \int_0^{\infty} z \cdot (f_X(z) + f_X(-z))\,dz = \int_{-\infty}^{\infty} |z| \cdot f_X(z)}$

↳ **POSLEDICA (ZGLEDA IN TRDITVE)**

Za "zvezno" s.s. X z gostoto f_X velja $E(|X|^p) = \int_{-\infty}^{\infty} |x|^p f_X(x)\,dx$

$E(|X|^p) = \int_0^{\infty} z^p \cdot (f_X(z) + f_X(-z))\,dz = \int_{-\infty}^{\infty} |z|^p \cdot f_X(z)$ → obrazložilo z uporabo zgornjega zgleda

↳ TRDITEV

Naj bo $h: \mathbb{R}^n \to \mathbb{R}$ "zvezna" (Borelova) in naj bo X „zvezen" s.v. z gostoto $f_X: \mathbb{R}^n \to [0,\infty)$. Tedaj velja

$$\boxed{E(h(X)) = \int_{\mathbb{R}^n} h(x)\cdot f_X(x)\,dx}$$ če obstaja

↳ ZGLED

Naj bo (X,Y) "zvezen" s.v. z gostoto $f_{(X,Y)}$; izračunajmo $E(X+Y)$. Vemo, da velja

$$f_{X+Y}(u) = \int_{-\infty}^{\infty} f_{(X,Y)}(u-v,v)\,dv$$

Po definiciji je zato $E(X+Y) = \int_{-\infty}^{\infty} u\cdot f_{X+Y}(u)\,du = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u\cdot f_{(X,Y)}(u-v,v)\,dv\,du = \int_{\mathbb{R}^2} \dots\, du\,dv = \int_{\mathbb{R}^2} (x+y)\, f_{(X,Y)}(x,y)\,dx\,dy$

uvedemo $u = x+y$, $v = y$

Nadaljujemo račun: $= \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,dy}_{f_X(x)}dx + \int_{-\infty}^{\infty} y \underbrace{\int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,dx}_{f_Y(y)}dy = E(X)+E(Y)$

↳ POSLEDICA TRDITVE $E(X\cdot Y) = \int_{\mathbb{R}^2} x\cdot y\cdot f_{(X,Y)}(x,y)\,dx\,dy$, če obstaja

D.N.: Če sta X in Y neodvisni in obstajata $E(X)$ in $E(Y)$, obstaja tudi $E(XY)$ in velja $E(XY) = E(X)\cdot E(Y)$

↳ DOKAZ TRDITVE

Za primer ko je $P((X,Y)\in D^{odp}) = 1$ in je $h|_D$ mogoče dopolniti do diferenciabilne bijekcije

$$g = (h,k): D \longrightarrow E$$

Tedaj ima $h(X,Y)$ (kot robna porazdelitev vektorja $g(X,Y)$) gostoto

$$f_{h(X,Y)}(u) = \int_{-\infty}^{\infty} f_{(X,Y)}(g^{-1}(u,v))\cdot |\det J g^{-1}(u,v)|\cdot \mathbb{1}_E(u,v)\,dv$$

Zato je $E(h(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u\cdot f_{(X,Y)}(g^{-1}(u,v))\cdot |\det J g^{-1}(u,v)|\; \mathbb{1}_E(u,v)\,dv\,du \overset{u=h(x,y),\ v=k(x,y)}{=} \int_D h(x,y)\, f_{(X,Y)}(x,y)\,dx\,dy = \int_{\mathbb{R}^2} h(x,y)\, f_{(X,Y)}(x,y)\,dx\,dy$

↳ POSLEDICA ADITIVNOSTI PRIČAKOVANE VREDNOSTI

Če je $X \le Y$ s.g., je $E(X) \le E(Y)$ (Tu $X \le Y$ s.g. pomeni $P(\{\omega \mid X(\omega) \le Y(\omega)\}) = 1$)

Iz $X \le Y$ namreč sledi $Y = X + (Y-X)$, kjer je $Y - X \ge 0$ s.g.

↳ TRDITEV

Naj bo X in Y s.s., za kateri obstajata $E(X^2)$ in $E(Y^2)$. Tedaj obstajata tudi $E(XY)$ in velja: $\boxed{|E(XY)| \le \sqrt{E(X^2)\,E(Y^2)}}$

t.i. CAUCHY – SCHWARZOVA NEENAKOST

DOKAZ: Iz $(u-v)^2 \ge 0$ sledi $u\cdot v \le \frac{1}{2}(u^2+v^2)$ ... KOT PRI VERJETNOSTI 1 → glej lanski VIS

DODATEK: Velja $E(X\cdot Y) = \sqrt{E(X^2)\cdot E(Y^2)} \iff \frac{Y}{\sqrt{E(Y^2)}} = \frac{X}{\sqrt{E(X^2)}}$ s.g.

in $E(XY) = -\sqrt{E(X^2)\cdot E(Y^2)} \iff \frac{Y}{\sqrt{E(Y^2)}} = -\frac{X}{\sqrt{E(X^2)}}$ s.g.

Pripomnimo, da iz trditve sledi sklep $E(X^2) < \infty \Rightarrow E(|X|) < \infty$ (in $E(|X|) \le \sqrt{E(X^2)}$)

DISPERZIJA, KOVARIANCA, VARIANČNO-KOVARIANČNA MATRIKA

↳ **DEFINICIJA**

Disperzija (varianca) s.s. X je $\boxed{D(X) = Var(X) = E\left((X-E(X))^2\right)}$

če obstaja. Vidimo, da obstajati pričakovana vrednost, da lahko obravnavamo obstoj variance.

Gre za pričakovano kvadratično odstopanje s.s. X od njene pričakovane vrednosti. To je mera "razpršenosti" s.s. X.

Iz $(X-E(X))^2 = X^2 - 2E(X)\cdot X + E(X)^2$ sledi da če obstaja $E(X^2)$ obstaja tudi varianca in velja $\boxed{D(X) = E(X^2) - E(X)^2}$

to ni def, ampak bolj "posledica"

↳ Dalje definiramo STANDARDNI ODKLON s.s. X kot $\boxed{\sigma = \sigma(X) = \sqrt{D(X)}}$; pišemo tudi $D(X) = \sigma(X)^2$

Če ima X enote, ima $\sigma(X)$ enake enote, torej je $\sigma(X)$ „neposredno" primerljiv z X.

"inspiracija iz prve norme"

↳ **OPOMBA:** Morda bi bila pričakovana vrednost $E(|X-E(X)|)$ intuitivnejša (in lažje razložljiva) mera za odstopanje, vendar je iz teoretičnega zornega kota kvadratično odstopanje mnogo preprostejše

OP: Inspiracija sledi druge norme, ki ima povezavo s skalarnim produktom

↳ **ZGLED**

- $X \sim U(-\frac{1}{2}, \frac{1}{2})$. Vemo, da je $E(X)=0$, zato je $D(X) = E(X^2) = \int_{-1/2}^{1/2} x^2\, dx = \frac{x^3}{3}\Big|_{-1/2}^{1/2} = \frac{\frac{1}{8}}{3} + \frac{\frac{1}{8}}{3} = \frac{2}{24} = \frac{1}{12}$
- $X \sim N(0,1)$. $E(X)=0 \implies D(X) = E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = \sqrt{\frac{2}{\pi}} \int_0^\infty \sqrt{2u}\, e^{-u}\, du = \frac{2}{\sqrt{\pi}} \int_0^\infty u^{1/2} e^{-u}\, du = \frac{2}{\sqrt{\pi}} \Gamma(3/2) = \frac{2}{\sqrt{\pi}} \frac{1}{2} \cdot \overbrace{\Gamma(1/2)}^{\sqrt{\pi}} = 1$

  $\frac{x^2}{2} = u$

↳ **LASTNOSTI DISPERZIJE**

- $D(X) \geq 0$ in $D(X) = 0 \iff X = \text{konst. s.g.} \iff X = E(X)$ s.g.
- $D(X)$ je minimum funkcije $a \longmapsto E((X-a)^2)$

  Res: $E((X-a)^2) = E((X - E(X) + E(X) - a)^2) = D(X) + (E(X)-a)^2$
- $D(aX+b) = a^2 D(X)$ [pravimo, da je disperzija kvadraten operator]
- $D(X+Y) = D(X) + D(Y)$ če sta X in Y neodvisni

  Res: $(X+Y-E(X)-E(Y))^2 = (X-E(X))^2 + 2\underbrace{(X-E(X))(Y-E(Y))}_{\text{neodvisna}} + (Y-E(Y))^2$ in če sta X in Y neodvisni je $E((X-E(X))(Y-E(Y)))=0$
- $D(X) = \int_{-\infty}^{\infty} (x-E(X))^2 f_X(x)\, dx$ za "zvezno" s.s. X z gostoto f_X

↳ **DEFINICIJA**

KOVARIANCA s.s. X in Y je pričakovana vrednost $\boxed{K(X,Y) = Cov(X,Y) = E\left((X-E(X))\cdot(Y-E(Y))\right)}$, če obstaja

Očitno morata „najprej" obstajati $E(X)$ in $E(Y)$

OP: Kovarianca meri jakost linearne povezanosti med s.s. X in Y.

OP: Meri korelacijo oz. odvisnost med X in Y

Če zmnožimo $(X-E(X))(Y-E(Y))$ sledi $K(X,Y) = E(X\cdot Y) - E(X)\cdot E(Y)$

(natančneje: če obstajajo $E(X), E(Y)$ in $E(XY)$ obstaja tudi $K(X,Y)$ in velja zgornja formula)

↳ LASTNOSTI KOVARIANCE

- če obstajata $E(X^2)$ in $E(Y^2)$, obstaja tudi $K(X,Y)$ in velja $|K(X,Y)| \leq \sqrt{D(X)\cdot D(Y)} = \sigma(X)\cdot\sigma(Y)$ To je Cauchy-Schwarzova neenakost

  (Izpeljava: uporabimo Cauchy-Schwarzova neenakost za $X-E(X)$ in $Y-E(Y)$).

  Dalje velja $K(X,Y)=\sqrt{D(X)\cdot D(Y)} \iff \frac{X-E(X)}{\sqrt{D(X)}} = \frac{Y-E(Y)}{\sqrt{D(Y)}}$ s.g. [pravimo, da sta X in Y s.g. „$\overset{X\uparrow \iff Y\uparrow}{\text{pozitivno}}$" linearno povezani]

  $K(X,Y) = -\sqrt{D(X)\cdot D(Y)} \iff -\frac{X-E(X)}{\sqrt{D(X)}} = \frac{Y-E(Y)}{\sqrt{D(Y)}}$ s.g. [pravimo, da sta s.g. „negativno" lineano povezani]

- Kovarianca je simetričen bilinearen operator

  $K(a_1X_1 + a_2X_2, Y) \underset{\text{linearnost v prvem faktorju}}{=} a_1K(X_1,Y) + a_2K(X_2,Y)$

  $K(X, b_1Y_1 + b_2Y_2) = b_1K(X,Y_1) + b_2K(X,Y_2)$

  RES: simetričnost je očitna, zato je dovolj dokazati linearnost v enem faktorju:

  $K(a_1X_1 + a_2X_2, Y) = E\big((a_1(X_1-E(X_1)) + a_2(X_2-E(X_2)))(Y-E(Y))\big) = a_1K(X_1,Y) + a_2K(X_2,Y)$

- $K(X,X) = D(X)$
- $K(X,Y) = \int_{\mathbb{R}^2}(x-E(X))(y-E(Y))\, f_{(X,Y)}(x,y)\,dxdy$, če je (X,Y) „zvezen" s.v. z gostoto $f_{(X,Y)}$
- $D(X+Y) = D(X) + 2K(X,Y) + D(Y)$ za splošni s.s. X in Y z $E(X^2)<\infty$ in $E(Y^2)<\infty$

  glej lastnosti disperzije

- Splošneje je: $D(X_1+X_2+\ldots+X_n) = \sum_{i=1}^{n}\sum_{j=1}^{n}K(X_i,X_j) = \sum_{i=1}^{n}D(X_i) + 2\cdot\sum_{i<j}K(X_i,X_j)$

VARIANČNO-KOVARIANČNA MATRIKA

↳ **DEFINICIJA**

Variančno-kovariančna (variančna) matrika slučajne s.v. $X=(X_1,\dots,X_n)$ je matrika z elementi $[Var(X)]_{ij} = K(X_i,X_j)$.

Vidimo, da je $Var(X)$ simetrična matrika z variancami komponent po diagonali

Velja $Var(X) = E\big(\underbrace{(X-E(X))}_{\text{stolpec}}\underbrace{(X-E(X))^T}_{\text{vrstica}}\big)$

[Naj bo A poljubna matrika. Tedaj je $A\cdot A^T$ simetrična matrika, ki je vedno pozitivna $\langle A\cdot A^T x, x\rangle = \|A^T x\|^2 \geq 0$]

Sledi $Var(X) \geq 0$

↳ **ZGLED**

Obravnavajmo $(X,Y) \sim N\left(\begin{bmatrix}\mu_X\\ \mu_Y\end{bmatrix}, \begin{bmatrix}\sigma_X^2 & \sigma_{XY}\\ \sigma_{XY} & \sigma_Y^2\end{bmatrix}\right)$

če je $A=\begin{bmatrix}1 & b\\ 0 & 1\end{bmatrix}$ (za neko realno število b), je $A\cdot\begin{bmatrix}\mu_X\\ \mu_Y\end{bmatrix} = \begin{bmatrix}X+bY\\ Y\end{bmatrix} \sim N\left(\begin{bmatrix}\mu_X+b\mu_Y\\ \mu_Y\end{bmatrix}, \underbrace{A\cdot\begin{bmatrix}\sigma_X^2 & \sigma_{XY}\\ \sigma_{XY} & \sigma_Y^2\end{bmatrix}A^T}_{}\right)$

uporaba transformacijske ?

$$\begin{bmatrix}\sigma_X^2+2b\sigma_{XY}+b^2\sigma_Y^2 & \sigma_{XY}+b\sigma_Y^2\\ \sigma_{XY}+b\sigma_Y^2 & \sigma_Y^2\end{bmatrix}$$

Vzemimo $b := -\frac{\sigma_{XY}}{\sigma_Y}$; za ta b dobimo matriko : $\begin{bmatrix}\sigma_X^2+2b\sigma_{XY}+b^2\sigma_Y^2 & 0\\ 0 & \sigma_Y^2\end{bmatrix}$

Vemo, da sta tedaj komponenti $X+bY$ in Y neodvisni. Iz faktorizacije gostot sledi $Y\sim N(\mu_Y,\sigma_Y^2)$

Dalje iz neodvisnosti $X+bY$ in Y sledi $0 = K(X+bY,Y) = K(X,Y)+b\,K(Y,Y) = K(X,Y)+b\cdot\sigma_Y^2 = K(X,Y) - \frac{\sigma_{XY}}{\sigma_Y^2}\cdot\sigma_Y^2$

Končno sledi $\sigma_{XY} = K(X,Y)$

(zaradi simetrije sledi $X\sim N(\mu_X,\sigma_X^2)$)

DODATEK: Vemo $a\cdot N(0,1)+b \sim N(b,a^2)$

$D(N(b,a^2)) = a^2\cdot D(N(0,1)) = a^2$

izračunali prej z gama funkcijo

↳ Splošno za s.v. $X\sim N(\mu,\Sigma)$ velja $E(X)=\mu$ in $Var(X)=\Sigma$

↳ **DEFINICIJA**: Pearsonov korelacijski koeficient s.s. X in Y je $\boxed{\overset{\text{ro}}{\rho}(X,Y) = \frac{K(X,Y)}{\sigma(X)\sigma(Y)}}$. Velja $\rho(X,Y)\in[-1,1]$

Vemo:
- $\rho(X,Y)=1 \iff$ X in Y sta s.g. pozitivno linearno povezani
- $\rho(X,Y)=-1 \iff$ X in Y sta s.g. negativno linearno povezani

↳ **DEFINICIJA**: s.s. X in Y sta nekorelirani $\iff \rho(X,Y)=0$

↳ Vemo, da sta X in Y nekorelirani, če sta neodvisni, obratno pa v splošnem ni res.

Če pa sta X in Y komponenti dvorazsežnega normalnega vektorja, velja tudi obratno.

5 POGOJNE PORAZDELITVE IN POGOJNA PRIČAKOVANA VREDNOST

↳ **PRIMER**

Naj bo $X := LDL$ (holesterol „nizke" gostote → "strupeni")

in $Y := SKT$ (sistolični krvni tlak)

Želimo formalno podlago za $P(Y \in [110,130] \mid X=3)$

GLAVNI PROBLEM POGOJNE VERJETNOSTI: Mi smo "elementarno" pogojno verjetnost že definirali pri verjetnosti kot $P(Y \in B \mid X=x) = \frac{P(Y \in B \cap X=x)}{P(X=x)}$, vedar če je X zvezna SS bo $P(X=x)=0$, torej bi imeli 0 v ulomku, zato smo v naslednji definiciji upeljali limito.

↳ Uvodoma definiramo pogojne porazdelitve takole.

Naj bo B borelova in naj bo x „verjetna" vrednost s.s. X.

DEFINICIJA #1: Pogojna verjetnost, da s.s. Y zavzame vrednost v B, pogojno na $X=x$, je $P(Y \in B \mid X=x) = \lim_{\delta \downarrow 0} P(Y \in B \mid X \in (x-\delta, x+\delta))$

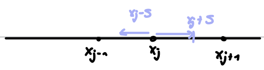
če obstaja. Tu je x „verjetna" vrednost, če je $P(X \in (x-\delta, x+\delta)) > 0$ za vsa št. $\delta > 0$.

Za "zvezno" s.s. X je $P(X \in (x-\delta, x+\delta)) = \int_{x-\delta}^{x+\delta} f_X(t)\,dt$, je to res npr.: če je f_X zvezna pri x in $f_X(x) > 0$

Ta def ni popolnoma matematično pravilna, zato pa jo bomo še malo dopolnili v nadaljevanju

↳ **ZGLED**

Naj bo X diskretna $\sim \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$ za $p_j > 0$ za vse j. Da si malo olajšamo lahko še vzamemo, da so x-i urejeni po vrsti



$P(Y \in B \mid X \in (x_j - \delta, x_j + \delta)) = P(X = x_j)$ za vse dovolj majhne $\delta > 0$ je $P(Y \in B \mid X \in (x_j - \delta, x_j + \delta)) = P(Y \in B \mid X = x_j)$ za vsa dovolj majhna $\delta > 0$

→ Torej nam ta nova def v primeru diskretne SS. vrne elementarno pogojno, kar je ravno kar smo želeli

↳ **TRDITEV**

Naj bo (X,Y) "zvezen" s.v. z gostoto $f_{(X,Y)}$ in naj bo B Borelova množica. Dalje naj bo f_X zvezna in pozitivna pri x. Tedaj velja

Tedaj velja $P(Y \in B \mid X=x) = \frac{1}{f_X(x)} \int_B f_{(X,Y)}(x,y)\,dy$ (pri dodatnih blagih predpostavkah)

DOKAZ: $$P(Y \in B \mid X \in (x-\delta, x+\delta)) = \frac{\frac{1}{2\delta}\int_{(x-\delta,x+\delta)} f_{(X,Y)}(u,y)\,du\,dy}{\frac{1}{2\delta}\int_{x-\delta}^{x+\delta} f_X(u)\,du}$$

→ OP: Dodali smo $\frac{1}{2\delta}$ zato da bomo lahko uporabili izrek o integralski srednji vrednosti (glej Analiza 2: posledica osnovnega izreka analize)

Imenovalec: $\frac{1}{2\delta}\left(F_X(x+\delta) - F_X(x-\delta)\right) \xrightarrow{\delta \downarrow 0} f_X(x)$

ko gre $\delta \to 0$ je to F_X' kar je ravno $f_X(x)$

osnovni izrek o analizi ima pogoj zveznosti, kar smo tudi zgoraj predpostavili

Števec: $\frac{1}{2\delta}\underbrace{\int_{x-\delta}^{x+\delta}\left(\int_B f_{(X,Y)}(u,y)\,dy\right)}_{g(u)}du \xrightarrow{\delta \downarrow 0} g(x)$

za osnovni izrek o analizi mora biti g zvezna, torej moramo imeti pogoj

$\boxed{\text{če je } g \text{ zvezna pri } x}$

↳ Zadnjič: iz uvodne definicije $P(Y\in B \mid X=x)$ smo izpeljali enakost $P(Y\in B \mid X=x) = \int_B \underbrace{\frac{f_{(X,Y)}(x,y)}{f_X(x)}}_{\text{funkcija spremenljivke } y} dy$ (pri pogojih f_X je zvezna in pozitivna v x in dodatnemu pogoju)

Ker je $\int_{\mathbb{R}} \frac{f_{(X,Y)}(x,y)}{f_X(x)} dy = \frac{1}{f_X(x)} \cdot f_X(x) = 1$

Smiselno je torej definirati pogojne gostote:

$$f_{(Y|X)}(y|x) = \begin{cases} \frac{f_{(X,Y)}(x,y)}{f_X(x)} & ; \text{ če } f_X(x) > 0 \\ f_Y(y) & ; \text{ če } f_X(x) = 0 \end{cases}$$

↳ **DROBNI TISK**: Začnemo pri $f_{(X,Y)}$ in definiramo $f_X(x) := \int f_{(X,Y)}(x,y)dy$ in $f_Y(y) := \int f_{(X,Y)}(x,y)dx$. Če sta X in Y neodvisni, lahko izpeljemo $f_X(x,y) = f_X(x) \cdot f_Y(y)$

↳ **DEFINICIJA #2**

Pogojna porazdelitev $P(Y\in \_ \mid X=x)$ je porazdelitev z gostoto $y \longmapsto f_{(Y|X)}(y|x)$

Eksplicitno je $P(Y\in B \mid X=x) \overset{(def)}{=} \int_{\mathbb{R}} f_{(Y|X)}(y|x)\,dy$

↳ **TRDITEV**

Velja $P(Y\in B) = \int P(Y\in B \mid X=x) \cdot f_X(x)\,dx$ [zakon popolne verjetnosti]

$P(Y\in B) = \sum_i P(Y\in B \mid X=x_i) \cdot P(X=x_i)$

DOKAZ: $\int_{\mathbb{R}} P(Y\in B \mid X=x) f_X(x) dx = \int_{\{f_X>0\}} \left(\int_B \frac{f_{(X,Y)}(y)dy}{f_X(x)} dy \right) f_X(x)\,dx = \int_{\mathbb{R}\times B} f_{(X,Y)}(x,y)\,dx\,dy = P((X,Y)\in \mathbb{R}\times B) = P(Y\in B)$

↳ **DEFINICIJA**

Pogojna pričakovana vrednost s.s. Y, pogojno na $X=x$, je $\underbrace{E(Y \mid X=x)}_{h(x)} = \int_{\mathbb{R}} y\, f_{(Y|X)}(y|x)\,dy$

Sledi $E(g(Y) \mid X=x) = \int g(y)\, f_{(Y|X)}(y|x)\,dy$ za poljubno (Borelovo) funkcijo g, če integral obstaja.

↳ **TRDITEV: ZAKON POPOLNE PRIČAKOVANE VREDNOSTI**

$$E(g(Y)) = \int E(g(Y) \mid X=x)\, f_X(x)\,dx$$

DOKAZ: $\int_{\mathbb{R}}\int_{\mathbb{R}} g(y)\, f_{(Y|X)}(y|x)\,dy\, f_X(x)\,dx = \int_{y\in\mathbb{R}} g(y) \overbrace{\int_{x\in\mathbb{R}} f_{(X,Y)}(x,y)\,dx}^{f_Y(y)}\,dy = E(g(Y))$

↳ DEFINICIJA

Pogojna pričakovana vrednost $E(g(Y)|X)$ je slučajna spremenljivka, podana s predpisom $E(g(Y)|X)(\omega) = E(g(Y) \mid X = X(\omega))$

↳ ZGLED:

Naj bo (X,Y) normalen S.V. s parametri $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho$. Tedaj je

$$f_{(Y|X)}(y|x) = \frac{\frac{1}{2\pi}\frac{1}{\sigma_X\sigma_Y\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)}}{\frac{1}{\sigma_X\sqrt{2\pi}}\cdot e^{-\frac{1}{(1-\rho^2)}\left((1-\rho^2)\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right)}}$$

$$= \frac{1}{\sigma_Y\sqrt{1-\rho^2}\cdot\sqrt{2\pi}}\, e^{-\frac{1}{2(1-\rho^2)}\left(\rho^2\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho(\dots)(\dots)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)}$$

$$= \frac{1}{\sqrt{2\pi\,\sigma_Y^2(1-\rho^2)}}\, e^{-\frac{1}{2(1-\rho^2)}\left(\frac{y-\mu_Y}{\sigma_Y} - \rho\,\frac{x-\mu_X}{\sigma_X}\right)^2}$$

Vidimo, da je $(Y|X=x) \sim N\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X),\ \sigma_Y^2(1-\rho^2)\right)$

prepoznaš iz korena, nato pa še to izpostaviš iz potence, da prepoznaš matematično upanje

Posebej je $E(Y|X=x) = \mu_Y + \rho\cdot\frac{\sigma_Y}{\sigma_X}(x-\mu_X)$ in $E(Y|X) = \mu_Y + \rho\cdot\sigma_Y\cdot\frac{X-\mu_X}{\sigma_X}$ (temu izrazu pravimo regresija S.S. Y na S.S. X)

"Porazdelitev je družina verjetnosti"

6 MOMENTI IN MOMENTNO-RODOVNA FUNKCIJA

↳ Naj bodo X s.s., $a \in \mathbb{R}$, $k \in \mathbb{N}$. k-ti moment s.s. X "okrog" a je pričakovana vrednost

$$\boxed{m_k(a) = E((X-a)^k),}$$

obstaja, torej če je $E(|X-a|^k) < \infty$

Momenti $m_k := E(X^k)$ so momenti okrog 0 (standardni momenti), če obstaja $E(X)$, momentom $m_k(E(X))$ pravimo CENTRALNI MOMENTI

Če je X „zvezna" s.s. z gostoto f_X, je $\boxed{m_k(a) = \int_{\mathbb{R}} (x-a)^k f_X(x)dx}$

če integral absolutno konvergira

↳ TRDITEV: Če obstaja $m_n(a)$ za neki $n>1$, obstajajo tudi $m_k(a)$ za $1 \leq k \leq n$

DOKAZ: $E(|X-a|^k) = E(\underbrace{\mathbb{1}_{[-1,1]}}_{\text{indikatorska funkcija}}(X-a) \cdot |X-a|^k + \mathbb{1}_{(\mathbb{R} \setminus [-1,1])}(X-a)|X-a|^k)$

ločimo na dva dela

$\leq 1 + E(|X-a|^n) < \infty$

kako smo prišli do tega

za zvezno: $\int_{-\infty}^{\infty} |x-a|^k f_X(x)dx = \int_{|x-a| \leq 1} |x-a|^k f_X(x)dx + \int_{|x-a|>1} |x-a|^k f_X(x)dx$

↳ TRDITEV: Če obstaja $m_n(a)$ za neki $a \in \mathbb{R}$, obstajajo tudi $m_n(b)$ za vse $b \in \mathbb{R}$

DOKAZ: $E(|X-b|^n) = E(|X-a+a-b|^n) \leq E((|X-a|+|a-b|)^n) = \sum_{k=0}^{n} \binom{n}{k} E(|X-a|^k |a-b|^{n-k}) < \infty$

trikotniška neenakost (abs. vsote je manjša od vsote abs. vrednosti)

Hkrati preberemo $E((X-b)^n) = \sum_{k=0}^{n} \binom{n}{k} (a-b)^{n-k} \cdot m_k(a)$

KOMENTAR: Seveda razumemo $m_0(a) = 1$.

↳ Motivacija za študij momentov pride iz naslednjega vprašanja:

Naj ima s.s. X in Y VSE momente ($\forall k: E(|X|^k)<\infty$, $E(|Y|^k)<\infty$). Če velja $E(X^k)=E(Y^k)$ za vse k, ali sta potem X in Y enako porazdeljeni, torej ali je $P(X \in B) = P(Y \in B)$ za vse Borelove množice B.

porazdelitvena f. je $F: \mathbb{R} \to [0,\infty]$

$\mathbb{R}$ – neštevna množica

če pa bomo gledali momente, ki so števna množica

↳ IZREK

[konvergenčni polmer je vsaj t]

Naj za s.s. X obstajajo vsi $E(X^k)$ in naj za neko število $t>0$ konvergira vrsta $\sum_{k=0}^{\infty} \frac{t^k}{k!} E(X^k)$. Potem je porazdelitev s.s. X določena z momenti, če za drugo s.s. velja $E(X^k)=E(Y^k)$ za vse k, sta X in Y enako porazdeljeni.

↳ DEFINICIJA

MOMENTNO RODOVNA FUNKCIJA s.s. X je funkcija $t \longmapsto E(e^{tX}) =: M_X(t)$

Definirana je za tista števila t, kjer PV obstaja (kjer je $E(e^{tX})<\infty$). Če je X "zvezna" s.s. z gostoto $f_X(x)$, je $M_X(t)=\int e^{tx} f_X(x)dx$

Če je X diskretna, je $M_X(t)=\sum_i e^{tx_i} P(X=x_i)$

Če so vse "operacije" dovoljene, je seveda

$$M_X(t)=E(e^{tX}) = E\Big(\underbrace{\sum_{k=0}^{\infty} \frac{(tX)^k}{k!}}_{\text{Taylorjeva vrsta za } e^{\square}}\Big) = \sum_{k=0}^{\infty}\frac{t^k}{k!}E(X^k)$$

↳ IZREK

dopušča razvoj v Taylorjevo vrsto okrog 0, ki je konvergentna] konvergenten Taylorjev razvoj

Naj $M_X(t)$ obstaja za $t\in(-\delta,\delta)$ za neki $\delta>0$. Tedaj je M_X analitična okrog 0 (s konvergenčnim polmerom $\geq \delta$) in velja

$$M_X(t)=\sum_{k=0}^{\infty}\frac{t^k}{k!}E(X^k) \qquad \text{za } t\in(-\delta,\delta)$$

Posebej je $M_X^{(k)}(0)=E(X^k)$ za vse k. (k-ti odvod)

Sledi, da je porazdelitev s.s. X v tem primeru določena z momenti.

OP: [porazdelitev je določena s števnim naborom kči]

↳ TRDITEV:

$$\boxed{M_{aX+b}(t)=e^{tb}\cdot M_X(at)}$$

DOKAZ: $E\left(e^{t(aX+b)}\right)=e^{tb}\cdot E\left(e^{(at)X}\right)$

↳ LEMA

Naj bosta $g:\mathbb{R}\to\mathbb{R}$ in $h:\mathbb{R}\to\mathbb{R}$ $\overbrace{\text{Borelovi funkciji}}^{\text{"zvezni"}}$. Če sta X in Y neodvisni s.s., sta tudi $g(X)$ in $h(Y)$ neodvisni.

DOKAZ: $P(g(X)\in C \text{ in } h(Y)\in D) = P(X\in g^{-1}(C) \text{ in } Y\in h^{-1}(D)) = P(X\in g^{-1}(C))\cdot P(Y\in h^{-1}(D))$

↳ TRDITEV:

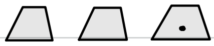
Če sta X in Y neodvisni, je $M_{X+Y}(t)=M_X(t)\cdot M_Y(t)$ (kje ima smisel)

DOKAZ: $M_{X+Y}(t)=E\Big(\overbrace{e^{tX}}^{g(X)}\,\overbrace{e^{tY}}^{g(Y)}\Big)=E(e^{tX})\cdot E(e^{tY})$

neodvisnost

7 LIMITNI IZREKI

↳ IGRA: Imamo 3 kozarčke, pod enim je kroglica, drugi kozarčke premeša, uganit moraš pod katerim kozarčkom je



če uganemo, je $X=10$, če zgrešiš je $X=-6$

$E(x) = 10 \cdot \frac{1}{3} + (-6) \cdot \frac{2}{3} = \frac{10}{3} - \frac{12}{3} = -\frac{2}{3}$

"Pričakujemo", da za veliko število n velja, da je $\frac{X(\omega_1) + X(\omega_2) + \dots + X(\omega_n)}{n}$ blizu $-\frac{2}{3}$

To želimo formalizirati kot $\lim_{x\to\infty} \frac{X(\omega_1)+\dots+X(\omega_n)}{n} = -\frac{2}{3}$

↳ Vidimo, da to lahko interpretiramo le za NESKONČNO ZAPOREDJE $\omega_1, \omega_2, \omega_3, \dots$

UPRAŠANJE: Ali se ne more zgoditi $X(\omega_1) = X(\omega_2) = X(\omega_3) = \dots = 10$ (Tedaj bi bila limita $=10$) — DA, lahko se

Ali se ne more zgoditi $X(\omega_1) = X(\omega_2) = \dots = -6$? — DA, lahko

→ zato uvedemo verjetnost, da bo veljalo formalno

↳ Rekli bomo: Verjetnost tistih neskončnih zaporedij $(\omega_1, \omega_2, \dots)$, za katero obstaja $\lim_{n\to\infty} \frac{X(\omega_1)+\dots+X(\omega_n)}{n}$ in je enaka $-\frac{2}{3}$ je 1.

28.3

↳ KONSTRUKCIJA

Naj bo P verjetnost na σ-algebri $\mathcal{F}$ podmnožic vzorčnega prostora Ω. Naj bo $S := \Omega \times \Omega \times \dots = \Omega^{\mathbb{N}} = \{\underbrace{(\omega_1, \omega_2, \dots)}_{s} \mid \omega_i \in \Omega \text{ za } i \in \mathbb{N}\}$

"prostor" neskončnih zaporedij "elementarnih izidov" iz Ω.

[P je def na vsaki Ω, ne pa na $\Omega^{\mathbb{N}}$]

kaj so (bi bili) dogodki $\subset S$?

Želeli bi, da so vsaj množice oblike $\{s \mid \omega_i \in A_i\}$ kjer je $A_i \in \mathcal{F}$

↓ dogodek na $\mathcal{F}$

Sledi, da morajo biti tudi $\{s \mid \omega_1 \in A_1, \dots, \omega_n \in A_n\} = A_1 \times A_2 \times \dots \times A_n \times \Omega \times \Omega \times \dots$ dogodki

↳ nismo vzeli neskončen, ker se v naravi "skoraj gotovo" ne pojavijo

Od želene verjetnosti $\mathbb{P}$ na „S" želimo, da je "pogoj" $\omega_n \in A_n$ neodvisen od prejšnjih členov $\omega_1, \omega_2, \dots, \omega_{n-1}$, hkrati pa želimo

[neodvisnost je definirana preko verjetnosti]

[tu je nimamo s.s. ampak imamo vzorčenje, katerega "porazdelitev" je ubistvu verjetnost]

da je vzorčenje $\omega_n \in \Omega$ narejeno po porazdelitvi P.

$$\mathbb{P}(\{s \mid \omega_n \in A_n\} \mid \omega_1, \dots, \omega_{n-1}) \underset{\text{neodvisnost}}{=} \mathbb{P}(\omega_n \in A_n) \underset{\text{vzorčimo iz } P}{=} P(A_n)$$

Formalno zahtevamo $\mathbb{P}(A_1 \times \dots \times A_n \times \Omega \times \Omega \times \dots) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_n)$

Izkaže se, da obstaja enolična verjetnost na najmanjši σ-algebri, ki vsebuje vse množice oblike $A_1 \times \dots \times A_n \times \Omega \times \Omega \times \dots$ in zadošča zgornjemu pogoju (izrek Kolmogorova)

↳ **TRDITEV**

Naj bo $X:\Omega\to\mathbb{R}$ poljubna s.s. in naj bodo $X_i: S\to\mathbb{R}$ definirane s predpisom $X_i(s)=X(\omega_i)$. Tedaj so X_i slučajne spremenljivke, ki so neodvisne in vse enako porazdeljene kot X.

KOMENTAR: To je formulacija neskončnega zaporedja neodvisnih replikacij danega slučajnega eksperimenta. !

OPOMBA: Dana množica ss. je neodvisna, če je vsaka njena končna podmnožica neodvisna.

DOKAZ:

- $\mathbb{P}(X_1\in B_1 \text{ in } \dots \text{ in } X_n\in B_n) = \mathbb{P}(X^{-1}(B_1)\times X^{-1}(B_2)\times\dots\times X^{-1}(B_n)\times\Omega\times\Omega\times\dots) =$

$X_1(s)\in B_1 \iff \omega_1\in X^{-1}(B_1)$

$\overset{\text{po def } \mathbb{P}}{\longrightarrow} = \mathbb{P}(X^{-1}(B_1))\cdot \mathbb{P}(X^{-1}(B_2))\cdot\ldots\cdot \mathbb{P}(X^{-1}(B_n)) =$

$\overset{\text{def „nazaj"}}{\longrightarrow} = \mathbb{P}(X_1\in B_1)\cdot\ldots\cdot \mathbb{P}(X_n\in B_n)$

- Zgoraj smo že uporabili $\mathbb{P}(X_i\in B_i)=\mathbb{P}(X\in B_i)$

↳ **IZREK**

Naj bo X s.s z pričakovano vrednostjo. Tedaj je

$$\boxed{\mathbb{P}\left(\{s \mid \exists \lim_{n\to\infty}\frac{X(\omega_1)+\ldots+X(\omega_n)}{n}=E(X)\}\right)=1}$$

Splošneje velja

IZREK (Krepki zakon velikih števil Kolmongorova)

Naj bodo $X_1, X_2, \dots$ NEP s.s. (definirane na nekem skupnem verjetnostnem prostoru) in naj obstaja $\mu := E(X_i)$. Tedaj je

$$P\left(\{s \mid \exists \lim_{n\to\infty}\frac{X_1(s)+\ldots+X_n(s)}{n}=\mu\}\right)=1$$

↳ **TRDITEV (Neenakost Markova)**

Naj bo X ss. s PV in $a>0$. Velja $P(|X|>a)\leq \frac{1}{a}E(|X|)$

DOKAZ

$|X|\geq a\cdot \mathbb{1}_{\{|X|>a\}} \implies E(|X|)\geq a\cdot P(|X|>a)$

↳ **TRDITEV (Neenakost Čebiševa)**

Naj bo X s.s. z $E(X^2)<\infty$ in $\varepsilon>0$. Tedaj je $P(|X-E(X)|>\varepsilon)\leq\frac{D(X)}{\varepsilon^2}$

DOKAZ: $P(|X-E(X)|>\varepsilon)=P(|X-E(X)|^2>\varepsilon^2)$, nato uporabiš neenakost Markova

↳ **IZREK (Šibki zakon velikih števil Markova)**

Naj bodo $X_1, X_2, \dots$ NEP ss. z disperzijo $\sigma^2<\infty$ in pričakovano vrednostjo μ. Tedaj za $\forall\,\varepsilon>0$ velja

$$\lim_{n\to\infty}\mathbb{P}\left(\left|\underbrace{\frac{X_1+\ldots+X_n}{n}}_{\text{ss.}}-\underbrace{\mu}_{\text{konstanta in tudi ss}}\right|>\varepsilon\right)=0$$

OPOMBA: s.s. $\bar{X} = \bar{X}^{(n)} = \frac{X_1 + \dots + X_n}{n}$ pravimo vzorčno povprečje

DOKAZ: Ker je $\mu = E(\bar{X})$ $\left(E(\bar{X}) = \frac{E(X_1)+\dots+E(X_n)}{n} = \frac{n\cdot\mu}{n} = \mu\right)$, je $P(|\bar{X}-\mu| > \varepsilon) \overset{\text{neenakost čebiševa}}{\leq} \frac{D(\bar{X})}{\varepsilon^2} = \frac{\sum_{i=1}^{n} D(X_i)}{n^2 \varepsilon^2} = \frac{n\cdot\sigma^2}{n^2\cdot\varepsilon^2} = \frac{\sigma^2}{n\cdot\varepsilon^2} \xrightarrow{n\to\infty} 0$

OPOMBA: Dovolj je, da so s.s. nekorelirane [neke majorante je omejiti]

↳ ZGLED

Naj bo $X_1, X_2, \dots \overset{NEP}{\sim} \mathrm{Ber}(1,p)$. Pišimo $\hat{p} := \bar{X}$: delež enic v n poskusih

Velja $D(X_i) = p(1-p)$ in $E(X_i) = p$ za vse. Ocena iz dokaza ŠZVŠ: $P(|\hat{p}-p|>\varepsilon) \leq \frac{p(1-p)}{n\,\varepsilon^2} \overset{\text{ocenimo s temenom parabole}}{\leq} \frac{1}{4n\,\varepsilon^2}$ → ocena brez neznanega parametra, GREAT

Privzamemo $n = 10000$ in $\varepsilon \doteq 0{,}05$. Dobimo

$P(|\hat{p}-p| > 0{,}05) \leq \frac{1}{100}$ z verjetnostjo vsaj 99%, se bo razlikovalo vrednost od pričakovane za vsaj 0,05 [npr. 0.45 ↔ 0.55]

ABSTRAKTNA KONVERGENCA ZAPOREDIJ SS.

Naj bodo $Y, Y_1, Y_2, \dots$ s.s. definirane na skupnem verjetnostnem prostoru. Pravimo, da $(Y_n)_{n\in\mathbb{N}}$ konvergira k Y

- VERJETNOSTNO [pišemo: $Y_n \xrightarrow[n\to\infty]{P} Y$], če $\forall \varepsilon > 0$: $\lim_{n\to\infty} P(|Y_n - Y| > \varepsilon) = 0$
- SKORAJ GOTOVO [pišemo $Y_n \xrightarrow{s.g.} Y$], če $P(\exists \lim_{n\to\infty} Y_n = Y) = 1$ → v analizi je to konvergenca po točkah !

  $P(\{\omega \mid \exists \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\}) = 1$

Vidimo, da KZVŠ pravi $\bar{X} \xrightarrow{s.g.} \mu$

ŠZVŠ pravi $\bar{X} \xrightarrow{P} \mu$

↳ LEMA: $\{\omega \mid \exists \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\}$ je dogodek

DOKAZ: $A = \{\omega \mid \forall \varepsilon > 0 \ \exists N \in \mathbb{N}$ da za $\forall n \in \mathbb{N}: |Y_n(\omega) - Y(\omega)| < \varepsilon\}$

$$= \bigcap_{\varepsilon>0} \bigcup_{N=1}^{\infty} \bigcap_{n\in\mathbb{N}}^{\infty} \{\omega \mid |Y_n(\omega) - Y(\omega)| < \varepsilon\}$$

$$= \bigcap_{k=1}^{\infty} \underbrace{\bigcup_{N=1}^{\infty} \underbrace{\bigcap_{n\in\mathbb{N}}^{\infty} \{\omega \mid \underbrace{|Y_n(\omega) - Y(\omega)|}_{\substack{h(Y_n,Y) \\ \text{zvezna funkcija s. spremenljivk} \\ \text{je ss.}}} < \tfrac{1}{k}\}}_{\text{je dogodek}}}_{\text{potem pa imaš števni presek, štetno unijo in števni presek in je to seveda dogodek}}$$ je dogodek

↳ TRDITEV: $(Y_n \xrightarrow{s.g.} Y) \Longrightarrow (Y_n \xrightarrow{P} Y)$

DOKAZ: Pišemo $A_k = *$. Velja $A_1 \supset A_2 \supset A_3 \supset \dots$ (padajoča veriga množic?) Pri privzetku s.g. konvergence sledi $\forall k: P(A_k) = 1$. Pri danem k to pomeni

$1 = P(\underbrace{\bigcup_{N=1}^{\infty} \bigcap_{n\geq N}^{\infty} \dots}_{\text{naraščajoča veriga v } N}) = \lim_{N\to\infty} P(\bigcap_{n\geq N} \{|Y_n - Y| < \tfrac{1}{k}\})$. $P(\text{presek}) \leq P(\text{člena}) \leq 1$ (presek: ↓ konv → 1; člena: ↓ sendvič → 1)

Po izreku o sendviču je $\lim_{N\to\infty} P(\{|Y_n - Y| < \frac{1}{k}\}) = 1$. To je konvergenca v verjetnosti. ■

↳ TRDITEV

Naj bodo $Y_1, Y_2, \ldots$ ss definirane na skupnem verjetnostnem prostoru in naj bo $g: \mathbb{R} \to \mathbb{R}$ zvezna (dovolj je, da je g zvezna na taki Borelovi B, da je zanjo $P(Y \in B) = 1$

(1) če $Y_n \xrightarrow{s.g.} Y$; $g(Y_n) \xrightarrow{s.g.} g(Y)$ — zvezne funkcije ohranjajo konvergenco in limito

(2) če $Y_n \xrightarrow{P} Y$; $g(Y_n) \xrightarrow{P} g(Y)$

DOKAZ: 1) zvezne funkcije ohranjajo konvergenco $\longrightarrow$ PREVERI

2) Najprej dokažemo, da $Y_n \xrightarrow{P} Y \iff$ vsako zaporedje zaporedja Y_n ima nadaljnje podzaporedje, ki k Y konvergira skoraj gotovo.

Potem pa lahko uporabimo 1)

↳ ZGLED – Monte Carlo integracija

Želimo "integrirati" $\int_0^1 f(x)dx$, kjer je f zvezna funkcija. Kadar je to numerično "težko" imamo na voljo M-C "integracijo". Prepoznamo $\int_0^1 f(x)dx = E(f(X))$ za $X \sim U(0,1)$. Če je $X_1, X_2, \ldots$ zaporedje NEP $U(0,1)$ ss., je tudi $f(X_1), f(X_2), \ldots$ zaporedje NEP s.s.

KZVŠ: $\frac{f(X_1)+\ldots+f(X_n)}{n} \xrightarrow[n\to\infty]{s.g.} E(f(U(0,1))) = \int_0^1 f(x)dx$

To lahko uporabimo v praksi s pomočjo simulacije NEP vzorčenja iz $U(0,1)$. Temu pravimo (konstrukcija) psevdoslučajnih števil.

CENTRALNI LIMITNI IZREK

↳ Naj bo Φ k.p.f. standardne normalne porazdelitve, $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt$. Naj bo $X_1, X_2, \ldots$ zaporedje NEP ss. z disperzijo σ^2 in pričakovano vrednostjo μ. Tedaj za $\forall x \in \mathbb{R}$ velja

$$\boxed{\lim_{n\to\infty} P\left(\frac{\frac{X_1+\ldots+X_n}{n} - \mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x) \quad (= P(N(0,1) \leq x))}$$

Ni važno, kakšna je porazdelitev ss. X_i, važni sta σ^2 in μ

↳ POSLEDICA: Za $\forall a < b \ (\in \mathbb{R})$ kjer $\lim_{n\to\infty} P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \in (a,b]\right) = \Phi(b) - \Phi(a) = P(N(0,1) \in (a,b])$

↳ $\frac{\frac{X_1+\ldots+X_n}{n} - \mu}{\sigma/\sqrt{n}}$ je standardizacija vzorčnega povprečja

↳ CLI pravi, da k.p.f standardiziranih vzorčnih povprečij po točkah konvergirajo k k.p.f. od $N(0,1)$

Ponovi si analizo konvergence po točkah!

↳ V praksi CLI uporabljamo za ocenjevanje verjetnosti $P(\bar{X} \in [a,b])$ (torej za ocenjevanje porazdelitve ss. $\bar{X}$) za "VELIKA" števila n: $P(\bar{X} \in (c,d]) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \in \left(\frac{c-\mu}{\sigma/\sqrt{n}}, \frac{d-\mu}{\sigma/\sqrt{n}}\right]\right) \doteq \Phi\left(\frac{d-\mu}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{c-\mu}{\sigma/\sqrt{n}}\right)$

Pravimo, da je za velika št. n vzorčno povprečje $\bar{X}$ porazdeljeno približno normalno, $\bar{X} \dot{\sim} N(\mu, \sigma^2/n)$

TRIVIALNO = dokaz je lahko največ v eni vrstici

Dokaz "šepa", saj je v naravi težko dobiš momentno rodovne funkcije in podobno ?

↳ IZREK O "ZVEZNOSTI"

Naj bo $Y_1, Y_2, \dots$ zaporedje s.s. z momentno-rodovnimi funkcijami (na $(-\delta,\delta)$). Če velja
$$t \in (-\delta,\delta): \lim_{n\to\infty} M_{Y_n}(t) = e^{t^2/2}$$
potem velja za $\forall x \in \mathbb{R}$: $\lim_{n\to\infty} P(Y_n \le x) = \Phi(x)$

OPOMBA: $M_{N(0,1)}(t) = e^{t^2/2}$ (za vse t)

Torej ker momentno rodovna funkcija enolično določa porazdelitev, lahko samo dokažemo da je $\lim_{n\to\infty} M_{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}(t) = e^{t^2/2}$

DOKAZ CLI (pri privzetku obstoja M_X, na nekem intervalu okrog 0.)

Pišimo $Y_n = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{X_1+X_2+\dots+X_n - n\mu}{\sqrt{n}\,\sigma} = \frac{1}{n}\sum_{i=1}^{n} \frac{X_i-\mu}{\sigma}$

$$\Rightarrow M_{Y_n}(t) = M_{\frac{1}{n}\sum_{i=1}^{n}\frac{X_i-\mu}{\sigma}}(t) \overset{M_{aX}(t) = M_X(at)}{=} M_{\sum\frac{X_i-\mu}{\sigma}}\left(\frac{t}{\sqrt{n}}\right) \underset{\text{neodvisnost } X_i}{=} \prod_{i=1}^{n} M_{\frac{X_i-\mu}{\sigma}}\left(\frac{t}{\sqrt{n}}\right) \underset{\text{enako porazdeljeni}}{=} \left(M_{\frac{X_1-\mu}{\sigma}}\left(\frac{t}{\sqrt{n}}\right)\right)^n = \left[1 + \overbrace{E\left(\frac{X_1-\mu}{\sigma}\right)}^{0}\frac{t}{\sqrt{n}} + \frac{1}{2}\overbrace{E\left(\left(\frac{X_1-\mu}{\sigma}\right)^2\right)}^{1}\left(\frac{t}{\sqrt{n}}\right)^2 + \left(\frac{t}{\sqrt{n}}\right)^3 g\left(\frac{t}{\sqrt{n}}\right)\right]^n =$$

saj je $E(X_1) = \mu$

saj je $E((X_1-\mu)^2) = \sigma^2$

momentna je analitična, torej ima Taylorjevo vrsto

$$= \left(1 + \frac{1}{2n}t^2 + \left(\frac{t}{\sqrt{n}}\right)^3 g\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

$$\boxed{\lim_{n\to\infty}(1+a_n)^{b_n} = e^{\lim_{n\to\infty}(a_n b_n)}}$$

$$= e^{\lim_{n\to\infty}\left(t^2/2 + \cancel{\frac{t^3}{\sqrt{n}}}^{\,0} g\left(\frac{t}{\sqrt{n}}\right)\right)}$$

g je zvezna v 0, torej ima g določeno vrednost v 0

$$= e^{t^2/2}$$

Uporabimo lahko izrek o zveznosti

↳ DEFINICIJA

Zaporedje s.s. konvergira k s.s. Y v porazdelitvi (v zakonu), če velja $\lim_{n\to\infty} F_{Y_n}(x) = F_Y(x)$ za vsak x, kjer je F_Y zvezna.

$\forall$: PPT- OPISNA STATISTIKA

SKLEPNA STATISTIKA

↳ Osnovni problem sklepne statistike je za "neznano" funkcijo $X: \Omega \to \mathbb{R}$ iz "vzorca" njenih vrednosti sklepati za njeno "globalno obnašanje".

↳ ZGLED

① Naj bo Ω množica točk na travnikih Golice in $X(\text{točka}) = \#$ narcis v krogu polmera m

Vzorec: naključno izbranih n točk: $\omega_1, \omega_2, \dots, \omega_n$

"Ocena": $\dfrac{X(\omega_1) + \dots + X(\omega_n)}{n}$

S to oceno ocenjamo pričakovano vrednost GLEDE NA PORAZDELITEV na Ω, s katero vzorcimo

② Naj bo Ω množica polnoletnih Zemljanov s previsokim LDL holesterolom in $X(\omega) = (\text{nivo "pod terapijo"} - \text{nivo „po terapiji"})(\omega)$

Z $\dfrac{X_1(\omega) + \dots + X(\omega_n)}{n}$ bomo ocenjevali pričakovano vrednost $E(X)$ GLEDE NA VZORČENJE P NA Ω

↳ Temeljna predpostavka je, da je Ω opremljen (s σ-algebro in) z verjetnostjo. Iz lastnosti vzorca potem sklepamo na lastnosti porazdelitvenega zakona P_X sl. sp. X. Za kvalitetno sklepanje potrebujemo VNAPREJŠNJE PREDPOSTAVKE na (obliko) P_X

1 STATISTIČNI MODEL

↳ Naj bo $X: \Omega \to \mathbb{R}$ proučevana s.s.. Statistični model je MNOŽICA DOPUSTNIH PORAZDELITVENIH ZAKONOV za X. Verjamemo (privzamemo), da je P_X ena izmed njih. Statistično sklepanje lahko predstavimo kot metodo za izbor tiste dopustne porazdelitve, ki je danemu vzorcu najbolje (...) prilega

↳ glede na neke pogoje??

↳ **ZGLED**

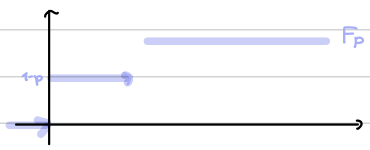
Pojavil se je ponarejen kovanec s $p_1 = 3/5$, uradno pa imajo $p_0 = 1/2$. Najdemo kovanec, ga 1000 krat vržemo in zabeležimo. $2/7$ cifer. SKLEP: kovanec je pošten

↳ Porazdelitev s.s. je enolično določena z njeno k.p.f., torej statistični model lahko enačimo z neko podmnožico $\mathcal{F}$ množice vseh (abstraktnih) k.p.f. Pri izbiri modela nas vodijo različni motivi (zgodovina, empirično preizkušanje), med njimi v praksi izstopa računska (in teoretična) preprostost. Po izbiri modela je sklepna statistika EKSAKTNA VEDA.

↳ **ZGLEDI MODELOV**

- BERNOULLIJEV MODEL

  Privzamemo, da za proučevano slučajno spremenljivko [PSS] X velja $X \sim Ber(1, p)$ za neki neznani $p \in (0,1)$. Torej je $\mathcal{F}$ družina vseh funkcij $F_p : \mathbb{R} \to [0,1]$



- NORMALNI MODEL

  Za PSS X privzamemo: $X \sim N(\mu, \sigma^2)$ za neko neznano $\mu \in \mathbb{R}$ in $\sigma \in (0, \infty)$

$$\mathcal{F} = \{ \mathcal{F}_{N(\mu,\sigma^2)} \mid \mu \in \mathbb{R},\ \sigma^2 \in (0,\infty) \}$$

$$F_{N(\mu,\sigma^2)}(x) = \int_{-\infty}^{x} (\sigma^2 \cdot 2\pi)^{-1/2}\, e^{-\frac{1}{2\sigma^2}(t-\mu)^2}\, dt$$

- ZVEZNI MODEL

  za PSS X privzamemo $P(X = x) = 0$ za vse x; $\mathcal{F}$ je množica vseh ZVEZNIH k.p.f.

- ABSOLUTNO ZVEZNI MODEL

  za PSS X privzamemo gostoto (glede na Lebesgueovo mero)

- POLNI MODEL

  $\mathcal{F}$ = vse k.p.f.

↳ **DEFINICIJA**

Model $\mathcal{F}$ je parametričen, če je $\mathcal{F}$ mogoče parametrizirati s končnim naborom realnoštevilskih parametrov. Tehnično to pomeni, da je $\mathcal{F}$ v bijektivni korespondenci z ustrezno podmnožico $\Theta \subset \mathbb{R}^d$ (za neki d): obstaja bijekcija $\mathcal{F} \leftrightarrow \Theta$

theta ($\equiv$ parametričen prostor)

Tipično želimo, da ima vložitev $\Theta \longleftrightarrow$ prostor vseh k.p.f. še dodatne lastnosti (gladka vložitev gladkih mnogoterosti)

Če model ni parametričen, je neparametričen.

$\Theta = (0,1)$ $\quad$ Θ $\mathbb{R} \times (0,\infty)$

↳ Parametrična modela sta Bernoullijev in Normalni. V parametričnih modelih lahko statistično sklepanje predstavimo kot metodo za izbiro tistega parametra $\vartheta \in \Theta$, pri katerem se ustrezna porazdelitev najbolje prilega danemu vzorcu

↳ Omenimo še parametrični model (1b).

Verjamemo, da je PSS $X \sim \begin{pmatrix} x_0 & x_1 & \cdots & x_d \\ p_0 & p_1 & \cdots & p_d \end{pmatrix}$ za fiksne $x_0 < x_1 < x_2 < \dots < x_r$ in neznane parametre $p_0, p_1, \dots, p_d$

Ta je $\Theta = \{(p_0, p_1, \dots, p_d) \mid \forall j : p_j > 0$ in $\sum_{j=0}^{d} p_j = 1\}$

To je STANDARDNI D-RAZSEŽNI SIMPLEKS V $(0,\infty)^{d+1}$

Temu modelu pravimo tudi kategorični model z $d+1$ kategorijami

2 TOČKOVNO OCENJEVANJE

↳ Naj bo $X:\Omega\to\mathbb{R}$ PSS. Recimo, da nas zanima pričakovana vrednost $E(X)$. Ocenili jo bomo (tipično) s povprečjem

$$\boxed{\frac{X(\omega_1)+\ldots+X(\omega_n)}{n}=\overline{X}(s)}$$

Na vzorcu $s=(\omega_1,\ldots,\omega_n)\in\Omega^n$, kar je prostor vzorcev (realizacij) velikosti n. Če pišemo $S=\Omega^n$ in Ω^n opremimo z običajno (produktno) verjetnostjo in definiramo $X_i:S\to\mathbb{R}$ s predpisom $X_i(s)=X(\omega_i)$, so $X_1,\ldots,X_n$ NEP SS., vse porazdeljene enako kot X. Slučajnemu vektorju $(X_1,\ldots,X_n)$ pravimo standardni neodvisni slučajni vzorec.

Slučajni spremenljivki $\overline{X}:S\to\mathbb{R}$, $\overline{X}=\frac{X_1+\ldots+X_n}{n}$ pravimo vzorčno povprečje. Z njim ocenjujemo pričakovano vrednost $E(X)=E(X_1)=\ldots=E(X_n)$

↳ **DEFINICIJA**

Naj bo c (neznana) karakteristika porazdelitve PSS X (na primer $E(X)$, $D(X)$, $E(X^2)$, $P(X\leq 2)$). Cenilka karakteristike c je (Borelova) funkcija slučajnega vzorca, s katero ocenjujemo c. Cenilka je določena s Borelovo funkcijo $T:\mathbb{R}^n\to\mathbb{R}$ (za $T(x_1,\ldots,x_n)=\frac{x_1+\ldots+x_n}{n}$ je $T(X_1,X_2,\ldots,X_n)=\overline{X}$)

↳ **DEFINICIJA**

Cenilka $T(X_1,\ldots,X_n)$ je NEPRISTRANSKA za c v modelu $\mathcal{F}$, če velja $E(T(X_1,\ldots,X_n))=c(F)$ za $\forall F\in\mathcal{F}$ in vsako n-terico $X_1,X_2,\ldots,X_n\overset{NEP}{\sim}F$.

↳ **ZGLEDI**

- V Poissonovem modelu, kjer $X\sim Poiss(\lambda)$ za nek neznani parameter $\lambda\in(0,\infty)$, je $\overline{X}$ nepristranska cenilka za $D(X)$.

  Reš: Naj bodo $X_1,\ldots,X_n\overset{NEP}{\sim}Poiss(\lambda)$

  $E(\overline{X})=\frac{1}{n}\cdot n\cdot E(Poiss(\lambda))=\lambda=D(Poiss(\lambda))$

  Drobni tisk: $P(Poiss(\lambda)=k)=e^{-\lambda}\cdot\frac{\lambda^k}{k!}$

- $\overline{X}$ je nepristranska cenilka za pričakovano vrednost v kateremkoli modelu s pričakovanimi vrednostmi

- $S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\overline{X})^2$ je nepristranska cenilka za disperzijo v kateremkoli modelu z drugimi momenti:

  $\sum_{i=1}^n(X_i-\overline{X})^2=\sum_{i=1}^n X_i^2-2\overline{X}\sum_{i=1}^n X_i+n\overline{X}^2=\sum_{i=1}^n X_i^2-n\overline{X}^2$

  Pišimo $\mu=E(X_i)$ in $\sigma^2=D(X_i)$

  $$E\Big(\sum_{i=1}^n(X_i-\overline{X})^2\Big)=n(\sigma^2+\mu^2)-n\cdot(D(\overline{X})+E(\overline{X})^2)=n(\sigma^2+\mu^2)-n\Big(\frac{\sigma^2}{n}+\mu^2\Big)=(n-1)\,\sigma^2$$

  $E(X^2)=D(X)+E(X)^2$

↳ Kar smo definirali za realnoštevilske cenilke, na očiten način posplošimo na vektorske cenilke.

Na primer: $T(X_1,\ldots,X_n)=(\overline{X},S^2)$ je nepristranska cenilka za $c(X)=(E(X),D(X))$

Kako bi primerjali dve nepristranski cenilki?

Naj bosta torej U in V nepristranski cenilki karakteristike $c(x) = c(F)$ v danem modelu

ZGLED: → $U(x_1, \dots, x_n) = \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$

→ $V(x_1, \dots, x_n) = x_{42}$ $(n \geq 42)$ → vzamemo samo en podatek

→ $W(x_1, \dots, x_n) = \begin{cases} \frac{x_1 + x_2 + \dots + x_n}{\frac{n+1}{2}} & ; n \text{ je liho število} \\ \frac{x_1 + x_2 + \dots + x_{n-1}}{\frac{n}{2}} & ; n \text{ je sodo število} \end{cases}$ → vzamemo samo pol podatkov

To so nepristranske cenilke za $c(x) = E(x)$

Za merjenje "kakovosti" cenilk nas zanima nepristranskost in tudi odstopanje vseh vrednosti od izbrane cenilke. Odstopanja "merimo" v verjetnosti z disperzijo.

Privzemimo $\mathcal{F}$ z disperzijo

$D(U) = D(\bar{x}) = \frac{D(\text{„}x\text{"})}{n} = \frac{D(F)}{n}$ če $x_i \overset{NEP}{\sim} F$

$D(V) = D(\text{„}x\text{"}) = D(F)$

$D(W) = \begin{cases} \frac{D(F)}{(n+1)/2} & ; n \text{ liho} \\ \frac{D(F)}{n/2} & ; n \text{ sodo} \end{cases}$

(a) ... normalni model

(b) ... $\mathcal{F} = \{ \chi^2_n \mid n \in \mathbb{N} \} = \{ \Gamma(\frac{n}{2}, \frac{1}{2}) \mid n \in \mathbb{N} \}$

$D(\chi^2_n) = 2 \cdot n$

Vidimo, da je $F \longmapsto D(F)$ funkcija $\mathcal{F} \longrightarrow [0, \infty)$ [spreminja se glede na n]

DEFINICIJA

U ima enakomerno manjšo disperzijo od V (na $\mathcal{F}$) če $D(U(x_1, \dots, x_n)) \leq D(V(x_1, \dots, x_n))$ za $\forall F \in \mathcal{F}$ in $\forall$ vzorec $x_1, \dots, x_n \overset{NEP}{\sim} F$

KOMENTAR: Imamo "dovolj velik" nabor standardnih modelov, pri katerih obstajajo nepristranske cenilke z enakomerno najmanjšo disperzijo (Statistika 1)

OBNAŠANJE CENILK ZA VELIKE VZORCE

· Privzemimo neskončno zaporedje $x_1, x_2, \dots \overset{NEP}{\sim} F \in \mathcal{F}$. Naj bo $c(F)$ ocenjevana karakteristika. Zaporedje cenilk $U_n(x_1, \dots, x_n)$ je nepristransko za c v limiti, če velja $\forall F$: $\lim_{n \to \infty} E(U_n(x_1, \dots, x_n)) = c(F)$

· ZGLED: $U_n = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$ (kot zaporedje cenilk za disperzijo) je nepristransko v limiti $(U_n = \frac{n-1}{n} S^2)$

· DEFINICIJA: Zaporedje cenilk U_n za c v modelu $\mathcal{F}$ je DOSLEDNO („consistent") če $\forall F \in \mathcal{F}$ $\forall x_1, \dots \overset{NEP}{\sim} F$:

$$U_n(x_1, \dots, x_n) \xrightarrow{P} c(F)$$

· ZGLED: v $\{ N(\mu, \sigma^2) \mid \mu > 0 \}$ ocenjujemo $\frac{1}{\mu}$; $\frac{1}{\bar{x}}$ je dosledna

METODE ZA KONSTRUKCIJO CENILK

① METODA MOMENTOV

Filozofija metode momentov je, da vse kar je mogoče izraziti z momenti, ocenjujemo preko (standardnih) cenilk za momente.

Spomnimo se : $m_k = E(X^k) = \int_{-\infty}^{\infty} x^k f_X(x)\,dx$

Standardna cenilka za m_k : $\boxed{\hat{m_k}(X_1, \dots, X_k) = \frac{1}{n}\sum_{i=1}^{n} X_i^k}$

To je nepristranska cenilka za m_k, ki tvori (krepko) dosledno zaporedje cenilk za m_k v vsakem modelu s k-timi momenti

· Če ocenjujemo $c(F) = g(m_1, m_2, \dots, m_r)$, bi po metodi momentov vzeli cenilko $g(\hat{m_1}, \hat{m_2}, \dots, \hat{m_r})$ $[1 \le k \le r]$. Če je g zvezna, sledi (krepko) dosledno zaporedje cenilk.

· Tipično (v praksi) postopamo takole:

Privzemimo, da $\mathcal{F} \leftrightarrow \Theta^{odp} \subset \mathbb{R}^r$

Tedaj za $1 \le k \le r$ velja : $m_k = m_k(\vartheta_1, \vartheta_2, \dots, \vartheta_r)$ [momenti so funkcije parametrov]

Privzemimo, da je sistem enačb

$$m_1 = m_1(\vartheta_1, \dots, \vartheta_r)$$
$$\vdots$$
$$m_r = m_r(\vartheta_1, \dots, \vartheta_r)$$

[tukaj iščemo "inverz", zato imamo za definicijsko območje odprto mn., da lahko uporabimo izrek o implicitni funkciji, da si inverz zagotovimo]

mogoče rešiti na $\vartheta = (\vartheta_1, \dots, \vartheta_r)$ · $\vartheta = g(m_1, \dots, m_r)$

funkcija: $\Theta^{odp} \subset \mathbb{R}^r \to \mathbb{R}^r$
inverz g : slika $\longrightarrow \Theta$

Potem po metodi momentov ocenjujemo ϑ s cenilko

$$\hat{\vartheta} = g(\hat{m_1}, \dots, \hat{m_r})$$

in vse druge funkcije parametra ϑ preko $\hat{\vartheta}$.

DOSLEDNOST sledi iz ŠZVŠ saj $m_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k \xrightarrow{P} E(X^k) = \mu_k$?

· **ZGLED**

Vzemimo $\mathcal{F} = \{\text{Beta}(a,b) \mid a, b \in (0,\infty)\}$

$$f_{\text{Beta}(a,b)}(x) = \mathbb{1}_{(0,1)}(x)\, \frac{1}{B(a,b)} \cdot x^{a-1}(1-x)^{b-1}$$

kjer je $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

S tem modelom pogosto modeliramo OMEJENE ZVEZNE s.s.

$$m_k = \frac{1}{B(a,b)} \int_0^1 x^{a+k-1}(1-x)^{b-1}\,dx = \frac{B(a+k,b)}{B(a,b)} = \frac{\Gamma(a+k)\Gamma(a+b)}{\Gamma(a+k+b)\Gamma(a)} = \frac{(a+k-1)\cdot\ldots\cdot a}{(a+k+b-1)\cdot\ldots\cdot(a+b)}$$

Posebej je $m_1 = \frac{a}{a+b}$ $m_2 = \frac{(a+1)\cdot a}{(a+b+1)\cdot(a+b)}$

$$m_1^{-1} = 1 + \frac{b}{a} \quad \text{in} \quad m_2^{-1} = m_1^{-1}\cdot\left(1 + \frac{b}{a+1}\right)$$

$$\Updownarrow \qquad\qquad \Updownarrow$$

$$b = a(m_1^{-1} - 1) \qquad (m_2^{-1} m_1 - 1)(a+1) = b$$

$$a(m_2^{-1} m_1 - m_1^{-1}) = 1 - m_2^{-1} m_1$$

$$\Rightarrow a = \frac{1 - m_2^{-1} m_1}{m_2^{-1} m_1 - m_1^{-1}} \qquad b = a\cdot(m_1^{-1} - 1)$$

$$a = \frac{m_1(m_2 - m_1)}{m_1^2 - m_2} = \frac{m_1(m_1 - m_2)}{m_2 - m_1^2}$$

$m_2 - m_1^2$ → disperzija Beta funkcije > 0

$$\Rightarrow \hat{a} = \frac{\hat{m_1}(\hat{m_1} - \hat{m_2})}{\hat{m_2} - \hat{m_1}^2} \qquad \hat{b} = \hat{a}\left(\frac{1}{\hat{m_1}} - 1\right)$$

② METODA NAJVEČJEGA VERJETJA (the method of maximum likelihood)

· UVODNI ZGLED

Pri našem vprašanju, ali bi pri $217/1000$ cifrah v Bernoullijevem eksperimentu raje podprli $p_0 = \frac{1}{2}$ ali $p_1 = \frac{3}{5}$, bi lahko rekli, da bi podprli $p_0 = \frac{1}{2}$,

ker je $P(B(1000, 1/2) = 217) > P(B(1000, 3/5) = 217)$

Temu rečemo metoda največjega verjetja

· Privzemimo polni Bernoullijev model, $\Theta = (0,1)$, in privzemimo, da imamo v vzorcu velikosti n k-enic. Za oceno za p po MNV bi uzeli tisti $\hat{p}$, pri katerem je $P(B(n,\hat{p}) = k) = L(\hat{p})$ maksimalna.

Seveda $L(p) = \binom{n}{k} p^k (1-p)^{n-k}$. Velja $L'(p) = \binom{n}{k}\left(k \cdot p^{k-1}(1-p)^{n-k} - (n-k)p^k(1-p)^{n-k-1}\right)$

Očitno je $L'(p) = 0 \iff \frac{k}{p} - \frac{n-k}{1-p} = 0 \iff k - np = 0 \iff p = \frac{k}{n}$ analiza, poglej najno❗ [mogoče, ker je 0 v obeh krajiščih in je L pozitivna, torej je to max ??]

Zlahka se prepričamo, da ima L v $\hat{p} = \frac{k}{n}$ maksimum (Prepričajte se, da je $\ln L$ konkavna)

Pravimo, da je cenilka $\hat{p} = \frac{k}{n}$ cenilka za p po MNV. Preberemo $\hat{p} = \bar{x}$ za $x_1, x_2, \dots, x_n \overset{NEP}{\sim}$ Bernoulli (p)

· DEFINICIJA

Privzemimo parametrični model s $\Theta \subset \mathbb{R}^r$ in naj bodo $f(\_; \vartheta)$ gostote (v zveznem primeru) oziroma verjetnostne funkcije (v diskretnem primeru) dopustnih porazdelitev. Funkcija verjetja za vzorec velikosti n je preslikava $L: \mathbb{R}^n \times \Theta \longrightarrow [0,\infty)$,

$$L(x_1, \dots, x_n; \vartheta) = f(x_1; \vartheta) \cdot \ldots \cdot f(x_n, \vartheta)$$

Kot funkcija (realizacija) vzorca je to seveda je to seveda gostota $(\dots)$ sl. vzorca $(X_1, \dots, X_n)$ pri parametru ϑ. V teoriji verjetja L obravnavamo kot funkcijo parametra ϑ.

· Pravimo, da je $\hat{\vartheta}$ OCENA za ϑ na vzorcu $x = (x_1, \dots, x_n)$ po MNV, če velja: $\boxed{L(x, \hat{\vartheta}) = \max_{\vartheta} L(x; \vartheta)}$

· ZGLED

① NEP-BERNOULLIJEV MODEL $\Theta = (0,1)$

$$\begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$$

$$f(x, \vartheta) = p^x (1-p)^{1-x}$$

$$L(x_1, \dots, x_n; p) = p^{\Sigma x_i} (1-p)^{n - \Sigma x_i}$$

Za $\hat{p}(x) = \frac{\Sigma x_i}{n}$ velja $L(\hat{p}(x)) = \max_{p \in [0,1]} L(p)$

razširjen na $[0,1] = \overline{\Theta}$

L moramo obravnavati kot funkcijo na zaprtem intervalu

<u>KOMENTAR</u>: · če $\Sigma x_i = 0$, $L(p) = (1-p)^n$

· če $\Sigma x_i = n$, je $L(p) = p^n$

Vidimo, da moramo dopustiti $\hat{\vartheta} \in \overline{\Theta}$, če je mogoče L razširiti na $\overline{\Theta}$. ❗

② NEP-NORMALNI MODEL

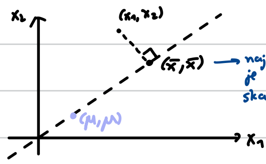
$\Theta = \mathbb{R} \times (0,\infty)$; $f(x; \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$

$\Rightarrow L(x_1, \dots, x_n; \mu, \sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\underbrace{\sum_{i=1}^n (x_i-\mu)^2}_{\mu \text{ je samo tu}}}$

DN: Obravnavajte maksimizacijski problem v dveh spremenljivkah

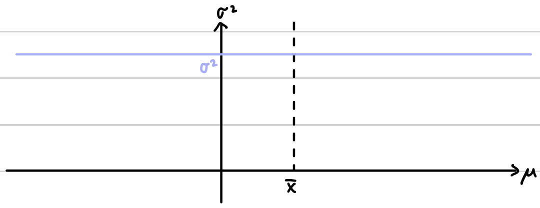
$$\sum_{i=1}^n (x_i - \mu)^2 = \sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2$$

Iz zapisa v Temenski obliki preberemo, da ima $\mu \longmapsto \sum_{i=1}^n (x_i - \mu)^2$ minimum $\sum_{i=1}^n (x_i - \bar{x})^2$, ki je dosežen v $\hat{\mu} = \bar{x} = \frac{\Sigma x_i}{n}$



$\longrightarrow$ najmanjša razdalja, ko je pravokotno, torej je skal. produkt $=0$, zato odpade srednji člen

$$\Longrightarrow L(x; \mu, \sigma^2) \leq (2\pi\sigma^2)^{-\frac{n}{2}} \; e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$



Maksimizacijski problem za $(\mu, \sigma^2) \longmapsto L(x; \mu, \sigma^2)$ je reduciran na maksimizacijski problem za

$$\sigma^2 \longmapsto (2\pi\sigma^2)^{-\frac{n}{2}} \; e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2} =: \Lambda(\sigma^2)$$

Vidimo, da je $\Lambda'(\sigma^2) = 0 \iff -\frac{n}{2}(\sigma^2)^{-\frac{n}{2}-1} + (\sigma^2)^{-\frac{n}{2}}\left(-\frac{1}{2}\cdot\frac{-1}{(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right) = 0$ [tu smo exp izpostavili in pokrajšali]

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

Sledi, da ima $(\mu, \sigma^2) \longmapsto L(x; \mu, \sigma^2)$ maksimum v $\widehat{(\mu, \sigma^2)} = \left(\bar{x}, \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$, če $\sum_{i=1}^{n}(x_i-\bar{x})^2 \neq 0$

Če je $\sum_{i=1}^{n}(x_i-\bar{x})^2 = 0$, L ni omejena. To je $\iff x_1 = x_2 = x_3 = \ldots = x_n$

To nas ne skrbi, ker je dogodek $\{X_1 = X_2 = \ldots = X_n\}$ neverjeten $(P(\ldots) = 0)$ (pri katerem koli $\vartheta = (\mu, \sigma^2)$)

DEFINICIJA

Funkcija $\hat{\vartheta}: D_{\hat{\vartheta}} \longrightarrow \Theta$ je cenilka največjega verjetja (CNV), če velja $P((X_1, \ldots, X_n) \in D_{\hat{\vartheta}}) = 1$ za $\forall \vartheta$ in $\forall X_1, \ldots, X_n \overset{NEP}{\sim} f(\_; \vartheta)$

Za $\forall\, x = (x_1, \ldots, x_n) \in D_{\hat{\vartheta}}$ velja $L(x; \hat{\vartheta}(x)) = \max_{\vartheta \in \Theta} L(x; \vartheta)$

(Tu privzamemo, da je L mogoče smiselno razširiti na Θ)

OP: v primeru da nimaš samo enega parametra, si pomagaš z Hessejevo matriko da poiščeš max

Ker je $\ln$ strogo naraščajoča funkcija je $\hat{\vartheta}$ ocena za ϑ po MNV $\iff \ln L$ ima v $\hat{\vartheta}$ maksimum

$\left(L = \prod_{i=1}^{n} f(x_i; \vartheta), \quad \ln L = \sum_{i=1}^{n} \ln f(x_i; \vartheta)\right)$

Tipično raje obravnavamo $\ln L$. V bolj pomembnih aplikacijah je L (dovoljkrat) zvezno diferenciabilna v $\vartheta \in \Theta^{odp} \subset \mathbb{R}^r$.

Enačbi $\boxed{\frac{\partial(\ln L)}{\partial \vartheta} = 0}$ pravimo LOGARITEMSKA ENAČBA VERJETJA.

$= EV$

REPARAMETRIZACIJA

Dani model lahko parametriziramo na različne načine. Če je $\{\tilde{f}(\_; \tilde{\vartheta}) \mid \tilde{\vartheta} \in \tilde{\Theta}\}$ alternativna parametrizacija, obstaja bijekcija $\tilde{\Theta} \longleftrightarrow \Theta$.

Sledi $(\ldots)$, da sta cenilki $\hat{\tilde{\vartheta}}$ in $\hat{\vartheta}$ povezani preko iste bijekcije.

DEFINICIJA

Naj bo $\Theta^{odp} \subset \mathbb{R}^r$ in naj bo $\hat{\vartheta}$ CNV za ϑ. Če je $g: \mathbb{R}^r \to \mathbb{R}^s$ $(s \leq r)$ Borelova funkcija, pravimo, da je $g(\hat{\vartheta})$ CNV za $g(\vartheta)$

↳ **ZGLED**

Recimo, da nas v NEP-normalnem modelu zanima $E(x^4)$

Izračunamo $E(x^4) = E((x-\mu+\mu)^4)$

$$= E\left(\sigma^2\left[\frac{(x-\mu)^2}{\sigma^2}\right]^2 + 4(x-\mu)^3\mu + 6(x-\mu)^2\mu^2 + 4(x-\mu)\mu^3 + \mu^4\right)$$

$X \sim N(\mu,\sigma^2)$

$$= \sigma^4 E((X_1)^4) + 0 + 6\mu^2\sigma^2 + 0 + \mu^4$$

$$= \sigma^2 \left(D(X_1^2) + E(X_1^2)^2\right) + 6\mu^2\sigma^2 + \mu^4$$

$$= 3\sigma^4 + 6\mu^2\sigma^2 + \mu^4 = g(\mu,\sigma^2)$$

$\Longrightarrow$ CNV za $E(x^4)$ v NEP-normalnem modelu je $3\cdot\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^2 + 6\bar{x}^2\cdot\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 + \bar{x}^4$

↳ Poseben primer definicije je: za $\vartheta = (\vartheta_1, \dots, \vartheta_n)$ uzamemo $\hat{\vartheta}_i = (\hat{\vartheta})_i$ (tu je $g = \mathrm{proj}_i$)

↳ **IZREK (o doslednosti CNV)**

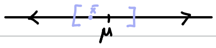
Privzemimo, da so gostote (..) $f(\_;\vartheta)$ dvakrat zvezna diferenciabilne v $\vartheta \in \Theta^{odp} \subset \mathbb{R}^r$, da je množica $\{x \mid f(x;\vartheta) > 0\}$ neodvisna od ϑ in da veljajo še nekateri drugi (blagi) regularnostni privzetki na gostote (...) $f(\_,\vartheta)$

Če imajo enačbe verjetja (EV) enolične rešitve za vsa dovolj velika števila n, potem njihove rešitve tvorijo dosledno zaporedje cenilk za ϑ.

je že vprašal
nekoga na
ustnem

3 INTERVALSKO OCENJEVANJE

↳ Privzemimo, da v nekem modelu ocenjujemo $E(X) =: \mu$. Imamo „najboljšo" nepristransko cenilko $\bar{X}$. Točkovno oceno $\bar{X}$ za μ želimo "razširiti" do intervalske ocene, npr. $[\bar{x} - \text{nekaj}, \bar{x} + \text{nekaj}]$, z željo, da bi dobili verjetnostno zagotovilo, da tak interval vsebuje μ.



Spomnimo se na neenakost Čebiševa $P(|\bar{X}-\mu| > \varepsilon) \leq \frac{D(\bar{X})}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$

Preoblikujmo: $P([\bar{x}-\varepsilon, \bar{x}+\varepsilon] \ni \mu) \geq 1 - \frac{\sigma^2}{n\varepsilon^2}$

Recimo, da želimo, da naša intervalska ocena vsebuje μ z verjetnostjo $\geq 1-\alpha$, kjer je α majhno število.

Zahtevajmo $\frac{\sigma^2}{n\varepsilon^2} = \alpha \iff \varepsilon = \sqrt{\frac{\sigma^2}{n\alpha}} = \frac{1}{\sqrt{\alpha}} \cdot \frac{\sigma}{\sqrt{n}}$

Sledi $P([\bar{x} - \frac{1}{\sqrt{\alpha}}\frac{\sigma}{\sqrt{n}}, \bar{x} + \frac{1}{\sqrt{\alpha}}\frac{\sigma}{\sqrt{n}}] \ni \mu) \geq 1-\alpha$

9.5

↳ **DEFINICIJA**

Privzemimo model $\mathcal{F}$ in naj bosta $\alpha \in (0,1)$ (majhno) in $n \in \mathbb{N}$ dani števili. Interval zaupanja z karakteristiko c za vzorec velikosti n in stopnjo zaupanja $1-\alpha$ je par preslikav $L, U: \mathbb{R}^n \to \mathbb{R}$ ($L \leq U$ povsod), ki zadošča zahtevi

$$\boxed{P([L(X_1, \dots, X_n), U(X_1, \dots, X_n)] \ni c(F)) \geq 1-\alpha}$$ ← samo funkcija vzorca

za vsako $F \in \mathcal{F}$ in vsak vzorec $X_1, \dots, X_n \overset{\text{nep}}{\sim} F$

↳ Zgoraj smo iz neenakosti Čebiševa izpeljali interval zaupanja za $c(X) = E(X)$: $\left[\bar{x} - \frac{1}{\sqrt{\alpha}}\frac{\sigma}{\sqrt{n}}, \bar{x} + \frac{1}{\sqrt{\alpha}}\frac{\sigma}{\sqrt{n}}\right]$ v poljubnem modelu z znano disperzijo

INTERVALI ZAUPANJA V NORMALNIH MODELIH

① IZ za μ za vzorec $X_1, \dots, X_n \overset{NEP}{\sim} N(\mu, \sigma^2)$ pri znani σ^2

porazdelitev neodvisna od μ !

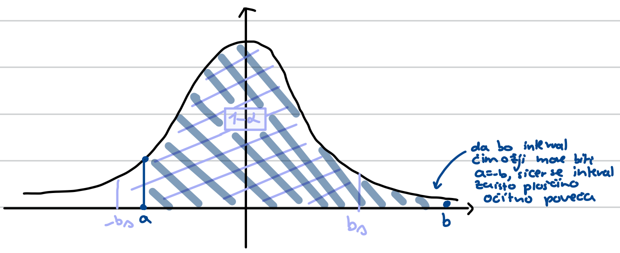
Vemo $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$. Če sta $a<b$ taki števili, da je zanje $P(N(0,1) \in [a,b]) = \Phi(b) - \Phi(a) = 1-\alpha$, velja $P\left(a \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le b\right) = 1-\alpha$

$$\Phi(b) - \Phi(a) = \int_a^b f_{N(0,1)}(t)\,dt$$

Prepišimo $\left\{a \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le b\right\} = \left\{\bar{X} - b\cdot\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} - a\cdot\frac{\sigma}{\sqrt{n}}\right\}$

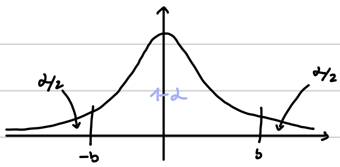
Sledi $P\Big(\big[\underbrace{\bar{X} - b\cdot\tfrac{\sigma}{\sqrt{n}}}_{L(x_1,\dots,x_n)}, \underbrace{\bar{X} - a\cdot\tfrac{\sigma}{\sqrt{n}}}_{U(x_1,\dots,x_n)}\big] \ni \mu\Big) = 1-\alpha$, torej imamo interval zaupanja za μ.

Imamo jih toliko, kot je parov $a<b$, ki zadoščajo zgornji zvezi



Takih parov je neskončno mnogo · za katerikoli a, za katerega je $P(N(0,1) \le a) < \alpha$, obstaja zahtevani b. Izbrali bi tisti par, pri katerem je intervalska ocena najbolj informativna, torej "najožjega". Širina zgornjega intervala je $(b-a)\cdot\frac{\sigma}{\sqrt{n}}$.
Zaradi simetričnosti standardne normalne porazdelitve je $b-a$ minimalna pri $-a = b = \Phi^{-1}(1-\alpha/2)$

$$\Phi^{-1}(1-\alpha/2) = z_{\alpha/2} \quad \text{majhen } z$$



Dobimo standardni IZ: $\left[\bar{X} - z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}\right]$ (majhen z)

↳ ZGLED: Primerjava z IZ iz neenakosti Čebiševa: za $\alpha = 0{,}05$ je $\frac{1}{\sqrt{\alpha}} \doteq 4{,}4$, $z_{\alpha/2} = 1{,}96$

KAJ PA V PRIMERU NEZNANE σ^2? σ^2 ocenimo npr. s $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

↳ **IZREK (Student-Fisher)**

Naj bodo $X_1, X_2, \dots, X_n \overset{NEP}{\sim} N(\mu, \sigma^2)$. Tedaj sta $\bar{X}$ in S^2 neodvisni in $\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{\sigma}\right)^2 = \frac{n-1}{\sigma^2} S^2$ ima porazdelitev $\boxed{\chi^2_{n-1} = \text{Gama}\left(\frac{n-1}{2}, \frac{1}{2}\right)}$

↳ **POSLEDICA** Slučajna spremenljivka $t = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ ima porazdelitev, ki je neodvisna od μ in σ^2

DOKAZ POSLEDICE

$$t = \frac{\overbrace{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}^{\sim N(0,1) \to \text{neodvisno od } \mu}}{\sqrt{\underbrace{\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2}_{\chi^2_{n-1} \to \text{neodvisno od } \sigma^2}/(n-1)}}$$

Torej bo funkcija neodvisnih, tudi neodvisna od μ in σ^2 ??

↳ **DEFINICIJA**

Naj bosta Z in H neodvisni s.s., $Z \sim N(0,1)$, $H \sim \chi^2_h = \text{Gama}\left(\frac{h}{2}, \frac{1}{2}\right)$. Tedaj ima: $\boxed{t = \frac{Z}{\sqrt{H/h}}}$ studentovo porazdelitev s h prostostnimi stopnjami, označimo jo t_h.

Očitno je t_h porazdelitev, ki lahko zavzame katerokoli realno število. Poleg tega je t_h simetrična $(-t_h \sim t_h)$. To sledi iz simetričnosti porazdelitve $N(0,1)$.

DN: Iz gostote $N(0,1)$ in χ^2_h izpeljite gostoto t_h.

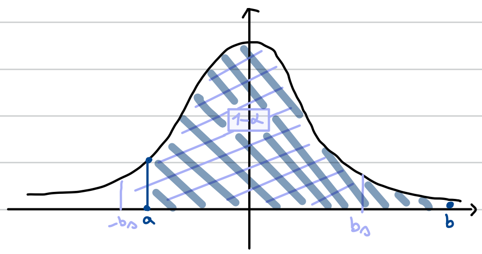
② **IZ za μ za vzorec $X_1, \dots, X_n \overset{NEP}{\sim} N(\mu, \sigma^2)$ pri neznani σ^2**

Vemo $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$. Če sta $a<b$ taki števili, da je zanje $P(t_{n-1} \in [a,b]) = F_{t_{n-1}}(b) - F_{t_{n-1}}(a) = 1-\alpha$ $\left(= \int_a^b f_{\chi^2_{n-1}}(u)\,du\right)$, velja $P\left(a \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq b\right) = 1-\alpha$

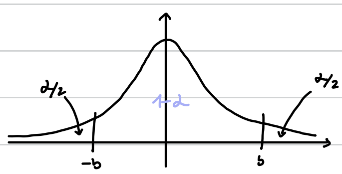
Prepišimo $\left\{a \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq b\right\} = \left\{\bar{X} - b\cdot\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} - a\cdot\frac{S}{\sqrt{n}}\right\}$

Sledi $P\left(\left[\underbrace{\bar{X} - b\cdot\frac{S}{\sqrt{n}}}_{L(x_1,\dots,x_n)}, \underbrace{\bar{X} - a\cdot\frac{S}{\sqrt{n}}}_{U(x_1,\dots,x_n)}\right] \ni \mu\right) = 1-\alpha$, torej imamo interval zaupanja za μ.

Imamo jih toliko, kot je parov $a<b$, ki zadoščajo zgornji zvezi



Takih parov je neskončno mnogo: za katerikoli a, za katerega je $P(t_{n-1} \leq a) < \alpha$, obstaja zahtevani b. Izbrali bi tisti par, pri katerem je intervalska ocena najbolj informativna, torej "najožjega". Širina zgornjega intervala je $(b-a)\cdot\frac{S}{\sqrt{n}}$. Ta širina je s.s.. Zaradi simetričnosti studentove porazdelitve je $b-a$ minimalna pri $-a = b = \underbrace{F_{t_{n-1}}^{-1}\left(\frac{1-\alpha}{2}\right)}_{= t_{n-1;\alpha/2}}$



Dobimo standardni IZ: $\left[\bar{X} - t_{n-1,\alpha/2}\cdot\frac{S}{\sqrt{n}}, \bar{X} + t_{n-1,\alpha/2}\cdot\frac{S}{\sqrt{n}}\right]$

↳ DOKAZ (STUDENT-FISHER)

Pišimo $Y_i = \frac{X_i - \mu}{\sigma}$. Seveda $Y_i \overset{NEP}{\sim} N(0,1)$. Zaradi linearnosti je $Y_i - \bar{Y} = \frac{X_i - \bar{X}}{\sigma}$ $\left(\Leftarrow \bar{Y} = \frac{\bar{X}-\mu}{\sigma}\right)$

Dovolj je dokazati, da sta $\bar{Y}$ in $\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ neodvisni in da je $\sum_{i=1}^{n}(Y_i-\bar{Y})^2 \sim \chi^2_{n-1}$. Izrazimo

$$Y_i - \bar{Y} = -\frac{1}{n}Y_1 - \frac{1}{n}Y_2 - \dots + \left(1-\frac{1}{n}\right)Y_i - \dots - \frac{1}{n}Y_n$$

Vpeljemo matriko $A_{ij} = -\frac{1}{n} + \underbrace{\mathbb{1}_{\{i=j\}}}_{\delta_{ij}}$ oziroma $A_{ij} = I - \frac{1}{n}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1-\frac{1}{n} & -\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & -\frac{1}{n} & 1-\frac{1}{n} & & -\frac{1}{n} \\ \vdots & \vdots & & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & -\frac{1}{n} & & 1-\frac{1}{n} \end{bmatrix}$

simetrične matrike lahko diagonaliziramo (Q je matrika ONB lastnih vektorjev)

Očitno je $\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \left\| A \cdot \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} \right\|_2$. Ker je A simetrična ima realne lastne vrednosti in jo je mogoče diagonalizirati (nad $\mathbb{R}$) v ON lastni bazi.

$$A = Q \cdot D \cdot Q^T$$

kjer je $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ matrika lastnih vrednosti Q pa ortogonalna matrika pridruženih lastnih vektorjev $(Q^TQ = QQ^T = I)$

lastne vrednosti: $A - \lambda I = I - \frac{1}{n}\begin{bmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \end{bmatrix} - \lambda I = -\frac{1}{n}\begin{bmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \end{bmatrix} - (\lambda - 1) I$

$$\underbrace{\det(A - \lambda I)}_{p_A(\lambda)} = (-1)^n \cdot \underbrace{\det(B - \overset{\mu}{\widetilde{(\lambda-1)}} I)}_{p_B(\mu)} \quad \text{kjer je} \quad B = \frac{1}{n}\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

karakteristični polinom

↳ rang $B = 1$ ⟶ rang $\ker(B) = n-1$ ⟶ lastna vr. 0 je $n-1$ kratna za B

Očitno je $\mu = 0$ lastna vr. stopnje $(n-1)$ za B

Dalje je $B\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, torej imamo še eno lastno vrednost $\mu = 1$ s pripadajočim lastnim vektorjem $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

Sledi $\underset{\mu = 1 - \lambda}{D} = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{bmatrix}$ in $A \cdot \begin{bmatrix} 1/\sqrt{n} \\ \vdots \\ 1/\sqrt{n} \end{bmatrix} = 0$, torej $Q = \left[\underbrace{\Big| \cdots \Big|}_{n-1 \text{ neodvisnih stolpcev}} \begin{matrix} 1/\sqrt{n} \\ \vdots \\ 1/\sqrt{n} \end{matrix} \right]$

$$\begin{aligned} f(y_1, \dots, y_n) &= \prod_{i=1}^{n} \left[(2\pi)^{-1/2} e^{-1/2 y_i^2} \right] \\ &= (2\pi)^{-n/2} e^{-1/2 \sum_{i=1}^{n} y_i^2} \\ &= (2\pi)^{-n/2} e^{-1/2 \langle y, y \rangle} \\ &= (2\pi)^{n/2} e^{-1/2 \langle I(y-0), y-0 \rangle} \end{aligned}$$

velja $\underbrace{Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, I_{n \times n}\right)}_{}$ n-razsežna standardna normalna

ortogonalna matrika ohranja skalarni produkt

$$\|AY\|^2 = \langle AY, AY \rangle = \langle QDQ^TY, QDQ^TY \rangle \underset{Q^TQ = QQ^T = I}{=} \langle DQ^TY, DQ^TY \rangle = \langle D(Q^TY), DQ^TY \rangle \underset{D^2 = D}{=} \|D(Q^TY)\|_2^2 = \sum_{i=1}^{n-1} Z_i^2$$

glej izrek na strani 11

kjer $Z = Q^TY \sim N(\underset{Q^T 0}{0}, \underbrace{Q^T I (Q^T)^T}_{I})$

Sledi $Z_i \overset{NEP}{\sim} N(0,1)$ in zato $\sum_{i=1}^{n-1} Z_i^2 \sim \chi^2_{n-1}$

Poleg tega sta $\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ in Z_n neodvisni. Po definiciji je $Z_n = \frac{1}{\sqrt{n}}(Y_1 + \dots + Y_n) = \sqrt{n}\,\bar{Y}$ ∎

saj so lin. kombinacije različnih lastnih vektorjev, ki pa so neodvisni

③ IZ za σ^2 pri znani μ v NEP-normalnem modelu

Za $X_1, X_2, \dots, X_n \overset{NEP}{\sim} N(\mu, \sigma^2)$ velja $\boxed{\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2_n}$

Naj bosta $a, b \in (0,\infty)$ taki št., da $a<b$ in: $P(\chi^2_n \in [a,b]) = F_{\chi^2_n}(b) - F_{\chi^2_n}(a) = 1-\alpha$

Tedaj je $P\left(a \le \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \le b\right) = 1-\alpha$

$$\parallel$$

$$P\left(\frac{1}{b}\sum_{i=1}^{n}(X_i-\mu)^2 \le \sigma^2 \le \frac{1}{a}\sum_{i=1}^{n}(X_i-\mu)^2\right)$$

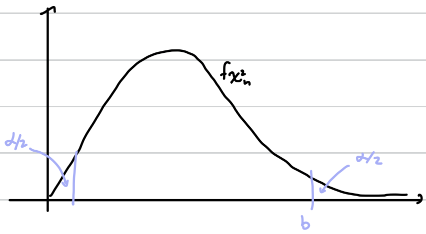
To pomeni, da je $\boxed{\left[\frac{1}{b}\sum_{i=1}^{n}(X_i-\mu)^2, \frac{1}{a}\sum_{i=1}^{n}(X_i-\mu)^2\right]}$ IZ za σ^2 za vzorec velikosti n stopnje zaupanja $1-\alpha$.

Zanima nas, katera izbira para (a,b) je optimalna v smislu širine IZ. Ta je $\left(\frac{1}{a}-\frac{1}{b}\right)\sum_{i=1}^{n}(X_i-\mu)^2$. To je s.s.!

Če sta a^* in b^* taki števili, da je $\frac{1}{a^*}-\frac{1}{b^*}$ minimalna pri zvezi $P(\chi^2_n \in [a,b]) = 1-\alpha$, ima pridruženi interval zaupanja ENAKOMERNO najmanjšo širino:

$$\forall (x_1, \dots, x_n): \left(\frac{1}{a^*}-\frac{1}{b^*}\right)\sum_{i=1}^{n}(x_i-\mu)^2 \le \left(\frac{1}{a}-\frac{1}{b}\right)\sum_{i=1}^{n}(x_i-\mu)^2$$

Izkaže se, da je ta minimizacijski problem $\left(\min\left(\frac{1}{a}-\frac{1}{b}\right) \text{ z vezjo } \int_a^b f_{\chi^2_n}(u)\,du = 1-\alpha\right)$ računsko intenziven, zato v praksi tipično vzamemo $\boxed{\tilde{a} = F^{-1}_{\chi^2_n}(\alpha/2) \quad \tilde{b} = F^{-1}_{\chi^2_n}(1-\alpha/2)}$ ENAKOREPA IZBIRA



Velja $\frac{a^*}{\tilde{a}} \xrightarrow{n\to\infty} 1$ in $\frac{b^*}{\tilde{b}} \xrightarrow{n\to\infty} 1$

④ IZ za σ^2 pri neznani μ v NEP-normalnem modelu

Za $X_1, X_2, \dots, X_n \overset{NEP}{\sim} N(\mu, \sigma^2)$ velja $\boxed{\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2_{n-1}}$ iz Student-Fisherjevega izreka

Naj bosta $a, b \in (0,\infty)$ taki št., da $a<b$ in: $P(\chi^2_{n-1} \in [a,b]) = F_{\chi^2_{n-1}}(b) - F_{\chi^2_{n-1}}(a) = 1-\alpha$

Tedaj je $P\left(a \le \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\bar{X})^2 \le b\right) = 1-\alpha$

$$\parallel$$

$$P\left(\frac{1}{b}\sum_{i=1}^{n}(X_i-\bar{X})^2 \le \sigma^2 \le \frac{1}{a}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)$$

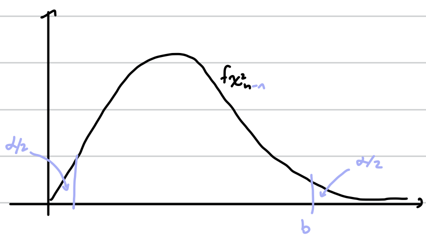
To pomeni, da je $\boxed{\left[\frac{1}{b}\sum_{i=1}^{n}(X_i-\bar{X})^2, \frac{1}{a}\sum_{i=1}^{n}(X_i-\bar{X})^2\right]}$ IZ za σ^2 za vzorec velikosti n stopnje zaupanja $1-\alpha$.

Zanima nas, katera izbira para (a,b) je optimalna v smislu širine IZ. Ta je $\left(\frac{1}{a}-\frac{1}{b}\right)\sum_{i=1}^{n}(X_i-\bar{X})$. To je s.s.!

Če sta a^* in b^* taki števili, da je $\frac{1}{a^*}-\frac{1}{b^*}$ minimalna pri zvezi $P(\chi^2_{n-1} \in [a,b]) = 1-\alpha$, ima pridruženi interval zaupanja ENAKOMERNO najmanjšo širino:

$$\forall (x_1, \dots, x_n): \left(\frac{1}{a^*}-\frac{1}{b^*}\right)\sum_{i=1}^{n}(x_i-\bar{x})^2 \le \left(\frac{1}{a}-\frac{1}{b}\right)\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Izkaže se, da je ta minimizacijski problem $\left(\min\left(\frac{1}{a}-\frac{1}{b}\right) \text{ z vezjo } \int_a^b f_{\chi^2_{n-1}}(u)\,du = 1-\alpha\right)$ računsko intenziven, zato v praksi tipično vzamemo $\boxed{\tilde{a} = F^{-1}_{\chi^2_{n-1}}(\alpha/2) \quad \tilde{b} = F^{-1}_{\chi^2_{n-1}}(1-\alpha/2)}$ ENAKOREPA IZBIRA



Velja $\frac{a^*}{\tilde{a}} \xrightarrow{n\to\infty} 1$ in $\frac{b^*}{\tilde{b}} \xrightarrow{n\to\infty} 1$

PREIZKUŠANJE DOMNEV

UVOD

Recimo, da obstajata dva tipa evrskih kovancev: "pošten" s $p_0 = 1/2$ in "nepošten" s $p_1 = 3/5$. Za kovanec v roki bi na podlagi 10 (neodvisnih slučajnih) metov radi bodisi podprli domnevo "poštenosti" ($p = p_0$) bodisi podprli domnevo "nepoštenosti" ($p = p_1$). Naj bo k število enic

k	1	2	3	4	5	6	7	8	9	10
odločitev	p_0	p_0	p_0	p_0	p_0	p_1	p_1	p_1	p_1	p_1

$\underbrace{1\dots5}_{\text{podpremo } p_0}$ $\underbrace{6\dots10}_{\text{podpremo } p_1}$

kaj gre lahko narobe?

1. Pošten kovanec lahko razglasimo za nepoštenega
   - Verjetnost te zmote je $P(B(10, 1/2) > 5) = 0{,}38$
   - Če bi podprli p_0 do vključno 8, bi imeli verjetnost zmote $P(B(10, 1/2) > 8) \doteq 0{,}01$
2. Nepošten kovanec "razglasimo" za poštenega
   - Verjetnost te zmote je $P(B(0, 3/5) \le 8) \doteq 0{,}95$

Splošneje bi $p = p_0$ podprli za $k \le C$, $p = p_1$ za $k > C$ za neko konstanto C

k	0	1	...	c	$c+1$	...	9	10
odločitev	p_0	p_0	...	p_0	p_1	...	p_1	p_1

$\underbrace{0\dots c}_{\text{podpremo } p_0}$ $\underbrace{c+1\dots10}_{\text{podpremo } p_1}$

Verjetnost napake 1. vrste: $P(B(10, 1/2) > C)$

Verjetnost napake 2. vrste: $P(B(10, 3/5) \le C)$

Vidimo, da ne moremo konstruirati verjetnost obeh napak.

SPLOŠNO O PREIZKUŠANJU DOMNEV

Privzemimo model $\mathcal{F}$ in naj bo $X = (X_1, \dots, X_n)$ slučajni vzorec z $X_i \overset{NEP}{\sim} F$ za neki $F \in \mathcal{F}$

STATISTIČNA DOMNEVA je izjava o PORAZDELITVI proučevanja vektorja X, ki je bodisi pravilna bodisi nepravilna, če je porazdelitev znana.

↳ ZGLEDI DOMNEV

- $E(X) \le 50$ — $\Theta = \mathbb{R} \times (0, \infty)$ $\{N(\mu, \sigma^2) \mid \mu \in \mathbb{R}, \sigma^2 \in (0,\infty)\}$, $H = (-\infty, 50] \times (0, \infty)$
- $P(X \le 7) = 1/2$ — $\{(\mu, \sigma^2) \mid P(N(\mu, \sigma^2) \le 7) = 1/2\}$
- $Var(X) > 42$ — $H = \mathbb{R} \times (42, \infty)$

↳ Posledično lahko domnevo enačimo s primerno podmnožico množice $\mathcal{F}$. Tipično jo označimo $\mathcal{H}$. (tiste porazdelitve, za katere je izjava pravilna)

↳ Če je model parametriziran $\mathcal{F} \leftrightarrow \Theta$, domnevo enačimo s primerno podmnožico $H \subset \Theta$

↳ Domnevi $\mathcal{H}$ je pridružena „alternativna domneva" $\mathcal{A}$, oziroma domnevi H je pridružena „alternativna domneva" $A \subset \Theta$. Tipično je $\mathcal{A} = \mathcal{H}^c$ oziroma $A = H^c$, včasih pa nas zanimajo $A \subsetneq H^c$

· V uvodnem zgledu $\mathcal{F} \Leftrightarrow \{1/2, 3/5\}$; $H = \{1/2\}$, $A = \{3/5\}$

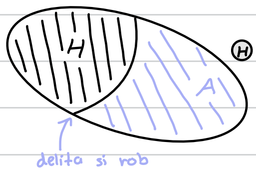
↳ **DEFINICIJA**

Preizkus domneve $\mathcal{H}$ je odločitveno pravilo, po katerem na podlagi realizacije slučajnega vzorca dane velikosti bodisi podpremo $\mathcal{H}$ bodisi podpremo $\mathcal{A}$.

↳ Kaj je lahko pri danem odločitvenem pravilu narobe?

odločitev \ dejansko stanje	drži $\mathcal{H}$	drži $\mathcal{A}$
podpremo $\mathcal{H}$ / ne zavrnemo $\mathcal{H}$	✓	NAPAKA 2. VRSTE
podpremo $\mathcal{A}$ / zavrnemo $\mathcal{H}$	NAPAKA 1. VRSTE	✓

↳ Želeli bi majhni verjetnosti napak. Izkaže se, da ne moremo kontrolirati verjetnosti napak obeh vrst. (Če je prostor parametrov Θ povezan in je H v njem lepa podmnožica, sta supremuma verjetnosti obeh vrst napak komplementarna: njuna vsota $=1$)



↳ V praksi se moramo za to odločiti (ali pa se za nas odloči „nadzorno telo") katera od napak je pomembnejša (po potrebi zamenjamo vlogi $\mathcal{H}$ in $\mathcal{A}$) in njeno verjetnost potem omejimo. Rečemo ji napaka 1. vrste (TYPE 1. ERROR)

Pravimo, da ima preizkus stopnje značilnosti α (stopnjo tveganja α), če je $\underbrace{\text{supremum verjetnosti napak 1. vrste}}_{= \text{velikost preizkusa}} \le \alpha$

(V praksi so tipične izbire za α: 0.05, 0.1, 0.01)

Preostali napaki pravimo napaka 2. vrste

↳ **KOMENTAR**

$P(\text{podprem } \mathcal{A} \mid \text{drži } \mathcal{H}) \le \alpha$. Če preizkus podpre $\mathcal{A}$ pravimo, da smo $\mathcal{A}$ „statistično dokazali"

$P(\text{podprem } \mathcal{H} \mid \text{drži } \mathcal{A})$ je lahko tudi $1-\alpha$

↳ Jezik: Napaka 1. vrste se zgodi, če zavrnemo $\mathcal{H}$, v resnici pa $\mathcal{H}$ drži

PREIZKUŠANJE STANDARDNIH DOMNEV V BERNOULLIJEVIH MODELIH

↳ To so domneve: • $p \le p_0$ $(H=(0,\infty], A=(p_0,1))$

• $p \ge p_0$ $(H=[p_0,1), A=(0,p_0))$

• $p=p_0$ $(H=\{p_0\}, A=(0,1)\setminus\{p_0\})$

① PREIZKUŠANJE H: $p \le p_0$ proti A: $p > p_0$

Preizkušamo na podlagi preizkusne statistike (za vzorec velikosti n)

$T = X_1 + \dots + X_n \sim B(n,p)$.

Edina smiselna oblika preizkusa je $\begin{cases} p \le p_0 \text{ zavrnemo, če je } T > C \\ p \le p_0 \text{ ne zavrnemo, če je } T \le C \end{cases}$ kjer moramo smiselno določiti C

• Verjetnost 1. napake

$$\boxed{P(T>C \mid p \le p_0) = P(B(n,p)>C \mid p \le p_0) = \sum_{i=C+1}^{n} \binom{n}{i} p^i (1-p)^{n-i}}$$

↳ To je funkcija definirana na $(0,p_0]$ in je naraščajoča in doseže maksimum v p_0.

Preizkus ima st. značilnosti $\alpha \iff P(B(n,p_0)>C) \le \alpha$

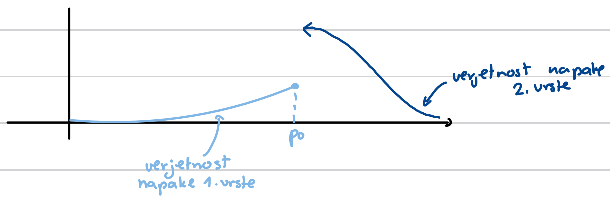
DN: dokaži, da je res naraščajoča

• Verjetnost napake 2. vrste

$$\boxed{P(T \le C \mid p > p_0) = \sum_{i=0}^{C} \binom{n}{i} p^i (1-p)^{n-i}}$$

funkcija, definirana na $(p_0,1)$. Je padajoča in njen supremum je dosežen, ko $p \downarrow p_0$.

Ta supremum znaša $P(B(n,p_0) \le C)$



• Vidimo $P(B(n,p_0)>C) + P(B(n,p_0) \le C) = 1$, torej sta supremuma komplementarna

• $\boxed{23.5}$ $P(B(n,p_0) > C+1) \le \underbrace{P(B(n,p_0)>C) \le \alpha}$

⇒ množica tistih C, ki zadoščajo neenakosti, je oblika $\{C_0, C_0+1, \dots\}$

PREIZKUŠANJE V NORMALNIH MODELIH

① PREIZKUŠANJE $\mu \le \mu_0$ PROTI $\mu > \mu_0$

Intuitivno je jasno, da bi $\mu \le \mu_0$ zavrnili, če $\bar{x} > C$.

a) Disperzija σ^2 je znana (konstantna)

ta je naraščajoča

padajoča

$$P(\bar{x} > C) = P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} > \frac{C-\mu}{\sigma/\sqrt{n}}\right) = 1 - \Phi\left(\frac{C-\mu}{\sigma/\sqrt{n}}\right)$$

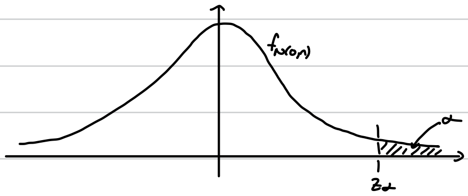
$\bar{x} \sim N(\mu, \sigma^2/n)$, $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

padajoča

Φ … k.p.f. od $N(0,1)$

To je naraščajoča funkcija μ. Verjetnost napake 1. vrste je zato maksimalna pri μ_0

Za C izberemo rešitev enačbe $1 - \Phi\left(\frac{C-\mu_0}{\sigma/\sqrt{n}}\right) = \alpha$. To pomeni $\frac{C-\mu_0}{\sigma/\sqrt{n}} = z_\alpha$ oziroma $C = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$



Zavrnitveni dogodek je $\{\bar{x} > C\} = \{\bar{x} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}\} = \left\{\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha\right\}$

Torej $\mu \le \mu_0$ (pri znani σ^2) zavrnemo $\iff \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha$

b) Pri neznani σ^2:

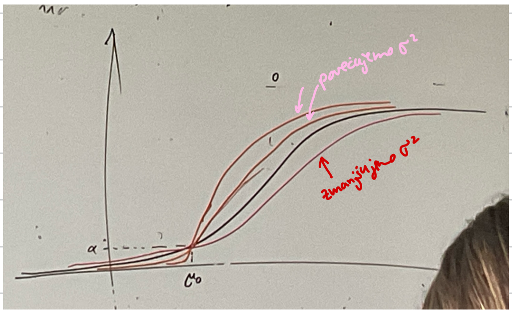
$$\frac{\bar{x}-\mu_0}{s/\sqrt{n}} > D$$

$\beta : \mathbb{R} \times (0,\infty) \to [0,1]$

$$\beta(\mu, \sigma^2) = P\left(\frac{\bar{x}-\mu_0}{s/\sqrt{n}} > D\right)$$

$\bar{x} \sim N(\mu, \sigma^2/n)$

$$\beta(\mu_0, \sigma^2) = P(t_{n-1} > D)$$



PROBLEM DVEH VZORCEV

① „PARNA PRIMERJAVA"

Meritvi X_{i1}, X_{i2} (dve časovni točki) za $1 \le i \le n$. Tipično domneve v zvezi z $\mu_1 - \mu_2$ prevedemo na domneve v zvezi s p.v. razlike $X_{i1} - X_{i2}$

Uporabimo preizkuse za pričakovano vrednost "ene" s.s.

② "NEPARNE PRIMERJAVE" oz. problem dveh neodvisnih vzorcev

X_{i1}, $1 \le i \le n_1$, NEP $N(\mu, \sigma_1^2)$

X_{i2}, $1 \le i \le n_2$, NEP $N(\mu, \sigma_2^2)$

Privzamemo, da sta vektorja $(X_{11}, \dots, X_{n_1 1})$ in $(X_{12}, \dots, X_{n_2 2})$ neodvisna

Pišimo $\overline{X_1} := \frac{1}{n_1} \sum_{i=1}^{n_1} X_{i1}$, $\overline{X_2} := \frac{1}{n_2} \sum_{i=1}^{n_2} X_{i2}$

$$Var(\overline{X_1} - \overline{X_2}) \underset{\uparrow \text{ neodvisnost}}{=} Var(\overline{X_1}) + Var(\overline{X_2}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$\frac{\overline{X_1} - \overline{X_2} - D_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$] na robu je to standardna normalna

a) Varianci σ_1^2 in σ_2^2 sta znani

Domnevo $\mu_1 - \mu_2 \le D_0$ zavrnemo (pri stopnji značilnosti α), če $\dfrac{\overline{X_1} - \overline{X_2} - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > z_\alpha$

$\mu_1 - \mu_2 \ge D_0$ —||— $< -z_\alpha$

$\mu_1 - \mu_2 = D_0$

nekaj manjka
↓
naslednja stran od koje

DVE NORMALNI POPULACIJI : DISPERZIJI POZNANI

Problem dveh vzorcev

① „PARNA PRIMERJAVA"

meritvi X_{i1}, X_{i2} (dve časovni točki) za $1 \leq i \leq n$. Tipično domneve v zvezi z $\mu_1 - \mu_2$ prevedemo na domneve v zvezi s p.v. razlike $X_{i1} - X_{i2}$.
Uporabimo preizkuse za pričakovano vrednost „ene" s.p.

② „NEPARNE PRIMERJAVE" oz. problem dveh neodvisnih vzorcev

X_{i1}, $1 \leq i \leq n_1$, NEP $N(\mu_1, \sigma_1^2)$

X_{i2}, $1 \leq i \leq n_2$, NEP $N(\mu_2, \sigma_2^2)$ Privzamemo, da sta vektorja $(X_{11}, \dots, X_{n_1 1})$ in $X(_{12}, \dots, X_{n_2 2})$ neodvisna.

Pišimo $\bar{X}_1 := \frac{1}{n_1} \sum_{i=1}^{n_1} X_{i1}$, $\bar{X}_2 := \frac{1}{n_2} \sum_{i=1}^{n_2} X_{i2}$

$$\mathrm{Var}(\bar{X}_1 - \bar{X}_2) = \mathrm{Var}(\bar{X}_1) + \mathrm{Var}(\bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$\mu_1 - \mu_2 \leq D_0 \to$ območje nad neko premico

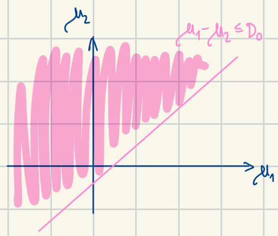
$\mu_1 - \mu_2 = D_0 \to$ rob območja je vzorec standardno normalen

(a) Varianci σ_1^2 in σ_2^2 sta znani

Domnevo • $\mu_1 - \mu_2 \leq D_0$ zavrnemo (pri stopnji značilnosti α), če $\dfrac{\bar{X}_1 - \bar{X}_2 - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > z_\alpha$

• $\mu_1 - \mu_2 \geq D_0$ zavrnemo (pri stopnji značilnosti α), če —″— $< -z_\alpha$.

• $\mu_1 - \mu_2 = D_0$ zavrnemo če $\left| \dfrac{\bar{X}_1 - \bar{X}_2 - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right| > z_{\alpha/2}$.



(b) Varianci neznani

(i) varianci neznani in RAZLIČNI ☹

(ii) varianci neznani in ENAKI (homoskedastični primer)

$$X_{i1} \overset{NEP}{\sim} N(\mu_1, \sigma^2), \quad X_{i2} \overset{NEP}{\sim} N(\mu_2, \sigma^2)$$

$$\frac{\bar{X}_1 - \bar{X}_2 - D_0}{\sqrt{\underbrace{\frac{1}{n_1+n_2-2}\left(\sum_{i=1}^{n_1}(X_{i1} - \bar{X}_1)^2 + \sum_{i=1}^{n_2}(X_{i2} - \bar{X}_2)^2\right)}_{\text{„}S^2\text{"}}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

zavračamo domneve, če $> t_{n_1+n_2-2\,;\,\alpha}$

$< t_{n_1+n_2-2\,;\,\alpha}$

$|\dots| > t_{n_1+n_2-2\,;\,\alpha/2}$

metoda za odločanje med dvema domnevama

PREIZKUŠANJE NA PODLAGI RAZMERJA VERJETIJ

- Razmerje verjetij (za dano domnevo H) je statistika, ki jo lahko v splošnem uporabimo kot preizkusno statistiko za preizkušanje H.
- Privzemimo parametričen model $\mathcal{F}$ s prostorom parametrom $\Theta \in \mathbb{R}^d$ in naj bo $H \subset \Theta$ domneva. Razmerje verjetij za H za vzorec velikosti n je funkcija: $\lambda : \mathbb{R}^n \to [0,1]$

$$\lambda(x_1,\dots,x_n) = \lambda(x) = \frac{\sup\{L(x,\vartheta) \mid \vartheta \in H\}}{\sup\{L(x,\vartheta) \mid \vartheta \in \Theta\}}$$

saj vzamemo ϑ iz podmnožice

kjer je $L(x;\vartheta) = f(x_1;\vartheta)\cdot\ldots\cdot f(x_n;\vartheta)$

op: [Definicijo je najlažje „prav" interpretirati, če primerjamo regularnostne privzetke za L iz razdelka o metodi največjega verjetja

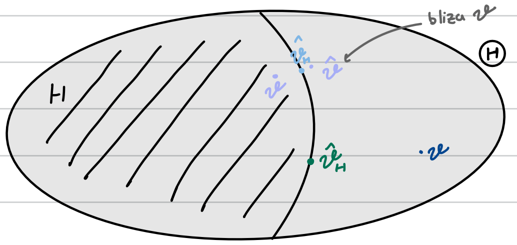
$\forall \vartheta : P(X=x) = 0$

se veča

KOMENTAR: Konkretno privzemimo, da je $S = \{x \in \mathbb{R}^n \mid L(x;\vartheta) > 0\}$ neodvisna od θ. Tedaj lahko razumemo, da je L definirana na S.

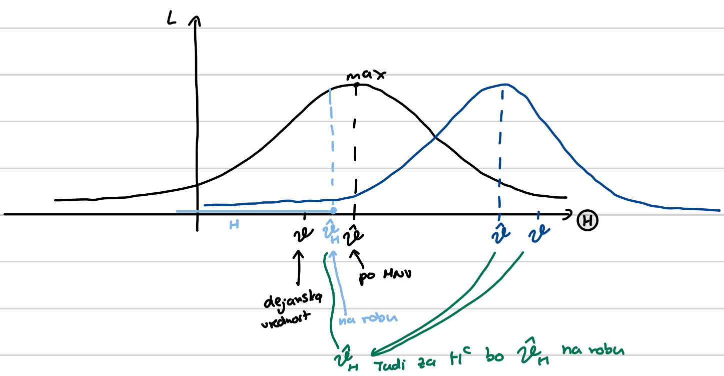
Seveda je $\lambda(x) = \dfrac{\sup\{L(x,\vartheta) \mid \vartheta \in H\}}{L(x,\hat{\vartheta}(x))}$ (vrednost v cenilki, ne cenilka), če obstaja CNV $\hat{\vartheta}$ za ϑ.

Pravimo, da je $\hat{\vartheta}_H(x)$ "restringirana" ocena po MNV za ϑ, če velja $\sup_{\vartheta \in H} L(x,\vartheta) = L(x,\hat{\vartheta}_H(x))$



★ Realizacija $(x_1,\dots,x_n)$ je prišla iz porazdelitve, ki jih ustreza $\vartheta \in H$
[$\hat{\vartheta}(x)$ je cenilka za ϑ]

$\lambda(x)$ bo "velik" "blizu" 1.



★★ - Realizacija je prišla iz porazdelitve, ki ji ustreza $\vartheta \in H^c$

Tu bo λ "blizu" 0

- IDEJA ZA PREIZKUS:

$$\begin{cases} H \text{ zavrnemo; če je } \lambda < D & \text{(izbrana konstanta)} \\ H \text{ ne zavrnemo; če je } \lambda \geq D \end{cases}$$

Konstanto D načeloma določimo z zahtevo z zahtevo velikosti: $\sup_{\vartheta \in H} P(\lambda(X_1,\dots,X_n) < D) \leq \alpha$

NEP

↳ ZGLED

① Preizkušanje $\mu = \mu_0$ proti $\mu \neq \mu_0$ v normalnem modelu z neznano disperzijo

$\Theta = \mathbb{R} \times (0,\infty) \to (\mu, \sigma^2)$, $H = \{\mu_0\} \times (0,\infty)$

$$L(x;\mu,\sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

že vemo: $\widehat{(\mu,\sigma^2)} = (\bar{x}, \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x}^2))$

Za "$\hat{\vartheta}_H$" iščemo maksimum funkcije $\sigma^2 \mapsto L(x;\mu_0,\sigma^2)$ ene spremenljivke.

Enako kot za funkcijo $\sigma^2 \mapsto L(x;\bar{x},\sigma^2)$ dobimo, da je maksimum dosežen pri $\hat{\sigma}_H^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu_0)^2$

op.: glej zgled pri MNV v normalnem modelu

Sledi $$\lambda(x) = \frac{L(x;\mu_0,\hat{\sigma}_H^2)}{L(x;\hat{\mu},\hat{\sigma}^2)} = \left(\frac{\hat{\sigma}_H^2}{\hat{\sigma}^2}\right)^{-\frac{n}{2}} = \left(\frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{-\frac{n}{2}} = \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2 + \mu(\bar{x}-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{-\frac{n}{2}} = \left(1 + \frac{n(\bar{x}-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{-\frac{n}{2}}$$

eksponentni del v $L(x;\mu,\sigma^2)$ se bo v celoti pokrajšal

ko smo $\sum_{i=1}^{n}(x_i-\mu)^2$ že gosdole minimizirali velja "pitagorov izrek"

Poglejmo kaj je zavrnitveni pogoj

$$\lambda(x) < D \iff 1 + \ldots > D' \iff \frac{\bar{x}-\mu_0}{\sum_{i=1}^{n}(x_i-\bar{x})^2} > D'' \iff \frac{(\bar{x}-\mu_0)^2}{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2} > D'''$$

$\lambda(x) = (1 + \text{bleh})^{-\frac{n}{2}}$

odštejemo 1

$$\iff \frac{(\bar{x}-\mu_0)^2}{(\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2)/n} > D'''' > 0$$ Torej lahko korenimo

konstante D bodo tudi $\in (0,1)$

$D' > 1$

D'' je pozitiven

t → t statistika

$$\iff \left| \frac{\bar{x}-\mu_0}{\underbrace{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2/n}}_{S/\sqrt{n}}} \right| > c$$

Ta preizkus je ekvivalenten „standardnemu" vzamemo $c = t_{n-1;\alpha/2}$

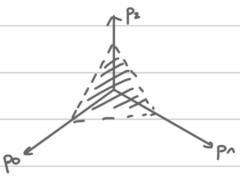
② DN: Prepričajte se, da podobno velja za $\mu \leq \mu_0$ proti $\mu > \mu_0$

③ Prilagoditveni preizkus čisto določene porazdelitve

Vzemimo diskretno porazdelitev na fiksnih vrednostih $\xi_0 < \xi_1 < \ldots < \xi_d$, torej $X_1, X_2, \ldots, X_n \overset{NEP}{\sim} \begin{pmatrix} \xi_0 & \xi_1 & \ldots & \xi_d \\ p_0 & p_1 & \ldots & p_d \end{pmatrix}$

Parametrični prostor je $\Delta^d = \{(p_0,p_1,\ldots,p_d) \mid p_j > 0, \sum_{j=0}^{d} p_j = 1\}$

primer $n=3$: $p_0 + p_1 + p_2 = 0$



Tu je dimenzija 2

To je t.i. d-razsežni simpleks

NPR: število pik pri slučajnem metu 6-strane kocke $\begin{pmatrix} 1 & 2 & \ldots & 6 \\ p_1 & p_2 & \ldots & p_6 \end{pmatrix}$

Tu je parametričen prostor 5-razsežen

Med drugim nas zanima domneva čisto konkretne porazdelitve: $(p_0,\ldots,p_d) = (\pi_0,\ldots,\pi_d)$ za neke fiksne verjetnosti $\pi_0,\ldots,\pi_d$

Domneva $H \subset \Theta$ je "enostavna": to je enoelementna množica

Funkcija verjetja: $L(x;p) = L(x_1,\ldots,x_n; p_0,\ldots,p_d) = \prod_{j=0}^{d} p_j^{T_j(x)}$, kjer je $T_j(x) = \sum_{i=1}^{n} \mathbb{1}_{\{\xi_j\}}(x_i)$ frekvenca pojavitve vrednosti ξ_j v vzorcu.

Seveda je $T_0(x) + \ldots + T_d(x) = n$

Gre za problem maksimizacije z vezjo $\boxed{\text{iščemo vezani ekstrem}}$

Tvorimo funkcijo $$G(p_0,\ldots,p_d,\lambda) = \sum_{j=0}^{d} T_j \ln p_j + \lambda\underbrace{\left(\sum_{j=1}^{d} p_j - 1\right)}_{\text{vez / en Lagrangev množitelj}}$$

Ekstremale? Sedaj lahko odvajamo po vseh p_j in λ

$$\frac{\partial G}{\partial p_j} = \frac{T_j}{p_j} + \lambda$$

$$\frac{\partial G}{\partial \lambda} = \sum_{j=0}^{d} p_j - 1$$

$$\forall j: T_j + \lambda p_j = 0 \quad , \quad \sum_{j=0}^{d} p_j = 1$$

$$\Big\{ \sum_{j=0}^{d}$$

$$n + \lambda = 0$$

$$\Longrightarrow \hat{p_j} = \frac{T_j}{n} \quad , \quad \hat{\lambda} = -n$$

↳ Ta Funkcija je Lock-konkavna → DN

Izkaže se, da gre za maksimum (DN : $ln(L) : \Theta \to \mathbb{R}$ je konkavna)

ker je konkavna, če je opt. v notranjosti ni problem, če pa je v robu se moramo še dodatno prepričati

Nastopanje ničelnih frekvenc

· $T_0 = 0 \Rightarrow L(x;p) = p_1^{T_1} p_2^{T_2} \dots p_d^{T_d}$ $\Longrightarrow$ očitno se lahko zožimo na $\hat{p_0} = 0$ (saj imamo samo naraščajoče funkcije)

Torej če imamo kakšno 0, se samo omejimo na manjši prostor in spet pridemo na "analogen" problem

Privzemimo: CNV je definirana na Δ^d ; zanjo velja $\hat{p_j} = \frac{T_j}{n}$

Sledi $\lambda(x) = \frac{\prod_{j=0}^{d} \pi_j^{T_j}}{\prod_{j=0}^{d} \left(\frac{T_j}{n}\right)^{T_j}} = \prod_{j=0}^{d} \left(\frac{n \pi_j}{T_j}\right)^{T_j}$

Preizkus: Domnevo $(p_0,\dots,p_d) = (\pi_0,\dots,\pi_d)$ zavrnemo, če $\lambda(x) < D$. konstanta je določena z zahtevo velikosti α:

$$P\Big(\prod_{j=0}^{d} \Big(\frac{n\pi_j}{\underbrace{T_j}_{\text{slučajni}}}\Big)^{T_j} < D\Big) \leq \alpha$$

če "velja" $(\pi_0,\dots,\pi_d)$

Tedaj ima vektor $(T_0, T_1, \dots, T_d)$ multinomsko porazdelitev

$[(T_0, T_1) \sim Bin]$

$$P((T_0,\dots,T_d) = (t_0,\dots,t_d)) = \underbrace{\frac{n!}{t_0!\, t_1! \dots t_d!}}_{\text{multinomski koeficient}} \pi_0^{t_0} \pi_1^{t_1} \dots \pi_d^{t_d}$$

Zaloga verjetnih vrednosti? $\{(t_0,\dots,t_d) \in \mathbb{Z}_+^{d+1} \mid t_0 + \dots + t_d = n\}$

$$\sum_{\substack{(t_0,\dots,t_d) \in \mathbb{Z}_+^{d+1} \\ t_0 + \dots + t_d = n}} \mathbb{1}\Big\{\prod_{j=0}^{d} \Big(\frac{n\pi_j}{t_j}\Big)^{t_j} < D\Big\} \cdot \frac{n!}{t_0! \dots t_d!} \pi_0^{t_0} \pi_1^{t_1} \dots \pi_d^{t_d}$$

Kako bi to sprogramirali: for $t_0 = 0$ to n

for $t_1 = 0$ to $n - t_0$

⋱

for $t_{d-1} = 0$ to $n - t_0 - \dots - t_{d-2}$

$$\lambda = \prod_{j=0}^{d} \left(\frac{n\hat{\pi}_j}{T_j}\right)^{T_j}$$

Preizkus: $\begin{cases} H \text{ zavrnemo; če } \lambda < D \\ H \text{ ne zavrnemo, če } \lambda \geq D \end{cases}$

Določitev konstante D z zahtevo, da naj bo velikost preizkusa $\leq D$, je v praksi lahko numerično zelo intenzivna ali celo neizvedljiva.

Glavna vrlina preizkušanja na podlagi neznanja vezij je asimptotično obnašanje statistike λ.

↳ **IZREK (Wilks)**

Naj bo Θ gladka mnogoterost razsežnosti d (npr. Θ je ⟂ v $\mathbb{R}^d$, Θ je praslika take gladke preslikave $V: \mathbb{R}^{d+n} \to \mathbb{R}^n$, katere Jacobijevka ima v vsaki točki v $V^{-1}(\sigma)$ poln rang

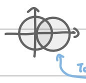
$v_1 = 0$
$\vdots$
$v_n = 0$

[Jacobijevka lin. funkcija je kar sama sebi]
Saj jo najbolje <u>gladko</u>? aproksimira

zgled: $f: \mathbb{R}^{d+1} \to \mathbb{R}$

$v_1(x_0, \ldots, x_d) = x_0^2 + \ldots + x_d^2 - 1$

ničla: $x_0^2 + \ldots + x_d^2 - 1 = 0$ → To je sfera



To je v prvih dveh razsežnostih

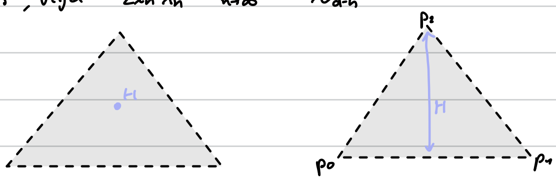
če dodamo še $v_2(x_0, \ldots, x_d) = (x_0 - 1)^2 + x_1^2 + \ldots + x_d^2 - 1$

↳ sfera premaknjena v prvi spremenljivki

in naj bo $H \subset \Theta$ gladka h-razsežna zaprta ???? podmnogoterost (brez robov) v Θ

Naj veljajo regularni privzetki na gostote $f(\_;\theta)$ iz razdelka o cenilki največjega verjetja

Če H drži, velja $-2\ln\lambda_n \xrightarrow[n\to\infty]{D} \chi^2_{d-h}$



[To je trorazsežni simpleks?]

$H: p_0 = p_1$
(znotraj $\Theta = \{(p_0, p_1, p_2) \mid p_j > 0, \sum p_j = 1\}$)

↳ Na podlagi Wilksnega izreka za "velike" vzorce PRIVZAMEMO $\underset{\text{strogo padajoča}}{-2\ln\lambda} \overset{\cdot}{\sim} \chi^2_{d-h}$, če H drži!

če predpostavke izreka držijo za velike vzorce preizkušamo takole: $\begin{cases} H \text{ zavrnemo; če } -2\ln\lambda > \chi^2_{d-h;\alpha} = F^{-1}_{\chi^2_{d-h}}(1-\alpha) \\ H \text{ ne zavrnemo; sicer} \end{cases}$

α percentil

napaka 1. reda

↳ KOMENTAR: Wilksov izrek zagotavlja $\lim_{n\to\infty} P(-2\ln\lambda_n > \chi^2_{d-h;\alpha}) = \alpha$, če H drži)

limitna velikost preizkusa je α

↳ Zapišimo $-2\ln\lambda = -2\sum_{j=0}^{d} T_j \ln\left(\frac{n\pi_j}{T_j}\right)$

$-2\ln\lambda = -2\sum_{j=0}^{d} T_j \ln\left(\underbrace{\frac{n\pi_j}{T_j} - 1}_{\xi} + 1\right)$

po Wilksonovem izreku velja $-2\ln\lambda \to \chi^2_{d-0}$ če imamo verjetnosti $(\pi_0, \pi_1, \dots, \pi_d)$, torej domneva drži

S pomočjo Taylorjeve formule $\ln(1+\xi) = \xi - \frac{\xi^2}{2} + o(\xi^2)$

$$-2\ln\lambda \approx -2\sum_{j=0}^{d} T_j \cdot \left(\underbrace{\left(\frac{n\pi_j}{T_j} - 1\right)} - \frac{1}{2}\left(\frac{n\pi_j}{T_j} - 1\right)^2\right)$$

$\sum_{j=0}^{d} n\pi_j - T_j = n\cdot 1 - n = 0$

$= \sum_{j=0}^{d} \frac{(T_j - n\pi_j)^2}{T_j}$: izkaže se (Pearson), da tudi $\sum_{j=0}^{d} \frac{(T_j - n\pi_j)^2}{T_j} \xrightarrow[n\to\infty]{D} \chi^2_{d-0}$, če "velja" $(\pi_0, \pi_1, \dots, \pi_d)$

Problem bi se lahko pojavil če bi bil $T_j = 0$: v teoriji za velika št. bo T_j konv proti ∞, kar ni nič

$\prod \underbrace{\left(\frac{n\pi_j}{T_j}\right)^{T_j}}$ smo dobili iz števca v imenovalcu pa je bilo

$\frac{\dots}{\prod_{j=0}^{d}\left(\frac{T_j}{n}\right)^{T_j}}$ To je bil rezultat maksimizacije $L = p_0^{T_0} \dots p_d^{T_d}$ def na odprtem simpleksu

če bi bil tu vsaj 1 eks. $= 0$, potem je to $= 1$, torej sploh ne nastopa v L, torej če bi imeli $T_j = 0$ ga moramo ignorirati ["izpade"]

$\prod \left(\frac{n\pi_j}{T_j}\right)^{T_j}$ To torej $T_j = 0$ ignoriramo

· v PRAKSI · če nimamo dovolj velikega vzorca bi lahko bil $T_j = 0$

Pearson je dokazal še da velja tudi

$$\sum_{j=0}^{d} \frac{(T_j - n\pi_j)^2}{n\pi_j} \xrightarrow[n\to\infty]{D} \chi^2_d$$

standardna pearsonova hi-kvadrat statistika

če je $(T_0, \dots, T_d)$ porazdeljen multinomsko s parametrom $(\pi_0, \pi_1, \dots, \pi_d)$