Seminar 13 – Inegalitatea Cebasev. Legea numerelor mari.

Tabel cu distributiile frecvente, probabilitatile, mediile si dispersiile acestora.

Distributie	Probabilitate = x / Densitate de probab.	Medie $E(X)$	Dispersie $D^2(X)$
Bernoulli (p)	p^x	p	$p(1-p)$
Binomiala (n,p)	$C_n^k p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson (λ)	$e^{-\lambda}\lambda^x/x!$	λ	λ
Geometrica (p)	$p(1-p)^{x-1}$	$1/p$	$(1-p)/p^2$
Binomiala negativa (k,p)	$C_{k-1}^{x-1} p^k (1-p)^{x-k}$	k/p	$k(1-p)/p^2$
Hipergeometrica (n extrageri, a bile albe si b bile negre)	$\frac{C_a^x C_b^{n-x}}{C_{a+b}^n}$	$np, p = \frac{a}{a+b}$	$np(1-p)\frac{a+b-n}{a+b-1}$
Uniforma discreta (a,b)	$h/(b-a+h), h = a, a+h, a+2h, \dots, b$	$(a+b)/2$	$(b-a)(b-a+2h)/12$
Normala (n,σ)	$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$	μ	σ^2
Exponentiala (λ)	$f(x) = \lambda e^{-\lambda x}, x \in \mathbb{R}$	$1/\lambda$	$1/\lambda^2$
Gamma (α,λ)	$f(x) = \lambda^\alpha x^{\alpha-1} e^{-\lambda x}/\Gamma(\alpha)$	α/λ	α/λ^2
Uniforma continua (a,b)	$f(x) = 1/(b-a)$	$(a+b)/2$	$(b-a)^2/12$
Beta (α,β)	$f(x) = \Gamma(\alpha+\beta)/(\Gamma(\alpha)\Gamma(\beta))$	$\alpha/(\alpha+\beta)$	$\alpha\beta/((\alpha+\beta)^2(\alpha+\beta+1))$

Teorema 1. *(**Inegalitatea lui Cebâşev**) Fie X o v.a. cu dispersia finită. Atunci $\forall \varepsilon > 0$ au loc*

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

sau echivalent

$$P(|X - E(X)| < \varepsilon) \geq 1 - \frac{D^2(X)}{\varepsilon^2}$$

Fie şirurile de variabile aleatoare definite pe acelaşi câmp de probabilitate $(X_n)_{n\geq 1}, (Y_n)_{n\geq 1}, (Z_n)_{n\geq 1}$, unde $X_n \sim \begin{pmatrix} -5n & 0 & 5n \\ \frac{1}{3n^2} & 1-\frac{2}{3n^2} & \frac{1}{3n^2} \end{pmatrix}$,

$Y_n \sim \begin{pmatrix} -n^2 & 0 & n^2 \\ \alpha^{-n} & 1-2\alpha^{-n} & \alpha^{-n} \end{pmatrix}$ şi Z_n are densitatea de repartiţie

$$f_n(x) = \begin{cases} \lambda^{-n} e^{-\frac{x}{\lambda^n}}, & x > 0 \\ 0, & \text{în rest} \end{cases}$$

$\forall n \geq 1, \lambda > 0$. Să se verifice aplicabilitatea teoremei lui Cebîşev celor trei şiruri.

Verificam daca mediile si dispersiile sirurilor de v.a. sunt finite.

$$E(X_n) = (-5n)\frac{1}{3n^2} + (5n)\frac{1}{3n^2} = 0$$

$$D^2(X_n) = (-5n)^2\frac{1}{3n^2} + (5n)^2\frac{1}{3n^2} - \left(E(X_n)\right)^2 = \frac{50}{3}$$

Deci X_n are medie si dispersie finita, deci putem sa aplicam inegalitatea lui Cebasev.

$$E(Y_n) = -n^2\alpha^{-n} + n^2\alpha^{-n} = 0$$

$$D^2(Y_n) = n^4\alpha^{-n} + n^4\alpha^{-n} - \left(E(Y_n)\right)^2 = 2n^4\alpha^{-n}$$

Avem $\lim\limits_{n\to\infty} D^2(Y_n) = 0$ daca $\alpha > 1$ si in acest caz putem aplica inegalitatea Cebasev.

$$\int_0^\infty f_n(x)\,dx = 1 = \int_0^\infty \lambda^{-n}\,e^{-\frac{x}{\lambda^n}}dx = \lambda^{-n}\left(\frac{e^{-\frac{x}{\lambda^n}}}{-\frac{1}{\lambda^n}}\right)\Bigg|_0^\infty = -(-1) = 1, doar\ daca\ \lambda > 0$$

Daca $\lambda < 0$ nu avem o densitate de probabilitate

$$E(Z_n) = \int_0^\infty x\lambda^{-n}\,e^{-\frac{x}{\lambda^n}}dx = \lambda^{-n}\int_0^\infty x\left(-\frac{e^{-\frac{x}{\lambda^n}}}{\frac{1}{\lambda^n}}\right)'dx = -\left(\left(xe^{-\frac{x}{\lambda^n}}\right)\Big|_0^\infty - \int_0^\infty e^{-\frac{x}{\lambda^n}}\,dx\right) = \int_0^\infty e^{-\frac{x}{\lambda^n}}\,dx$$
$$= \lambda^n$$

$$D^2(Z_n) = \int_0^\infty x^2\lambda^{-n}\,e^{-\frac{x}{\lambda^n}}dx - \lambda^{2n} = \lambda^{-n}\int_0^\infty x^2\left(-\frac{e^{-\frac{x}{\lambda^n}}}{\frac{1}{\lambda^n}}\right)'dx - \lambda^{2n}$$
$$= -\left(\left(x^2e^{-\frac{x}{\lambda}}\right)\Big|_0^\infty - 2\int_0^\infty xe^{-\frac{x}{\lambda}}\,dx\right) - \lambda^{2n} = 2\lambda^{2n} - \lambda^{2n} = \lambda^{2n}$$

Se poate aplica inegalitatea Cebasev doar pentru $\{\lambda \in \mathbb{R} | \lambda > 0\} \cap \{\lambda \in \mathbb{R} | \lambda \leq 1\} => \lambda \in (0,1]$.

Ex: Fie $X \sim Exp(\lambda)$. Folosind inegalitatea Cebasev gasiti o limita maxima pentru valoarea probabilitatii $P(|X - E(X)| \geq b)$, unde $b > 0$.

$$P(|X - E(X)| \geq b) \leq \frac{D^2(X)}{b^2} = \frac{1}{(b\lambda)^2}$$

Ex: Limita (marginea) superioara a probabilitatii ca abaterile, in modul, ale valorilor v.a. X fata de medie, sa fie mai mare decat 3, este 0,96. Sa se afle $D^2(X)$.

$$P(|X - E(X)| \geq 3) \leq \frac{D^2(X)}{9} = 0{,}96$$

$$D^2(X) = 8{,}64.$$

Ex: Se observa ca durata de functionare fara defectiuni a computerelor produse de o companie este de 36 luni, cu o dispersie de 2 luni. Care este procentul de calculatoare care functioneaza fara defectiuni un numar de luni cuprins intre 30 si 42 de luni?

Fie X v.a. egala cu durata de functionare fara defectiuni a computerelor, in luni, $E(X) = 36, D^2(X) = 2$. Aplicam inegalitatea Cebasev

$$P(30 \leq X \leq 42) = P(|X - 36| \leq 6) = P(|X - E(X)| \leq 6) \geq 1 - \frac{D^2(X)}{6^2} = 1 - \frac{2}{36} = \frac{17}{18}$$

Ex: Durata medie a unei curse facute cu tramvaiul 16, din orasul XYZ, este de 50 de minute, iar dispersia este de 4 minute. Primaria doreste sa faca publicitate noului sistem de dirijare a traficului, zicand ca: „in 95% din cazuri, cursa cu tramvaiul 16 va dura intre ... si ... minute". Ce numere ar trebui sa completeze in materialul promotional?

Fie X v.a. egala cu durata cursei cu tramvaiul, exprimata in minute, $E(X) = 50, D^2(X) = 4$. Trebuie sa-l calculam pe t a.i. sa avem

$$P(|X - E(X)| \leq t) = 95\%$$

Aplicand inegalitatea Cebasev avem

$$P(|X - E(X)| \leq t) \geq 1 - \frac{D^2(X)}{t^2} = 1 - \frac{4}{t^2} = 95\%$$

$$\frac{t^2 - 4}{t^2} = \frac{19}{20} => t^2 = 80 => t = \sqrt{80} \simeq 9\, minute$$

Deci fraza trebuie sa fie: in 95% din cazuri, cursa cu tramvaiul 16 va dura intre 41 si 59 minute.

Ex: Aplicand inegalitatea Cebasev, sa se gaseasca limita inferioara a probabilitatii inegalitatii

$$\left|\frac{\alpha}{10^5} - \frac{1}{6}\right| < \frac{1}{100},$$

unde α reprezinta numarul de aparitii ale fetei 6 din 100.000 de aruncari ale unui zar.

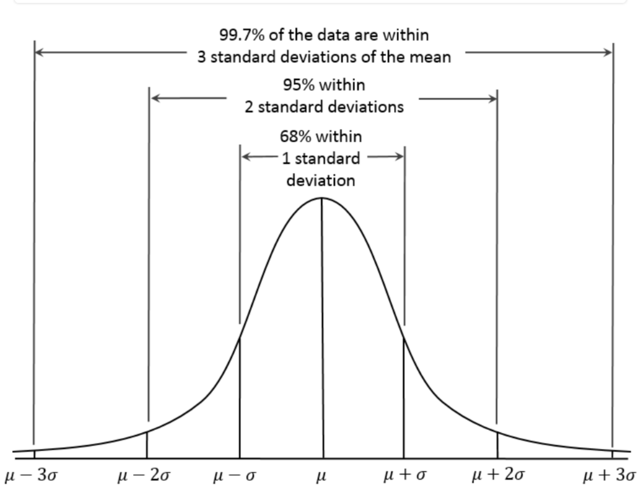
V.a. α are o repartitie binomiala $n = 10^5, p = \frac{1}{6}$.

$$E(\alpha) = np = \frac{10^5}{6}, D^2(\alpha) = np(1-p) = \frac{5 * 10^5}{36}$$

$$P\left(\left|\frac{\alpha}{10^5} - \frac{1}{6}\right| < \frac{1}{100}\right) = P\left(\left|\alpha - \frac{10^5}{6}\right| < 10^3\right) \geq 1 - \frac{D^2(\alpha)}{10^6} = 1 - \frac{\frac{5 * 10^5}{36}}{10^6} = 1 - \frac{5}{360} = 1 - \frac{1}{72} = \frac{71}{72}$$

Teorema limita centrala – aplicatii

Teorema limita centrala: distributia unui esantion de variabile aleatoare tinde catre cea normala, indiferent de distributia initiala a esantionului. Pe masura ce crestem esantionul, distributia acestuia va deveni mai apropiata de cea normala si mai ingusta.



Aplicatii:

- asteptarea de viata a populatiei unei tari
- amplitudinea zgomotului termic din circuitele electrice
- intensitatea fluxului de laser
- numarul de gauri din jurul centrului unei table de darts
-

Teorema 4. *(**Teorema limită centrală**) Fie $(X_n)_n$ un şir de v.a. independente, identic repartizate. Presupunem că $E(X_i) = m, D^2(X_i) = \sigma^2, i = \overline{1,n}$ există şi notăm $Z_n = \frac{Y_n - E(Y_n)}{\sqrt{D^2(Y_n)}}$, unde $Y_n = \sum_{i=1}^{n} X_i$. Atunci $\forall x \in \mathbb{R}, \lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P(Z_n < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du$, adică şirul $(Z_n)_n$ converge în repartiţie către o v.a. $Z \sim N(0,1)$.*

Un caz particular al teoremei limită centrală este:

Teorema 6. *(**Teorema integrală a lui Moivre-Laplace**) Fie un şir de experimente independente astfel încât în fiecare experiment probabilitatea de realizare a unui eveniment A este p. Dacă ν_n este numărul de apariţii ale lui A în primele n experimente, atunci*

$$\forall x \in \mathbb{R}, \lim_{n\to\infty} P\left(\frac{\nu_n - np}{\sqrt{npq}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du = \Phi(x)$$

sau

$$\forall x_1, x_2 \in \mathbb{R}, \lim_{n\to\infty} P\left(x_1 < \frac{\nu_n - np}{\sqrt{npq}} < x_2\right) = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{u^2}{2}} du$$

Ex: Care este probabilitatea ca, din 100 de aruncari ale unei monede, stema sa apara de un numar de ori cuprins intre 40 si 60?

Fie X_n v.a. egala cu evenimentul aparitiei stemei la aruncarea unei monede de 100 ori. Aplicam teorema Moivre-Laplace si gasim ca

$$Z_n = \frac{X_n - np}{\sqrt{npq}} \sim N(0,1)$$

$$P(40 < X_n < 60) = P\left(\frac{40 - np}{\sqrt{npq}} < Z_n < \frac{60 - np}{\sqrt{npq}}\right) = \Phi\left(\frac{60 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{40 - np}{\sqrt{npq}}\right) = \Phi(2) - \Phi(-2)$$
$$= 2\Phi(2) - 1 = 2 * 0{,}9773 - 1 = 0{,}9546$$

Ex: Se arunca un zar de 500 de ori. Cu ce probabilitate putem afirma ca obtinem fata 6 de cel mult 90 de ori?

Fie X_n un sir de v.a. reprezentand evenimentul ca sa apara fata 6 la aruncarea unui zar. Aplicam teorema Moivre-Laplace

$$\frac{X_n - np}{\sqrt{npq}} \sim N(0,1)$$

$$P(X_n \leq 90) = P\left(\frac{X_n - np}{\sqrt{npq}} < \frac{90 - np}{\sqrt{npq}}\right) = \Phi\left(\frac{90 - 500 * \frac{1}{6}}{\sqrt{500 * \frac{5}{36}}}\right) = \Phi\ (0{,}8) = 0{,}7881$$

Ex: Se arunca de 360 de ori o pereche de zaruri. Cu ce probabilitate ne putem astepta sa apara 12 puncte (dubla 6) de un numar de ori cuprins intre 8 si 12?

Fie X_n un sir de v.a. egale cu evenimentul aparitiei dublei 6 la aruncarea a 2 zaruri, $n = 360, p = \frac{1}{36}$. Aplicand teorema Moivre-Laplace avem ca

$$Z_n = \frac{X_n - np}{\sqrt{npq}} \sim N(0,1)$$

$$P(8 \leq X_n \leq 12) = P\left(\frac{8 - np}{\sqrt{npq}} \leq Z_n \leq \frac{12 - np}{\sqrt{npq}}\right) = \Phi\left(\frac{12 - 10}{\sqrt{\frac{350}{36}}}\right) - \Phi\left(\frac{8 - 10}{\sqrt{\frac{350}{36}}}\right)$$
$$= \Phi(0{,}64) - \Phi(-0{,}64) = 2 * \Phi(0{,}64) - 1 = 2 * 0{,}7389 - 1 = 0{,}477$$

Ex: Probabilitatea ca o anumita operatie chirurgicala sa fie un succes este 0,8. Daca operatia a fost facuta la 120 de persoane, gasiti probabilitatea ca mai mult de 90 de operatii sa fie efectuate cu succes.

Fie X_n sir de v.a. egale cu evenimentul ca operatia sa fie un succes. Aplicam teorema Moivre-Laplace si gasim ca

$$Z_n = \frac{X_n - np}{\sqrt{npq}} \sim N(0,1)$$

$$P(X_n > 90) = P\left(Z_n > \frac{90 - np}{\sqrt{npq}}\right) = 1 - \Phi\left(\frac{90 - 96}{4{,}3812}\right) = 1 - \big(1 - \Phi(1{,}37)\big) = \Phi(1{,}37) = 0{,}9147$$

Observatie: Fie X_n sir de v.a. independente si identic repartizate cu medie μ si abatere medie patratica σ. Vom nota cu $\bar{X}_n$ v.a. egala cu media aritmetica a v.a. $X_1, \dots, X_n$.

$$E(\bar{X}_n) = \frac{n * E(X)}{n} = E(X) = \mu,$$

$$D^2(\bar{X}_n) = D^2\left(\frac{\sum_{i=1}^{n} X_n}{n}\right) = E\left(\frac{\sum_{i=1}^{n} X_n}{n}\right)^2 - \big(E(X)\big)^2 = \frac{1}{n^2}\left(\sum_{i=1}^{n} E(X_n^2) + \sum_{i\neq j} E\big(X_i \cdot X_j\big)\right) - \mu^2$$
$$= \frac{1}{n^2}\left(nE(X^2) + \sum_{i\neq j} E\big(X_j\big)E\big(X_j\big)\right) - \mu^2 = \frac{1}{n^2}(nE(X^2) + 2C_n^2\mu^2) - \mu^2$$
$$= \frac{1}{n}E(X^2) + \frac{1}{n^2}\left(2\frac{n!}{2!\,(n-2)!} - n^2\right)\mu^2 = \frac{1}{n}E(X^2) + \frac{1}{n^2}(n(n-1) - n^2)\mu^2$$
$$= \frac{1}{n}E(X^2) - \frac{1}{n}\mu^2 = \frac{1}{n}D^2(X) = \frac{\sigma^2}{n}$$
$$=> \sigma_{\bar{X}_n} = \frac{\sigma}{\sqrt{n}}$$

Aplicand teorema limita centrala gasim ca

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

Ex: Presupunem ca X are o distributie discreta egala cu

$$f(x) = \begin{cases} \frac{1}{3}, daca\ x = 1,2,3 \\ 0, in\ rest \end{cases}$$

Se considera un esantion de $n = 36$ de obiecte cu aceasta distributie. Gasiti probabilitatea ca media esantionului sa fie cuprinsa intre 2,1 si 2,5.

Observam ca X are o repartitie uniforma discreta $a = 1, b = 3, h = 1$

$$E(X) = 2, D^2(X) = \frac{(b-a)(b-a+2h)}{12} = \frac{2}{3}, D(X) = \sqrt{\frac{2}{3}}$$

Fie X_n un sir de v.a. cu repartitia de mai sus si fie $\bar{X}_n$ media aritmetica a acestora. Vom avea $\mu = E(X) = 2, \sigma = D(X) = \sqrt{\frac{2}{3}}$

Conform teoremei limita centrala gasim ca

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

$$P(2{,}1 \le \bar{X}_n \le 2{,}5) = P\left(\frac{2{,}1-2}{\frac{\sqrt{\frac{2}{3}}}{\sqrt{36}}} < \bar{Z}_n < \frac{2{,}5-2}{\frac{\sqrt{\frac{2}{3}}}{\sqrt{36}}}\right) = \Phi(3{,}674) - \Phi(0{,}734) \simeq 1 - 0{,}7673 = 0{,}2327$$

Ex: Un cercetator doreste sa estimeze media unei populatii folosind un esantion suficient de mare, astfel incat sa aiba o probabilitate de 99% ca media esantionului sa nu difere de media populatiei cu mai mult de 25%. Cat de mare trebuie sa fie acest esantion?

Fie X_n un sir de v.a. cu medie μ si abatere medie patratica σ si fie $\bar{X}_n$ media aritmetica a acestor variabile. Trebuie sa calculam pe n a.i. sa avem indeplinita inegalitatea din enunt, adica:

$$P\left(|\bar{X}_n - \mu| < \frac{\sigma}{4}\right) = 99\%$$

Conform teoremei limita centrala

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

$$P\left(|\bar{X}_n - \mu| < \frac{\sigma}{4}\right) = P\left(-\frac{\sigma}{4} < \bar{X}_n - \mu < \frac{\sigma}{4}\right) = P\left(-\frac{\frac{\sigma}{4}}{\frac{\sigma}{\sqrt{n}}} < \bar{Z}_n < \frac{\frac{\sigma}{4}}{\frac{\sigma}{\sqrt{n}}}\right) = \Phi\left(\frac{\sqrt{n}}{4}\right) - \Phi\left(-\frac{\sqrt{n}}{4}\right)$$

$$= 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1 = 0{,}99$$

$$\Phi\left(\frac{\sqrt{n}}{4}\right) = \frac{1{,}99}{2} = 0{,}995 = \Phi(2{,}575) \Rightarrow$$

$$\sqrt{n} = 4 * 2{,}575 = 10{,}3 \Rightarrow n \simeq 106$$

Asadar trebuie sa considere un esantion de cel putin 106 persoane.

Ex: Probabilitatea obtinerii unei piese rebut din productia unei masini automate este $p = 0{,}005$. Sa se determine probabilitatea ca din 10000 piese fabricate la aceeasi masina, numarul rebuturilor sa fie:

a) intre 60 si 70

b) cel putin 70

Fie X_n un sir de v.a. egale cu numarul rebuturilor. Aplicam teorema Moivre-Laplace,

$$P(60 < X_n < 70) = P\left(\frac{60-50}{7{,}05} < Z_n < \frac{70-50}{7{,}05}\right) = \Phi\left(\frac{20}{7{,}05}\right) - \Phi\left(\frac{10}{7{,}05}\right) = \Phi(2{,}84) - \Phi(1{,}42)$$
$$= 0{,}9977 - 0{,}9222 = 0{,}0755$$

$$P(X_n \leq 70) = \Phi(2{,}84) = 0{,}998$$

La un restaurant pot servi pranzul 75 de clienti. Din practica, proprietarul stie ca 20% dintre clientii care au rezervat un loc nu se mai prezinta.

a) Presupunand ca au fost 90 de rezervari, care este probabilitatea ca sa se prezinte mai mult de 50 de clienti?

b) Cate rezervari trebuie sa accepte proprietarul pentru ca, cu o probabilitate mai mare sau egala cu 0,9, sa poata servi toti clientii care se prezinta? ($\Phi(1{,}281) = 0{,}9$).

a) Fie X_n un sir de v.a. egale cu numarul de clienti care se prezinta la restaurant, $n = 90, p = 0{,}8$. Aplicam teorema Moivre-Laplace

$$P(X_n > 50) = P\left(Z_n > \frac{50-72}{3{,}79}\right) = P(Z_n > -5{,}8) = 1$$

b)

$$Z_n = \frac{X_n - np}{\sqrt{npq}} \sim N(0,1)$$

$$P(X_n \leq 75) = 0{,}9$$

$$P\left(Z_n \leq \frac{75 - n * 0{,}8}{\sqrt{n}0{,}4}\right) = \Phi\left(\frac{75 - n * 0{,}8}{\sqrt{n}0{,}4}\right) = 0{,}9 \Rightarrow$$

$$\frac{75 - n * 0{,}8}{\sqrt{n}0{,}4} = 1{,}281$$

$$75 - n * 0{,}8 = \sqrt{n} * 0{,}5124$$

$$\sqrt{n} = y$$

$$0{,}8y^2 + 0{,}5124 * y - 75 = 0$$

$$y_{1,2} = \frac{-0{,}5124 \pm \sqrt{0{,}5124 + 4 * 0{,}8 * 75}}{1{,}6} = 9{,}37$$

$$\Rightarrow n = 9{,}37^2 = 87{,}84.$$

Deci numarul maxim de rezervari pe care le poate accepta este egal cu 87.