VJEROJATNOST I STATISTIKA

SAŽETAK BY Cope

- napravljeno po predavanjima A.A.A (prvi dio) i Burića (drugi dio), 2015./2016.
- dodatno: - dokazi koji su se pojavljivali prošlih godina

1. DISKRETNA VJEROJATNOST

1. VJEROJATNOST

1.1. ALGEBRA DOGAĐAJA

→ def. SLUČAJNI POKUS – svaki pokus kojem unaprijed ne znamo ishod

→ def. ELEMENTARNI DOGAĐAJ – ishod pokusa ($\omega_1, \omega_2, \omega_3, \dots$)

Ω → skup svih elem. događaja $|\Omega|$ – KARDINALAN BROJ skupa – broj elemenata

→ def. DOGAĐAJ – podskup od Ω (oznaka $A, B, C, D, \dots$)

UNIJA: $A \cup B$ ili $A + B$ PRESJEK: $A \cap B$ ili $A \cdot B$ RAZLIKA: $B \setminus A$ ili $B - A$

KOMPLEMENT: $\bar{B}$ ili $B^c = \Omega \setminus B$

$\boxed{A \setminus B = A \cap \bar{B}}$

DEMORGAN: $\boxed{\begin{aligned} \overline{A \cup B} &= \bar{A} \cap \bar{B} \\ \overline{A \cap B} &= \bar{A} \cup \bar{B} \end{aligned}}$ → DOKAZIĆ: $\omega \in \overline{A \cup B}$

$\omega \notin A \cup B \Rightarrow \omega \notin A$ i $\omega \notin B \Rightarrow$

$\Rightarrow \omega \in \bar{A}$ i $\omega \in \bar{B} \Rightarrow \omega \in \bar{A} \cap \bar{B}$

→ ALGEBRA DOGAĐAJA – familija F podskupova od Ω:

i) $\Omega \in F$, $\emptyset \in F$

ii) $A \in F \Rightarrow \bar{A} \in F$

iii) $A, B \in F \Rightarrow A + B \in F$

1.2. VJEROJATNOST

→ def. VJEROJATNOST – preslikavanje $P : F \to [0,1]$ def. na algebri događaja F koja ima svojstva:

1) $P(\Omega) = 1$, $P(\emptyset) = 0$ → NORMIRANOST

2) ako je $A \subset B \Rightarrow P(A) \leq P(B)$ → MONOTONOST

3) ako su A i B DISJUNKTNI ($A \cap B = \emptyset$), tada $\boxed{P(A \cup B) = P(A) + P(B)}$ → ADITIVNOST

→ VJER. KOMPLEMENTA: $\boxed{P(\bar{A}) = 1 - P(A)}$ DOKAZIĆ: $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$

$\Rightarrow P(\bar{A}) = 1 - P(A)$

→ VJER. UNIJE – ako događaji nisu disjunktni $\boxed{P(A\cup B) = P(A) + P(B) - P(A\cap B)}$

1.3. KONAČNI VJER. PROSTOR

def. VJEROJATNOSNI PROSTOR Ω koji posjeduje konačno mnogo događaja nazivamo KONAČNI VJER. PROSTOR

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\} \quad P(\omega_1) = p_1, \dots, P(\omega_n) = p_n$$

$$\boxed{\sum_{i=1}^{n} p_i = 1}$$

→ svi element. događaji su jednako vjerojatni: $p_1 = p_2 = \dots = p_n$

$$\boxed{P(A) = \frac{m}{n} = \frac{\text{broj povoljnih}}{\text{broj ukupnih}}} \qquad \sum p_i = n p = 1 \quad \boxed{p = \frac{1}{n}}$$

→ KOMBINATORIKA

1) PRODUKTNO PRAVILO □□□□ npr. $5 \cdot 4 \cdot 7 \cdot 2$ → npr. faktorijeli ili potencije

2) KOMBINACIJE $\binom{n}{k}$ – odabir k elemenata od n

→ ponavljanje pokusa – pazi: u kojem smo povlačenju izvukli! $\binom{n}{k}$ (bitan redoslijed)

1.4. BESKONAČNI VJER. PROSTOR

def. Algebra događaja F je σ-algebra ako vrijedi: $\boxed{A_1, A_2, \dots \in F \Rightarrow \bigcup_{n=1}^{\infty} A_n \in F}$

– uvjet σ-ADITIVNOSTI: $\boxed{P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)}$ za disj. $A_1, A_2, \dots$

→ svojstvo neprekinute vjerojatnosti:

$$\boxed{A = \bigcup_{n=1}^{\infty} A_n \Rightarrow \lim_{n\to\infty} P(A_n) = P(A)}$$

→ Ω beskonačan → 1. PREBROJIV → $\Omega = \{\omega_1, \omega_2, \dots\}$ → $n \in \mathbb{Z}$

↘ 2. NEPREBROJIV → $\Omega = \mathbb{R}$ → geometrijska

$$\boxed{\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \text{ za } |x| < 1} \to \text{kad suma ide od 1: } \boxed{\sum_{n=1}^{\infty} x^n = \frac{1}{1-x} - 1}$$

$\left(\frac{5}{6}\right)^n < 0{,}1 \;/\; \ln\left(\frac{5}{6}\right)$ → BAZA < 1 → mijenjamo znak nejednakosti !!!

$n > 12{,}63$

1.5. GEOMETRIJSKA VJER.

def. GEOM. VJER. definiramo kao $\boxed{P(A) = \frac{m(A)}{m(\Omega)}}$

-> vjerojatnost točke ili pravca ako je $\Omega \to 2D$ je 0!!!

-> $x^2+ax+b>0 \Rightarrow \boxed{D<0} \Rightarrow \boxed{a^2-4b<0}$

2. UVJETNA VJEROJATNOST

2.1. DEFINICIJA I PRIMJERI

def. Neka je B događaj takav da je $P(B)>0$. UVJETNA VJEROJATNOST događaja A uz uvjet B je $F \to [0,1]$ definiramo s:

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

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

2.2. NEZAVISNOST DOGAĐAJA

def. Za A i B kažemo da su nezavisni ako je:

$$\boxed{P(A) = P(A|B) \quad \text{ili} \quad P(B) = P(B|A)}$$

-> nužan i dovoljan uvjet za nezavisnost: $\boxed{P(A \cdot B) = P(A) \cdot P(B)}$

def. Događaji $A_1, A_2, \dots, A_n$ su nezavisni ako za svaki izbor $A_{i_1}, \dots, A_{i_k}$ vrijedi $P(A_{i_1} \cdot \dots \cdot A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$

2.3. POTPUNA VJEROJATNOST I BAYESOVA FORMULA

-> potpuna vjerojatnost: $\boxed{P(A) = P(H_1) \cdot P(A|H_1) + \dots + P(H_n) \cdot P(A|H_n)}$

-> Bayesova formula: $\boxed{P(H_i|A) = \frac{P(H_i) \cdot P(A|H_i)}{\sum_{i=1}^{n} P(H_i) \cdot P(A|H_i)}}$

3. DISKRETNE SLUČAJNE VARIJABLE

3.1. DEFINICIJE I OSNOVNA SVOJSTVA

def. Preslikavanje $X: \Omega \to S$ je DISK. SLUČ. VAR. ako je za

$\forall x_k \in S$ skup $A_k = \{\omega \in \Omega : X(\omega) = x_k\}$ događaj

$\to \boxed{p_k = P(X = x_k)}$, $p_k > 0$, $\boxed{\sum p_k = 1}$ $\to$ ZAKON RAZDIOBE:

$$X \sim \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_k \dots \\ p_1 & p_2 & p_3 & \dots & p_k \dots \end{pmatrix}$$

def. Slučajne varijable $X, Y : \Omega \to S$ su NEZAVISNE ako ~~za~~ sve $x_k, y_j \in S$ vrijedi:

$$\boxed{P(X = x_k, Y = y_j) = P(X = x_k) \cdot P(Y = y_j)}$$

$\to$ FUNKCIJE SLUČ. VARIJABLI:

Neka je X neka sluč. var, $\Psi : \mathbb{R} \to \mathbb{R}$. Kolika je razdioba od varijable $Y = \Psi(X)$? $\to$ fja. djeluje na x_k, NE i na p_k !!!

3.2. DISKRETNI SLUČ. VEKTORI

def. Razdioba slučajnog vektora (X, Y) se definira sa:

$$\boxed{p_{ij} = P(X = x_i, Y = y_j)}$$

X \ Y	y_1	y_2	...	y_m	
x_1	p_{11}	p_{21}	...	p_{1m}	p_1
x_2	p_{21}	p_{22}	...	p_{2m}	p_2
⋮	⋮	⋮		⋮	⋮
x_n	p_{n1}	p_{n2}	...	p_{nm}	p_n
	q_1	q_2	...	q_m	1

$p_1 \dots p_n$: marg. razdioba od X

$q_1 \dots q_m$: marg. razdioba od Y

$\to \boxed{\sum_i \sum_j p_{ij} = 1}$ $p_{ij} \geq 0$

$\to$ MARGINALNE RAZDIOBE – sume u nekom retku ili stupcu

$\to X$ i Y su nezavisni ako i samo ako $\boxed{p_{ij} = p_i \cdot p_j \text{ za } \forall i, j}$

$\to$ ako je 0 u tablici, sigurno su zavisni!

3.3. NUMERIČKE KARAKTERISTIKE

def. OČEKIVANJE (očekivani ishod) sluč. varijable definiramo kao

$$\boxed{E(X) = \sum_k x_k \cdot p_k}$$

ako konvergira. Ako suma divergira, očekivanje ne postoji.

$$\boxed{\sum_{n=1}^{\infty} x^n = \frac{1}{1-x}} \Big/ ' \Rightarrow \boxed{\sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2}} \Big/ '' \Rightarrow \boxed{\sum_{n=1}^{\infty} n(n-1) x^{n-2} = \frac{2}{(1-x)^3}}$$

→ pazi na prvi član! → ako je član manji od prvog člana nemoguć, onda ništa ne oduzimamo (tj. oduzimamo 0)

TM SVOJSTVO OČEKIVANJA

Neka su X, Y na istom prostoru Ω, a $s, t \in \mathbb{R}$:

i) $\boxed{E(sX + tY) = sE(X) + tE(Y)}$ → LINEARNOST

ii) ako su X, Y nezavisne: $\boxed{E(X \cdot Y) = E(X) \cdot E(Y)}$

DOKAZIĆI:

$$E(s \cdot X) = \sum_k s \cdot x_k \cdot p_k = s \cdot \sum_k x_k \cdot p_k = s \cdot E(X) \quad //$$

$$E(X+Y) = \sum_{jk} (x_j + y_k) p_{jk} = \sum_{jk} x_j p_{jk} + \sum_{jk} y_k p_{jk} = \sum_j x_j \underbrace{\sum_k p_{jk}}_{p_j} + \sum_k y_k \underbrace{\sum_j p_{jk}}_{p_k} =$$

$$= \sum_j x_j p_j + \sum_k y_k p_k = E(X) + E(Y) \quad //$$

$$E(X \cdot Y) = \sum_{jk} x_j \cdot y_k \cdot p_{jk} = \sum_{jk} x_j \cdot y_k \cdot p_j \cdot p_k = \sum_j x_j p_j \sum_k y_k p_k = E(X) \cdot E(Y) \quad //$$

$$\rightarrow \boxed{E(\Psi(X)) = \sum_k \Psi(x_k) p_k}$$

def. ISHODIŠNI MOMENT REDA n od sluč. var. X definira se kao

$$\boxed{E(X^n) = \sum_k x_k^n p_k}$$

def. CENTRALNI MOMENT REDA n definiramo sa

$$\boxed{E\left[(X - E(X))^n\right] = \sum_k (x_k - m_X)^n \cdot p_k} \qquad m_X = E(X)$$

def. DISPERZIJA (VARIJANCA, RASIPANJE) slučajne varijable X je

$$\boxed{D(X) = E\left[(X - E(X))^2\right]}$$

(centralni moment za $n=2$)

$\boxed{D(X) \geq 0}$!

→ mjera raspršenosti

def. STANDARDNA DEVIJACIJA ILI ODSTUPANJE: $\boxed{\sigma = \sqrt{D(X)}}$

→ $\boxed{D(X) = E(X^2) - E(X)^2 = \sum_k x_k^2 p_k - m_X^2}$

DOKAZ / IZVOD:

$$D(X) = E\left[(X - E(X))^2\right] = E\left[X^2 - 2XE(X) + E(X)^2\right] =$$

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

TM SVOJSTVO DISPERZIJE

i) $\boxed{D(sX) = s^2 D(X)}$

ii) ako su X i Y nezavisne tada: $\boxed{D(X+Y) = D(X) + D(Y)}$ $\boxed{D(X-Y) = D(X) + D(Y)}$

DOKAZ

$$D(X+Y) = E\left((X+Y)^2\right) - \left(E(X+Y)\right)^2 = E(X^2) + 2E(X \cdot Y) + E(Y^2) - \left(E(X) + E(Y)\right)^2 =$$

$$= \underbrace{E(X^2) - E(X)^2}_{D(X)} + \underbrace{E(Y^2) - E(Y)^2}_{D(Y)} + \underbrace{2E(XY) - 2E(X)E(Y)}_{=0 \text{ ako su nezavisni}} =$$

$$= D(X) + D(Y)$$

$$\rightarrow \sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2} \Big/ \cdot x \Rightarrow \boxed{\sum_{n=1}^{\infty} n x^n = \frac{x}{(1-x)^2}} \Big/ ' \Rightarrow \boxed{\sum_{n=1}^{\infty} n^2 x^{n-1} = \frac{1+x}{(1-x)^3}} \Big/ \cdot x \Rightarrow$$

$$\Rightarrow \boxed{\sum_{n=1}^{\infty} n^2 x^n = x \cdot \frac{1+x}{(1-x)^3}}$$

*MEĐUSOBNI ODNOS 2 VARIJABLE:

def. KOVARIJACIJSKI MOMENT od X i Y se definira kao

$$\boxed{\mathrm{cov}(X,Y) = E\left[(X - m_X)(Y - m_Y)\right]} \qquad \boxed{\mathrm{cov}(X,Y) = E(XY) - E(X) \cdot E(Y)}$$

⇒ kod zavisnih varijabli!

→ KOEFICIJENT KORELACIJE : $\boxed{r(X,Y) = \dfrac{cov(X,Y)}{\sqrt{D(X)\cdot D(Y)}}}$

→ ako su X i Y nezavisni, tada je $cov(X,Y)=0$, pa je i $r(X,Y)=0$, tj. X i Y NISU KOLINEARNI. Obrat ne vrijedi !

TM DISPERZIJA ZBROJA

$$\boxed{D\left(\sum_i X_i\right) = \sum_i D(X_i) + 2\sum_{i,j} cov(X_i, X_j)}$$

DOKAZ: $D\left(\sum_i X_i\right) = E\left[\sum_i (X_i - m_{X_i})\right]^2 = \sum_i E(X_i - m_{X_i})^2 + \sum_{i\neq j} E\left((X_i - m_{X_i})(X_j - m_{X_j})\right) =$

$$= \sum_i D(X_i) + 2\sum_{i<j} cov(X_i, X_j)$$

→ ako su X_i nezavisni : $\boxed{D\left(\sum_i X_i\right) = \sum_i D(X_i)}$

*CENTRIRANJE I NORMIRANJE VARIJABLI

$$\boxed{\begin{aligned} E(X-a) &= E(X) - a \\ D(X-a) &= D(X) \end{aligned}}$$

→ $D(X)$ se ne mijenja, pa se ne mijenjaju ni $cov(X,Y)$ ni $r(X,Y)$

→ odaberemo $a = m_X$: $\boxed{\dot{X} = X - m_X}$ je CENTRIRANA varijabla

→ $\boxed{X^* = \dfrac{X - m_X}{\sigma_X}}$ je NORMIRANA sluč. varijabla $E(X^*) = 0$, $D(X^*) = 1$

→ ni centriranje ni normiranje ne mijenjaju $r(X,Y)$

DOKAZ: $r(X^*, Y^*) = \dfrac{E(X^* \cdot Y^*) - \overbrace{E(X^*)}^{0}\overbrace{E(Y^*)}^{0}}{\underbrace{\sigma(X^*)}_{=1}\underbrace{\sigma(Y^*)}_{=1}} = \overbrace{E\left(\dfrac{X - m_X}{\sigma_X} \cdot \dfrac{Y - m_Y}{\sigma_Y}\right)}^{cov(X,Y)} =$

$$= \frac{cov(X,Y)}{\sigma_X \sigma_Y} = r(X,Y)$$

TM Vrijedi $\boxed{|r(X,Y)| \leq 1.}$ Jednakost vrijedi samo ako je $y = ax + b$.

DOKAZ:

znamo: $r(X^*, Y^*) = r(X,Y)$

$$D(X^* \pm Y^*) = D(X^*) + D(Y^*) \pm 2\operatorname{cov}(X^*, Y^*) = 2[1 \pm r(X,Y)] \geq 0$$

$$1 \pm r(X,Y) \geq 0$$

$$\pm r(X,Y) \leq 1$$

$$\boxed{|r(X,Y)| \leq 1}$$

//

* KARAKTERISTIČNE FUNKCIJE

def. KARAKTER. FJU. varijable X definiramo s

$$\boxed{\vartheta_X(t) = E\left[e^{itX}\right] = \sum_k e^{itx_k} \cdot p_k}$$

i - imaginarna jedinica

TM SVOJSTVA KAR. FJE.

i) jednoznačno određuje sluč. varijablu

ii) ako su $X_1, \dots, X_n$ nezavisne, tada je: $\boxed{\vartheta_{X_1+X_2+\dots+X_n}(t) = \vartheta_{X_1} \cdot \vartheta_{X_2} \cdot \dots \cdot \vartheta_{X_n}}$

iii) $\boxed{E(X^n) = \dfrac{\vartheta^{(n)}(0)}{i^n}}$ → za $n=1$: $\boxed{E(X) = -i\vartheta'(0)}$

$$\boxed{D(X) = -\vartheta''(0) + \vartheta'(0)^2}$$

$\cos t = \frac{1}{2}\left(e^{it} + e^{-it}\right)$ $\quad$ $\sin t = \frac{1}{2i}\left(e^{it} - e^{-it}\right)$

4. PRIMJERI DISKRETNIH RAZDIOBA

4.1. GEOMETRIJSKA RAZDIOBA

→ ponavljanje pokusa do prve realizacije nekog događaja → sa vraćanjem!

→ neka je X broj ponavljanja a p vjerojatnost realizacije događaja A

$$X \sim \begin{pmatrix} 1 & 2 & 3 & \dots & k \\ p & (1-p)p & (1-p)^2 p & \dots & (1-p)^{k-1} \cdot p \end{pmatrix} \Rightarrow X \sim G(p)$$

$$\boxed{P(X=k) = (1-p)^{k-1} \cdot p \;, \quad k = 1, 2, 3 \dots}$$

$$E(X)=\sum_{k=1}^{\infty} k(1-p)^{k-1}\cdot p = p\frac{1}{(1-(1-p))^2} = \frac{p}{p^2} = \boxed{\frac{1}{p}=E(X)}$$

-> KARAKT. F-JA:

$$\vartheta_X(t)=\sum_{k=1}^{\infty} e^{itk}(1-p)^{k-1}\cdot p = p(1-p)^{-1}\sum_{k=1}^{\infty}\left[e^{it}(1-p)\right]^k = \frac{p}{1-p}\left(\frac{1}{1-e^{it}(1-p)}-1\right)$$

$$\boxed{\vartheta_X(t)=\frac{pe^{it}}{1-(1-p)e^{it}}} \qquad \boxed{E(X)=-i\vartheta'(0)=\frac{1}{p}}$$

TM ODSUSTVO PAMĆENJA

X ima geom. razdiobu ako i samo ako vrijedi za sve $k,m \geq 1$

$$\boxed{P(X=k+m \mid X>k)=P(X=m)}$$

DOKAZ:

$$P(X=k+m\mid X>k)=\frac{P(X=k+m,\ X>k)}{P(X>k)}=\frac{(1-p)^{k+m-1}\cdot p}{(1-p)^k}=$$

$$=(1-p)^{m-1}\cdot p = P(X=m)$$

$$\boxed{P(X>k)=(1-p)^k}$$

4.2. BINOMNA RAZDIOBA

-> X - koliko puta se dogodio događaj (od n)

$$\boxed{X\sim B(n,p)} \qquad \boxed{P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}},\ k=0,1,2,3,\dots$$

$$\boxed{E(X)=n\cdot p} \qquad \boxed{\vartheta_X(t)=(pe^{it}+1-p)^n} \qquad \boxed{D(X)=np(1-p)}$$

-> SVOJSTVO STABILNOSTI

$$\left.\begin{aligned} X_1&\sim B(n_1,p)\\ X_2&\sim B(n_2,p)\end{aligned}\right\}\ \text{nezavisne} \rightarrow \boxed{X_1+X_2\sim B(n_1+n_2,p)}$$

DOKAZ: $\vartheta_{X_1+X_2}(t)=\vartheta_{X_1}(t)\cdot\vartheta_{X_2}(t)=(pe^{it}+1-p)^{n_1}\cdot(pe^{it}+1-p)^{n_2}=$

$$=(pe^{it}+1-p)^{n_1+n_2} \Rightarrow X_1+X_2\sim B(n_1+n_2,p)$$

→ BERNOULLIJEVA SLUČAJNA VARIJABLA

→ slučajna varijabla X_i koja bilježi je li se nešto dogodilo ili nije, ishod jednog jedinog pokusa

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

$$\boxed{\begin{aligned} E(X_i) &= p \\ D(X_i) &= p(1-p) \end{aligned}}$$

→ Binomna razdioba je suma Bernoullijevih slučajnih varijabli:

$$X \sim B(n,p) \rightarrow \boxed{X = \sum_{i=1}^{n} X_i}$$

$$E(X) = E(\Sigma X_i) = \sum_{i=1}^{n} E(X_i) = np$$

$$D(X) = D(\Sigma X_i) = \Sigma D(X_i) = np(1-p)$$

4.3. POISSONOVA RAZDIOBA

$$\boxed{E(X) = \lambda}$$

λ - intenzitet dolazaka (pojavljivanja) $\boxed{\lambda = np}$

TM Neka je n velik a p malen. Ako je $\lambda = np$, onda binomna razdioba

$$\boxed{\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}}$$

Dokaz:

$$p = \frac{\lambda}{n} \rightarrow \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} \cdot \frac{n(n-1)\cdots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!} \underbrace{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots}_{1} \underbrace{\left(1-\frac{\lambda}{n}\right)^{n}}_{e^{-\lambda}} \underbrace{\left(1-\frac{\lambda}{n}\right)^{-k}}_{1} = / \lim_{n\to\infty}$$

$$= \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{za } n \to \infty \quad //$$

$$\boxed{B(n,p) \approx P(\lambda)}$$

$$\boxed{P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0,1,2,3,\ldots,\infty}$$

→ KARAKT. FJA:

$$\vartheta(t)=\sum_{k=0}^{\infty} e^{itk}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{(e^{it}\lambda)^k}{k!}=e^{-\lambda}\cdot e^{e^{it}\lambda}=e^{\lambda(e^{it}-1)}$$

$$\boxed{\sum_{k=0}^{\infty}\frac{x^k}{k!}=e^x}$$

$$\boxed{\vartheta(t)=e^{\lambda(e^{it}-1)}} \quad \boxed{E(X)=\lambda} \quad \boxed{D(X)=\lambda}$$

→ SVOJSTVO STABILNOSTI

$$\left.\begin{array}{l} X_1\sim P(\lambda_1) \\ X_2\sim P(\lambda_2)\end{array}\right\}\text{nezavisne} \Rightarrow \boxed{X_1+X_2\sim P(\lambda_1+\lambda_2)}$$

DOKAZ:

$$\vartheta_{X_1+X_2}(t)=\vartheta_{X_1}(t)\vartheta_{X_2}(t)=e^{\lambda_1(e^{it}-1)}e^{\lambda_2(e^{it}-1)}=e^{(\lambda_1+\lambda_2)(e^{it}-1)} \quad //$$

2. SLUČAJNE VARIJABLE

5. NEPREKINUTE SLUČAJNE VARIJABLE

5.1. DEFINICIJE I SVOJSTVA

def. Preslikavanje $X:\Omega\to\mathbb{R}$ je sluč. varijabla ako za $\forall x\in\mathbb{R}$

$$\Delta_x=\{\omega\in\Omega : X(\omega)<x\}\in F$$

→ FUNKCIJA RAZDIOBE varijable X je fja. $F:\mathbb{R}\to[0,1]$ definirana s

$$\boxed{F(x)=P(X<x)}$$

TM SVOJSTVA FJE. RAZDIOBE

i) $\lim_{x\to-\infty}F(x)=0 \quad \lim_{x\to+\infty}F(x)=1$

ii) $F(x)$ je rastuća fja. (1. derivacija >0) → što je veći x sve više je obuhvaćen X

iii) $F(x)$ je neprekinuta s lijeva, ali samo s lijeva

iv) $\boxed{P(x_1\le X\le x_2)=F(x_2)-F(x_1)}$!

def. Za sluč. var. X kažemo da je **NEPREKINUTA** ako postoji **NENEGATIVNA** fja. $f: \mathbb{R} \to \mathbb{R}^+$ takva da je

$$\boxed{F(x) = \int_{-\infty}^{x} f(t)\,dt}$$

odnosno $f(x) \neq F'(x)$

→ $f(x)$ → **FUNKCIJA GUSTOĆE VJEROJATNOSTI**

$f(x) \geq 0$

$$\boxed{P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)\,dx} \qquad \boxed{\int_{-\infty}^{\infty} f(x)\,dx = 1}$$

def. **OČEKIVANJE**: $\boxed{E(X) = \int_{-\infty}^{\infty} x f(x)\,dx}$ ako integral konvergira.

def. **DISPERZIJA**: $\boxed{D(X) = \int_{-\infty}^{\infty} x^2 f(x)\,dx - E(X)^2}$ $\qquad \boxed{\sigma = \sqrt{D(X)}}$

→ **SVOJSTVA**:

i) $E(sX + tY) = sE(X) + tE(Y)$

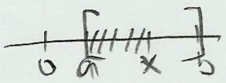
ii) $D(sX) = s^2 D(X)$

iii) ako su X i Y nezavisni: $E(X \cdot Y) = E(X) \cdot E(Y)$

$D(X \pm Y) = D(X) + D(Y)$

def. **KARAKT. FJA.**: $\boxed{\vartheta_X(t) = E(e^{itx}) = \int_{-\infty}^{\infty} e^{itx} f(x)\,dx}$ → Fourier

→ **JEDNOLIKA (UNIFORMNA) RAZDIOBA**



↳ $P(X < x)$

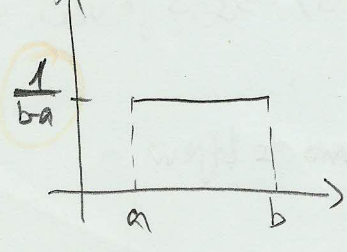
$$\boxed{X \sim U[a,b]}$$

$$\boxed{F(x) = P(X < x) = \frac{x-a}{b-a}}$$

$$\boxed{f(x) = \frac{1}{b-a}}$$

$$\boxed{E(X) = \bar{x} = \frac{a+b}{2}}$$

$$\boxed{D(X) = \frac{(b-a)^2}{12}}$$



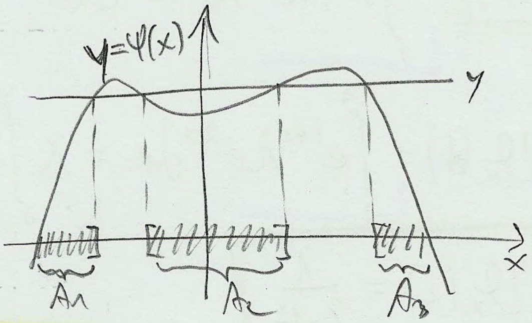
5.2. FUNKCIJE SLUČ. VARIJABLI

→ neka je $\Psi : \mathbb{R} \to \mathbb{R}$, X je sluč. varijabla, $Y=\Psi(X)$

F(x) – razdioba od X

G(y) – razdioba od Y

$$\boxed{G(y) = P(Y<y) = P(\Psi(x)<y)}$$



→ kad nije bijekcija – intervali:

$$\boxed{G(y)=P(Y<y)=P(x\in A_1)+P(x\in A_2)+P(x\in A_3)}$$

→ određivanje G(y) (f-je koje nisu injekcije)

1. ako nije zadana f(x) → odredimo ju
2. nacrtamo $y=\Psi(x)$
3. nađemo za koje x je $Y<y$ i podijelimo na intervale ako treba
4. računamo vjer. od f/h x preko inverza
   ↳ preko: $\int_a^b f(x)\,dx$
5. paziti gdje živi y! (ako smo dijelili na intervale)

npr. $P((x-2)^2<y) = P(2-\sqrt{y}<x<2+\sqrt{y})$

$y=(x-2)^2 \;/\sqrt{\ }$

$\pm\sqrt{y} = x-2$ → $x_1 = 2-\sqrt{y}$, → $x_2=2+\sqrt{y}$

→ za injektivne f-je: (ili intervale)

$$\boxed{g(y) = G'(y) = \underbrace{f(x)\left|\frac{dx}{dy}\right| = f(\Psi^{-1}(y))\left|\frac{d\Psi^{-1}(y)}{dy}\right|}_{\text{bijekcije}}}$$

→ kad traže G(y), bolje na prvi način (jer ovo može biti teško integrirati)

6. PRIMJERI NEPREKINUTIH RAZDIOBA

6.1. EKSPONENCIJALNA RAZDIOBA

→ vrijeme do sljedeće pojave događaja

→ $Z\sim \mathcal{P}(\lambda)$, Z_x broj pojavljivanja u $[0,x]$, $Z_x \sim \mathcal{P}(\lambda \cdot x)$

X – vrijeme do pojave događaja

$$\boxed{F(x)} = P(X<x) = 1-P(Z_x=0) = 1-\frac{(\lambda x)^0}{0!}e^{-\lambda x} = \boxed{1-e^{-\lambda x}}, \quad x>0$$

$$\boxed{f(x)=\lambda e^{-\lambda x}, \ x>0} \qquad \boxed{X\sim \mathcal{E}(\lambda)}$$

$$\boxed{E(X)} = \int_0^\infty x\lambda e^{-\lambda x}dx = \left|\begin{matrix}\text{parc.}\\ \text{int.}\end{matrix}\right| = \boxed{\frac{1}{\lambda}} \qquad E(X^n) = \frac{n!}{\lambda^n}$$

$$\boxed{D(X)} = \int_0^\infty x^2\lambda e^{-\lambda x}dx - E(X)^2 = \left|\begin{matrix}\text{dvaput}\\ \text{parc.}\\ \text{int}\end{matrix}\right| = \boxed{\frac{1}{\lambda^2}}$$

$$\vartheta_X(t) = \int_0^\infty e^{itx}\lambda e^{-\lambda x}dx = \lambda\int_0^\infty e^{(it-\lambda)x}dx = \frac{\lambda}{it-\lambda}e^{(it-\lambda)x}\Big|_0^\infty = \frac{-\lambda}{it-\lambda}$$

$$\boxed{\vartheta_X(t) = \frac{\lambda}{\lambda-it}}$$

TM **ODSUSTVO PAMĆENJA**

X ima eksp. raspodjelu ako i samo ako vrijedi:

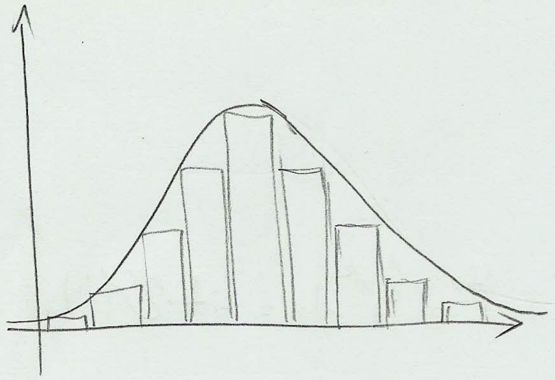
$$\boxed{P(X<x+t \mid X>t) = P(X<x)}$$

DOKAZ:

$$P(X<x+t \mid X>t) = \frac{P(t<X<x+t)}{P(X>t)} = \frac{F(x+t)-F(t)}{1-F(t)} =$$

$$= \frac{\cancel{1}-e^{-\lambda(x+t)}\cancel{-1}+e^{-\lambda t}}{\cancel{1}\cancel{-1}+e^{-\lambda t}} = 1-e^{-\lambda x} = F(x) = P(X<x)$$

6.2. NORMALNA (GAUSSOVA) RAZDIOBA



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

μ = očekivanje

σ^2 = varijanca (disperzija)

σ = stddev

-> veća σ -> veća raspršenost -> rastegnutija krivulja

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

- jedinična: $X^* \sim N(0,1)$: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$

$$\phi(x) = \frac{1}{2} + \frac{1}{2}\phi^*(x)$$

Izvod:

$$\phi(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}t^2}dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{-\frac{1}{2}t^2}dt + \frac{1}{\sqrt{2\pi}}\int_{0}^{x} e^{-\frac{1}{2}t^2}dt =$$

$$= \frac{1}{2} + \frac{1}{2}\cdot\frac{1}{\sqrt{2\pi}}\int_{-x}^{x} e^{-\frac{1}{2}t^2}dt = \frac{1}{2} + \frac{1}{2}\phi^*(x)$$

$$\phi(x) = \int_{-\infty}^{x} f(t)dt = P(X < x)$$

$$\phi^*(x) = \int_{-x}^{x} f(t)dt = P(-x < X^* < x)$$ -> TABLICA

- za zadatke:

$$P(-a < X^* < a) = \phi^*(a)$$
$$P(X^* < a) = \frac{1}{2} + \frac{1}{2}\phi^*(a)$$
$$P(X^* > a) = \frac{1}{2} - \frac{1}{2}\phi^*(a)$$
$$P(a < X^* < b) = \frac{1}{2}\left[\phi^*(b) - \phi^*(a)\right]$$

-> normiranje (u $N(0,1)$):

$$X^* = \frac{X-\mu}{\sigma}$$

$$\phi^*(-a) = -\phi^*(a)$$

-> neparna fja.

PRAVILO 3σ

$\to 1\sigma \Rightarrow p=0{,}6827$ $\quad \to 2\sigma \Rightarrow p=0{,}9545$ $\quad \to 3\sigma \Rightarrow p=0{,}9973$

KARAKT. FJA.

Izvod:

$$\vartheta(t)=\int_{-\infty}^{\infty} e^{itx}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx \quad \Big/ \frac{d}{dt}$$

$$\vartheta'(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ixe^{-\frac{1}{2}x^2}e^{itx}dx \overset{\text{parc. int}}{=} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} -te^{itx}e^{-\frac{1}{2}x^2}dx = -t\cdot\vartheta(t)$$

$$\vartheta'(t)=-t\,\vartheta(t) \qquad (y'=x\cdot y \to \text{separacija var})$$

$$\boxed{\vartheta_{x^*}(t)=e^{-\frac{1}{2}t^2}} \qquad \vartheta_x(t)=\vartheta_{\mu+\sigma x^*}(t)=e^{it\mu}\vartheta_{x^*}(t)$$

$$\boxed{\vartheta_x(t)=e^{it\mu-\frac{1}{2}\sigma^2t^2}}$$

TM SVOJSTVO STABILNOSTI

Neka su $X_1\sim N(\mu_1,\sigma_1^2)$ i $X_2\sim N(\mu_2,\sigma_2^2)$ NEZAVISNE!

Tada: $\boxed{\alpha X_1+\beta X_2\sim N(\alpha\mu_1+\beta\mu_2,\ \alpha^2\sigma_1^2+\beta^2\sigma_2^2)}$

DOKAZ:

$$\vartheta_{\alpha X_1+\beta X_2}(t)=\vartheta_{\alpha X_1}(t)\cdot\vartheta_{\beta X_2}(t)=e^{it\alpha\mu_1-\frac{1}{2}\sigma_1^2\alpha^2t^2}\cdot e^{it\beta\mu_2-\frac{1}{2}\sigma_2^2\beta^2t^2}=$$

$$=e^{it(\alpha\mu_1+\beta\mu_2)-\frac{1}{2}(\sigma_1^2\alpha^2+\sigma_2^2\beta^2)t^2} \quad //$$

APROKSIMACIJA BINOMNE NORMALNOM

TM „MOIVRE-LAPLACE - PRVI C.G.T."

$\boxed{B(n,p)\approx N(np,\ np(1-p))}$ za dovoljno velik n

$\to$ za $n<50$ proširimo interval za 0,5 sa svake strane

7. SLUČAJNI VEKTORI

7.1. NEPREKINUTI SLUČ. VEKTORI

- uređena n-torka sluč. varijabli: $X = (X_1, X_2, \dots, X_n)$
- FJA. RAZDIOBE: $F(x_1, \dots, x_n) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n)$
- 2D: $\boxed{F(x,y) = P(X<x, Y<y)}$

def. Za sluč. vektor kažemo da je NEPREKINUT ako postoji nenegativna funkcija $f: \mathbb{R}^2 \longrightarrow [0, +\infty)$ takva da je: $\boxed{F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)\,dx\,dy}$

- FJA. GUSTOĆE: $\boxed{f(x,y) = \frac{\partial^2 F}{\partial x \partial y}}$
- vjerojatnost: $\boxed{P\big((x,y)\in G\big) = \iint_G f(x,y)\,dx\,dy}$ $\boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1}$
- MARGINALNE RAZDIOBE

$$F_X(x) = P(X<x) = P(X<x, -\infty < Y < +\infty) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = \int_{-\infty}^{x} f_X(x)\,dx$$

$$\boxed{f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy} \quad \boxed{f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx}$$

TM X i Y su NEZAVISNI ako i samo ako $\boxed{f(x,y) = f_X(x) \cdot f_Y(y)}$

Dokaz:

$$F(x,y) = P(X<x, Y<y) \overset{nez.}{=} P(X<x)\cdot P(Y<y) = F_X(x)\cdot F_Y(y) \Big/ \frac{\partial x}{\partial y}$$

$\Rightarrow$ $f(x,y) = f_X(x)\cdot f_Y(y)$ //

$\Leftarrow$ $P(X\in A, Y\in B) = \iint_{A\,B} f(x,y)\,dx\,dy = \iint_{A\,B} f_X(x)\cdot f_Y(y)\,dx\,dy =$

$$= \int_A f_X(x)\,dx \int_B f_Y(y)\,dy = P(X\in A)\cdot P(X\in B) //$$

- **OČEKIVANJE**: $$\boxed{E(X)=\int_{-\infty}^{\infty} x\cdot f_X(x)\,dx}$$

- **jednolika**: $X\sim U[a,b] \Rightarrow \boxed{f(x)=\frac{1}{b-a}}$

$$\boxed{f(x,y)=\frac{1}{P}}$$

- **eksp**: $X\sim E(\lambda) \Rightarrow$ očekivanje: $\boxed{E(X)=\frac{1}{\lambda}}$ $\boxed{f(x)=\lambda e^{-\lambda x}}$

OČEKIVANJE SLUČ. VEKTORA

$$\boxed{E(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, f(x,y)\,dx\,dy}$$

TM **SVOJSTVA OČEKIVANJA**

a) $E(sX+tY)=sE(X)+tE(Y)$

b) $E(X\cdot Y)=E(X)\cdot E(Y)$ samo ako su nezavisni

DOKAZIĆ:

$$E(X\cdot Y)=\int_{-\infty}^{+\infty}dx\int_{-\infty}^{+\infty} xy\, f_X(x)\cdot f_Y(y)\,dy=\int x f_X(x)\,dx\int y f_Y(y)\,dy=E(X)\cdot E(Y)$$

7.2. UVJETNE RAZDIOBE

def. Neka je $f(x,y)$ GUSTOĆA VEKTORA (X,Y) te neka je poznata razdioba od Y. Tada se uvjetna gustoća od X uz uvjet $Y=y$ definira s:

$$\boxed{f_{X|Y=y}(x)=\frac{f(x,y)}{f_Y(y)}}$$

→ odavde

$$\boxed{f(x,y)=f_{X|Y=y}(x)\cdot f_Y(y)} \qquad f_X(x)=\int_{-\infty}^{\infty} f_{X|Y=y}(x)\,f_Y(y)\,dy$$

def. **UVJETNO OČEKIVANJE**: $$\boxed{E(X|Y=y)=\int_{-\infty}^{\infty} x\cdot f_{X|Y=y}(x)\,dx}$$

→ vrijedi: $$\boxed{E(X)=\int_{-\infty}^{+\infty} E(X|Y=y)\,f_Y(y)\,dy}$$

def. VJER. DOGAĐAJA A KOJI OVISI O REALIZACIJI OD X

$$\boxed{P(A)=\int_{-\infty}^{+\infty} P(A|X=x)f_X(x)\,dx}$$

8. FJE. SLUČ. VEKTORA

$$\boxed{g(y_1,\dots,y_n)=f(x_1,\dots,x_n)\left|\frac{\partial(x_1,\dots,x_n)}{\partial(y_1,\dots,y_n)}\right|}$$

JAKOBIJAN! stare var.

$$J=\frac{\partial(x_1,\dots,x_n)}{\partial(y_1,\dots,y_n)}=$$

nove var.

$$=\begin{vmatrix}\frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n}\\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n}\end{vmatrix}$$

→ za $n=2$: $x=X$

$y=\Psi(x,z)$

- Izvod J za $n=2$:

$$\frac{\partial(x,y)}{\partial(x,z)}=\begin{vmatrix}\frac{\partial x}{\partial x} & \frac{\partial x}{\partial z}\\ \frac{\partial y}{\partial x} & \frac{\partial y}{\partial z}\end{vmatrix}=\begin{vmatrix}1 & 0\\ \frac{\partial y}{\partial x} & \frac{\partial y}{\partial z}\end{vmatrix}=\boxed{\left|\frac{\partial y}{\partial z}\right|}$$

$$\boxed{g(x,z)=f(x,y)\cdot\left|\frac{\partial y}{\partial z}\right|,\quad y=\Psi(x,z)}$$

- marginalna od z:

$$\boxed{g(z)=\int_{-\infty}^{+\infty} f(x,y)\cdot\left|\frac{\partial y}{\partial z}\right|dx}$$

→ postupak:

1) Koje vrijednosti može poprimiti z? → $z\in(a,b)$
2) Što je y? → $y=\Psi(x,z)$
3) $J=\left|\frac{\partial y}{\partial z}\right|$
4) Nacrtati $y=\Psi(x,z)$ → za granice od x
5) $g(z)=\int_{-\infty}^{+\infty} f(x,y)|J|\,dx$

9. TEORIJA VJEROJATNOSTI

- indikatorska varijabla: $I_k \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$
- n puta ponavljamo pokus: $X_n = I_1 + I_2 + \dots + I_n = \sum I_k \Rightarrow X_n \to np$
- test W $\frac{X_n}{n} \xrightarrow{?} p$ ↳ koliko puta se dogodio događaj u n bacanja

9.1. NEJEDNAKOSTI I ZAKONI VELIKIH BROJEVA

def. Niz X_n KONVERGIRA PO VJEROJATNOSTI k sluč. varijabli X ako $\forall \varepsilon > 0$ vrijedi $\boxed{\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0}$.

→ pišemo: $X_n \xrightarrow{P} X$

TM a) „NEJEDNAKOST MARKOVA"

Ako X poprima nenegativne vrijednosti, onda za $\forall \varepsilon > 0$ vrijedi da je $\boxed{P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}}$.

b) „ČEBIŠEVLJEVA NEJEDNAKOST"

Za svaku sluč. var. s konačnim očekivanjem vrijedi da je

$$\boxed{P(|X - E(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}}$$

SLABI ZAKON VELIKIH BROJEVA

def. $X_n \xrightarrow{P} Y$ ako $\forall \varepsilon > 0$ $\boxed{\lim_{n\to\infty} P(|X_n - Y| \geq \varepsilon) = 0}$

→ konvergencija po vjerojatnosti

def. Niz $(X_n)_{n\in\mathbb{N}}$ zadovoljava **SZVB** ako $\frac{1}{n}\sum_{k=1}^{n}[X_k - E(X_k)] \xrightarrow{P} 0$

TM „DOVOLJNI UVJETI ZA SZVB"

Ako varijable $(X_n)_{n\in N}$ zadovoljavaju uvjet $\boxed{\lim \frac{1}{n^2} D\left(\sum_{k=1}^{n} X_k\right)=0}$,

tada $(X_n)_{n\in N}$ zadovoljava SZVB.

JAKI ZAKON VELIKIH BROJEVA

def. $(X_n)_{n\in Z}$ konvergira GOTOVO SIGURNO slučajnoj var. Y ako

$\boxed{P\left(\lim_{n\to\infty} X_n = Y\right) = 1}$ te označavamo $\boxed{X_n \xrightarrow{(g.s.)} Y}$

def. Kažemo da $(X_n)_{n\in Z}$ zadovoljava JZVB ako $\boxed{\frac{1}{n}\sum_{k=1}^{n}[X_k - E(X_k)] \xrightarrow{(g.s.)} 0}$.

TM $X_n \xrightarrow{(g.s.)} Y \Rightarrow X_n \xrightarrow{P} Y$ — JZVB povlači SZVB

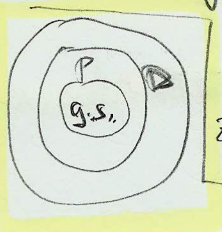
TM Neka je $(X_n)_{n\in Z}$ niz nez. jednako distribuiranih sluč. var. s očekivanjem m.

Tada vrijedi JZVB, tj.: $\boxed{\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{(g.s.)} m}$

KONVERGENCIJA PO DISTRIBUCIJI

def. Niz $(X_n)_{n\in N}$ KONVERGIRA PO DISTRIBUCIJI prema sluč. var. Y ako za pripadni niz funkcija distribucije vrijedi: $\boxed{\lim_{n\to\infty} F_{X_n}(x) = F_Y(x)}$

- oznaka: $\boxed{X_n \xrightarrow{D} Y}$



za svaki x u kojem je F_Y neprekinuta

TM KARAKT. FJA.

$X_n \xrightarrow{D} X \iff \vartheta_{X_n}(t) \longrightarrow \vartheta_X(t)$, za $\forall t$, $\boxed{\vartheta_X(t) = E[e^{itx}]}$

CENTRALNI GRANIČNI TEOREM (CGT)

TM $(X_n)_{n\in N}$ niz nezavisnih jednako distribuiranih sluč. varijabli s očekivanjem m i disperzijom σ^2. Tada za normirani zbroj vrijedi:

$$\boxed{\frac{\sum_{i=1}^{n} X_i - n\cdot m}{\sigma\sqrt{n}} \xrightarrow{D} N(0,1)}$$

- spec. slučaj: MOIVRE-LAPLACEOV TEOREM: $\boxed{\frac{B(n,p)-n\cdot p}{\sqrt{npq}} \xrightarrow{D} N(0,1)}$

- jednolika: $E(X) = \frac{b-a}{2}$, $D(X) = \frac{(b-a)^2}{12}$

- pogreška zbroja: $\sum [y_i] - \sum y_i = \sum(\underbrace{[y_i] - y_i}_{X_i \to \text{greške}})$

10. STATISTIKA

10.1. TOČKASTE PROCJENE

- predmet proučavanja: populacija X
- proučavamo podatak: obilježje

def. Za $(X_1, \dots, X_n)$ kažemo da je **UZORAK** ako su X_i međusobno nezavisni i uzeti iz iste populacije X

def. Slučajna varijabla $\Theta = g(X_1, \dots, X_n)$ se naziva statistika.
Za statistiku Θ kažemo da je procjenitelj parametra ϑ, a vrijednost $\hat{\vartheta} = g(x_1, \dots, x_n)$ nazivamo **PROCJENA** od ϑ.

- kriterij za odabir dobrih statistika:
    1) Θ je **NEPRISTRANA** statistika (procjenitelj) za ϑ ako je $\boxed{E(\Theta) = \vartheta}$.
    2) Ako niz statistika Θ_n konv. po vjerojatnosti prema parametru ϑ, tada je Θ_n **VALJANA** statistika. $\boxed{\forall \varepsilon > 0 \quad P(|\Theta_n - \vartheta| \geq \varepsilon) \xrightarrow{P} 0}$, $n \to \infty$

3) Θ_1 i Θ_2 dvije nepristrane statistike $\to$ Θ_1 bolja ako $\boxed{D(\Theta_1) < D(\Theta_2)}$

TM Nepristrana statistika je valjana ako joj disperzija teži k 0.

STATISTIKA ZA PROCJENU OČEKIVANJA a: $\boxed{\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}}$

- nepristranost: $E(\overline{X}) = E\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{E(X_1) + \ldots + E(X_n)}{n} = \frac{n \cdot a}{n} = a$
- valjanost: $D(\overline{X}) = D\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{\sum D(X_i)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \xrightarrow[\text{za } n \to \infty]{} 0$

STATISTIKA ZA PROCJENU DISPERZIJE

1) a poznat: $\boxed{D^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - a)^2}$

- nepristranost:

$$E(D^2) = \frac{1}{n}\sum_{i=1}^{n}\underbrace{E(X_i - a)^2}_{D(X_i)} = \frac{1}{n} n\sigma^2 = \sigma^2$$

$$D(D^2) = \frac{1}{n^2}\sum_{i=1}^{n} D(X_i - a)^2 = \frac{1}{n^2} n D(X-a)^2 = \frac{1}{n}\left(E(X-a)^4 - (E(X-a)^2)^2\right) =$$

$$= \frac{\mu^4 - \sigma^4}{n} \xrightarrow{n\to\infty} 0$$

2) a nepoznat: $\boxed{S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$

- nepristranost: $\Theta = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \Rightarrow E(\Theta) = \frac{n-1}{n}\sigma^2$

$$E\left[\frac{n}{n-1}\Theta\right] = \sigma^2$$

$$E(S^2) = \sigma^2$$

- češće: $\boxed{S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)}$

- min uvjet $>$
MAX uvjet $<$

10.2. KRITERIJ NAJVEĆE IZGLEDNOSTI (KNI)

MAXIMUM INTENSITY LIKELYHOOD FUNCTION (MILF)

def. Neka je $x_1, \dots, x_n$ realizacija uzorka iz neke populacije X čija funkcija gustoće f ovisi o parametru ϑ. FUNKCIJA IZGLEDNOSTI se definira kao

$$\boxed{L(\vartheta, x_1, \dots, x_n) = f(\vartheta, x_1) \cdot \dots \cdot f(\vartheta, x_n)}$$

→ za diskretne varijable: $\boxed{= P(X=x_1) \cdots P(X=x_n)}$ → množimo vjerojatnosti

- za procjenu parametra ϑ uzima se vrijednost za koju L poprima GLOBALNI MAXIMUM → derivaciju izjednačiti s nulom

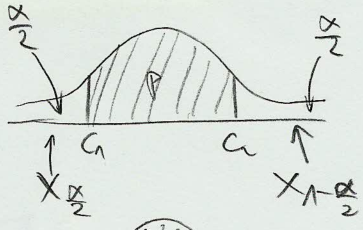
- ln: $\ln(\lambda^n) = n \ln \lambda \;/ \frac{\partial}{\partial \lambda} \rightarrow \frac{n}{\lambda}$

ln umnoška = zbroj ln-ova

11. INTERVALNE PROCJENE

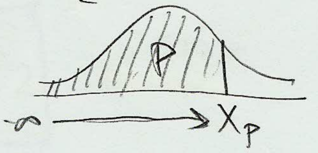
11.1. KVANTILI I UVOD

def. Interval $[c_1, c_2]$ za koji vrijedi $P(c_1 < X < c_2) = p$ naziva se INTERVAL POVJERENJA (POUZDANOSTI) reda p.



def. Veličina $\alpha = 1 - p$ se naziva NIVO ZNAČAJNOSTI.

def. Realan broj x_p za koji vrijedi da je $F(x_p) = p$, tj. $\boxed{\int_{-\infty}^{x_p} f(t)\,dt = p}$ se naziva KVANTIL reda p.



11.2. INT. PROCJENE ZA PARAMETRE NORMALNE RAZDIOBE

→ formule i tablice!

→ studentova: $t_{n-1, 1-\frac{\alpha}{2}}$ → u tablici gledamo baš $\boxed{\alpha}$ → trebamo 95%-tni interval, onda je $\alpha = 5\% = 0{,}05$

$n-1 = \boxed{n}$ u tablici

→ χ^2 razdioba: $\chi^2_{n-1, 1-\frac{\alpha}{2}}$ → u tablici je za $1-\frac{\alpha}{2}$, ne α → normalna: isto $1-\frac{\alpha}{2}$ (u_p)

11.3. INTERVALNE PROCJENE ZA VJEROJATNOST

$$\boxed{\hat{p} = \frac{m}{n} = \frac{\text{broj povoljnih}}{\text{ukupan broj}}}$$

→ za zadani kvantil računat preko Φ^* → $\boxed{x_p = \Phi^*[2\cdot(0{,}5-u_p)]}$

12. TESTIRANJE HIPOTEZA

12.1. UVOD I OSNOVNI POJMOVI

H_0 ... NULTA HIPOTEZA

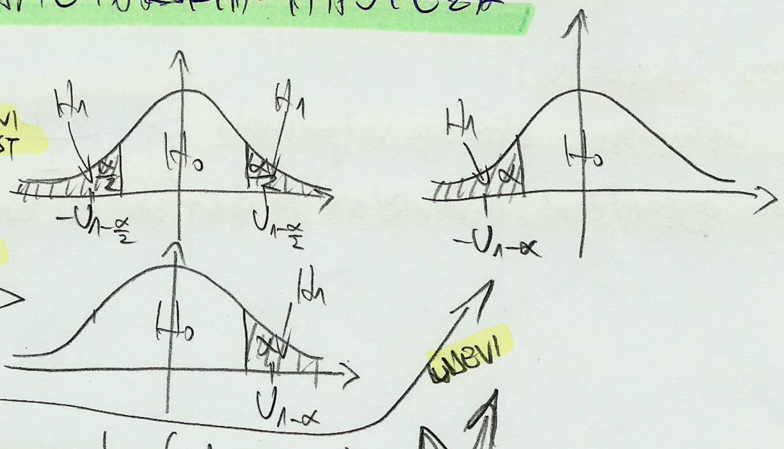
H_1 ... ALTERNATIVNA HIPOTEZA

def. POGREŠKA PRVE VRSTE se definira kao $\alpha = \sup S(\vartheta)$ ako je $\vartheta \in H_0$.
Vjerojatnost da smo prihvatili H_1, a istinita je H_0! → OZBILJNIJA

POGREŠKA DRUGE VRSTE se definira kao $\beta = \sup(1 - S(\vartheta))$ ako je $\vartheta \in H_1$.
Vjer. da prihvatimo H_0 ako je istinita H_1.

12.2. TESTIRANJE PARAMETARSKIH HIPOTEZA

1. $H_0 \dots \vartheta = \vartheta_0$
   $H_1 \dots \vartheta \neq \vartheta_0$ → DVOSTRANI TEST
   $\vartheta > \vartheta_0$ → DESNI
   $\vartheta < \vartheta_0$ → LIJEVI



2. Izračunati statistiku iz uzorka $(U, T, \dots)$
3. Usporediti sa kvantilom → testovi

- kod JEDNOSTRANOG t-testa iz tablice gledamo $\alpha' = 2\alpha$
  → duplo veći α! jer imamo $t_{n-1}(1-\alpha)$

12.3. USPOREDBA 2 POPULACIJA

→ nema ništa tricky

12.4. TEST PRILAGODBE RAZDIOBAMA (χ^2-test)

x_j	n_j	p_j	np_j	$\frac{(n_j - np_j)^2}{np_j}$
⋮	⋮	⋮	⋮	⋮
$\Sigma =$	N	1	N	$\Sigma = \chi_q^2$

↳ tu spajamo tek

→ svaki razred sa $n_j < 5$ spajamo sa susjednim → uvijek odozdo

→ χ_q^2 usporedimo sa $\boxed{\chi^2_{m-r-1,\,1-\alpha}}$

m = broj razreda nakon spajanja

r = broj procjenjenih parametara razdiobe

→ podsjetnik – Poisson: $\boxed{P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad E(X) = \lambda = \overline{x}}$

– Eksp: $P(a < X < b) = F(b) - F(a) = 1 - e^{-b\lambda} - (1 - e^{-a\lambda}) = e^{-a\lambda} - e^{-b\lambda}$

$P(X < x) = 1 - e^{-\lambda x}$

$f(x) = \lambda e^{-\lambda x}$

$E(X) = \frac{1}{\lambda} \rightarrow \boxed{\lambda = \frac{1}{\overline{x}}}$

→ kad imamo intervale:

- za x_j uzimamo srednje vrijednosti $\boxed{\overline{x}_j = \frac{a+b}{2}}$
- ako nisu ograničeni, ograničimo ih tako da svi budu iste duljine

DOKAZI, IZVODI, ISKAZI

2010.

- svojstvo stabilnosti normalne (iskaz)
- def. kvantil reda p (za $p \in (0,1)$ i fju razdiobe F)

2011.

- dokaz: Markova

$$P(X \geq \varepsilon) = \int_{\varepsilon}^{\infty} f(x)\,dx \leq \int_{\varepsilon}^{\infty} \frac{x}{\varepsilon} f(x)\,dx \leq \frac{1}{\varepsilon}\int_{0}^{\infty} x f(x)\,dx = \frac{E(X)}{\varepsilon}$$

- def. kvantil reda p

2012. /

2013. - procjena disperzije $\hat{s}^2$

$$\hat{s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n} x_i^2 - 2x_i\bar{x} + \bar{x}^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\bar{x}^2\right) =$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 \underbrace{- 2\bar{x}\cdot n\bar{x} + n\bar{x}^2}_{-2n\bar{x}^2}\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right) //$$

- def. kvantil
- dokaz: da za kvantile jed. norm razdiobe vrijedi: $u_p = -u_{1-p}$

$$\underbrace{\int_{0}^{x_p} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx = -\int_{1-x_p}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}}_{\downarrow}$$

parna fja. → jednakost vrijedi

2014.

neprístranost $\bar{X}$

— 11 — $\hat{s}^2$

$\rightarrow \hat{s}^2$:

$$E(s^2) = \frac{1}{n-1}\sum_{i=1}^{n} E(x_i - \bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n} D(x_i - \bar{x}) = \frac{1}{n-1}\sum_{i=1}^{n} D\left(x_i - \frac{\sum x_i}{n}\right) =$$

$$= \frac{1}{n-1}\sum_{i=1}^{n} D\left(\frac{n-1}{n}x_i - \frac{1}{n}\sum_{j\neq i} x_j\right) = \frac{1}{n-1}\sum\left[\left(\frac{n-1}{n}\right)^2 D(x_i) + \frac{1}{n^2} D\left(\sum_{j\neq i} x_j\right)\right] =$$

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

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

2015. /