Datum

OVS AV 24. 2. 2009

- Martin Raič
  FMF, Jadranska 19,
  5. nadstropje, soba 5.12
- martin.raic@fmf.uni-lj.si

OSNOVE KOMBINATORIKE

Manca: 12 kap, 3 klobuke
- lahko se pokrije na 15 načinov (12+3)

Nace: 4 puloverji, 2 srajci, 3 hlače.
$4 \cdot 2 \cdot 3 = 24$
- lahko se obleče na 24 načinov (1 srajca, 1 pulover, 1 hlače)
- če lahko tudi brez puloverja: $5 \cdot 2 \cdot 3 = 30$ načinov (lahko tudi brez puloverja)

Olga: 4 vrste parketa
3 vrste nelesnih talnih oblog
5 garnitur pohištva
$(4+3) \cdot 5 = 35$ → dnevno sobo lahko opremi na 35 načinov

Peter: Lj - Polje - Zalog - Laze - Jernica - Kresnice - Litija
- 7 možnih štartov, 6 možnih ciljev
- možnih vozovnic = $7 \cdot 6 = 42$ (enosmernih)
- možnih različnih cen = $42/2 = 21$

Datum

Rezka: – 5! besed $= 1\cdot 2\cdot 3\cdot 4\cdot 5 = 120$

Tatjana:

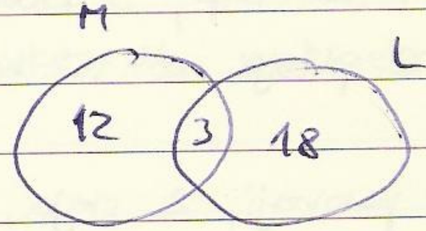
- $T_1 A_1 T_2 J A_2 N A_3$: $7! = 5040$
- Sovpade po $3!\cdot 2! = 12$ možnosti
- vseh permutacij $= \frac{7!}{3!\cdot 2!} = 420$ možnosti

Urban:

15 : modelarski krožek

21 : likovni krožek

3 : oba krožka



– koliko najmanj učencev? $15+21-3 = 33$ učencev

10 športnikov: (variacije – V^3_{10})

- na koliko načinov si lahko razdelijo medalje (zlato, srebro, bron), če je delitev mest izključena?

$\frac{10!}{7!} = 10\cdot 9\cdot 8 = 720$ možnosti

10 učencev, 3-članska delegacija:

- ni pomembno, kateri delegat je prvi, drugi, tretji. $\frac{10!}{7!\cdot 3!} = \frac{10\cdot 9\cdot 8}{1\cdot 2\cdot 3} = 120 = \binom{10}{3}$ → binomski simbol

$$\binom{m}{k} = \frac{m(m-1)(m-2)\cdot\ldots\cdot(m-k+1)}{k!} = \frac{m!}{k!\,(m-k)!} = \binom{m}{m-k}$$

* choose (m, k) * kalkulatorji: m $\boxed{nCr}$ k

→ C^k_m – kombinacije

ELEMENTARNA VERJETNOST

Osnova: $\boxed{P = \frac{\text{št. ugodnih izidov}}{\text{št. vseh izidov}}}$

Ta formula velja, če so vsi izidi enako verjetni.

Pošten kovanec: $P(grb) = \frac{1}{2} = 0{,}5$

Vržemo dva kovanca:

- kolikšna je verjetnost, da pade 1 cifra, 1 grb?

$P(\text{ena } C, \text{ en } G) = \frac{2}{4} = \frac{1}{2}$

→ CC, CG, GC, GG

→ ločimo objekte!

10 kovancev:

$P(1 \times G, 9 \times C) = \frac{10}{2^{10}} = 0{,}00977$

$$\left.\begin{array}{l} \to GCCCCCCCCC \\ CGCCCCCCCC \\ \vdots \\ CCCCCCCCCG \end{array}\right\} \text{10 ugodnih izidov}$$

Poštena kocka:

$P(6 \text{ pik}) = \frac{1}{6}$

$P(5 \text{ pik}) = \frac{1}{6}$

$P(\text{sodo št. pik}) = \frac{1}{2} = \frac{3}{6}$

moč množice

$P(\text{najmanj 3 pike}) = \frac{|\{3,4,5,6\}|}{6} = \frac{4}{6} = \underline{\underline{\frac{2}{3}}}$

Datum

2 kocki

a) $P(\text{enako mnogo pik}) = \frac{6}{6 \cdot 6} = \frac{1}{6}$

b) $P(\text{skupaj vsaj 10 pik}) = \frac{6}{36} = \frac{1}{6}$

46, 64, 56, 65, 55, 66

32 kart (4·8) → izvlečemo eno karto

$P(\text{izvlečena pik}) = \frac{1}{4}$

$P(\text{as}) = \frac{1}{8}$

$P(\text{pik ali as}) = \frac{11}{32}$

4 asi

8 pikov

− pikov as

	PIK	KRIŽ	SRCE	KARO
7	*			
8	*			
9	*			
10	*			
J	*			
Q	*			
K	*			
A	*	*	*	*

štejemo 2x, odštejemo

32 kart, izlečemo jih 5, kart ne vračamo
P(vsaj en as) = ?
bolje računati ~~vrednost~~ verjetnost nasprotnega dogodka - kjer ni nobenega asa → P(med njimi ni asa) = ?

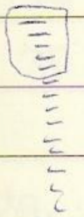
$$\underbrace{AAAA}_{4}\ \underbrace{-\ -\ -\ -\ -\ -\ -\ -}_{28}$$

$$P(\text{med njimi ni asa}) = \frac{28\cdot27\cdot26\cdot25\cdot24}{32\cdot31\cdot30\cdot29\cdot\cancel{28}} \doteq 0{,}488$$

$$P(\text{med izvlečenimi vsaj 1 as}) \doteq 1-0{,}488 = \underline{0{,}512}$$

VEČ NAČINOV:

1) $P(\text{ni asa}) = \dfrac{\binom{28}{5}}{\binom{32}{5}}$

2) 

$$\cancel{\binom{32}{4}} \qquad \frac{\binom{27}{4}}{\binom{32}{4}} = \frac{27\cdot26\cdot25\cdot24}{32\cdot31\cdot30\cdot29}$$

Loto: 39 številk

Na kombinacijskem listku ~~pr~~ prekrižamo 7 številk. Izžreba se 7 navadnih številk in še ena dodatna.

Dobitki: sedmica
šestica
petica
štirica
(zadevajo le navadne izžrebane)

6+1: med prekrižanimi je 6 navadnih izžrebanih in tudi dodatna

Datum

$P(\text{sedmica}) = \cancel{\ldots} \frac{1}{\binom{39}{7}} = \frac{7!\,32!}{39!} = \cancel{\ldots}$

→ loterija: na koliko načinov lahko igralec prekriža 7 številk?

→ $\frac{1}{15\,380\,937} = 6{,}5 \cdot 10^{-8}$

$P(\text{štirica}) = \frac{\binom{7}{4}\cdot\binom{32}{3}}{\binom{39}{7}}$

→ Pogled loterije: 7 izžrebanih, neizžrebane (32)

$\underbrace{\checkmark\checkmark\checkmark\checkmark\checkmark\checkmark\checkmark}\ \underbrace{\text{- - - - - - -}}$

→ Pogled igralca: $\underbrace{\times\times\times\times\times\times\times}_{\substack{\text{prekrižane}\\(7)}}\ \underbrace{\text{- - - - - - - - - - -}}_{\substack{\text{neprekrižane}\\(32)}}$

$P(6+1) = \frac{7}{\binom{39}{7}} = 7 \cdot P(\text{sedmice})$

→ Pogled loterije:

$\underbrace{\checkmark\checkmark\checkmark\checkmark\checkmark\checkmark\checkmark}_{\text{izžrebane (7)}}\ \underset{\substack{\uparrow\\ \text{dodatne}\\(1)}}{D}\ \underbrace{\text{- - - - - - - - - - - -}}_{\substack{\text{neizžrebane}\\(31)}}$

Datum

Sistem najmanj koliko števil bi morali vplačati, če naj bo $P(\text{sedmice}) \geq \frac{1}{2}$?

Pogled igralca:

$\underbrace{x x x \dots x}_{k \text{ prekrižanih}} \underbrace{\_\ \_\ \_\ \_\ \_\ \_\ \_}_{39-k \text{ neprekrižanih}}$

$$P(\text{sedmice}) = \frac{\binom{k}{7}}{\binom{39}{7}}$$

k	$P(\text{sedmice})$
39	1
38	0·821
37	0·669
36	0·543
35	0·437

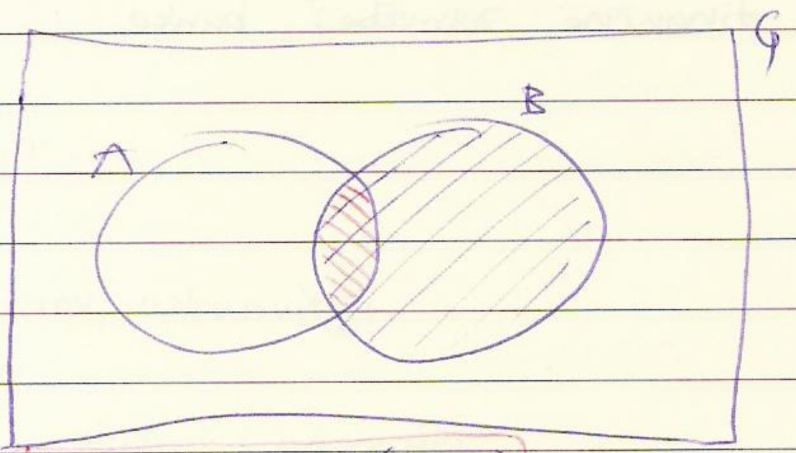
Odg.: sistem najmanj 36-ih števil bi morali vplačati.

$\boxed{\text{OVS AV } 3.3.2009}$

POGOJNA VERJETNOST

• $P(A|B)$

↑ vemo, da se je B zgodil



$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

$$P(A \cap B) = P(A) \cdot P(B|A)$$

• Če so vsi izidi enako verjetni, je: $P(A) = \frac{|A|}{|G|}$

$$\boxed{P(A|B) = \frac{|A \cap B|}{|B|}}$$

10-krat vržemo kovanec, vsi izidi so enako verjetni.

$A = \{$v vseh metih pade C$\}$

$B = \{9 \times C, 1 \times g\}$

$C = \{$v prvih 9 metih pade c$\}$

$P(A), P(B), P(A|C), P(B|C) = ?$

$P(A) = \frac{1}{2^{10}} = \frac{1}{1024}$

$P(B) = \frac{10}{2^{10}} = \frac{10}{1024}$

$P(A|C) = \frac{|A \cap C|}{|C|} = \frac{1}{2}$

$P(B|C) = \frac{1}{2}$

cccccccccc

cccccccccg

Datum

V posodi je 5 rdečih in 2 zeleni kroglici. Iz posode na slepo in brez vračanja vlečemo kroglice.

$A = \{\text{prva izvlečena rdeča}\}$

$P(A) = \frac{5}{7}$

$B = \{\text{prvi dve izvlečeni kroglici rdeči}\}$

$P(B) = \frac{5}{7} \cdot \frac{4}{6}$

$\boxed{\begin{matrix} 5R \\ 2Z \end{matrix}}$ $\boxed{\begin{matrix} 4R \\ 2Z \end{matrix}}$

$C = \{\text{prva izvlečena rdeča, druga zelena}\}$

$P(C) = \frac{5}{7} \cdot \frac{2}{6} = \frac{10}{42} = \frac{5}{21}$

enaki verjetnosti

$D = \{\text{prva zelena, druga rdeča}\}$

$P(D) = \frac{2}{7} \cdot \frac{5}{6} = \frac{10}{42} = \frac{5}{21}$

$E = \{\text{prve tri izvlečene rdeče}\}$

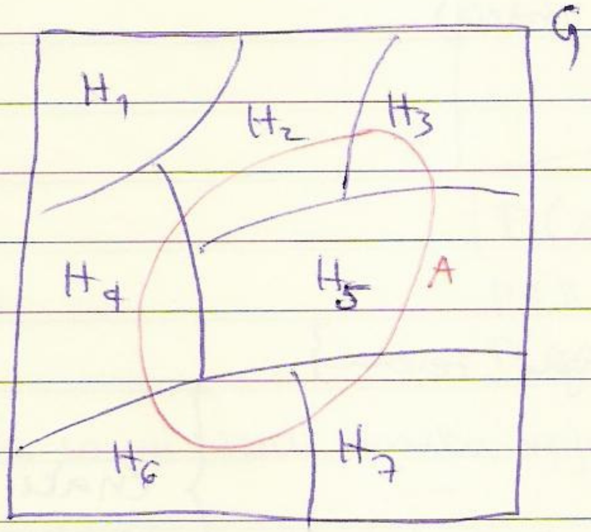
$P(E) = \frac{5}{7} \cdot \frac{4}{6} \cdot \frac{3}{5}$

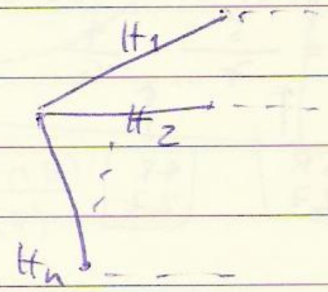
$F = \{\text{prve tri izvlečene zelene}\}$

$P(F) = \frac{2}{7} \cdot \frac{1}{6} \cdot \frac{0}{5} = 0$; F je nemogoč dogodek

IZREK O POPOLNI VERJETNOSTI:

- Dogodki $H_1, H_2, \ldots, H_n$ tvorijo popoln sistem dogodkov, če so ~~dogodki~~ paroma nezdružljivi ($H_i \cap H_j = \emptyset$ za $i \neq j$), njihova unija pa je gotovi dogodek G. Gotovo se zgodi natanko eden izmed teh dogodkov





Če je $H_1, H_2, \ldots, H_n$ popoln sistem dogodkov $\Rightarrow$

$$\Rightarrow P(A) = P(H_1)\cdot P(A/H_1) + P(H_2)\cdot P(A|H_2) + \ldots$$
$$\ldots + P(H_n)\cdot P(A/H_n).$$

Enaka naloga kot prej; 5 R, 2 Z

$R_n = \{$ n-ta izvlečena rdeča $\}$
$Z_n = \{$ n-ta izvlečena zelena $\}$

$\cancel{Z}\, P(Z_1)\cdot P(R_2|Z_1) =$

$$P(R_2) = P(R_1)\cdot P(R_2|R_1) \cancel{+ P(R_2)\cdot P(R_2|R_2)}$$

↳ odvisno od tega, ~~katerega~~ kakšne je bila prva

$$= \frac{5}{7}\cdot\frac{4}{6} + \frac{2}{7}\cdot\frac{5}{6} = \frac{20}{42} + \frac{10}{42} = \frac{30}{42} = \frac{15}{21} = \frac{5}{7} = \cancel{P(R_1)}$$

$P(R_1)$, glej glej ! :)

• „pogojna verjetnost zgodovine"

$$P(R_1 | R_2) = \frac{P(R_1 \cap R_2)}{P(R_2)} = \frac{P(R_1) \cdot P(R_2 | R_1)}{P(R_2)} = \frac{\frac{5}{7} \cdot \frac{4}{6}}{\frac{5}{7}} = \underline{\underline{\frac{2}{3}}}$$

Opazili smo:

$\rightarrow$ $P(R_1) = P(R_2)$

$P(R_1 \cap Z_2) = P(Z_1 \cap R_2)$

$P(R_1 | R_2) = P(R_2 | R_1)$

Zgodi naključje? Nikakor ne!

$G :=$ vse permutacije kroglic R1 R2 R3 R4 R5 Z6 Z7

Nekaj možnih izidov:

R2	Z6	R1	R3	Z7	R5	R4
R1	Z7	R3	R2	Z6	R4	R5
Z6	R2	Z7	R5	R1	R4	R3

Vsi izidi ($7! = 5040$) so enako verjetni.
Če zamenjamo prvo in drugo pozicijo, še vedno dobimo vse permutacije.

Zanima nas pogojna verjetnost, da je bila prva izvlečena rdeča, če vemo, da sta bili prvi izvlečeni rdeči.

$P(R_1 | R_1 \cap R_2) = 1$

$$P(R_1 | \overbrace{R_1 \cup R_2}^{\text{vsaj ena rdeča}}) = \frac{P(R_1 \cap (R_1 \cup R_2))}{P(R_1 \cup R_2)} = \frac{P(R_1)}{1 - P(Z_1 \cap Z_2)} = \frac{\frac{5}{7}}{1 - \frac{2}{7} \cdot \frac{1}{6}} =$$

$$= \frac{\frac{5}{7}}{\frac{20}{21}} = \frac{5 \cdot 21 \cdot 3}{7 \cdot 20 \cdot 4} = \frac{3}{4}$$

nasprotna dogodka

Datum

$$P\Big(R_1 \,\Big|\, \overbrace{(R_1 \cap Z_2) \cup (R_2 \cap Z_1)}^{\text{natanko ena rdeča}}\Big) = \frac{P(R_1 \cap Z_2)}{P(R_1 \cap Z_2) + P(Z_1 \cap R_2)} =$$

→ dogodka sta nezdružljiva → verjetnost unije lahko zapišemo kot vsoto verjetnosti

$= \underline{\underline{\frac{1}{2}}}$

ti dve verjetnosti sta enaki

Kup 32-ih kart: pik, križ, srce, karo
Vrednosti: AS, K, D, B, 10, 9, 8, 7
Izvlečemo 2 karti
A = {prva karta as}
$P(A) = \frac{4}{32} = \frac{1}{8}$

$P(\text{prva karta večja od druge}) = \frac{\text{št. ugodnih}}{32 \cdot 31} = \ldots$

- uporabimo simetrijo: zamenjamo 1. in 2.

$= P(\text{druga karta večja od prve}) =$
$= \frac{1}{2} \cdot P(\text{karti imata različno vrednost}) =$
$= \frac{1}{2}\left(1 - P(\text{karti imata enako vrednost})\right) =$
$= \frac{1}{2}\left(1 - \frac{3}{31}\right) = \frac{14}{31}$

↳ Izrek o popolni verjetnosti

Pogojna verjetnost
$P(\text{prva karta večja od druge} \mid \text{prva as}) = 1 - \frac{3}{31} =$

$= \underline{\underline{\frac{28}{31}}}$

Datum

P(prva karta as | prva karta večja od druge) =

$$= \frac{P(\text{prva as} \cap \text{prva večja od druge})}{P(\text{prva večja od druge})} = \frac{\frac{1}{8} \cdot \frac{28}{31}}{\frac{14}{31}} =$$

← že izračunali

$$= \frac{28 \cdot 31 \cdot 4}{8 \cdot 31 \cdot 14 \cdot 2} = \frac{2}{8} = \frac{1}{4}$$

Nezdružljivost:

A, B sta nezdružljiva, če je $A \cap B = \emptyset$.

A, B nezdružljiva $\Rightarrow P(A \cup B) = P(A) + P(B)$.

Neodvisnost:

verjetnost preseka   verjetnost produkta

A, B neodvisna, če je $P(A \cap B) = P(A) \cdot P(B)$.

$P(A | B) = P(A)$

$P(B | A) = P(B)$

- Dogodek, ki ima vrednost 0 ali 1, je neodvisen od kateregakoli dogodka, tudi od samega sebe.

$A_1, A_2, A_3, \ldots, A_n$ so neodvisni, če za poljubne različne indekse $i_1, i_2, \ldots, i_k$ velja

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_k})$$

2 kovanca, vsi izidi enako verjetni

$A = \{$prvič cifra$\}$

$B = \{$drugič cifra$\}$

$C = \{$obakrat cifra$\}$

$D = \{$obakrat različno$\}$

- C in D sta edina nezdružljiva dogodka.
- Neodvisna: A, B
  - $\rightarrow A \cap B = \{cc\}$
  - $P(A \cap B) = \frac{1}{4}$
  - $P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
- ~~Ne~~Odvisna: A, C
  - $\rightarrow A \cap C = \{cc\}$
  - $P(A) = \frac{1}{2}, \; P(C) = \frac{1}{4}$
  - Sta odvisne.
  - $C \subseteq A$.
- Neodvisna: A, D
  - $\rightarrow A \cap D = \{cg\}$
  - $P(A) = \frac{1}{2}, \; P(D) = \frac{1}{2}$
  - A in D sta neodvisna.

D.N. B in D ?

A, B in D ?

Datum

OVS AV 10.3.2009

A in B sta neodvisna, če je: $P(A \cap B) = P(A) \cdot P(B)$

A, B in C so neodvisni, če je: $P(A \cap B) = P(A) \cdot P(B)$, $P(A \cap C) = P(A) \cdot P(C)$, $P(B \cap C) = P(B) \cdot P(C)$, $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

PRIMER: dva kovanca, vsi izidi enako verjetni

$A = \{$prvič cifra$\}$, $B = \{$drugič cifra$\}$, $D = \{$obakrat različno$\}$

$G = \{\overset{1/4}{cc}, \overset{1/4}{cg}, \overset{1/4}{gc}, \overset{1/4}{gg}\}$ $\quad A \cap B = \{cc\}$

$A = \{cc, cg\}$ $\quad A \cap D = \{cg\}$

$B = \{cc, gc\}$ $\quad B \cap D = \{gc\}$

$D = \{cg, gc\}$ $\quad A \cap B \cap D = \emptyset = N$

$P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$, $P(D) = \frac{1}{2}$, $P(A \cap B) = \frac{1}{4}$, $P(A \cap D) = \frac{1}{4}$, $P(B \cap D) = \frac{1}{4}$,

$P(A \cap B \cap D) = 0$

- A in B sta neodvisna
- B in D sta neodvisne
- A in D sta neodvisne
- A, B in D niso odvisni $\left(\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \neq 0\right)$

→ A, B in D so paroma neodvisni, kljub temu pa so odvisni, ker je $P(A \cap B \cap D) \neq P(A) \cdot P(B) \cdot P(D)$

NEODVISNOST IZVEDENIH DOGODKOV:

$A_1, A_2, \dots, A_m, B_1, B_2, \dots, B_n$ neodvisni

A dogodek, izveden iz $A_1, A_2, \dots, A_m$

B dogodek, izveden iz $B_1, B_2, \dots, B_n$

Tedaj sta A in B tudi neodvisna.

vinogradniki, šmarnica, cviček

P(Janez ponudi šmarnico) = 60%

P(Lojz ponudi šmarnico) = 30%

P(boli glava, ne da bi pil šmarnico) = 40%

P(kozarec šmarnice) = 80%

P(po 2h kozarcih šmarnice) = 100%

P(Miho boli glava)?

* XOR $\Leftrightarrow$ simetrična razlika $\triangle$

$J = \{$Janez dal šmarnico$\}$, $L = \{$Lojz dal šmarnico$\}$,

$B = \{$Miho boli glava$\}$

$P(J) = 0'6$

$P(L) = 0'3$

$P(B \mid \bar{J} \cap \bar{L}) = 0'4$

$P(B \mid (J \setminus L) \cup (L \setminus J)) = 0'8$

$P(B \mid (J \cap \bar{L}) \cup (\bar{J} \cap L)) = 0'8$

$P(B \mid \cancel{\bar{J}}^{\,J} \cap L) = 1$

$P(B)$?

Hipoteze:

$\bar{J} \cap \bar{L}$ → 0 kozarcev šmarnice

$(J \cap \bar{L}) \cup (J \cap \bar{L})$ → 1 kozarec šmarnice

$J \cap L$ → 2 kozarca šmarnice.

popoln sistem dogodkov

$P(\bar{J} \cap \bar{L}) \overset{\text{neodvisnost}}{=} P(\bar{J}) \cdot P(\bar{L}) = 0'4 \cdot 0'7 = 0'28$

$P((J \cap \bar{L}) \cup (\bar{J} \cap L)) \overset{\text{nezdružljivost}}{=} P(J \cap \bar{L}) + P(\bar{J} \cap L) \overset{\text{neodvisnost}}{=} 0'6 \cdot 0'7 + 0'4 \cdot 0'3 = 0'54$

$P(J \cap L) = P(J) \cdot P(L) = 0'6 \cdot 0'3 = 0'18$

→ če sta dogodka neodvisna, je tudi komplement neodvisen

• po izreku popolne verjetnosti:

$P(B) = 0'28 \cdot 0'4 + 0'54 \cdot 0'8 + 0'18 \cdot 1 = \underline{0'724}$

BAYESOVA FORMULA:

$H_1, H_2, \dots, H_n$ popoln sistem dogodkov

$$P(H_i \mid A) = \frac{P(H_i) \cdot P(A \mid H_i)}{P(H_1) P(A \mid H_1) + \dots + P(H_n) \cdot P(A \mid H_n)} = \boxed{\frac{P(H_i \cap A)}{P(A)}}$$

→ vzroki za katastrofe

$P(\text{Francka}) = 40\%$

$P(\text{Micka}) = 60\%$

$F(\text{nagnita}) = 10\%$ $\quad M(\text{nagnita}) = 20\%$

privzamemo, da branjevki izbirata solato na slepo

P(mož okregan po pravici) = ?

$F = \{\text{kupil pri Francki}\}$ $\qquad P(F) = 0'4$

$M = \{\text{kupil pri Micki}\}$ $\qquad P(M) = 0'6$

$S = \{\text{solata nagnita}\}$ $\qquad P(S|F) = 0'1 \quad P(S|M) = 0'2$

$P(M|S) = ?$

hipotezi: F, M ; dogodek: S

$$P(M|S) = \frac{P(M)\cdot P(S|M)}{P(M)\cdot P(S|M) + P(F)\cdot P(S|F)} = \frac{0'6\cdot 0'2}{0'6\cdot 0'2 + 0'4\cdot 0'1} = \frac{0'12}{0'16} =$$
$$= \frac{3}{4} = \underline{\underline{75\%}}$$

vinogradnika

Janez, Lojz, šmarnica

$P(J) = 0'6$

$P(L) = 0'3$

$P(B|\bar{J}\cap\bar{L}) = 0'4$

- verjetnost, da Miho boli glava po kozarcu šmarnice, ~~ne~~, ~~glede na to, po čigavem~~ je 0'8

$P(B|(J\cap\bar{L})\cup(\bar{J}\cap L)) = 0'8$

- verjetnost, da Miha boli glava po kozarcu šmarnice, ne glede na to, čigave, je 0'8 → privzame, da sta šmarnici enakovredni

$P(B|J\cap\bar{L}) = P(B|\bar{J}\cap L) = 0'8$ *

⇒ pogojna verjetnost, da mu je Janez dal šmarnico?

$P(J|B) = ?$

hipoteze: $\underbrace{\bar{J}\cap\bar{L},\ \bar{J}\cap L,\ J\cap\bar{L},\ J\cap L}_{\text{popoln sistem dogodkov, nezdružljivi}}$; dogodek: B

$P(J|B) = P(J\cap\bar{L}|B) + P(J\cap L|B)$ *

$$P(J\cap\bar{L}|B) = \frac{P(J\cap\bar{L})\cdot P(B|J\cap\bar{L})}{P(\bar{J}\cap\bar{L})\cdot P(B|\bar{J}\cap\bar{L}) + P(\bar{J}\cap L)\cdot P(B|\bar{J}\cap L) + P(J\cap\bar{L})\cdot P(B|J\cap\bar{L}) + P(J\cap L)\cdot P(B|J\cap L)} =$$

→ imenovalec = P(B) $\qquad = \frac{0'6\cdot 0'7\cdot 0'8}{0'724}$ * $= 0'336/0'724$

$P(J \cap L \mid B) = \frac{P(J \cap L) \cdot P(B \mid J \cap L)}{P(B)} = \frac{0{,}6 \cdot 0{,}3 \cdot 1}{0{,}724} = \cancel{0{,}713}$

$P(J \mid B) = P(J \cap \bar{L} \mid B) + P(J \cap L \mid B) = \underline{0{,}713}$

Janez, Francej, Tone streljajo zajce

$P(J) = 0{,}1$
$P(F) = 0{,}2$
$P(T) = 0{,}3$
} neodvisni dogodki J, F, T

$Z = \{\text{zajec je zadet}\}$

$P(J \mid Z) = ?$

→ po definiciji pogojne verjetnosti:

$P(J \mid Z) = \frac{P(J \cap Z)}{P(Z)} =$

• če bi bili J, F, T, nezdružljivi, bi lahko šteli, vendar so neodvisni!

$Z = J \cup F \cup T$

$P(Z) = 1 - P(\bar{J} \cap \bar{F} \cap \bar{T})$

$P(Z) = 1 - 0{,}9 \cdot 0{,}8 \cdot 0{,}7 = 0{,}496$

$P(J \cap Z) = P(J) = 0{,}1$

$P(J \mid Z) = \frac{0{,}1}{0{,}496} = \underline{0{,}202}$

• ko pridejo do zajca, se izkaže, da ga je zadel natanko eden. Kolikšna je verjetnost da je bil to Janez?

$Z_1 = \{\text{zadel ga je natanko eden}\}$

$P(J \mid Z_1) = \frac{P(J \cap Z_1)}{P(Z_1)} = \frac{P(J \cap \bar{F} \cap \bar{T})}{\underbrace{P(J \cap \bar{F} \cap \bar{T}) + P(\bar{J} \cap F \cap \bar{T}) + P(\bar{J} \cap \bar{F} \cap T)}_{\text{dogodki so nezdružljivi}}} =$

$= \frac{0{,}1 \cdot 0{,}8 \cdot 0{,}7}{0{,}1 \cdot 0{,}8 \cdot 0{,}7 + 0{,}9 \cdot 0{,}2 \cdot 0{,}7 + 0{,}9 \cdot 0{,}8 \cdot 0{,}3} =$

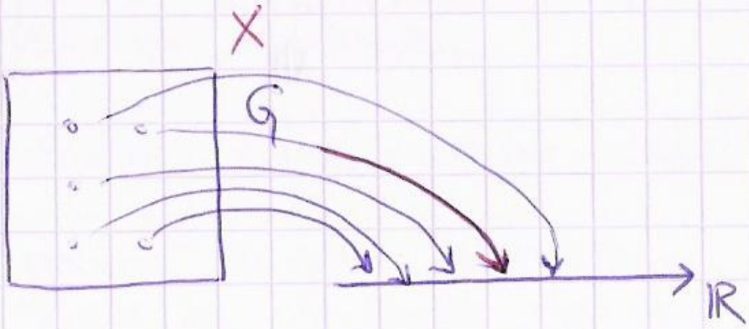
$= \underline{0{,}141}$

OVS AV 17. 3. 2009

1. kolokvij:

TOR 14.4. ob 19. uri v PR15

Slučajne spremenljivke



$$X \sim \begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ p_1 & p_2 & p_3 & \dots & p_n \end{pmatrix}$$

$P(X=a_1) = p_1$ $\quad p_1 + p_2 + \dots + p_n = 1$

$P(X=a_2) = p_2$

$P(X=a_3) = p_3$

$\vdots$

Vržemo dva neodvisna kovanca.

a) 1€ 2€

b) 2x 2€

$G = \{gg, gc, cg, cc\}$

ω	gg	gc	cg	cc
$P(\omega)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
(1,2) $S(\omega)$ primer a)	0	2	1	3
(2,2) $S(\omega)$ primer b)	0	2	2	4

Porazdelitev slučajne spremenljivke:

a) $S \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$

b) $S \sim \begin{pmatrix} 0 & 2 & 4 \\ \frac{1}{4} & \frac{2}{4} & \frac{1}{4} \end{pmatrix}$

$\frac{2}{4} = \frac{1}{2}$

5x kocka

S ... št. šestic, ki padejo.

Porazdelitev slučajne spremenljivke:

$S \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ \left(\frac{5}{6}\right)^5 & & & & & \end{pmatrix}$

→ 5 metov

$P(S=1) = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^4 \cdot 5$ → na koliko načinov se to lahko zgodi

$P(6 \text{ pade prvič in samo prvič}) = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^4$

↑ za 1. ↑ za 2., 3., 4., 5.

→ verjetnost, da drugič ne pade 6-ica = 5/6

→ verjetnost, da prvič pade 6, drugič ne = $\frac{1}{6} \cdot \frac{5}{6}$ zaradi neodvisnosti

$$P(S=2) = \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3 \cdot \binom{5}{2}$$

$$\binom{5}{2} = \frac{5 \cdot 4}{1 \cdot 2} = 10$$

$$P(S=3) = \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^2 \cdot \binom{5}{3}$$

$$\binom{5}{3} = \frac{5 \cdot 4 \cdot 3}{1 \cdot 2 \cdot 3} = 10$$

$\binom{5}{3} = \binom{5}{2}$, ker lahko izberemo mete, ko šestica pade, ali pa, ko šestica ne pade

$$P(S=k) = \binom{5}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{5-k},$$

S je porazdeljena binomsko;

$S \sim B(5, \frac{1}{6})$

BINOMSKA PORAZDELITEV:

$S \sim B(n, p)$;

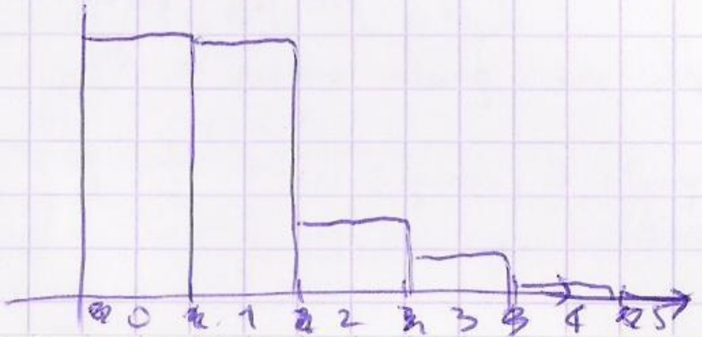
$$P(S=k) = \binom{n}{k} p^k (1-p)^{n-k};$$

$$k = 0, 1, 2, \ldots, n$$

Naključno n neodvisnih poskusov, vsak uspe z verjetnostjo p. S = število uspešnih

$$S \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 0.402 & 0.402 & 0.161 & 0.0322 & 0.00322 & 0.000129 \end{pmatrix}$$

HISTOGRAM:



- Verjetnost, da je S sodo!

$$P(S=\text{sodo}) = P(S=0) + P(S=2) + P(S=4) \doteq$$

$$\doteq 0.402 + 0.161 + 0.003 = \underline{0.566}$$

- Verjetnost, da je $S \geq 3$!

$$P(S \geq 3) = P(S=3) + P(S=4) + P(S=5) \doteq$$

$$\doteq 0.0322 + 0.0032 + 0.0001 = \underline{0.0355}$$

- Pogojna verjetnost, da je S sodo, če vemo, da je S večje ali enako 3.

$$P(S=\text{sodo} \mid S \geq 3) = \frac{P(S=4)}{P(S \geq 3)} \doteq \frac{0.00322}{0.0355} =$$

$$= \underline{0.0906}$$

Pošten kovanec mečemo, dokler ne pade cifra, ampak največ 4x.

M... število metov → slučajna spremenljivka

Napiši porazdelitev slučajne spremenljivke.

$$M \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \end{pmatrix}$$

c   gc   ggc $\frac{2}{16}$ → ggc ali gggg

Matematično upanje (pričakovana vrednost)

- $E(X) = \sum_{x} x \cdot P(X=x)$

$$X \sim \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} \Rightarrow$$

$$\Rightarrow E(X) = a_1 \cdot p_1 + a_2 \cdot p_2 + \cdots + a_n \cdot p_n$$

- $E[f(X)] = \sum_{x} f(x) \cdot P(X=x)$

$$= f(a_1) \cdot p_1 + f(a_2) \cdot p_2 + \cdots + f(a_n) \cdot p_n$$

Disperzija (varianca)

$$D(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2$$

$var(X)$

- disperzija je ~~vedno~~ vedno pozitivna ali 0, ko je X skoraj konstanta - trivialna vrednost

Standardni odklon (deviacija):

$$\sigma(X) = \sqrt{D(X)}$$

$$X \sim \begin{pmatrix} -1 & 0 & 1 & 4 \\ 0.1 & ? & 0.1 & 0.5 \end{pmatrix}$$

- Izračunaj verjetnost, da je $X=0$; matematično upanje $E(X)$; disperzijo $D(X)$; standardni odklon $\sigma(X)$.

$$X \sim \begin{pmatrix} -1 & 0 & 1 & 4 \\ 0.1 & 0.3 & 0.1 & 0.5 \end{pmatrix} \rightarrow \text{vsota} = 1$$

$P(X=0) = \underline{0.3}$

$E(X) = -0.1 + 0 + 0.1 + 2 = \underline{2}$

$D(X) = (-1-2)^2 \cdot 0.1 + (0-2)^2 \cdot 0.3 + (1-2)^2 \cdot 0.1 +$
$(4-2)^2 \cdot 0.5 =$
$= 0.9 + 1.2 + 0.1 + 2 = \underline{4.2}$

drugi način:

$E(X^2) = (-1)^2 \cdot 0.1 + 0^2 \cdot 0.3 + 1^2 \cdot 0.1 + 4^2 \cdot 0.5 =$
$= 0.1 + 0.1 + 8 = 8.2$

$D(X) = 8.2 - 2^2 = \underline{4.2}$

$\sigma(X) = \sqrt{D(X)} = \sqrt{4.2} \doteq \underline{2.05}$

$X \sim \begin{pmatrix} 1001 & 1002 & 1003 & 1004 \\ 0.2 & u & v & 0.1 \end{pmatrix}$

$E(X) = 1002$

$u, v = ?$

U - metoda

$E(X) = a + E(X-a)$
$D(X) = E[(X-a)^2] - (E(X-a))^2$

$a \to$ poljubna vrednost

$a = 1002$

$E(X) = 1002 + (-1) \cdot 0.2 + 0 \cdot u + 1 \cdot v + 2 \cdot 0.1 = 1002$
$-0.2 + v + 0.2 = 0$
$\underline{v = 0, \ u = 0.7}$

$$X \sim \begin{pmatrix} 1001 & 1002 & 1003 & 1004 \\ 0{\cdot}2 & u & v & 0{\cdot}1 \end{pmatrix}$$

- Kolikšno je minimalno in kolikšno je maksimalno matematično upanje?
- $D(X)$?

- $E(X) = 1002 - 0{\cdot}2 + v + 0{\cdot}2 = 1002 + v$

→ v je lahko 0, toda ne manjši od nič, ker je v verjetnost

$\min E(X) = 1002$

$\max v = 0{\cdot}7 \Rightarrow$
$\max E(X) = 1002{\cdot}7$

- $D(X) = E((X-1002)^2) - (E(X-1002))^2$

$$E((X-1002)^2) = (-1)^2 \cdot 0{\cdot}2 + 0^2 \cdot u + 1^2 \cdot v + 2^2 \cdot 0{\cdot}1 = 0{\cdot}6 + v$$

$E(X-1002) = v$

$D(X) = 0{\cdot}6 + v - v^2$

$v \in [0, 0{\cdot}7]$

$\frac{d}{dv} D(X) = -2v + 1$

$v = \frac{1}{2}$

v	$D(X)$	
0	0·6	← $\min D(X)$
0·5	0·85	← $\max D(X)$
0·7	0·81	

$\boxed{\text{OVS AV 24.3.2009}}$

3-je kovanci po 1 SIT
6 x 2 SIT
1 x 5 SIT
izvlečemo 2 kovanca. X_1 vrednost 1., X_2 vrednost 2. kovanca.

$X_1 :=$ vrednost prvega
$X_2 :=$ vrednost drugega
$S := X_1 + X_2 =$ skupna vrednost

$E(X_1)$, $D(X_1)$, $E(S)$, $D(S) = ?$

- $X_1 \sim \begin{pmatrix} 1 & 2 & 5 \\ \frac{3}{10} & \frac{6}{10} & \frac{1}{10} \end{pmatrix}$

- $E(X_1) = 1 \cdot \frac{3}{10} + 2 \cdot \frac{6}{10} + 5 \cdot \frac{1}{10} = \underline{\underline{2}}$

- $D(X_1) = E[(X_1 - E(X_1))^2] = E(X_1^2) - (E(X_1))^2$

$$E(X_1^2) = 1 \cdot \frac{3}{10} + 4 \cdot \frac{6}{10} + 25 \cdot \frac{1}{10} = \frac{52}{10} = 5{,}2$$

$$D(X_1) = 5{,}2 - 2^2 = \underline{\underline{1{,}2}}$$

- Direkten način: $D(X_1) = (1-2)^2 \cdot \frac{3}{10} + (5-2)^2 \cdot \frac{1}{10} =$
$$= \frac{3}{10} + \frac{9}{10} = \underline{\underline{\frac{12}{10}}}$$

- Če vračamo je: $\rightarrow$ neodvisno

$$S \sim \begin{pmatrix} 2 & 3 & 4 & 6 & 7 & 10 \\ \frac{9}{100} & \frac{36}{10} & \frac{36}{10} & \frac{6}{10} & \frac{12}{10} & \frac{1}{100} \end{pmatrix}$$

$\left(\frac{3}{10}\right)^2 = \frac{9}{100}$ $\quad \frac{6}{10} \cdot \frac{6}{10}$ $\quad \frac{3}{10} \cdot \frac{6}{10} \cdot 2$ $\quad \frac{3}{10} \cdot \frac{1}{10} \cdot 2$ $\quad \frac{6}{10} \cdot \frac{1}{10} \cdot 2$ $\quad \left(\frac{1}{10}\right)^2$

$$S \sim \begin{pmatrix} 2 & 3 & 4 & 6 & 7 & 10 \\ 0{,}09 & 0{,}36 & 0{,}36 & 0{,}06 & 0{,}12 & 0{,}01 \end{pmatrix}$$

$$E(S) = 2 \cdot 0'09 + 3 \cdot 0'36 + 4 \cdot 0'36 + 6 \cdot 0'06 + 7 \cdot 0'12 + 10 \cdot 0'01 =$$
$$= 0'18 + 1'08 + 1'44 + 0'36 + 0'84 + 0'1 =$$
$$= \underline{4}$$

$$D(S) = \overset{(2-4)^2}{4} \cdot 0'09 + \overset{(3-4)^2}{1} \cdot 0'36 + \overset{(6-4)^2}{4} \cdot 0'06 + \underset{(7-4)^2}{9} \cdot 0'12 + \underset{(10-4)^2}{36} \cdot 0'01 = \underline{\underline{2'4}}$$

↳ gledamo kvadrate odmikov od matematičnega upanja

- Če ne vračamo, je:

$$S \sim \begin{pmatrix} 2 & 3 & 4 & 6 & 7 \\ \frac{6}{90} & \frac{36}{90} & \frac{6}{10} \cdot \frac{5}{9} & \frac{1}{10} \cdot \frac{3}{9} \cdot 2 & \frac{1}{10} \cdot \frac{6}{9} \cdot 2 \end{pmatrix}$$

$$\frac{3}{10} \cdot \frac{2}{9} \qquad \frac{3}{10} \cdot \frac{6}{9} \cdot 2$$

lahko okrog obrnemo

$$S \sim \begin{pmatrix} 2 & 3 & 4 & 6 & 7 \\ \frac{1}{15} & \frac{6}{15} & \frac{5}{15} & \frac{1}{15} & \frac{2}{15} \end{pmatrix}$$

$$E(S_2) = 2 \cdot \frac{1}{15} + 3 \cdot \frac{6}{15} + 4 \cdot \frac{5}{15} + 6 \cdot \frac{1}{15} + 7 \cdot \frac{2}{15} =$$
$$= \frac{2}{15} + \frac{18}{15} + \frac{20}{15} + \frac{6}{15} + \frac{14}{15} =$$
$$= \frac{60}{15} = \underline{4}$$

$$D(S_2) = \frac{4}{15} + \frac{6}{15} + \frac{4}{15} + \frac{18}{15} = \underline{\underline{\frac{32}{15}}} \doteq \underline{\underline{2'13}}$$

- Vedno velja: $E(X_1 + X_2) = E(X_1) + E(X_2)$ za poljubni slučajni spremenljivki X_1 in X_2

- X_1, X_2 neodvisni $\Rightarrow D(X_1 + X_2) = D(X_1) + D(X_2)$

Kovarianca

$$K(X,Y) = E\left[(X-E(X))(Y-E(Y))\right] = E(XY) - E(X)E(Y)$$

Korelacijski koeficient

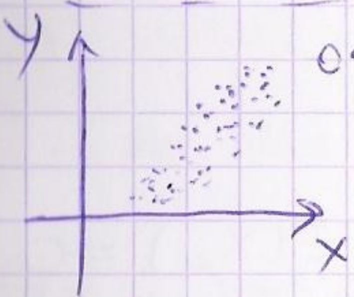
ro

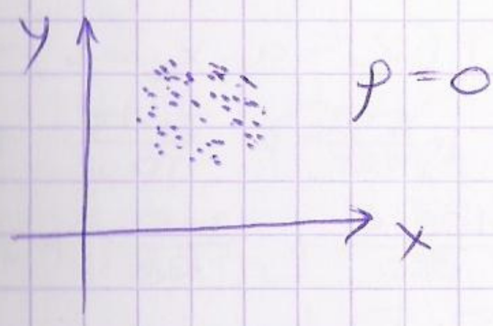
$$\rho(X,Y) = \frac{K(X,Y)}{\sigma(X)\sigma(Y)}$$

~~X, Y nekorelirani~~

$$-1 \leq \rho \leq 1$$

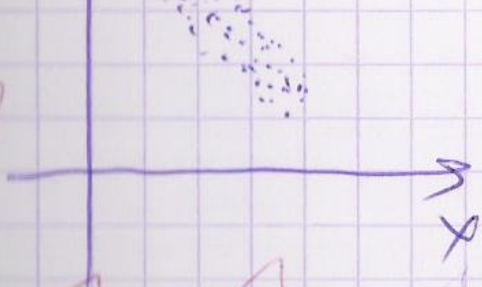
X, Y neodvisni $\underset{\not\Leftarrow}{\Rightarrow}$ nekorelirani

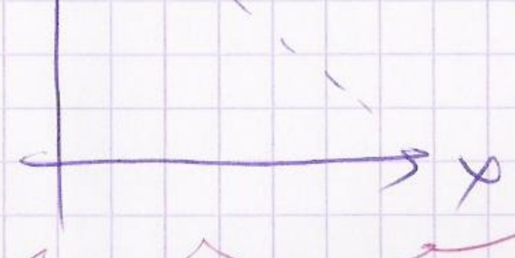




- X in Y sta nekorelirani

- čim sta neodvisni, sta korelirani





Za prejšnji primer:

$K(X_1, X_2) = ?$

- potrebujemo navzkrižno porazdelitev:

	$X_2 = 1$ ($3/10$)	$X_2 = 2$ ($6/10$)	$X_2 = 5$ ($1/10$)
$X_1 = 1$	$\frac{9}{100}$	$\frac{18}{100}$	$\frac{3}{100}$
$X_1 = 2$	$\frac{18}{100}$	$\frac{36}{100}$	$\frac{6}{100}$
$X_2 = 5$	$\frac{3}{100}$	$\frac{6}{100}$	$\frac{1}{100}$

navzkrižne verjetnosti

↑ Če vračamo!

$$K(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2) = 0$$
$$\Rightarrow \rho(X_1, X_2) = 0$$

$$E(X_1 X_2) = \sum_a \sum_b a \cdot b \cdot P(X_1 = a, X_2 = b) =$$
$$= \frac{9}{100} + \frac{2 \cdot 18}{100} + \frac{5 \cdot 3}{100} + \frac{2 \cdot 18}{100} +$$
$$+ \frac{4 \cdot 36}{100} + \frac{10 \cdot 6}{100} + \frac{3 \cdot 5}{100} +$$
$$+ \frac{5 \cdot 2 \cdot 6}{100} + \frac{5 \cdot 5 \cdot 1}{100} =$$
$$= \underline{\underline{4}}$$

* X, Y neodvisni ⇒ E(X)E(Y) ⇔ X, Y nekorelirani

• Če ne vračamo!

	$X_2=1$	$X_2=2$	$X_2=5$
$X_1=1$	$\frac{6}{90}=\frac{2}{30}$	$\frac{18}{90}=\frac{6}{30}$	$\frac{3}{90}=\frac{1}{30}$
$X_1=2$	$\frac{18}{90}=\frac{6}{30}$	$\frac{30}{90}=\frac{10}{30}$	$\frac{6}{90}=\frac{2}{30}$
$X_1=5$	$\frac{3}{90}=\frac{1}{30}$	$\frac{6}{90}=\frac{2}{30}$	0

Velja:

1) vsota vseh mora biti 1

2) ~~za~~ robne porazdelitve → seštevek v vrstici = 1 in v stolpcu = 1

$$E(X_1X_2) = \frac{2}{30}+\frac{12}{30}+\frac{5}{30}+\frac{12}{30}+\frac{40}{30}+\frac{20}{30}+$$
$$+\frac{5}{30}+\frac{20}{30}+0 = \frac{116}{30} = \underline{\underline{\frac{58}{15}}}$$

$$K(X_1,X_2) = \frac{58}{15} - \underset{2\cdot 2}{4} = -\frac{2}{15}$$

→ negativno korelirani

$$\rho(X_1,X_2) = \frac{K(X_1,X_2)}{\sqrt{D(X_1)\cdot D(X_2)}}$$

$$= \frac{-\frac{2}{15}}{\sqrt{1{,}2\cdot 1{,}2}} = \frac{-\frac{2}{15}}{\frac{12}{10}} = -\frac{20}{15\cdot 12} =$$

$$= \underline{\underline{-\frac{1}{9}}}$$

Kovarianca je bilinearna operacija (obnaša se tako, kot množenje).

$$K(X_1 + X_2, Y) = K(X_1, Y) + K(X_2, Y)$$

$$K(X, Y) = K(Y, X)$$

$$K(X, X) = D(X) \geq 0$$

→ Kovarianca je skalarni produkt na vektorskem prostoru slučajnih spremenljivk z matematičnim upanjem nič.

$$D(X_1 + X_2) = K(X_1 + X_2, X_1 + X_2) =$$

$$\cancel{= K(X_1, X_2) + K(X_1, X_2}$$

$$= K(X_1, X_1) + K(X_1, X_2) + K(X_2, X_1) + K(X_2, X_2) =$$

$$= D(X_1) + 2K(X_1, X_2) + D(X_2)$$

Primer: X_1, X_2 nekorelirani (npr. če sta neodvisni) → $D(X_1 + X_2) = D(X_1) + D(X_2)$

↳ Pitagorov izrek

Primer: vaši kovanci ki jih ne vračamo.

$$D(X_1 + X_2) = \frac{32}{15}, \quad D(X_1) = D(X_2) = 1 \cdot 2 = \frac{6}{5},$$

$$K(X_1, X_2) = -\frac{2}{15}$$

$$\frac{32}{15} = \frac{6}{5} - 2 \cdot \frac{2}{15} + \frac{6}{5} \quad ✓$$

	$Y=0$	$Y=1$	$Y=2$	
$X=0$	0·2	0·2	0·2	0·6
$X=1$	0·2	0	0·2	0·4
	0·4	0·2	0·4	1

← robne porazdelitve

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

→ izračunamo robne porazdelitve

$$X \sim \begin{pmatrix} 0 & 1 \\ 0{\cdot}6 & 0{\cdot}4 \end{pmatrix} \quad Y \sim \begin{pmatrix} 0 & 1 & 2 \\ 0{\cdot}4 & 0{\cdot}2 & 0{\cdot}4 \end{pmatrix}$$

$$E(XY) = 0 + 0{\cdot}4 = 0{\cdot}4$$

$$E(X) = 0\cdot 0{\cdot}6 + 1\cdot 0{\cdot}4 = 0{\cdot}4$$

$$E(Y) = 0{\cdot}2 + 2\cdot 0{\cdot}4 = \cancel{0{\cdot}1}\ 1$$

$$K(X,Y) = 0{\cdot}4 - 0{\cdot}4 = \underline{\underline{0}}$$

⇒ X, Y sta ~~nekoleni~~ nekorelirani

Ali sta X in Y neodvisni ?

- Ali za poljubna x in y velja

$$P(X=x, Y=y) = P(X=x)\cdot P(Y=y) \; ?$$

$$P(X=0, Y=0) = 0{\cdot}2$$

$$P(X=0)\cdot P(Y=0) = 0{\cdot}6 \cdot 0{\cdot}4 = 0{\cdot}24$$

⇒ X in Y sta odvisni.

	Y=1	Y=2	Y=3	
X=1	0·1	0·2	0·1	0·4
X=2	0·15	0·3	0·15	0·6
	0·25	0·5	0·25	1

Dopolni tabelo tako, da bosta X in Y neodvisni!

$$P(X=1, Y=1) = \underset{\text{robna verjetnost}}{P(X=1)} \cdot P(Y=1)$$

$$0{\cdot}1 = 0{\cdot}4 \cdot P(Y=1)$$

$$P(Y=1) = \frac{0{\cdot}1}{0{\cdot}4} = 0{\cdot}25$$

Slučajni spremenljivki X in Y s končno zalogo vrednosti sta neodvisni n.t.k. ima ~~njun~~ matrika njune ~~po~~ navzkrižne porazdelitve rang 1.

OVS AV 31.3.2009

ZVEZNE SLUČAJNE SPREMENLJIVKE

Diskretne porazdelitve:

$$P(a \leq X \leq b) = \sum_{x;\, a \leq x \leq b} P(X = x)$$

(recimo $P(X=2) + P(X=3) + P(X=4) + P(X=5)$)

$$P(X=x) \geq 0, \quad \sum_x P(X=x) = 1$$

$$E[f(X)] = \sum_x f(x) \cdot P(X=x)$$

Zvezne porazdelitve:

$$P(a \leq X \leq b) = \int_a^b g_X(x)\,dx$$

$$g_X(x) \geq 0$$

$$\int_{-\infty}^{\infty} g_X(x)\,dx = 1$$

$$E[f(X)] = \int_{-\infty}^{\infty} f(x)\, g_X(x)\,dx$$

$$g_X(x) = \begin{cases} cx^2; & 0 \leq x \leq 3 \\ 0; & \text{sicer} \end{cases}$$

$c = ?$

$P(1 \leq X \leq 2) = ?$

$P(X \leq 1) = ?$

$E(X) = ?$

$D(X) = ?$

$$\int_{-\infty}^{\infty} g_x(x)\,dx = \int_0^3 cx^2 \cdot dx = 1$$

$$c \cdot \frac{x^3}{3} \Big|_0^3 = 1$$

$$\frac{c \cdot 27}{3} = 1$$

$$9c = 1$$

$$\underline{\underline{c = \frac{1}{9}}}$$

$$P(1 < x < 2) = \int_1^2 \frac{1}{9} x^2 \cdot dx = \frac{1}{27} x^3 \Big|_1^2 =$$

$$= \frac{8}{27} - \frac{1}{27} = \underline{\underline{\frac{7}{27}}}$$

$$P(x \leq 1) = \int_{-\infty}^{1} g_x(x)\,dx = \int_0^1 \frac{x^2}{9}\,dx = \frac{x^3}{27} \Big|_0^1 = \underline{\underline{\frac{1}{27}}}$$

$$E(X) = \int_{-\infty}^{\infty} x \cdot g_x(x)\,dx = \int_0^3 \frac{x^3}{9}\,dx = \frac{x^4}{36} \Big|_0^3 =$$

$$= \frac{81}{36} = \underline{\underline{\frac{9}{4}}}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot g_x(x)\,dx = \int_0^3 \frac{x^4}{9}\,dx = \frac{x^5}{45} \Big|_0^3 =$$

$$= \frac{243}{45} = \underline{\underline{\frac{27}{5}}}$$

$$D(X) = E(X^2) - E(X)^2 = \frac{27}{5} - \frac{81}{16} = \frac{432}{80} - \frac{405}{80} =$$

$$= \underline{\underline{\frac{27}{80}}}$$

$$\begin{array}{r} 16 \cdot 27 \\ \hline 32 \\ 112 \\ \hline 432 \end{array}$$

Kumulativna porazdelitvena funkcija

- $F_X(x) = P(X < x)$

Za zvezne slučajne spremenljivke:

$$F_X(x) = \int_{-\infty}^{x} g_X(t)\, dt$$

$$g_X(x) = F_X'(x)$$

- Vedno: F_X je naraščajoča

$$\lim_{x \to -\infty} F_X(x) = 0,$$

$$\lim_{x \to \infty} F_X(x) = 1$$

F_X zvezna, odsekoma zvezno odvedljiva $\Rightarrow$

$\Rightarrow$ F_X absolutno zvezna $\Leftrightarrow$ X je zvezno porazdeljena $\Rightarrow$

$\Rightarrow$ F_X zvezna

$$g_X(x) = \begin{cases} c \cdot x^2 ; & 0 \le x \le 3 \\ 0 ; & \text{sicer} \end{cases}$$

(prejšnji primer)

$F_X = ?$ ($\leftrightarrow$ kum. porazd. funkc.)

$$F_X(x) = \int_{-\infty}^{x} g_X(t)\, dt$$

(1) $x \leq 0$: $\quad F_X(0) = 0$

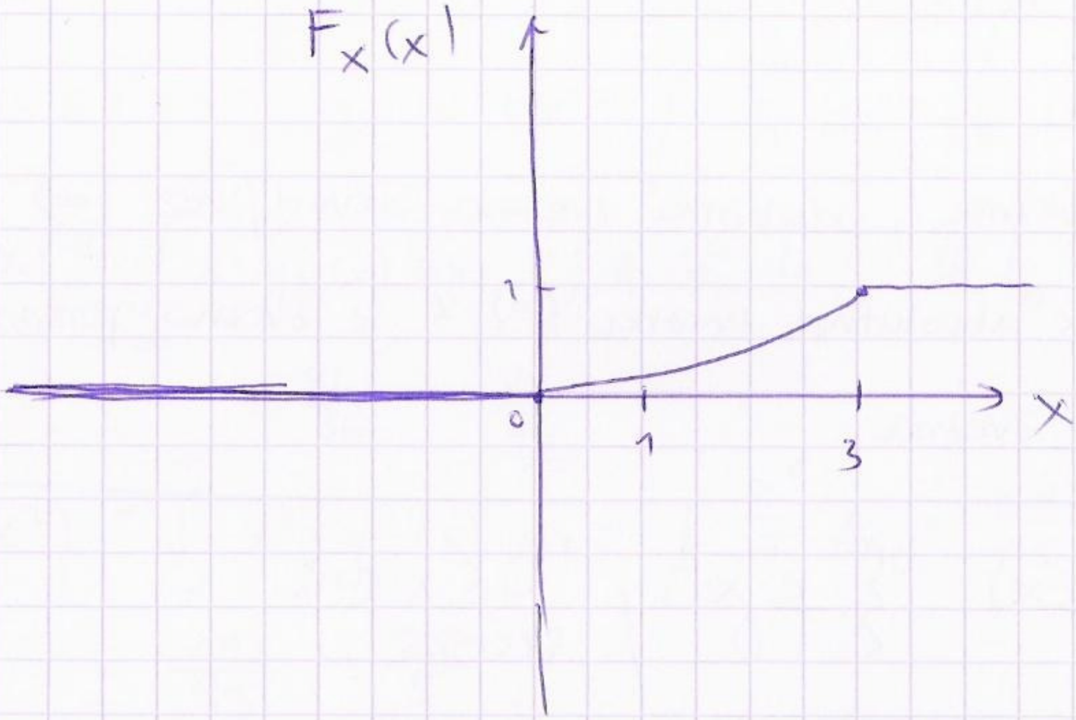
(2) $0 \leq x \leq 3$:

$$F_X(x) = \int_0^x \frac{1}{9} t^2 dt = \frac{1}{9} \cdot \frac{t^3}{3} \Big|_0^x = \frac{x^3}{27}$$

(3) $X \geq 3$:

$$F_X(x) = P(X < x) \geq P(X \leq 3) = 1$$

$$\rightarrow F_X(x) = \begin{cases} 0; & x \leq 0 \\ \frac{x^3}{27}; & 0 \leq x \leq 3 \\ 1; & x \geq 3 \end{cases}$$



$$a \leq b \Rightarrow P(a \leq X < b) = F_X(b) - F_X(a)$$

$$P(1 < X < 2) = P(1 \leq X < 2) =$$

$$= F_X(2) - F_X(1) = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

$$g_x(x) = \begin{cases} ae^{bx}\ ; & x \geq 0 \\ 0; & \text{sicer} \end{cases}$$

$\uparrow$

gostota slučajne spremenljivke x, veljati mora $\rightarrow g_x = g$

$E(x) = 4.$

$a, b = ?$

$$\int_{-\infty}^{\infty} g_x(x)\,dx = \int_0^{\infty} a \cdot e^{bx}\,dx = \frac{1}{b}\, ae^{bx}\Big|_0^{\infty}$$

$$\boxed{\int e^{bx}\,dx = \frac{1}{b}e^{bx} + C}$$

če $b \geq 0$; ne bo v redu

$b < 0$; ~~zg.~~ meja na zg. meji je ∞

$b = 0$; ni definirano

$$b < 0 \Rightarrow \int_{-\infty}^{\infty} g_x(x)\,dx = \frac{a}{b} \cdot e^{bx}\Big|_0^{\infty} = \frac{a}{b}(0-1) = \frac{-a}{b}$$

g gostota $\Leftrightarrow$ $b < 0$ in $-\frac{a}{b} = 1$

$b = -a$

$$g_x(x) = \begin{cases} ae^{-ax}\ ; & x \geq 0 \\ 0; & \text{sicer} \end{cases} \qquad (a > 0)$$

$$E(x) = \int_{-\infty}^{\infty} x \cdot g_x(x)\,dx =$$

$$t = ax$$
$$dt = a\,dx$$
$$dx = \frac{dt}{a}$$

$$= \int_0^\infty x \cdot a e^{-ax}\,dx = \int_0^\infty t \cdot e^{-t}\,\frac{dt}{a} =$$

$$= \frac{1}{a} \cdot \int_0^\infty t e^{-t}\,dt =$$

$$\int_a^b u\,dv = uv\Big|_a^b - \int_a^b v\,du$$

$$u = t, \quad dv = e^{-t}\,dt$$
$$du = dt, \quad v = -e^{-t}$$

$$= \frac{1}{a} \cdot \Big(\underbrace{-te^{-t}\Big|_0^\infty}_{0} + \int_0^\infty e^{-t}\,dt\Big) = -\frac{1}{a}\,e^{-t}\Big|_0^\infty =$$

$$= \underline{\underline{\frac{1}{a}}}$$

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

$$a = \frac{1}{4}, \quad b = -\frac{1}{4}$$

- X je porazdeljena eksponentno $X \sim \mathrm{Exp}(a)$

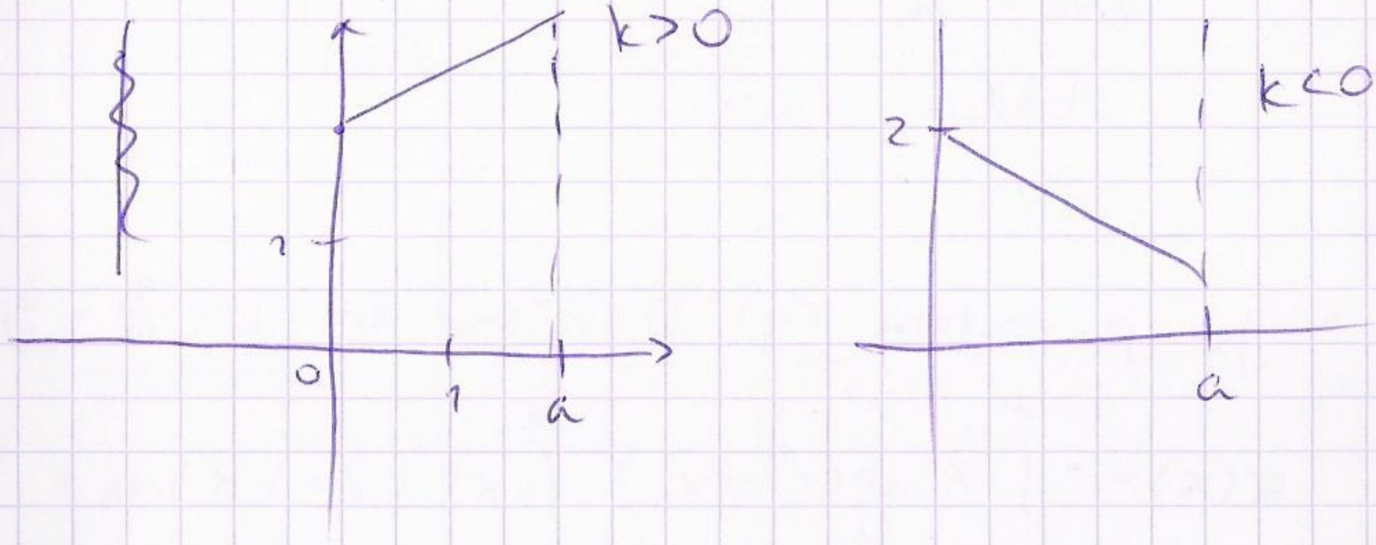
- $$\boxed{\int_0^\infty t^n \cdot e^{-t}\,dt = n!}$$

$$g_X(x) = \begin{cases} 2 - kx \,; & 0 \leq x \leq a \\ 0 \,; & \text{sicer} \end{cases}$$

$E(X) = \max$; $a, k = ?$

Da bo g_X sploh gostota: $a > 0$, $\int_{-\infty}^{\infty} g_X(x)\,dx = 1$,
$g_X(x) \geq 0$

GRAF GOSTOTE:



Za $g_X(x) \geq 0$ bo dovolj, $g_X(a) \geq 0$.

$2 - k \cdot a \geq 0$

$2 \geq k \cdot a$

$\underline{ka \leq 2}$

$$\int_{-\infty}^{\infty} g_X(x)\,dx = \int_0^a (2 - kx)\,dx = \left(2x - \frac{kx^2}{2}\right)\Big|_0^a =$$

$$= 2a - \frac{ka^2}{2}$$

$$2a - \frac{k \cdot a^2}{2} = 1$$

$$\frac{ka^2}{2} = 2a - 1$$

$$k = \frac{4a-2}{a^2} = \frac{4}{a} - \frac{2}{a^2}$$

$ka \leq 2$:

$$\left(\frac{4}{a} - \frac{2}{a^2}\right) \cdot a \leq 2$$

$$4 - \frac{2}{a} \leq 2$$

$$2 \leq \frac{2}{a}$$

$$1 \geq a \qquad a > 0$$

$$a \leq 1$$

- g_x je gostota $\Leftrightarrow$ $0 < a \leq 1$ in $k = \frac{4}{a} - \frac{2}{a^2}$

$$E(x) = \int_{-\infty}^{\infty} x \cdot g(x)\,dx = \int_0^a x \cdot (2 - kx)\,dx =$$

$$= \int_0^a (2x - kx^2)\,dx = \left[x^2 - \frac{kx^3}{3}\right]_0^a =$$

$$= a^2 - \frac{ka^3}{3}$$

$$E(x) = a^2 - \frac{a^3}{3} \cdot \left(\frac{4}{a} - \frac{2}{a^2}\right) =$$

$$= a^2 - \frac{4a^2}{3} + \frac{2a}{3} =$$

$$= \frac{2a - a^2}{3} \; ; \quad 0 < a \leq 1$$

$\rightarrow$ globalni max!

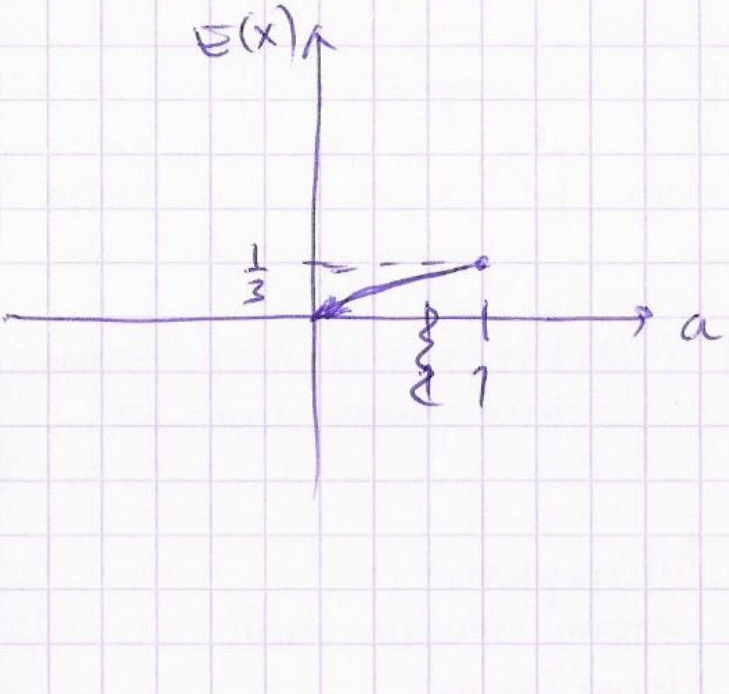
a	$\frac{2a-a^2}{3}$	
0	0	
1	$\frac{1}{3}$	MAX

stacionarne točka!

$$\cancel{\frac{2a-a^2}{3}} \quad \frac{d}{da} E(X) = \frac{2-2a}{3} = 0 \Rightarrow \underline{a = 1}$$

max

$\underline{\underline{k = 2}}$



OVS AV 7. 4. 2008

NORMALNA ali GAUSSOVA PORAZDELITEV

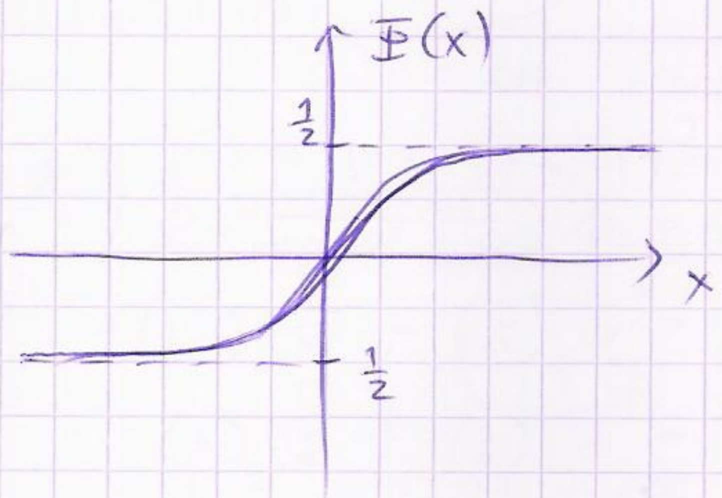
$\mu \in \mathbb{R}, \sigma > 0$

$N(\mu, \sigma)$ je porazdelitev z gostoto:

$$g(x) = \frac{1}{\sigma\sqrt{2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$a \leq b \Rightarrow P(a \leq X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}}\, dt$$



$$\Phi(-x) = -\Phi(x)$$

STANDARDNA NORMALNA PORAZDELITEV

$N(0,1)$ ima gostoto

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

in porazdelitveno funkcijo $\frac{1}{2} + \Phi(x)$.

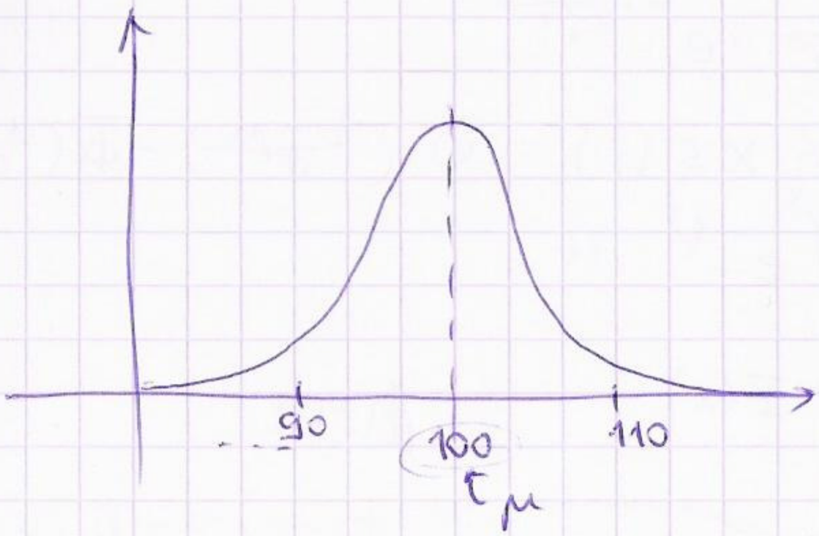
$$\mathrm{erf}(x) = 2\,\Phi(x \cdot \sqrt{2})$$

↳ „error function"

$N(\overset{\mu}{100}, \overset{\sigma}{10})$

$P(90 \leq X \leq 120)$?

$P(X \leq 0)$?



$$P(90 \leq X \leq 120) = \Phi\left(\frac{120-100}{10}\right) - \Phi\left(\frac{90-100}{10}\right) =$$

$$= \Phi(2) - \Phi(-1) =$$

$$= \Phi(2) + \Phi(1) =$$

$$= 0{,}4772 + 0{,}3413 =$$

$$= \underline{0{,}8185}$$

$$P(X \leq 0) = P(-\infty \leq X \leq 0) =$$

$$= \Phi\left(\frac{0-100}{10}\right) - \Phi\left(\frac{-\infty-100}{10}\right) =$$

$$= \Phi(-10) + \Phi(\infty) =$$

$$= \frac{1}{2} - \Phi(10) \doteq \underline{0}$$

$\doteq \underline{7{,}62 \cdot 10^{-24}}$ na najmanj 4 dec. natančno

↳ bolj točen rezultat

$$x > 0 \Rightarrow \frac{x}{1+x^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} < \frac{1}{2} - \Phi(x) < \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

(veliki odkloni)

→ znotraj 3 sigem: približno 99 %

- avtobus spelje točno ob 7:00
- sam pridem na postajo:

čas prihoda

$T \sim N(6{:}59, 3 \text{ min})$

- kolikšna je verjetnost, da še ujamem avtobus?

↳ $P(T < 7{:}00)$

$$P(T < 7{:}00) = P(-\infty < T < 7{:}00) =$$

$$= \Phi\left(\frac{7{:}00 - 6{:}59}{3 \text{ min}}\right) - \Phi\left(\frac{-\infty - 6{:}59}{3 \text{ min}}\right) =$$

$$= \Phi\left(\frac{1}{3}\right) + \Phi(\infty) =$$

$$= 0{,}1293 + 0{,}5 =$$

$$= \underline{\underline{0{,}6293}}$$

Natančnejši rezultat: 0,63056

Zakaj ne recimo: $\frac{1}{\pi(1+x^2)}$

↳ integral = arctg

↳ Cauchyjeva porazdelitev

→ centralni limitni izrek

CENTRALNI LIMITNI IZREK

Če je S vsota veliko neodvisnih slučajnih spremenljivk, z dovolj lepimi porazdelitvami, od katerih nobena posebej ne izstopa, je približno $S \sim N(\mu, \sigma)$, kjer je $\mu = E(S)$ in $\sigma^2 = D(S)$.

$$S = X_1 + X_2 + \ldots + X_n$$

$X_1, X_2, \ldots, X_n$ neodvisne:

$$\mu = E(X_1) + E(X_2) + \ldots + E(X_n)$$

$$\sigma = D(X_1) + D(X_2) + \ldots + D(X_n)$$

Zavarovalnica pri neki polici izplača:

z verjetnostjo 15% ... 5 €,

z verjetnostjo 10% ... 10 €,

sicer ne izplača ~~..~~ nič.

Zavarovalna premija je 2 €.

Kolikšna je verjetnost, da ima zavarovalnica izgubo – za 1, 3, 5 in 500 polic.

→ predpostavimo, da so „škodni dogodki" neodvisni.

$P(5) = 0{,}15$

$P(10) = 0{,}1$

$P(0) = 0{,}75$

<u>$n=1$:</u>

$P(\text{izguba})$

$\underline{P(I_1) = 0{,}25}$

$n=3$: 6 € premij

škoda je lahko imela 1. ali 2. ali 3. polica

$$P(I_3) = 1-(0,75^3 + 0,75^2 \cdot 0,15 \cdot 3) =$$
$$= \underline{0,325}$$

$n=5$: 10 € premij

$$P(I_5) = 1-\left(0,75^5 + 0,75^4 \cdot 0,25 \cdot 5 + 0,75^3 \cdot 0,15^2 \cdot \binom{5}{2}\right) =$$
$$= \underline{0{'}272}$$

$n=500$: 1000 € premij

S = skupna škoda

$$S = \underset{\text{škode pri posameznih policah}}{X_1} + X_2 + \ldots + X_{500}$$

$$X_i \sim \begin{pmatrix} 0 & 5 & 10 \\ 0{'}75 & 0{'}15 & 0{'}1 \end{pmatrix}$$

$$E(X_i) = 1,75$$

$$E(X_i^2) = 25 \cdot 0{'}15 + 100 \cdot 0{'}1 = 13{'}75$$

$$D(X) = E(X_i^2) - E(X_i)^2 = 13,75 - 1,75^2 =$$
$$= 10,6875$$

$$E(S) = 500 \cdot 1,75 = 875 = \mu$$

$$D(S) = 500 \cdot 10,6875 = 5343,75 = \sigma^2$$

$$P(I_{500}) = P(S \geq 1000) =$$

$$= P(1000 < S < \infty =$$

$$= \Phi\left(\frac{\infty - 875}{\sqrt{5343{,}75}}\right) - \Phi\left(\frac{1000 - 875}{\sqrt{5343{,}75}}\right) =$$

$$= \frac{1}{2} - \Phi(1{,}78) = \frac{1}{2} - 0'4625 =$$

$$= \underline{\underline{0'0375}}$$ → preveč! prevelika napaka

→ točen rezultat: $\underline{0'04364}$

APROKSIMACIJA BINOMSKE PORAZDELITVE

→ POSEBEN PRIMER CLI

$S \sim B(n,p)$

* n neodvisnih poskusov, vsak uspe z verjetnostjo p
  S šteje uspešne poskuse

$\underbrace{n \to \infty,\ p,\ 1-p \gg \frac{1}{n}}$ → pomeni bistveno ~~manjše~~ večje, npr. za faktor 10

$\Leftrightarrow \sigma := \sqrt{n \cdot p \cdot (1-p)} \to \infty$

<u>Laplaceova lokalna formula:</u>

$$|k - n \cdot p| \ll \sigma^{4/3} \Rightarrow P(S=k) \sim \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-n\cdot p)^2}{2\sigma^2}}$$

<u>Laplaceova integralska formula:</u>

$$|a - np| \ll \sigma^{4/3} \text{ ali } |b - np| \ll \sigma^{4/3}$$

$$\Rightarrow P(a \leq S \leq b) \sim \Phi\left(\frac{b - np + \frac{1}{2}}{\sigma}\right) - \Phi\left(\frac{a - np - \frac{1}{2}}{\sigma}\right)$$

||

$$\parallel$$

$$P\left(a-\tfrac{1}{2} < S < b+\tfrac{1}{2}\right)$$

$b-a \gg 1 \Rightarrow$ smemo polovičke izpustiti

- V tovarni vsak dan 1600 izdelkov.
- verjetnost, da je okvarjeno = 0,1

$P(\text{okvarjeno}) = 0{,}1$
$E(S) = 160$

$S :=$ št. okvarjenih
$S \sim B(\underset{n}{\underset{\parallel}{1600}}, \underset{p}{\underset{\parallel}{0{,}1}})$

$$P(S = \overset{k}{\overset{\parallel}{160}}) \approx \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(k-np)^2}{2\sigma^2}} =$$

$$= \frac{1}{12\cdot\sqrt{2\pi}}\cdot e^{-\frac{(160-160)^2}{2\cdot 12^2}} =$$

$$\sigma = \sqrt{n\cdot p\cdot(1-p)} = \sqrt{1600\cdot 0{\cdot}1\cdot 0{\cdot}9} = 12$$

$$= \frac{1}{12\cdot\sqrt{2\pi}}\cdot e^{0} = \frac{1}{12\cdot\sqrt{2\pi}} \doteq \underline{0{\cdot}033245}$$

$\rightarrow$ točen rezultat: $0{\cdot}033228$

naslednjič nadaljujemo primer

OVS - AV - 14.4.2009 14.4.2009

$S \sim B(n,p)$

n neodvisnih poskusov

P(posamezen poskus uspe) $= p$

$S :=$ število uspešnih poskusov

$$X_i = \begin{cases} 1; & i\text{-ti poskus uspe} \\ 0; & i\text{-ti poskus ne uspe} \end{cases}$$

$S = X_1 + X_2 + \dots + X_n$

n velik, p ne preblizu 0 ali 1 $\Rightarrow S \sim N(\overset{E}{\mu}, \overset{D}{\sigma})$

$$X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix} \qquad E(X_i) = p$$

$$\mu = E(S) = E(X_1) + \dots + E(X_n) = np$$

$$E(X_i^2) = p$$

$$D(X_i) = p - p^2 = p(1-p)$$

$$D(S) = D(X_1) + \dots + D(X_n) = np(1-p)$$

$$\sigma = \sqrt{np(1-p)}$$

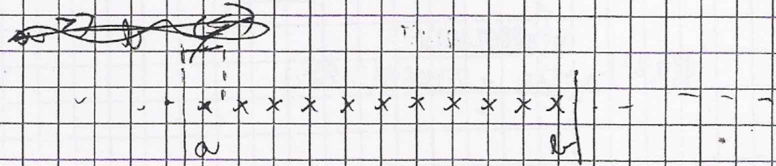
$n \to \infty$, $p, 1-p \gg \frac{1}{n}$ (ali, ekvivalentno $\sigma \to \infty$)

$|k - \mu| \ll \sigma^{4/3} \Rightarrow$ Laplaceova lokalna formula:

$$P(S=k) \sim \underbrace{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-\mu)^2}{2\sigma^2}}}_{\text{gostota normalne porazdelitve}} \qquad k \in \mathbb{Z}$$

LAPLACEOVA INTEGRALSKA FORMULA:

$$P(a \le S \le b) \sim \Phi\left(\frac{b-\mu+\frac{1}{2}}{\sigma}\right) - \Phi\left(\frac{a-\mu-\frac{1}{2}}{\sigma}\right) \qquad (a \le b)$$



$b - a \gg 1 \Rightarrow$ mamo polovičke izpustit

V tovarni vsak dan proizvedejo 1600 izdelkov. Za vsakega je verjetnost, da je pokvarjen 10%. Pokvarjenih natanko 160 izdelkov in 175 izdelkov

1600 izdelkov

$p = 0.1$

S := št. pokvarjenih

$P(S=160) = ?$

$P(S=175) = ?$

$S \sim B(1600, 0.1)$

$\sigma = \sqrt{n \cdot p(1-p)} = \sqrt{1600 \cdot 0.1 \cdot 0.9} = 12$

$\mu = np = 160$

$P(S=160) \sim \frac{1}{12\sqrt{2\pi}} = 0.033245$

Točen rezultat: 0.03328

$k = 175 \quad k - \mu = 15$

$\sigma^{4/3} = 27.4$

$P(S=175) \sim \frac{1}{12\sqrt{2\pi}} e^{-\frac{15^2}{2 \cdot 12^2}} = 0{,}015221$

Točen rezultat: 0.014929

$P(S \geq 175)$ uporabimo Laplaceovo INTEGRALSKO FORMULO

$P(175 < S < \infty) = P(176 \leq S < \infty) \sim \Phi\left(\frac{\infty - 160 + \frac{1}{2}}{12}\right) - \Phi\left(\frac{176 - 160 + \frac{1}{2}}{12}\right) = \Phi(\infty) - \Phi\left(\frac{33}{24}\right) =$

$\frac{33}{2} : 12 = \frac{33}{2} \cdot \frac{1}{12} = \frac{33}{24}$

$= \frac{1}{2} - \Phi(1.29) = \frac{1}{2} - 0{,}4015 = 0{,}0985$

Točen rezultat: 0.09944

Verjetnost da se pokvari manj kot 150 izdelkov.

$P(S < 150) = P(-\infty < S < 150) = P(-\infty < S \leq 149) = \Phi\left(\frac{149 - 160 + \frac{1}{2}}{12}\right) + \frac{1}{2} =$

$= \Phi\left(\frac{-10.5}{12}\right) + \frac{1}{2} = -\Phi(0.875) + \frac{1}{2} = \underbrace{-0.3092}_{\text{INTERPOLACIJA}} + \frac{1}{2} = \underline{0.1908}$

$\Phi(0{,}87) = 0{,}3078 \quad \frac{0{,}3078 + 0{,}3106}{2} \quad 0{,}3106 = \Phi(0{,}88)$

Točen rezultat = 0.19147

Alternativna predstava o $\frac{1}{2}$:

$P(S > 175) = P(175.5 < S < \infty) \sim -\Phi\left(\frac{175.5 - 160}{12}\right) + \Phi(\infty) = \frac{1}{2} - \Phi\left(\frac{15.5}{12}\right)$

Zavarovalnica zavaruje 1000 oseb. Verjetnost nezgode osebe 0,15%.
Kolikšna je verjetnost da se nobeden ne ponesreči, več kot štirje?

$S := \#$ ponesrečenih $\sim B(1000,\ 0{,}0015)$

$$P(S=0) = P(-\infty < S \leq 0) = +\Phi\left(\frac{0-1{,}5+\frac{1}{2}}{1{,}224}\right) + \Phi(\infty) = -\Phi(0{,}82) + \frac{1}{2} = -0{,}2939 + 0{,}5 = \underline{0{,}2061}$$

$\mu = np = 1{,}5$

$\sigma = \sqrt{np(1-p)} = \sqrt{1{,}5 \cdot 0{,}9985} = 1{,}224$

Točen rezultat:

$$P(S=k) \sim \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-\mu)^2}{2\sigma^2}}$$

$$P(S=0) \approx \frac{1}{1{,}224\sqrt{2\pi}} e^{-\frac{(0-1{,}5)^2}{2\cdot 1{,}224^2}} \doteq 0{,}1538$$

Točen rezultat: $0{,}9985^{1000} = 0{,}2229$

$0{,}9985 \cdot 0{,}9985 \cdot 0{,}9985 \ldots$

1. oseba ostane cela; 2. oseba; 3. oseba

Naj bosta X in Y neodvisni porazdeljeni standardno normalno. Verjetnost da bo $2X - Y$ manjši kot 3.

$P(2X - Y < 3)$? $\quad \sigma = \sqrt{D(X,Y)}$

$X, Y \sim N(0,1)$, neodvisni

$\underline{U \text{ in } V}$ neodvisni in normalni $\Rightarrow U + V$ normalna

$2X - Y \sim N(0, \sqrt{5})$

nato ker sta standardno normalno $\qquad D(aX) = a^2 D(X)$

$E(2X-Y) = 2E(X) - E(Y) = 2\cdot 0 - 0 = 0$

če sta standardno normalni vzamemo ko 1

$D(2X-Y) = D(2X) + D(-Y) = 4D(X) + D(Y) = 4 + 1 = 5$

$$P(2X-Y<3) = P(-\infty < 2X-Y < 3) = \Phi\left(\frac{3-0}{\sqrt{5}}\right) - \Phi(-\infty) = \Phi(1{,}34) + \frac{1}{2} = 0{,}4099 + \frac{1}{2} = \underline{0{,}9099}$$

21.4.2009

KVANTILI

$0 < \alpha < 1$:

q_α je α-ti kvantil slučajne spr. X, če velja:

$$P(X < q_\alpha) \leq \alpha, \quad P(X \leq q_\alpha) \geq \alpha$$

Če je X zvezno porazdeljena in ima v okolici točke q_α strogo

OVS AV 21.4.2009

Kvantili

$0 < \alpha < 1$:

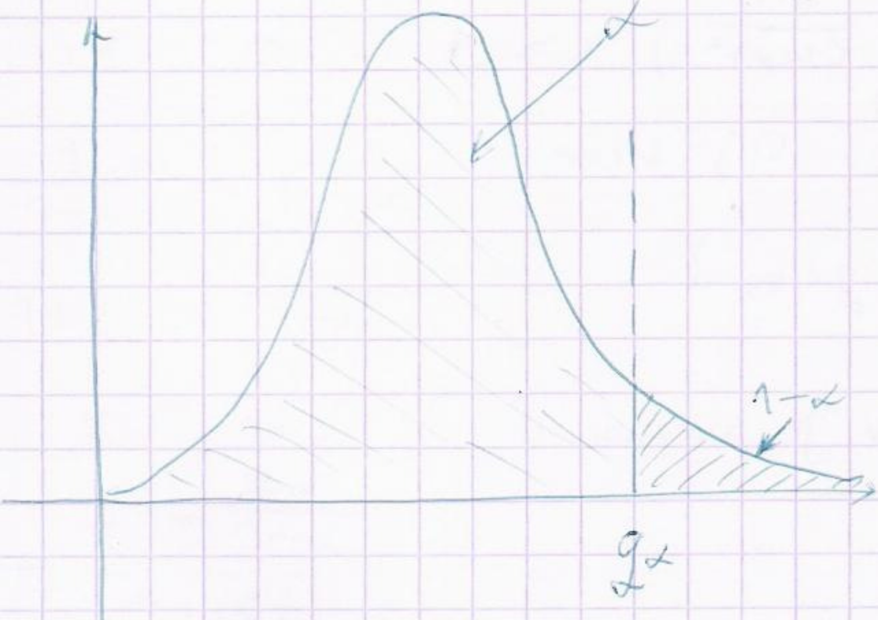
g_α je α-ti kvantil sluč. spr. X, če velja:

$P(X > g_\alpha) \le 1-\alpha$

$P(X < g_\alpha) \le \alpha$, $P(X \le g_\alpha) \ge \alpha$.

Če je X zvezno porazdeljena in ima v okolici točke g_α strogo pozitivno gostoto, pa velja:

$$F_X(g_\alpha) = P(X < g_\alpha) = P(X \le g_\alpha) = \alpha$$



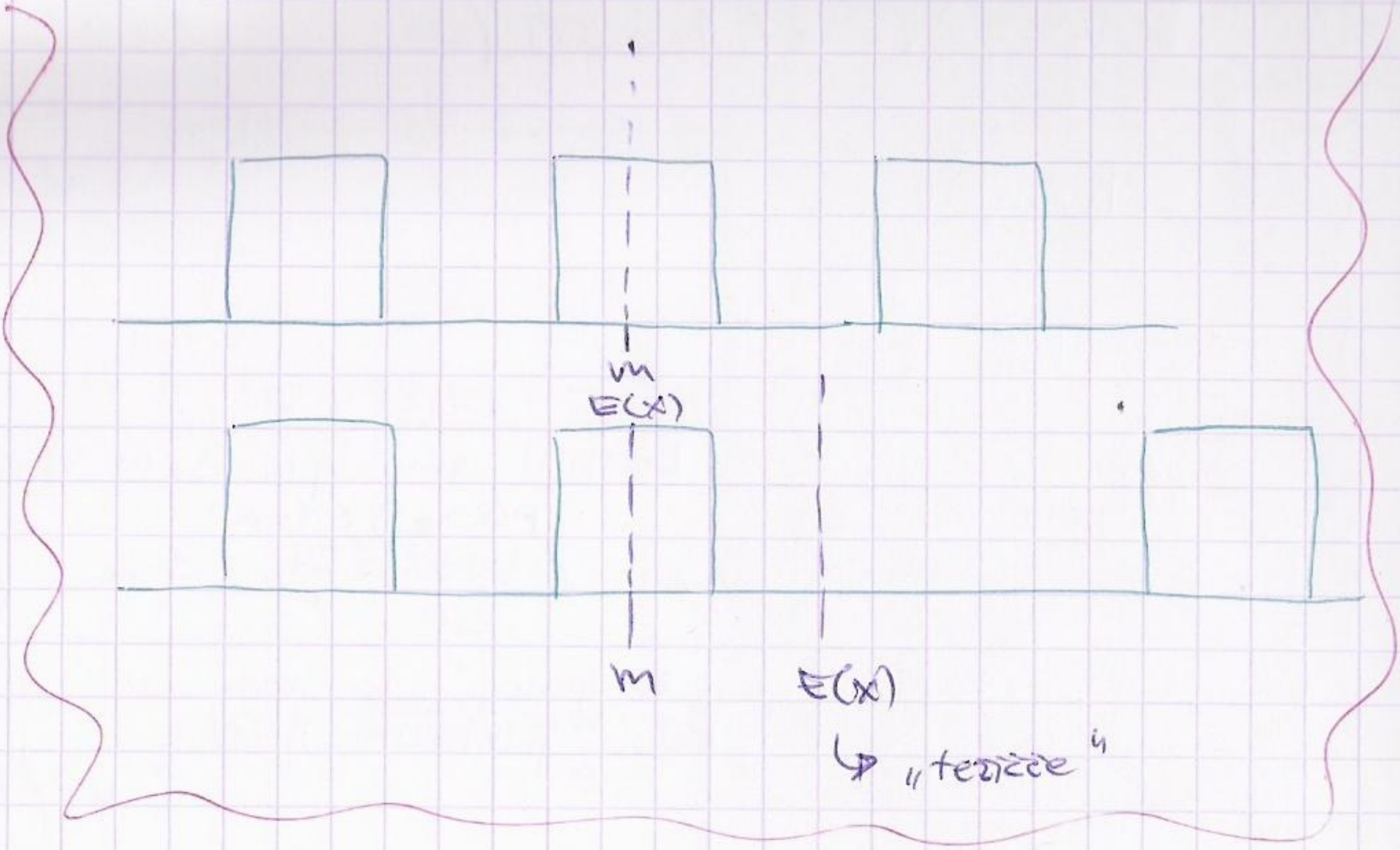
Mediana: $m = g_{1/2}$

Tercila: $g_{1/3}$, $g_{2/3}$

Kvartili: $g_{1/4}$, $g_{1/2} = m$, $g_{3/4}$

Decili: $g_{0.1}, g_{0.2}, \dots, g_{0.9}$

Percentili: $g_{0.01}, g_{0.02}, \dots, g_{0.99}$
Centili

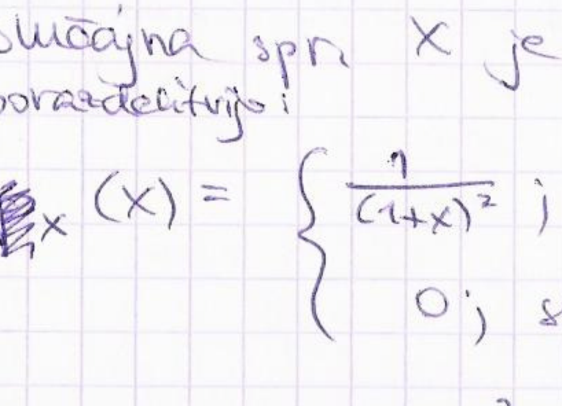


Slučajna spr. X je porazdeljena z naslednjo porazdelitvijo:

$$g_x(x) = \begin{cases} \frac{1}{(1+x)^2}; & x \geq 0 \\ 0; & \text{sicer} \end{cases}$$

m, $q_{1/4}$, $q_{0.9}$, $q_{0.99}$?

→ porazdelitev je zvezna:



$$\boxed{F_X(x) = \int_{-\infty}^{x} g_X(x)\, dt}$$

$$x > 0 \Rightarrow F_X(x) = \int_0^x \frac{1}{(1+t)^2}\, dt = \left.\frac{(1+t)^{-1}}{-1}\right|_0^x =$$

$$= \frac{(1+x)^{-1}}{-1} - \frac{(1+0)^{-1}}{-1} = \underline{\underline{\frac{-1}{1+x} + 1}}$$

$$\boxed{\begin{array}{ll} F_X(g_\alpha) = \alpha & m \\ F_X(x) = \alpha & (x = g_\alpha) \end{array}}$$

$$\frac{-1}{1+x} + 1 = \alpha$$

$$\frac{-1}{1+x} = \alpha - 1$$

$$\frac{1}{1+x} = 1 - \alpha$$

$$1 + x = \frac{1}{1-\alpha}$$

$$x = \frac{1}{1-\alpha} - 1$$

$$x = \frac{\alpha}{1-\alpha}$$

	α	g_α
$m =$	$\frac{1}{2}$	1
	$\frac{1}{4}$	$\frac{1}{3}$
	0'9	9
	0'99	99

$$x = \frac{1\,(\frac{1}{2})}{2} = 1$$

$$x = \frac{1 \cdot (\frac{3}{4})}{4} = \frac{1 \cdot 4}{4 \cdot 3} = \frac{1}{3}$$

$$x = \frac{0'9}{0'1} = 9$$

$$x = \frac{0'99}{0'01} = 99$$

Določite
(Izračunaj) 99.-ti percentil standardne normalne porazdelitve.

$Z \sim (0, 1)$

ker ima Z vedno pozitivno gostoto

$F_Z(z) = P(Z < z) = 0.99$

$F_Z(z) = P(-\infty < Z < z)$

$= \Phi\left(\frac{z-0}{1}\right) - \Phi\left(\frac{-\infty-0}{1}\right) =$

$= \Phi(z) - \Phi(-\infty) =$

$= \frac{1}{2} + \Phi(z)$

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

$P(a \leq X \leq b) =$

$= \Phi\left(\frac{b-\mu}{\sigma}\right) -$

$- \Phi\left(\frac{a-\mu}{\sigma}\right)$

$\Phi(z) = 0.49$

$z = 2.33$

Tovarna, vsak dan 1600 izdelkov.

P(okvarjen izdelek) = 0·1

P(skladišče premajhno) $\leq$ 0·05

Najmanj kako veliko mora biti skladišče?

$X :=$ potrebna velikost skladišča

$S :=$ št. okvarjenih

X : najmanjše celo število, za katerega je

$$P(S > X) \leq 0.05$$

→ Laplaceova integralska formula:

$a, b \in \mathbb{Z}$:

$$P(a \leq S \leq b) \approx \Phi\left(\frac{b + \frac{1}{2} - \mu}{\sigma}\right) - \Phi\left(\frac{a - \frac{1}{2} - \mu}{\sigma}\right)$$

$$\mu = n \cdot p$$

$$\sigma = \sqrt{n \cdot p \cdot (1-p)} \qquad P(X+1 \leq S \leq \infty)$$

$$\mu = 1600 \cdot 0.1 = 160$$

$$\sigma = \sqrt{1600 \cdot 0.9} = \sqrt{144} = 12$$

$$P(X \leq S \leq \infty) = \cancel{\Phi\left(\frac{X + \frac{1}{2} - 160}{12}\right) + 0.5 =}$$

$\sim$

$$= 0.5 - \Phi\left(\frac{x+1-\frac{1}{2}-160}{12}\right)$$

$$0.5 - \Phi\left(\frac{x-159.5}{12}\right) \leq 0.05$$

$$-\Phi\left(\frac{x-159.5}{12}\right) \leq 0.05 - 0.5$$

$$\Phi\left(\frac{x-159.5}{12}\right) \geq 0.45$$

x := najmanjše celo število, za katerega je

$$\Phi\left(\frac{x-159.5}{12}\right) \geq 0.45$$

$x = \lceil y \rceil$, kjer je $\Phi\left(\frac{y-159.5}{12}\right) = 0.45$

$$\Phi(1.645) = 0.45$$

$$\frac{y-159.5}{12} = 1.645$$

$$y = 1.645 \cdot 12 + 159.5$$

$$y = 179.24$$

$$\underline{\underline{x = 180}}$$

V resnici:

$P(S > 179) \doteq 0.0539$

$P(S > 180) \doteq 0.0457$

$\Downarrow$

Bolj površno:

$$P(S > x) \approx \Phi(\infty) - \Phi\left(\frac{x-160}{12}\right)$$

$$= \frac{1}{2} - \Phi\left(\frac{x-160}{12}\right)$$

$$\Phi\left(\frac{x-160}{12}\right) = 0'45$$

$$x = 1'695 \cdot 12 + 160 = 179'74 \nearrow 180$$

$$\frac{x-160}{12} = 1'645$$

$$P(\text{prvovrsten}) = 0'6$$

$$P(\text{min}) = 0'99$$

59% naročenih je prvovrstnih

$S :=$ št. prvovrstnih v pošiljki $\quad n =$ velikost pošiljke

$$P(S \geq 0'59n) \geq 0'99 \qquad p = 0'6$$

$$P(0'59n \leq S \leq \infty)$$

$$0'5 - \Phi\left(\frac{0'59n - \mu}{\sigma}\right) = 0'99$$

$$0'5 - \Phi\left(\frac{0'59n - 0'6n}{\sqrt{0'6 \cdot 0'4n}}\right) = 0'99$$

$$0'5 + \Phi\left(\frac{0'01n}{\sqrt{0'6n \cdot 0'4}}\right) = 0'99$$

$$\Phi\left(\frac{0'01n}{\sqrt{0'6n \cdot 0'4}}\right) = 0'49$$

$$\frac{0'01n}{\sqrt{0'6 \cdot 0'4n}} = 2'33$$

$$0{,}01n = 2{,}33 \cdot \sqrt{0{,}6 \cdot 0{,}4n} \quad /^2$$

$$0{,}0001n^{\cancel{2}} = 2{,}33^2 \cdot 0{,}24\cancel{n}$$

$$n = 10^4 \cdot 2{,}33^2 \cdot 0{,}24$$

$$n = 13\,029{,}36 \rightarrow \boxed{13030}$$

$n = 12921$: $P(S \geq 7624) \doteq 0{,}9897146436$

$\underline{n = 12922}$: $P(S \geq 7624) \doteq 0{,}9900021378$

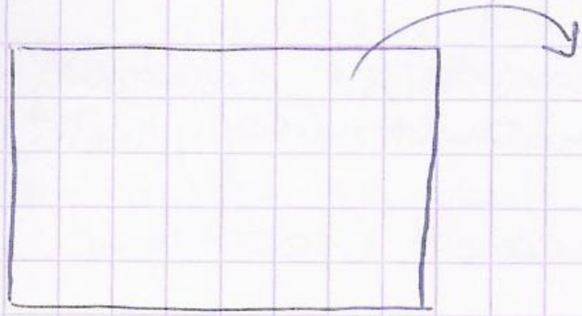
$n = 13095$: $P(S \geq 7727) \doteq 0{,}9899715692$

$\underline{n = 13096}$: $P(S \geq 7727) \doteq 0{,}9902505306$

$\downarrow$ nad 13096 je uredu, umes niha.

OVS AV 5.5.2009

Vzorčenje



Piščanci v populaciji imajo teže v gramih:
900, 920, 930, 950, 950, 980, 1000, 1010*, 1010**, 1050. ~~[illegible]~~ Vzamemo vzorec 2 piščancev.
$P(\bar{X} > 1000)$?

- 1. <u>S ponavljanjem:</u> $10^2 = 100$ možnih vzorcev
  (tukaj moramo pri posameznem vzorcu upoštevati vrstni red)

Prvi: 1050, drugi: 5 piščancev (980, 1000, 1010, 1010, 1050)
Prvi: 1010, drugi: 4 piščanci (1000, 1010, 1010, 1050) $\times 2$
Prvi: 1000, drugi: 3 piščanci (1010, 1010, 1050)
Prvi: 980, drugi: 1 piščanec (1050)

17 vzorcev

$P(\bar{X} > 1000) = \underline{\underline{0{,}17}}$

- 2. Brez ponavljanja: vrstni red ne bo pomemben. → tukaj lahko vrstni red v paramezniem vzorcu zanemarimo.

(težji) → uredimo, ker zanemarimo vrstni red

Prvi: 1050, drugi (lažji): 4 možnosti (980, 1000, 1010, 1010)

Prvi: 1010, drugi: 2 možnosti (1000, 1010*)

Prvi*: 1010*, drugi: 1 možnost (1000)

Ugodnih vzorcev je 7.

Vseh vzorcev je $\binom{10}{2} = \frac{10 \cdot 9}{2} = 45$

$P(\bar{X} > 1000) = \frac{7}{45} = \underline{\underline{0.156}}$

- Izračunaj 95. percentil; ponavljanje.

95. percentil: $P(\bar{X} < 20.95) \leq 0.95$

$P(\bar{X} \leq 20.95) \geq 0.95$

Vzorce uredimo po velikosti glede na vzorčno povprečje.

1: 1050, 1050: 1050
2: 1050, 1010: 1030
3: 1050, 1010: 1030
4: 1010, 1050: 1030
5: 1010, 1050: 1030
6: 1050, 1000: 1025

1050, 980

~~1015~~ ~~1025~~ 1025 | 1030 1030 1030 1030 | 1050

0.93 0.95 0.99

100-5

~~Pa~~ Karkoli z intervala [1025, 1030] je 95. percentil.

↳ interval za dani percentil je vedno zaprt.

<u>1030</u> je 95 percentil, ker je:

$P(\overline{X} < 1030) = 0.95$ ✓

$P(\overline{X} \leq 1030) = 0.99 \geq 0.95$ ✓

- <u>Izračunaj 95. percentil za vzorce brez ponavljanja!</u>

Vrstni red ne bo pomemben, vzorce razvrstimo po velikosti

1050, 1010 : 1030
1050, 1010 : 1030
1050, 1000 : 1025
1050, 980 : 1015

... 1015 | 1025 | 1030 1030

$\frac{42}{45}$ $\frac{43}{45}$

0.9333 0.956

Edini možni 95. percentil je <u>1025</u>.

Velika populacija.

$$X \sim \begin{pmatrix} 1 & 2 & 4 \\ 1/2 & 1/4 & 1/4 \end{pmatrix}$$

Iz populacije vzamemo enostavni slučajni vzorec velikosti 100.

P(vzorčno povprečje odstopa od populacijskega za manj kot 0'1) = ?

vzorčno povprečje: $\overline{X} = \frac{X_1+X_2+\dots+X_{100}}{100}$

populacijsko povprečje: $E(X) = 2$

$$P(1'9 < \overline{X} < 2'1) \approx$$

$$\approx \Phi\left(\frac{2,1-\mu}{\sigma(\overline{X})}\right) - \Phi\left(\frac{1,9-\mu}{\sigma(\overline{X})}\right)$$

$\uparrow$

CLI

$$\sigma(\overline{X}) = \sqrt{D(X)}$$

$$D(\overline{X}) = \frac{D(X_1+X_2+\dots+X_{100})}{100^2}$$

neodvisnost $\rightarrow$ $= \frac{D(X_1)+D(X_2)+\dots+D(X_{100})}{100^2}$

$$= \frac{100 \cdot D(X)}{100^2} = \frac{D(X)}{100} = 0'015$$

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

$$= \left(1\cdot\frac{1}{2} + 4\cdot\frac{1}{4} + 16\cdot\frac{1}{4}\right) - 2^2 = 1'5$$

$$P(1,9 < \overline{X} < 2,1) \approx \Phi\left(\frac{0'1}{\sqrt{0'015}}\right) - \Phi\left(\frac{-0'1}{\sqrt{0'015}}\right) =$$

$$= 2\cdot\Phi\left(\frac{0'1}{\sqrt{0'015}}\right) \doteq \underline{\underline{0'5878}}$$

Točen rezultat: 0.56133.

Popravek: računamo ~~vez~~ $P(1.905 < \overline{X} < 2.095)$
≈ 0.56206

Izračunaj 95. percentil za vzorčno povprečje!

$$P(\overline{X} < q) = 0.95$$

$$\Phi\left(\frac{q-\mu}{\sigma(\overline{X})}\right) - \Phi(-\infty) = \Phi\left(\frac{q-2}{\sqrt{0.015}}\right) + \frac{1}{2} =$$
$$= 0.95$$

$$\Phi\left(\frac{q-2}{\sqrt{0.015}}\right) = 0.95 - \frac{1}{2} = 0.45$$

$$\frac{q-2}{\sqrt{0.015}} = 1.645 \quad / \cdot \sqrt{0.015}$$

$$q - 2 = 1.645 \cdot \sqrt{0.015}$$

$$q = 1.645 \cdot \sqrt{0.015} + 2 = \underline{\underline{2.201}}$$

V resnici je edini ~~možni~~ 95. percentil enak 2.2.

$$\ldots \quad 2.19 \;\big|\; 2.2 \;\big|\; 2.21 \quad \ldots$$

$0.94\ldots$ $0.95\ldots$

belih = 60%

vzorec = 50 cvetov

P(delež belih odstopa od populacijskega za manj kot absolutnih 5%)?

Delež v populaciji: $p = 0.6$

Delež v vzorcu: $\hat{p} = \frac{B}{50}$ → št. belih v vzorcu

znak za oceno

$B \sim B(50, 0.6)$

$E(B) = 50 \cdot 0.6$

$\rightarrow E(\hat{p}) = \underline{0.6}$

$D(B) = n \cdot p \cdot (1-p) = 50 \cdot 0.6 \cdot 0.4$

$D(\hat{p}) = \frac{0.6 \cdot 0.4}{50^2} = 0.0048$

$\sigma(\hat{p}) = \sqrt{0.0048}$

$P(|\hat{p} - p| < 0.05) =$

$= P(0.55 < \hat{p} < 0.65) \approx$

$\approx \Phi\left(\frac{0.65 - 0.6}{\sqrt{0.0048}}\right) - \Phi\left(\frac{0.55 - 0.6}{\sqrt{0.0048}}\right) =$

$2 \cdot \Phi(0.72) =$

$= 0.2642 \cdot 2 = \underline{0.5284}$

Točen rezultat: $\underline{\underline{0.52914}}$

<u>2. možnost:</u> računamo

$P(50 \cdot 0.55 < B < 50 \cdot 0.65)$

→ DN: izračunaj 95% percentil.

OVS AV 19.5.2009

P_θ statistični model

STATISTIČNO SKLEPANJE

Intervali zaupanja

Zanima nas neznan parameter a.
Interval zaupanja: $a_{min} < a < a_{max}$

morata se dati izračunati iz naših opažanj

$$\min P(a_{min} < a < a_{max}) \geq \beta$$

β... stopnja zaupanja

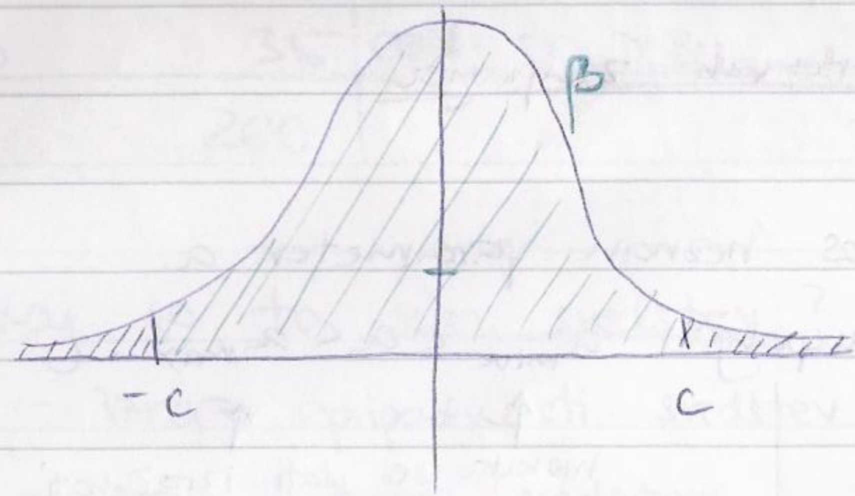
Tipično $\beta = 0'9, 0'95, 0'99$

Interval zaupanja za delež

θ delež enot v populaciji, ki imajo določeno lastnost. Vzamemo vzorec $n > 30$ enot. Naj jih ima k našo lastnost.

- $\hat{\Theta} = \frac{k}{n}$ – točkasta ocena za Θ
- Poiščemo c, za katerega velja $\phi(c) = \frac{\beta}{2}$.

$c = z_{\frac{1+\beta}{2}} = t_{\frac{1+\beta}{2}}$ pri $df = \infty$ (tabela).



- $\Delta = c\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$

- Približni interval zaupanja za Θ:

$\hat{\Theta} - \Delta < \Theta < \hat{\Theta} + \Delta$

zaokrožimo navzdol (za $\hat{\Theta} - \Delta$)

zaokrožimo navzgor (za $\hat{\Theta} + \Delta$)

(40/100)

Janez je dobil 40% glasov na volitvah. Izračunajte 95% interval zaupanja!

$n = 100$

$k = 40$ • $\hat{\Theta} = \frac{k}{n} = \frac{40}{100} = \underline{0{,}4}$

$\beta = 0{,}95$

• Poiščemo c:

$*\ \phi(c) = \frac{0{,}95}{2} = 0{,}475$

$\underline{c = 1{,}96}$ } 1. način

$*$ 2. način: $c = t_{\frac{1+0{,}95}{2}} = t_{0{,}975} = \underline{\underline{1{,}96}}$

$\cancel{\frac{1+0{,}95}{2} = 0{,}975}$

• $\Delta = 1{,}96 \sqrt{\frac{0{,}4(1-0{,}4)}{100}} = 0{,}096\ldots = 0{,}097$ → zaokrožimo navzgor ↓ „razširimo interval"

• $0{,}4 - 0{,}097 < \Theta < 0{,}4 + 0{,}097$

$0{,}303 < \Theta < 0{,}497$ → naš interval zaupanja

$n = 10000$

št. grbov $= 5.048$

$\beta = 0\text{'}9$

Določi interval zaupanja.

- $\hat{\Theta} = \frac{k}{n} = \frac{5048}{10000} = \underline{0\text{'}5048}$

- $\Phi(c) = \frac{\beta}{2} = 0\text{'}45$

  $\underline{c = 1\text{'}645}$

- $\Delta = c\sqrt{\frac{0\text{'}5048 \cdot 0\text{'}4952}{10000}} = 0\text{'}0082$

- Interval zaupanja:

  $0\text{'}5048 - 0\text{'}0082 < \Theta < 0\text{'}5048 + 0\text{'}0082$

  $$\boxed{0\text{'}4966 < \Theta < 0\text{'}5130}$$

Interval zaupanja za μ pri $N(\mu, \sigma)$

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

σ je znan, μ pa ne.

Vzamemo vzorec $X_1, X_2, \ldots, X_n$

- $\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$ - ocena za μ
- $c = z_{\frac{1+\beta}{2}}$ $\left(= t_{\frac{1+\beta}{2}} \text{ pri } df = \infty\right)$
- $\Delta = \frac{c\sigma}{\sqrt{n}}$

- Interval zaupanja: $\overline{X} - \Delta < \mu < \overline{X} + \Delta$

101, 91, 93, 103, 91, 101, 103, 95, 95

vzorec iz $N(\mu, 5)$

95% interval zaupanja za μ?

- $\overline{X} = \frac{101 + 91 + \cdots + 95}{9} = \underline{97}$
- $c = z_{\frac{1+0{,}95}{2}} = \underline{1{,}96}$
- $\Delta = \frac{1{,}96 \cdot 5}{\sqrt{9}} = 3{,}27$ (zaokrožimo navzgor)
- Interval: $\boxed{93{,}73 < \mu < 100{,}27}$

σ navadno ni znan, zato ga ocenimo.

- $\underline{S} = \sqrt{\frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \ldots + (x_n-\bar{x})^2}{n-1}}$

- $c = t_{\frac{1+\beta}{2}}$ pri $df = \underline{n-1}$

- $\Delta = \frac{c \cdot S}{\sqrt{n}}$

- Interval zaupanja: $\bar{x} - \Delta < \mu < \bar{x} + \Delta$

Recimo, da σ ni znan v prejšnji nalogi - ocenimo ga.

- $S = \sqrt{\frac{(101-97)^2 + (91-97)^2 + \ldots + (95-97)^2}{8}} = \underline{5}$

- $c = t_{\frac{1+0{,}95}{2}} = t_{0{,}975} = \underline{2{,}31}$

- $\Delta = \frac{2{,}31 \cdot 5}{\sqrt{9}} = \frac{11{,}55}{3} = \underline{3{,}85}$

- Interval: $97 - 3{,}85 < \mu < 97 + 3{,}85$

$$93{,}15 < \mu < 100{,}85$$

Interval zaupanja za σ

- $S = \sqrt{\frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \dots + (x_n-\bar{x})^2}{n-1}}$

$$\left.\begin{aligned} c_1 &= \chi^2_{\frac{1-\beta}{2}} \\ c_2 &= \chi^2_{\frac{1+\beta}{2}} \end{aligned}\right\} df = n-1$$

- Interval zaupanja za σ:

- $S\sqrt{\frac{n-1}{c_2}} < \sigma < S\sqrt{\frac{n-1}{c_1}}$

oziroma

- $\sqrt{\frac{1}{c_2}\sum_{i=1}^{n}(x_i-\bar{x})^2} < \sigma < \sqrt{\frac{1}{c_1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$

124, 129, 126, 122, 124

σ na okoi 3

- točkasta ocena - s?
- 90% interval zaupanja

- $\overline{X} = \frac{124+129+126+122+124}{5} = 125$

- $S = \sqrt{\frac{1+16+1+9+1}{4}} = \sqrt{\frac{28}{4}} = 2,646$

- $c_1 = \chi^2_{\frac{1-\beta}{2}} = \chi^2_{0.05} = 0.711$

- $c_2 = \chi^2_{\frac{1+\beta}{2}} = \cancel{0.711}\ \chi^2_{0.95} = 9.49$

- interval: $2.646 \cdot \sqrt{\frac{4}{9.49}} \leq \sigma \leq 2.646 \cdot \sqrt{\frac{4}{0.711}}$

$$1.72 < \sigma < 6.28$$

PON: 25.5.2009 od 1h do 2h

TOR: 26.5.2009

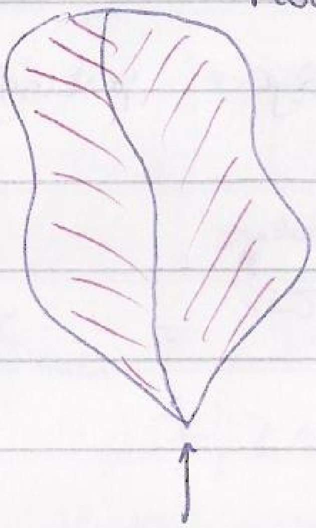
TOR: 2.6.2009

OVS AV 25. 5. 2009

Kolokvij: 9. junij 2009

Testi značilnosti

Model



H_0 - ničelna hipoteza
H_1 - alternativna hipoteza

Možna statistična sklepa:

* H_0 zavrnemo
* $\emptyset$

$\max P(H_0 \text{ zavrnemo, čeprav velja}) \leq \alpha$

α ... stopnja značilnosti

$\alpha = 0'05$: H_0 zavrnemo = odstopanja so statistično značilna

$\alpha = 0'01$: H_0 zavrnemo = odstopanja so statistično zelo značilna

Postopek:

1. Testiranje deleža / verjetnosti uspeha poskusa

- $p :=$ delež v celi populaciji / verjetnost uspeha poskusa
- $H_0 : p = p_0$
- Vzamemo vzorec velikosti n / izvedemo n neodvisnih poskusov.

Naše opažanje: k enot v vzorcu ustreza / k poskusov je uspelo

$n > 30$, $H_0 \Rightarrow Z = \sqrt{\frac{n}{p_0(1-p_0)}}\left(\frac{k}{n} - p_0\right) \sim N(0,1)$.

- $H_1 : p \neq p_0 : K_\alpha = (-\infty, -z_{1-\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, \infty)$.

$H_1 : p < p_0 : K_\alpha = (-\infty, -z_{1-\alpha})$

$H_1 : p > p_0 : K_\alpha = (z_{1-\alpha}, \infty)$

$\hookrightarrow \phi(z_{1-\alpha}) = \frac{1}{2} - \alpha$

$\hookrightarrow z_{1-\alpha} = t_{1-\alpha}$ pri $df = \infty$.

H_0 zavrnemo, če $Z \in K_\alpha$.

Tovarna jamči, da je delež izdelkov z napako enak 20%. Vzamemo vzorec 100 izdelkov, 23 jih ima napako..

$p_0 = 20\%$
$k = 23$

$n = 100$

$H_0: p = 0.2$
$H_1: p > 0.2$

$\alpha = 0.05$ → Hipotezo po krivici zavrnemo v enem od 20-ih primerov.

$$Z = \sqrt{\frac{100}{0.2 \cdot 0.8}}\left(\frac{23}{100} - 0.2\right) =$$
$$= \sqrt{\frac{100}{0.16}} \cdot 0.03 =$$
$$= 25 \cdot 0.03 =$$
$$= \underline{0.75}$$

$z_{0.95} = t_{0.95}$, $df = \infty$
$= 1.645$

Kritično območje za test: $K_{0.05} = (1.645, \infty)$
→ ni kritična. ⇒ Odstopanja niso statistično značilna. Ne navedimo ničesar.

Če vzamemo vzorec 1000 izdelkov, 230 napak. → Bolj sumljivo, kot 23 od 100, ker je vzorec bolj reprezentativen.

Variacija: $n = 1000$, $k = 230$

$$Z = \sqrt{\frac{1000}{0{,}2 \cdot 0{,}8}} \left(\frac{230}{1000} - 0{,}2\right) =$$

$$= \sqrt{\frac{1000}{0{,}16}} \left(0{,}23 - 0{,}2\right) =$$

$= \ldots$ poveča se za faktor $\sqrt{10}$.

$\sqrt{10} \cdot 0{,}75 = \underline{2{,}37 = Z} \in K_{0{,}05}$

↳ H_0 zavrnemo (odstopanja so značilna).

$K_{0{,}01} = (2{,}33, \infty)$ → odstopanja so zelo značilna.

* Če je „p vrednost" manjša od 0,05, so odstopanja značilna, če manjša od 0,01, so zelo značilna.

2. Test sredine

$$\text{Model: } X \sim (\underset{\text{ne poznamo}}{\mu}, \underset{\text{poznamo}}{\sigma_0}) \text{ — porazdelitev na populaciji}$$

Vzamemo vzorec: $X_1, X_2, \dots, X_n$

Testiramo $H_0: \mu = \mu_0$

$$\overline{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

$$Z := \frac{\overline{X} - \mu_0}{\sigma_0}\sqrt{n} \overset{H_0}{\sim} N(0,1)$$

H_0 zavrnemo, če $Z \in K_\alpha$.

$H_1: \mu \neq \mu_0 : K_\alpha = (-\infty, -z_{1-\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, \infty)$

$H_1: \mu < \mu_0 : K_\alpha = (-\infty, -z_{1-\alpha})$

$H_1: \mu > \mu_0 : K_\alpha = (z_{1-\alpha}, \infty)$

$\phi(z_{1-\alpha}) = \frac{1}{2} - \alpha$

$z_{1-\alpha} = t_{1-\alpha}$ pri $df = \infty$.

Model $X \sim N(\mu, 5)$

101, 91, 93, 103, 91, 101, 103, 95, 95

$H_0 \dots \mu = 100$

$H_1 \dots \mu \neq 100$

$\alpha = 0.05$

$\overline{X} = 97$

$Z = \frac{97-100}{5} \cdot \sqrt{9} = \underline{-1.8}$

poišči

Kritično območje.

$1 - \frac{0.05}{2} = 0.975$

$z_{0.975} = 1.96$

$K_\alpha = (-\infty, -1.96] \cup (1.96, \infty)$

→ Odstopanja niso statistično značilna.

OVS AV 26.5.2009

2. kolokvij: TOREK, 9.6. ob 11. uri v PR15

Test populacijskega povprečja pri neznanem σ
(model: $X \sim N(\mu, \sigma)$)

$H_0: \mu = \mu_0$

$X_1, X_2, \dots, X_n$ i $\overline{X} = \frac{X_1 + \dots + X_n}{n}$

$S = \sqrt{\frac{(X_1-\overline{X})^2 + \dots + (X_n-\overline{X})^2}{n-1}}$

$T = \frac{\overline{X} - \mu_0}{S}\sqrt{n} \underset{H_0}{\sim} \text{Student}(n-1)$

$H_1: \mu \neq \mu_0: K_\alpha = (-\infty, -t_{1-\frac{\alpha}{2}}) \cup (t_{1-\frac{\alpha}{2}}, \infty)$

$H_1: \mu > \mu_0: K_\alpha = (t_{1-\alpha}, \infty)$

$H_1: \mu < \mu_0: K_\alpha = (-\infty, -t_{1-\alpha})$

$df = n-1$

109, 107, 95, 121, 107, 97, 111, 121, 95

- vzorec iz $N(\mu, \sigma)$

$H_0: \mu = 100$

$H_1: \mu \neq 100$

$\alpha = 0.05$

$\overline{X} = 107$

$s = 10$

$T = \frac{107-100}{10}\sqrt{9} = \frac{21}{10} = 2.1$

$df = 8$, $t_{0.975} = 2.31$

Hipoteze ne moremo ne zavrniti, ne sprejeti

Odstopanja niso statistično značilna.

- Kaj pa, če bi vedeli, da je $\sigma = 10$?

$Z = 2.1$

$K_\alpha = (-\infty, -1.96) \cup (1.96, \infty)$

H_0 bi zavrnili.

Testiramo zdravilo proti nespečnosti. Koliko so spali prej, ~~k~~ koliko po tem, ko so vzeli tablete?

PRED: 4·7, 4·5, 5·2, 4·1, 3·9, 4·4, 4·5, 4·1, 4·2, 4·8

PO ZDRAVLJENJU: 5·5, 4·3, 5·6, 4·9, 3·5, 5·0, 4·3, 5·3, 4·8, 5·2

H_0 : zdravilo ne deluje (isto kot prej)

H_1 : zdravilo podaljša spanec → enostranski test

RAZLIKA: 0·8, −0·2, 0·4, 0·8, −0·4, 0·6, −0·2, 1·2, 0·6, 0·4

$\alpha = 0{\cdot}05$

$\Delta = 10 \cdot$ (po zdravljenju − pred zdr.)

$\bar{\Delta} = \frac{40}{10} = 4$

$S = \sqrt{\ldots} \quad 5{\cdot}16$

$T = \frac{4-0}{5{\cdot}16}\sqrt{10} = 2{\cdot}45$

$df = 9, \; t_{0{\cdot}95} = 1{\cdot}83$

$K_{0{\cdot}05} = (1{\cdot}83, \infty)$

Statističen sklep: sprejmemo H_1, H_0 pa zavrnemo.

Test enakosti sredin različnih populacij

$X \sim N(\mu_1, \sigma)$
$Y \sim N(\mu_2, \sigma)$ → homoskedastičnost (enaki sigmi)

(Heteroskedastična varianta: $N(\mu_1, \sigma_1)$ in $N(\mu_2, \sigma_2)$)

Vzamemo 2 vzorca:

$X_1, \dots, X_m \to \overline{X}$

$Y_1, \dots, Y_n \to \overline{Y}$

$$S = \sqrt{\frac{(X_1-\overline{X})^2 + \dots + (X_m-\overline{X})^2 + (Y_1-\overline{Y})^2 + \dots + (Y_n-\overline{Y})^2}{m+n-2}}$$

$$T = \frac{\overline{X}-\overline{Y}}{S}\sqrt{\frac{m \cdot n}{m+n}} \sim \text{Student}(m+n-2)$$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2: K_\alpha = (-\infty, -t_{1-\frac{\alpha}{2}}) \cup (t_{1-\frac{\alpha}{2}}, \infty)$

$H_1: \mu_1 > \mu_2: K_\alpha = (t_{1-\alpha}, \infty)$

$H_1: \mu_1 < \mu_2: K_\alpha = (-\infty, -t_{1-\alpha})$

$$S = \sqrt{\frac{(m-1) \cdot S_X^2 + (n-1) \cdot S_Y^2}{m+n-2}} \qquad df = m+n-2$$

Pred tablo !

Inf : 25, 16, 23, 17, 22, 18, 18, 21, 20

Po : 19, 21, 23, 21, 25, 21, 24

$\alpha = 0{\cdot}05$

Dvostranski test

H_0 : oboji enako dobri

H_1 : niso enako dobri

$\overline{X} = 20$ $\qquad m = 9$

$\overline{Y} = 22$ $\qquad n = 7$

$S = 2{\cdot}65$

$T = -1{\cdot}5$

$df = 14$, $t_{0{\cdot}975} = 2{\cdot}14$

Ker je dvostranski test :

$K_{0{\cdot}05} = (-\infty, -2{\cdot}14) \cup (2{\cdot}14, \infty)$.

t ne spada v kritično območje ⇒

H_0 zavrnemo.

Test z znaki

X in Y, definirani na isti populaciji.

$H_0: P(X>Y) = P(X<Y)$ (na populaciji)

$S^+ :=$ število enot, na katerih je $X>Y$

$S^- :=$ število enot, na katerih je $X<Y$.

$$Z = \frac{S^+ - S^-}{\sqrt{S^+ + S^-}} \overset{\cdot}{\underset{\underset{H_0}{\uparrow}}{\sim}} N(0,1)$$

$H_1: P(X>Y) \neq P(X<Y)$:

$K_\alpha = (-\infty, -z_{1-\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, \infty)$

$H_1: P(X>Y) > P(X<Y)$;

$K_\alpha = (z_{1-\alpha}, \infty)$

$H_1: P(X>Y) < P(X<Y)$:

$K_\alpha = (-\infty, -z_{1-\alpha})$

2 kolokvija: X prvi, Y drugi.

X: 66, 37, 63, 70, 76, 62, 97, 51, 14, 83, 65, 95, 68, 87, 94

Y: 24, 51, 21, 51, 6, 77, 5, 48, 100, 31, 23, 61, 61, 74, 76

\+ − + + + − + + − + + + + + +

X: 62, 68, 62, 65, 83, 19, 61, 45, 70,
24, 52, 81, 80, ~~22~~, 18, 81

Y: 97, 22, 16, 20, 18, 34, 54, 86, 22,
78, 15, 27, 35, 39, ~~42~~ 79

− + + + + + − − + − − + − + + − +

$\alpha = 0.05$

H_0: na obeh enak odziv

H_{alt}: različno težak kolokvij

$S^- = 8$, $S^+ = 22$

$z = \frac{14}{\sqrt{30}} = 2.56$

$df = \infty$

1.96

H_0 zavrnemo.

Wilcoxon - Mann - Whitneyev test

Primerjamo urejenostni spr. X in Y na dveh populacijah.

H_0: X in Y sta enako porazdeljeni.

Vzamemo vzorca velikosti m in n, ju združimo in elemente uredimo po velikosti.

$R_1, R_2, \ldots, R_m$ - mesta (rangi) enot prvega vzorca.

$$z := \sqrt{\frac{3}{mn(m+n+1)}} \cdot \left[2 \cdot \sum_{i=1}^{m} R_i - m(m+n+1)\right] \underset{H_0}{\sim} N(0,1)$$

H_1: X in Y sta porazdeljeni različno.

$K_\alpha = (-\infty, -z_{1-\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, \infty) \qquad df = \infty$

Tek, programska op. in inf.

~~Rhirik pritale na lciy CP, salugi~~

P, P, î, P, P, î, P, P, î, P, î, î, î, P, î, î, î, î, P, î

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20.

H_0: oboji enako dobri $\qquad \alpha = 0{,}05$

H_1: niso enako dobri

P: $m = 9$

î: $n = 11$

$$\sum_{i=1}^{9} R_i = 1 + 2 + 4 + 5 + 7 + 8 + 10 + 14 + 19 = 70$$

$$Z = \sqrt{\frac{3}{99(9+11+1)}} \cdot \left[2 \cdot 70 - 9 \cdot 21\right] =$$

$$= -1{,}86$$

$$K_{0{,}05} = (-\infty, -\cancel{1{,}86}\ 1{,}96) \cup (\cancel{1{,}86}\ 1{,}96, \infty)$$

Hipotezo zavrnemo; ni dovolj značilna.

OVS AV 9.6.2009

Kolokvij! 23.6.2009

Test χ^2

① Mečemo kocko:

m vrednosti

k	1	2	3	4	5	6
N_k	12	14	17	8	5	4
$\tilde{N}_k$	10	10	10	10	10	10

$\Sigma = 60$
št. posameznih pojavitev rezultata k

Testiramo: ali je kocka poštena?

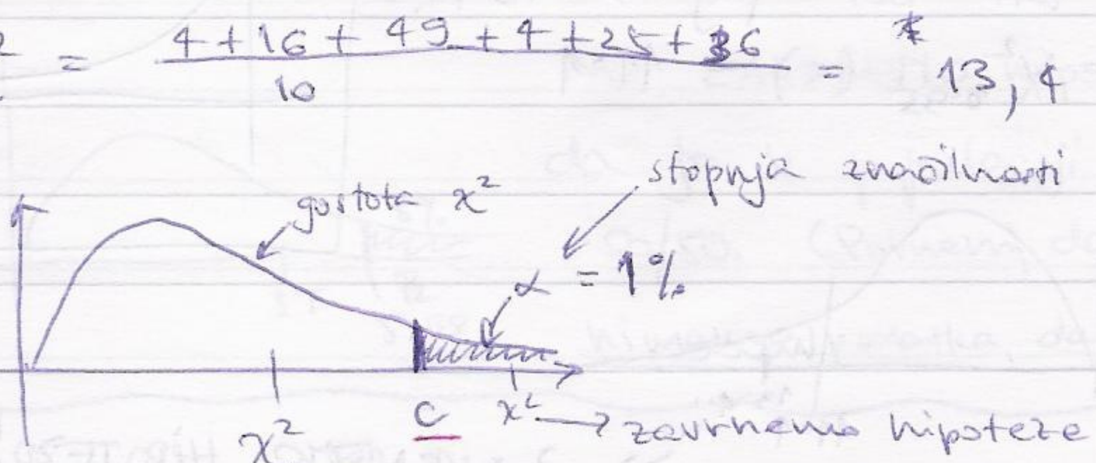
poštena: $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$.

$\tilde{N}_k = n \cdot p_k$

teoretična vrednost

$$\boxed{\chi^2 = \sum \frac{(N_k - \tilde{N}_k)^2}{\tilde{N}_k}}$$ → definiran samo za pozitivne

$$\chi^2 = \frac{4+16+49+4+25+36}{10} = 13,4$$

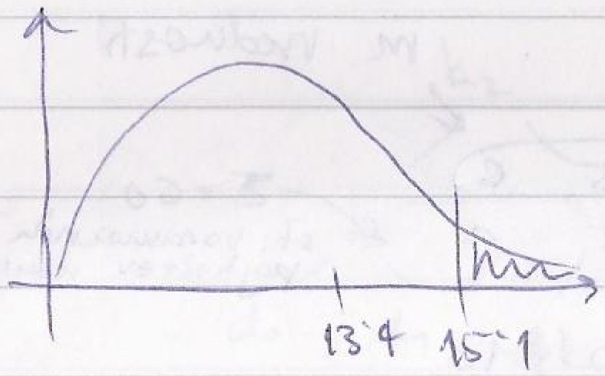


podemics
št. vrednosti

Tabela χ^2:

$$\boxed{c = \chi^2_{1-\alpha}(m-1)}$$

$$\chi^2_{0.99}(5) = 15'1$$



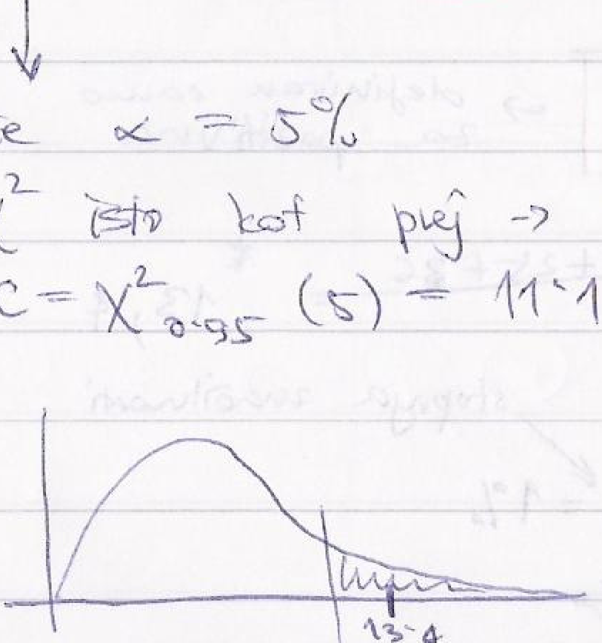
NE ZAVRNEMO! ~~trdila~~
da je kocka poštena.

Ne trdimo, da je kocka poštena, ker je premalo podatkov.

če $\alpha = 5\%$

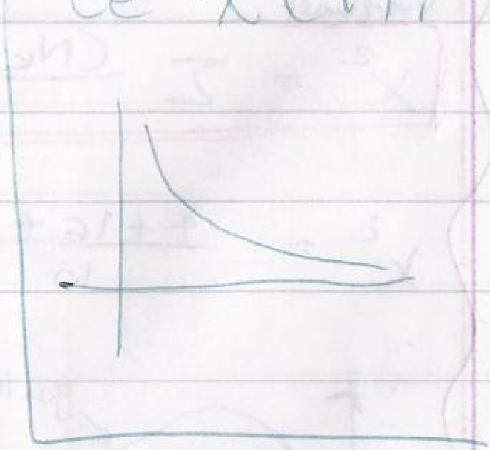
χ^2 isto kot prej → 13'4.

$$c = \chi^2_{0.95}(5) = 11'1$$



→ ZAVRNEMO HIPOTEZO!

če $\chi^2(1)$,



(2.) Vzorci 20 osebkov tipa RR

45 x Rr

35 x rr

H_0: V populaciji je 50%, 50% r genov.

	RR	Rr	rr
N_k	20	45	35
$\tilde{N}_k$	25	50	25

↑ teoretični

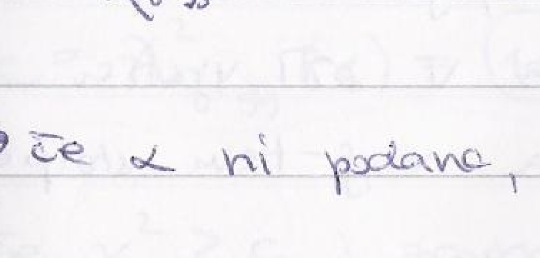
$$\begin{pmatrix} RR & Rr & rr \\ 1/4 & 1/2 & 1/4 \end{pmatrix}$$ → podobno kot met 2 kovancev

$$\chi^2 = \frac{25}{25} + \frac{25}{50} + \frac{100}{25} = 5'5$$

$\alpha = 0'05$

$c = \chi^2_{0'95}(2) = 5'99$

→ če α ni podana, si jo zmislimo (1% ali 5%)



NE ZAVRNEMO hipoteze, da je v populaciji 50/50. (Pomeni da nimamo podatka, da ni tako.)

TEST NEODVISNOSTI

(3) 100 oseb „klasificiramo" (barva oči, barva las).

lasje \ oči	modra		zelena		rjava		
blond	15	6'25	8	12'5	2	6'25	25 ↗ če bi bilo neodvisni
rdeči	5	2'5	2	5	3	2'5	10
rjavi	3	10	30	20	7	10	40
črni	2	6'25	10	12'5	13	6'25	25
	25		50		25		100

$\alpha = 1\%$

Testiramo: H_0: lasje in oči (barva) sta neodvisni.

N_{ij} ... št. v i-ti vrstici, j-tem stolpcu

$N_{i\cdot} = \sum_j N_{ij}$ ... vsota v i-ti vrstici

$N_{\cdot j} = \sum_i N_{ij}$ ... vsota v j-tem stolpcu

$n = \sum_{i,j} N_{ij}$

$100 \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{100}{16}$ (*)

$\frac{1}{4} = \frac{25}{100}$

$$\boxed{\chi^2 = \sum_i \sum_j \frac{\left(N_{ij} - \frac{N_{i\cdot} \cdot N_{\cdot j}}{n}\right)^2}{\frac{N_{i\cdot} \cdot N_{\cdot j}}{n}}}$$

$$\chi^2 = \boxed{\frac{8.75^2}{6.25}} + \frac{4.5^2}{12.5} + \frac{4.25^2}{6.25} +$$

$$+ \frac{2.5^2}{2.5} + \frac{3^2}{5} + \frac{0.5^2}{2.5} +$$

$$\boxed{+ \frac{7^2}{10}} + \boxed{\frac{10^2}{20}} + \frac{3^2}{10} +$$

$$+ \frac{4.25^2}{6.25} + \frac{2.5^2}{12.5} + \boxed{\frac{6.75^2}{6.25}} > 8+5+6 = 19$$

$$= \underline{42.6}$$

$$c = \chi^2_{1-\alpha}((r-1)(s-1))$$

r: št. vrstic, s: št. stolpcev

$$c = \chi^2_{0.99}(6) = 16.8$$

Če $\chi^2 > c$: zavrnemo

$\chi^2 > c \Rightarrow$ hipotezo zavrnemo!

Barvi oči in las **ništa neodvisni!!**

100 semen, vzklije 80. Poiščite 95% interval za verjetnost, da seme vzklije.

$\hat{p} = \frac{80}{100} = 0.8$

$\Delta = \frac{c \cdot \sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} = \frac{1.96 \cdot \sqrt{0.8 \cdot 0.2}}{\sqrt{100}} \doteq \underline{0.08}$

za c: $\phi(c) = \beta/2$

$\phi(c) = 0.475$

$c = 1.96$

Interval zaupanja:

$[\hat{p} - \Delta, \hat{p} + \Delta]$

$[0.72, 0.88]$

Z verjetnostjo $\beta = 0.95$ lahko trdimo, da je verjetnost vzklitja semena med 72% in 88%.

X, $g_X(x) = \begin{cases} e^{-x} ; & x > 0 \\ 0 ; & \text{sicer} \end{cases}$

99. centil?

$\Rightarrow P(X < c) = 0.99$

$P(X < c) = \int_0^c e^{-x}\,dx = -e^{-x}\Big|_0^c =$

$= 1 - e^{-c} = \underline{0.99}$

$0.01 = e^{-c}$ $100 = e^{c}$ $\ln 100 = c$

10000 hmčk cepimo, cepič posuši z verjetnostjo 0.15. Kolikšna je verjetnost da se posuši več kot 1600 cepičev?

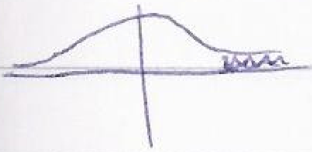
$n = 10000$

$p = 0.15$

$X \sim Bin(n, p) \approx N(n \cdot p, \sqrt{n \cdot p \cdot (1-p)})$

$P(X > 1600)$?

$\hookrightarrow \approx \frac{1}{2} - \phi\left(\frac{1600 - 10000 \cdot 0.15}{\sqrt{10000 \cdot 0.15 \cdot 0.85}}\right) =$



$= \frac{1}{2} - \phi\left(\frac{100}{100 \cdot \sqrt{0.15 \cdot 0.85}}\right) =$

$= \frac{1}{2} - \phi(2.80) = 0.5000 -$

$- 0.4974 = \underline{\underline{0.0026}}$

$N(\mu, \sigma)$

Vzorec: 34, 29, 33, 35, 31, 30, 34, 38

$\beta = 95\%$. Interval zaupanja za μ!

$$\bar{X} = \frac{24}{8} + 30 = 33$$

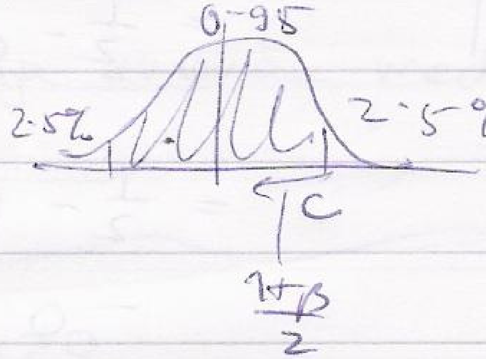
$$s^2 = \frac{(x_1-\bar{x})^2 + \dots + (x_n-\bar{x})^2}{n-1} =$$

$$= \frac{1+16+0+4+4+9+1+25}{7} = \frac{60}{7}$$

$$\Delta = \frac{c \cdot s}{\sqrt{n}} = \frac{2.36 \cdot \sqrt{\frac{60}{7}}}{\sqrt{8}} = \underline{\underline{2.44}}$$

$$c = t_{1+\beta/2}\,(n-1) = t_{0.975}\,(7) = 2.36$$

↑
tabela Student,
ker σ ni znana!



95% interval zaupanja:

$$[\bar{X} - \Delta, \bar{X} + \Delta] = [30.56, 35.44]$$

↑ z verjetnostjo 95% trdimo ...