INEGALITATEA LUI MARKOV

Dacă X este o variabilă aleatoare ce ia doar valori nenegative, atunci pentru oricare a>0

$$P\{X \geq a\} \leq (1/a) \cdot M[X]$$

•<u>Demonstrație</u>.

$$M[X] = \int_0^\infty x \cdot f(x)dx = \int_0^a x \cdot f(x)dx + \int_a^\infty x \cdot f(x)dx \geq$$

$$\geq \int_a^\infty x \cdot f(x)dx \geq a \cdot \int_a^\infty f(x)dx = a \cdot P\{X \geq a\}$$

INEGALITATEA LUI CEBÂŞEV

- <u>Consecință</u>. Dacă X este o variabilă aleatoare cu media μ şi dispersia σ^2 , atunci pentru oricare k>0: $$P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$
- <u>Demonstrație</u>.

Aplicăm Markov pentru v.a. $|(X-\mu)^2|$ şi a=k^2

$$P\{(X-\mu)^2 \geq k^2\} \leq \frac{M[(X-\mu)^2]}{k^2} = \frac{\sigma^2}{k^2}; \quad P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

- Inegalitățile Markov şi Cebâşev dau margini pentru probabilități când nu se ştie distribuția v.a., ci doar media / dispersia ei.

LEGEA TARE A NUMERELOR MARI

- “Media aritmetică a unui şir de v.a. independente, de aceeaşi distribuție, converge la media distribuției cu probabilitate 1”.
- Teoremă. Fie X_1, X_2,… un şir de v.a. independente cu aceeaşi distribuție şi fie $M[X_i]=\mu$, i=1,2,… Atunci:

$$P\{\lim_{n\to\infty}\frac{X_1+X_2+...+X_n}{n}=\mu\}=1$$

- Selecții repetate, independente, din aceeaşi populație
- Experiment repetat independent: media tinde la parametru

DEFINIȚIILE CU FRECVENȚE ŞI AXIOMATICĂ

- Fie **E** un eveniment, P{**E**} probabilitatea ataşată.
- Fie X_i = *if* (**E** se produce la a i-a repetare) *then 1 else* **0** *endif*.
- Legea tare a numerelor mari asigură, **cu probabilitate 1**, că $\dfrac{X_1+\ldots+X_n}{n} \to M[X_i] = P\{E\}$
- $X_1+\ldots+X_n$ e nr. de apariții ale lui **E** în n repetări, deci, la limită, E se produce <u>cu frecvența P{**E**}.</u>
- Cu frecvențe: probabilitatea $\lim((X_1+\ldots+X_n)/n)$
- Axiomatic: P**{E}**, consistent cu axiomele.

TEOREMA LIMITĂ CENTRALĂ (1)

•Calculul aproximativ al probabilităților pentru sume de v.a. independente.

•De ce atât de multe populații dau curbe normale?

•<u>Teoremă</u>. Fie X_1, X_2, … un şir de v.a. independente identic distribuite fiecare cu media μ şi dispersia σ^2. Atunci distribuția v.a.

$$\frac{\frac{X_1 + X_2 + ... + X_n}{n} - \mu}{\sqrt{\sigma^2 / n}}, n \in \mathbf{N},$$ tinde la **N**(0,1).

•Sau: $$\lim_{n\to\infty} P\{\frac{X_1 + X_2 + ... + X_n - n\mu}{\sigma\sqrt{n}} \le a\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx$$

TEOREMA LIMITĂ CENTRALĂ (2)

- Oricare ar fi distribuția comună!
- Chiar pentru distribuții diferite, dacă nici una nu domină!
- Chiar pentru v.a. ne-independente, dacă acestea au corelație mică!

APROXIMAREA NORMALĂ A BINOMIALEI

- $X = B(n,p)$ are aceeaşi distribuție ca $\sum_{1..n} X_i$, cu $X_i = B(1,p) =$ Bernoulli(p) şi X_i v.a. independente
- Atunci distribuția lui

$$Y = \frac{X - M[X]}{\sqrt{D^2(X)}} = \frac{X - np}{\sqrt{np(1-p)}} = \frac{n \cdot \left(\frac{X_1 + \ldots + X_n}{n} - M[X_i]\right)}{n \cdot \sqrt{\frac{D^2(X_i)}{n}}}$$

<u>tinde la</u> **N**(0,1) *<u>când</u>* $n \to \infty$ **(Y normală!)**

- Aproximarea

N(0,1) este bună pentru: $np(1-p) \geq 10$.

UN EXEMPLU

- Fie X numărul de apariții ale feței “ban” la 40 de aruncări independente ale unei monede.
- Cât este P{X=20}?

$$P\{X=20\}\approx P\{19{,}5\le X_c\le 20{,}5\}=P\{\frac{19{,}5-20}{\sqrt{10}}\le\frac{X_c-20}{\sqrt{10}}\le\frac{20{,}5-20}{\sqrt{10}}\}=$$

$$P\{-0{,}16\le\frac{X_c-20}{\sqrt{10}}\le+0{,}16\}\approx\Phi(0{,}16)-\Phi(-0{,}16)\qquad (D^2(X)=np(1-p)=10)$$

- Φ funcția de distribuție a lui **N**(0,1): $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-(y^2/2)}dy$
- Φ(-0,16) = 1- Φ(0,16) = 1- 0,5636
- P{X=20} ≈ 2·Φ(0,16) - 1 ≈ 0,1272
- P{X=20} = $C_{40}^{20}\cdot(\frac{1}{2})^{20}\cdot(\frac{1}{2})^{20}=0{,}1268$

PROCESE STOCHASTICE

- <u>Definiție</u>. Un *proces stochastic* $\{X(t), t \in \mathbf{T}\}$ este o colecție de variabile aleatoare. ($\forall t$, $X(t)$ e o v.a.).
- $X(t)$ este *starea* procesului la momentul t

(un proces stochastic descrie evoluția în timp a unui proces).

- <u>Exemple</u>. $X(t)$ poate număra:
  - clienții intrați în magazin până la momentul t;
  - clienții aflați în magazin până la momentul t;
  - banii încasați până la momentul t.
- **T** numărabilă – *proces discret*;
- **T** interval real – *proces continuu*.

PROCESE STOCHASTICE FINITE

- **T**={1,…,n}. Procesul este un arbore.
- Proces independent: X(t) nu depinde de X(t-k)
- Proces Markov (finit): X(t) depinde doar de X(t-1)
- Lanț Markov (finit): un proces Markov pentru care probabilitățile de trecere dintr-o stare în alta nu depind de t (matricea probabilităților de trecere).
- Proces cu legături complete: clase de stări. Din cele *finale* procesul nu mai iese.

RĂSPUNSURI RANDOMIZATE

- “Ați copiat la vreo lucrare (proiect) în facultate?”
- Procedura de răspuns.

Fiecare student aruncă o monedă în secret.
Dacă e “BAN” şi nu a copiat, spune “NU”.
În orice altă situație, spune “DA” (a copiat sau “STEMA”).

- Ştiind că 30% din studenți au copiat măcar o dată şi presupunând moneda corectă, să se construiască un arbore de decizie probabilistă. (0,5 – “BAN”)
- (0,7 – “NU”) (0,3 – “DA”)

DA 0,35 DA 0,15 NU 0,35 DA 0,15

- Dar dacă “NU” reprezintă 39% din răspunsuri, câți studenți putem estima că nu au copiat?

$$P\{N\} = P\{N/B\} = P\{N \cap B\} / P\{B\} = 0{,}39 / 0{,}5 = 0{,}78$$