9. TEORIJA VJEROJATNOSTI

⇝ uvijek dolazi 1 zad. na ISPITU i na školskoj

MOTIVACIJA: indikatorska varijabla ... I_k → je li se nešto dogodilo ili nije

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

→ n puta ponavljamo pokus $\quad X_n = I_1 + I_2 + \dots + I_n = \sum I_k$

↳ koliko puta se dogodio događaj u n bacanja ($X_n \to n \cdot p$)

$$\frac{X_n}{n} \xrightarrow{?} p$$

9.1. NEJEDNAKOSTI I ZAKONI VELIKIH BROJEVA

def.

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

Pišemo: $X_n \xrightarrow{P} X$

„ZNALO JE DOĆI NA ISPITU"

2 $\boxed{\text{TM}}$ a) „NEJEDNAKOST MARKOVA"

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

b) „ČEBIŠEVLJEVA NEJEDNAKOST"

Za svaku slučajnu varijablu s konačnim očekivanjem vrijedi da je

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

U PRAKSI SE UVIJEK KORISTI

6.DZ-2.) „sličan i na ispitu i na šk"

Očekivana devijacija brzine vjetra na nekoj visini je 25 km/h, a brzina 4.5 km/h. Kolika se brzina vjetra može očekivati s vjerojatnosti ne manjom od 90% ?

$E(X) = 25$ km/h

$\sigma(X) = 4.5$ km/h

$$P(|X - 25| \geq \varepsilon) \leq \frac{4.5^2}{\varepsilon^2}$$

$$\underbrace{1 - P(|X - 25| \leq \varepsilon)}_{\text{RJEŠITI!}} \geq \underbrace{1 - \frac{4.5^2}{\varepsilon^2} = 0.9}$$

koliko se neće razlikovati

$\Rightarrow \underline{\varepsilon = 14.23}$

$\Rightarrow |X-25| \leq 14.23$

$$\boxed{10.77 \leq X \leq 39.23}$$

⇝ zanima nas: aritmetička sredina rezultata $(X_n)_{n\to N}$

$$(*) \quad \frac{X_1+\dots+X_n}{n} \rightsquigarrow \text{konvergira?}$$

$\Rightarrow$ ZAKON VELIKIH BROJEVA objašnjava konvergenciju ovog tipa

- ovisi o tipu konvergencije:

1) konvergencija po vjerojatnosti (SLABI)
2) konvergencija gotovo sigurno (JAKI)

- cilj: naći uvjete na sluč. var. (X_n) tako da imamo konvergenciju $(*)$ po vjerojatnosti ili gotovo sigurno

... SLABI ZAKON VELIKIH BROJEVA

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

⇝ konvergencija po vjerojatnosti

def. Kažemo da niz $(X_n)_{n\in\mathbb{N}}$ zadovoljava SLABI ZAKON VELIKIH BROJEVA ako

$$\frac{1}{n}\sum_{k=1}^{n}[X_k - E(X_k)] \xrightarrow{P} 0$$

TM

„Dovoljni uvjeti za slabi zakon velikih brojeva"

Ako varijable $(X_n)_{n\in\mathbb{N}}$ zadovoljavaju uvjet

$$\lim \frac{1}{n^2}\cdot D\left(\sum_{k=1}^{n} X_k\right) = 0$$

tada $(X_n)_{n\in\mathbb{N}}$ zadovoljava slabi zakon velikih brojeva.

⇝ tražimo da $D(X_n)$ postoji

⇝ Čebiševljeva nejednakost za dokaz, tražimo da $P\left(\left|\sum_{k=1}^{n}(X_k - E(X_k))\right| \geq \varepsilon\right) \to 0$

- specijalni slučajevi kada SZVB vrijedi:
  1) (X_n) nekoreliranih s ograničenom varijancom
  2) (X_n) nezavisnih s istom varijancom $D(X_n) = \sigma^2$
  3) (X_n) nezavisnih i jednako distribuiranih s konačnom varijancom

DOKAZ:
- ako su nezavisne (X_n) tada je disperzija sume jednaka sumi disperzija
- budući da su jednako distribuirane sve imaju jednaku $D(X_n) = \sigma^2$

$$\lim_{n\to\infty} \frac{1}{n^2} D\left(\sum_{k=1}^{n} X_n\right) = \lim_{n\to\infty} \frac{1}{n^2}\cdot \overbrace{\sum_{k=1}^{n} D(X_n)}^{n\text{-disperzija}} = \lim_{n\to\infty} \frac{1}{n^2}\cdot n\cdot\sigma^2 = \lim_{n\to\infty} \frac{\overset{\text{konst.}}{\sigma^2}}{n} = 0$$

Primjer: Promotrimo jednostavan slučaj:

$$X_n = I_n \sim \begin{pmatrix} 0 & 1 \\ Q & p \end{pmatrix} \longrightarrow \text{niz nezavisnih indikatorskih sluč. varijabli}$$

$$S_n = \sum_{k=1}^{n} I_k = \text{broj uspjeha u } n \text{ pokusa}$$

$$\frac{1}{n}\cdot S_n = \frac{1}{n}\sum_{k=1}^{n} I_k = \text{prosječan broj uspjeha}$$

- za veliki n: ⇝ $(I_n)_{n\in\mathbb{N}}$ zadovoljava specijalni slučaj 3)

$$m(I_n) = p \quad D(I_n) = p\cdot Q$$

⇝ vrijedi SZVB: $\frac{1}{n}\sum_{k=1}^{n}[I_k - \underbrace{E(I_k)}_{p}] \xrightarrow{p} 0$

$$\frac{1}{n}\sum_{k=1}^{n} I_k - p \xrightarrow{p} 0$$

$$\boxed{\frac{S_n}{n} \xrightarrow{p} p}$$

$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) = 0$ → vjerojat. da je odstupanje od teoretske vjerojatnosti 0

⇝ možemo dobiti ocjene na odstupanja prosječnog broja uspjeha od vjerojatnosti uspjeha

→ koristimo Čebiševljevu nejednakost

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{D\left(\frac{1}{n}\cdot S_n\right)}{\varepsilon^2}$$

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

→ zbog nezavisnosti: $D(\Sigma) = \Sigma D$

$$D\left(\frac{1}{n}\cdot S_n\right) = \frac{1}{n^2} D(S_n) = \frac{1}{n^2}\cdot\sum_{k=1}^{n} D(I_k) = \frac{1}{n^2}\cdot \cancel{n}\cdot p\cdot Q$$

↳ $P\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{pQ}{n\cdot\varepsilon^2} \leq \frac{1}{4\cdot n\cdot\varepsilon^2}$

$pQ = p(1-p) \leq \frac{1}{4}$ za $p \in \langle 0, 1\rangle$ ⇝ max: $p = \frac{1}{2}$

$\Rightarrow p\cdot(1-p) = \frac{1}{4}$ max

Primjer: obavili smo pokus 1000 puta i tražimo poklapanje do na 1. decimalu

$n = 1000 \ , \ \varepsilon = 0.1$

$$P\Big(\Big|\frac{S_n}{n} - \underbrace{p}_{\text{nepoznato}}\Big| \geq 0.1\Big) \leq \frac{1}{4 \cdot 1000 \cdot 0.1^2} = 0{,}025$$

komplement: $P\left(\left|\frac{S_n}{n} - p\right| < 0.1\right) \geq 1 - 0{,}025 = 0.975$

↳ u barem 97,5 % slučajeva će relativni uspjeh odstupati od vjerojatnosti uspjeha za manje od 0.1

SZVB $\Rightarrow$ vjerojatnost odstupanja $\longrightarrow 0$

⇝ želimo jači rezultat: $\frac{S_n}{n} \longrightarrow p$: <u>JZVB</u>

def. Kažemo da $(X_n)_{n \in \mathbb{Z}}$ konvergira GOTOVO SIGURNO slučajnoj varijabli Y ako:

$$\boxed{P\left(\lim_{n\to\infty} X_n = Y\right) = 1}$$

te označavamo $X_n \xrightarrow{(g.s.)} Y$.

! NAPOMENA:

$$K = \left\{\lim_{n\to\infty} X_n = Y\right\} = \left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = Y(\omega)\right\}$$

$$K \in \mathcal{F} \ \ i \ \ P(K) = 1$$

JAKI ZAKON VELIKIH BROJEVA

def. Kažemo da $(X_n)_{n\in\mathbb{N}}$ zadovoljava JAKI ZAKON VELIKIH BROJEVA ako

$$\frac{1}{n}\sum_{k=1}^{n}\left[X_k - E(X_k)\right] \xrightarrow{(g.s.)} 0$$

$\boxed{\text{TM}}$ „Odnos konvergencije po vjerojatnosti i gotovo sigurno"

$$X_n \xrightarrow{(g.s.)} Y \Rightarrow X_n \xrightarrow{P} Y$$

Posljedica: JZVB povlači SZVB

$\boxed{\text{TM}}$ Neka je $(X_n)_{n\in\mathbb{Z}}$ niz nezavisnih jednako distribuiranih slučajnih varijabli s očekivanjem m. Tada vrijedi JZVB, tj.

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{g.s.} m$$

Primjer. $(I_n)_{n\in\mathbb{N}}$ nezavisne indikatorske jednako distribuirane varijable

$$I_n \sim \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix} \qquad E(I_n) = p = m$$

$\longrightarrow$ JZVB vrijedi:

$$\frac{1}{n}\sum_{k=1}^{n} I_k \xrightarrow{g.s.} p$$

$\frac{S_n}{n}$ = prosječan broj uspjeha

$$P\left(\left\{\omega\in\Omega : \frac{S_n(\omega)}{n} \to p\right\}\right) = 1$$ $\rightsquigarrow$ skoro svugdje ima konv. $\frac{S_n}{n}$

$\rightsquigarrow$ zanimaju nas vjerojatnosti da svaki od X_n upada u A:

$\lim\limits_{n\to\infty} P(X_n\in A)$ (limes nekih distrib.) $\underset{A=\langle-\infty,x\rangle}{\longleftrightarrow}$ $P(X_n\in\langle-\infty,x\rangle) = P(X_n < x) = F_{X_n}(x)$

... KONVERGENCIJA PO DISTRIBUCIJI I KARAKTERISTIČNE FUNKCIJE ...

def.

Kažemo da niz $(X_n)_{n\in\mathbb{N}}$ slučajnih varijabli **KONVERGIRA PO DISTRIBUCIJI** prema slučajnoj varijabli Y ako za pripadni niz funkcija distribucije vrijedi

$$\boxed{\lim_{n\to\infty} F_{X_n}(x) = F_Y(x)}$$

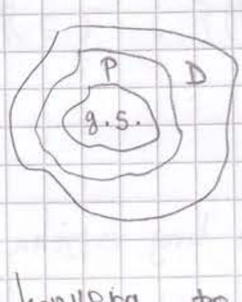
za svaki x u kojem je F_Y neprekinuta.

⇝ označavamo: $X_n \xrightarrow{D} Y$

TM

"ODNOS KONVER. PO DISTRIBUCIJI S OSTALIM KONVERGENCIJAMA:"

$$X_n \xrightarrow{P} Y \implies X_n \xrightarrow{D} Y$$



(IDEJA: Povezati konverg. po distribuciji s karakterističnim funkcijama

$\boxed{TM}$

$$X_n \xrightarrow{D} X \iff \vartheta_{X_n}(t) \longrightarrow \vartheta_X(t), \text{ za } \forall t$$

$$\vartheta_X(t) = E\left[e^{itX}\right] \rightsquigarrow \text{karakteristična funkcija}$$

Primjer:

Aproksimacija binomne slučajne varijable Poissonovom:

$$X_{n,p_n} \sim B(n, p_n), \quad p_n \xrightarrow{n\to\infty} 0, \quad n \cdot p_n \xrightarrow{n\to\infty} \lambda > 0$$

$$\Rightarrow X_{n,p_n} \xrightarrow{D} P(\lambda)$$

$\rightsquigarrow$ pokažimo da ovo vrijedi:

$$Y \sim P(\lambda)$$

$$X_{n,p_n} \xrightarrow{D} Y \iff \vartheta_{X_{n,p_n}}(t) \xrightarrow{n\to\infty} \vartheta_Y(t)$$

$$\vartheta_{X_{n,p_n}}(t) = \left(Q_n + p_n e^{it}\right)^n = \left(1 - p_n + p_n e^{it}\right)^n =$$

$$= \left[1 + p_n(e^{it} - 1)\right]^n =$$

$$= \left\{\left[1 + p_n(e^{it} - 1)\right]^{\frac{1}{m}}\right\}^{\overset{\to\lambda}{n p_n}} \longrightarrow \left(e^{e^{it}-1}\right)^{\lambda} =$$

$$= e^{\lambda(e^{it}-1)} = \vartheta_Y(t), \text{ za } \forall t$$

CENTRALNI GRANIČNI TEOREM (CGT)

$\boxed{\text{TM}}$

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

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

!

SPECIJALNI SLUČAJ: MOIVRE LAPLACEOV TEOREM:

$$\boxed{\frac{B(n,p) - n\cdot p}{\sqrt{npQ}} \xrightarrow{D} N(0,1)}$$

DOKAZ:

⇝ uzmemo $X_i = I_i$ indikatorske var. s parametrom p i nezavisne

$$\sum_{i=1}^{n} I_i \sim B(n,p) \qquad \mathbb{E}I_i = p = m \qquad D(I_i) = pQ = \sigma^2$$

$$\underbrace{\frac{\sum_{i=1}^{n} X_i - nm}{\sigma\sqrt{n}}}_{\downarrow D \; N(0,1)} \overset{D}{=} \underbrace{\frac{B(n,p) - np}{\sqrt{npQ}}}_{\downarrow D \; N(0,1)}$$

6.DZ. - 5.) $X_1, \dots, X_{10}$ nezavisne sl. var. s jednolikom razdiobom na $[0,1]$.

Pomoću CGT izračunati: $P\left(\sum_{i=1}^{10} X_i > 6\right) = ?$

$$E(X_1) = \frac{b-a}{2} = \frac{1-0}{2} = \frac{1}{2}$$

$$D(X_1) = \frac{(b-a)^2}{12} = \frac{1}{12} \qquad \sigma = \frac{\sqrt{12}}{12}$$

$$\frac{\sum_{i=1}^{10} X_i - 10 \cdot \frac{1}{2}}{\frac{\sqrt{12}}{12} \cdot \sqrt{10}} \xrightarrow{D} N(0,1)$$

ODUZELI OBJE STRANE -5, A ZATIM PODIJELILI S $\frac{\sqrt{10}}{\sqrt{12}}$ DA NAŠTIMAMO NA CGT

$$P\left(\sum_{i=1}^{10} X_i > 6\right) = 1 - P\left(\sum_{i=1}^{10} X_i < 6\right) =$$

TREBA NAŠTIMATI D i L NA CGT

$$= 1 - P\left(\frac{\sum_{i=1}^{10} X_i - 5}{\frac{\sqrt{10}}{\sqrt{12}}} < \frac{6-5}{\frac{\sqrt{10}}{\sqrt{12}}}\right) =$$

$$= 1 - \Phi\left(\frac{6-5}{\frac{\sqrt{10}}{\sqrt{12}}}\right) =$$

TABLICA!

$$= 1 - \frac{1}{2}\left[1 + \Phi^*(\sqrt{12})\right] =$$

$$= 0{,}1379$$

21-11-1.) b)

$$\left.\begin{matrix} y_1, \dots, y_{100} \\ [y_1], \dots, [y_{100}] \end{matrix}\right\} \quad [y_i] - y_i \sim U(-0.5, 0.5)$$

nezavisne

$\rightsquigarrow$ pogreška zbroja:

$$\sum_{i=1}^{100} [y_i] - \sum_{i=1}^{100} y_i = \sum_{i=1}^{100} \underbrace{([y_i] - y_i)}_{X_i \text{ greške}}$$

$$P\left(\sum_{i=1}^{100} X_i \in \langle -t, t \rangle\right) \geq 0.95$$

$$P\left(\left|\sum_{i=1}^{100} X_i\right| < t\right)$$

(X_i) nezavisne, jednaka distribucija $\sim U(-0.5, 0.5)$

$m = 0 \qquad \sigma^2 = \frac{1}{12} \implies$ CGT

$$\frac{\sum_{i=1}^{100} X_i - 100 \cdot 0}{\sqrt{\frac{1}{12}} \cdot \sqrt{100}} \approx N(0,1)$$

$$P\left(\left|\sum_{i=1}^{100} X_i\right| < t\right) = P\left(\left|\frac{\sum_{i=1}^{100} X_i}{\frac{10}{\sqrt{12}}}\right| < \frac{t}{\frac{10}{\sqrt{12}}}\right) \approx \phi^*\left(\frac{t}{\frac{10}{\sqrt{12}}}\right) \geq 0.95$$

$$\phi^*(1.96) = 0.95$$

$$\frac{t}{\frac{10}{\sqrt{12}}} = 1.96$$

$$\Rightarrow \boxed{t = 5.658}$$