Matematică 3

Notițe de seminar

Adrian Manea

Curs: Luminița Costache

15 mai 2021

Cuprins

SEMINAR 1

NUMERE ȘI FUNCȚII COMPLEXE — RECAPITULARE

1.1 Numere complexe – Noțiuni de bază

Începem cu cîteva noțiuni esențiale și recapitulative privitoare la mulțimea numerelor complexe. Amintim definiția:

$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}, i^2 = -1\},$$

precum și faptul că avem, în general, șirul de incluziuni:

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

Dat un număr complex $z = a + bi$, se numește *partea reală*, notată $\mathrm{Re}(z)$, numărul real a, iar *partea imaginară*, notată $\mathrm{Im}(z)$, numărul real b.

De asemenea, *conjugatul* numărului complex z de mai sus este $\overline{z} = z^* = a - bi$, iar *modulul* numărului complex z este $|z| = \sqrt{a^2 + b^2}$.

Există mai multe forme de reprezentare a numerelor complexe:

- *forma algebrică*, dată mai sus, $z = a + bi$, $a, b \in \mathbb{R}$;
- *forma trigonometrică*, $z = r(\cos\theta + i\sin\theta)$, unde $r = |z|$, iar $\theta = \mathrm{Arg}(z)$ se numește *argumentul principal*;
- *forma polară*, $z = re^{i\theta} = r\exp(i\theta)$, unde r și θ au înțelesul din forma trigonometrică;
- *forma geometrică*, în care $z = a + bi$ reprezintă *afixul* punctului din plan $A(a, b)$. Pentru acest caz, menționăm că $|z| = \sqrt{a^2 + b^2}$ reprezintă lungimea vectorului de poziție al punctului A, iar θ reprezintă unghiul pe care îl face acest vector de poziție cu axa OX, măsurat în sens trigonometric.

Operațiile cu numere complexe se fac în modul uzual, ținînd seama de proprietatea $i^2 = -1$. Mai amintim *formula lui Moivre*, utilă în special atunci cînd numărul complex a fost scris sub formă trigonometrică. Fie, așadar, numerele complexe:

$$z_1 = r_1(\cos\theta_1 + i\sin\theta_1), \quad z_2 = r_2(\cos\theta_2 + i\sin\theta_2).$$

Atunci înmulțirea acestora se face cu formula:

$$z_1 \cdot z_2 = r_1 r_2 \left(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)\right),$$

formulă care se generalizează ușor în forma:

$$z^n = r^n(\cos(n\theta) + i\sin(n\theta)), \forall n.$$

Tot folosind numere complexe, putem reprezenta și curbe:

- *Cercul* centrat în punctul de afix z_0 și de rază R are ecuația:

$$|z - z_0| = r;$$

- *Elipsa* cu focarele în punctele de afixe z_0 și w_0, iar axa mare are lungimea d are ecuația:

$$|z - z_0| + |z - w_0| = d.$$

Amintim și *formele canonice* ale ecuațiilor acestor conice:

- *Cercul* centrat în (x_0, y_0) și de rază R are ecuația:

$$(x - x_0)^2 + (y - y_0)^2 = R^2;$$

- *Elipsa* de semiaxe a, b are ecuația:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Tot din punct de vedere geometric, mai amintim și că *distanța* între două puncte $A(z)$ și $B(w)$ este $AB = |z - w|$.

1.2 Funcții complexe elementare

Cea mai simplă funcție complexă este *funcția exponențială*, definită prin:

$$\exp : \mathbb{C} \longrightarrow \mathbb{C}, \quad \exp z = e^z = \sum_{n\geq 0} \frac{z^n}{n!}.$$

Se observă imediat că au loc proprietățile așteptate:

- $\exp(0) = 1$;
- $\exp(z_1 + z_2) = \exp(z_1) \cdot \exp(z_2), \forall z_{1,2} \in \mathbb{C}$;
- $\exp(iy) = \cos y + i \sin y, \forall y \in \mathbb{R}$ (formula lui Euler).

Mai departe, putem calcula simplu *logaritmul complex*, dacă numărul complex a fost adus în forma polară. Fie, așadar, $z = re^{i\theta}$. Rezultă:

$$\ln z = \ln r + i\theta.$$

În general, cum argumentul unui număr complex nu este unic (θ de mai sus reprezintă *argumentul principal*, dar $\theta + 2k\pi$ este argumentul general), spunem că funcția logaritm este *multiformă*, deoarece valoarea ei generală este:

$$\mathrm{Ln} z = \{\ln |z| + i(\mathrm{Arg} z + 2k\pi) \mid k \in \mathbb{Z}\},$$

iar $\ln z$ se numește *valoarea principală* a logaritmului.

Funcția putere se definește acum simplu pornind de la formula:

$$a^b = \exp\left(b \ln a\right).$$

Rezultă:

$$z^m = \exp\left(m \ln z\right) = \exp\left(m(\ln |z| + i\mathrm{Arg} z)\right),$$

folosind valoarea principală.

Similar se definește și puterea rațională, adică *funcția radical*:

$$\sqrt[n]{z} = z^{\frac{1}{n}} = \exp\left(\frac{1}{n} \ln z\right).$$

Folosind funcțiile exponențiale și identitatea lui Euler, putem defini și *funcții trigonometrice complexe:*

$$\cos z = \frac{\exp(iz) + \exp(-iz)}{2}$$
$$\sin z = \frac{\exp(iz) - \exp(-iz)}{2i}$$
$$\tan z = -i\frac{\exp(iz) - \exp(-iz)}{\exp(iz) + \exp(-iz)}$$

Mai avem nevoie și de *funcțiile trigonometrice hiperbolice*:

$$\sinh z = \frac{\exp(z) - \exp(-z)}{2}$$
$$\cosh z = \frac{\exp(z) + \exp(-z)}{2}$$
$$\tanh z = \frac{\sinh z}{\cosh z}$$

și au loc legăturile:

$$\sinh z = -i \sin(iz), \quad \cosh z = \cos(iz).$$

Funcțiile trigonometrice inverse se pot obține din rezolvarea unor ecuații trigonometrice[1]:

$$w = \mathrm{arcsin} z \Longrightarrow z = \sin w \Longrightarrow z = \frac{\exp(iw) - \exp(-iw)}{2i} \Longleftrightarrow \exp(2iw) - 2iz\exp(iw) - 1 = 0,$$

care se rezolvă (pentru w) ca o ecuație de gradul al doilea și se obține:

$$w = \mathrm{arcsin} z = -i \ln(iz \pm \sqrt{1 - z^2})$$

și se procedează similar pentru arccos și arctan.

1.3 Funcții olomorfe

Fie o funcție complexă oarecare $f : A \longrightarrow \mathbb{C}$, cu $A \subseteq \mathbb{C}$. Pentru orice $z \in A \subseteq \mathbb{C}$, deoarece z are o parte reală și o parte imaginară, și imaginea sa prin f se poate separa. Deci, în general, orice funcție complexă f ca mai sus poate fi scrisă sub forma

$$f = P + iQ, \quad P, Q : \mathbb{R}^2 \longrightarrow \mathbb{R}.$$

Rezultă că noțiunile de *limită, continuitate, derivabilitate* pot fi puse pe componente.

Definiție 1.1: O funcție complexă $f : \mathbb{C} \longrightarrow \mathbb{C}$ se numește *olomorfă* dacă este derivabilă în orice punct din domeniul de definiție.

Nu intrăm în detalii, deoarece nu vom rezolva exerciții cu derivate complexe.

Vor fi foarte importante, însă, rezultatele:

Teoremă 1.1: *Funcție $f : \mathbb{C} \longrightarrow \mathbb{C}, f = P + iQ$ este olomorfă dacă $P, Q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ sunt diferențiabile, iar derivatele lor parțiale verifică* condițiile Cauchy-Riemann, *adică:*

$$\frac{\partial P}{\partial x} = \frac{\partial Q}{\partial y}, \quad \frac{\partial Q}{\partial x} = -\frac{\partial P}{\partial y}.$$

[1] Pentru corectitudine, ar trebui să notăm funcțiile trigonometrice cu inițială mare, Arcsin, Arccos etc. Dar vom folosi de fiecare dată doar valoarea principală, astfel că, prin abuz de notație, folosim scrierea din cazul real. Însă trebuie reținut că majoritatea funcțiilor complexe sînt *multiforme*, i.e. pot avea mai multe valori!

Observaţie 1.1: Pentru simplitate, vom mai nota derivatele parțiale cu indici, adică, de exemplu:

$$\frac{\partial f}{\partial x} \overset{\text{not.}}{=} f_x$$

și similar $f_y, f_{xx}, f_{xy}, f_{yx}$ etc.

Corola 1.1: *Dacă funcția complexă $f = P + iQ$ este olomorfă, atunci P și Q sînt armonice, adică:*

$$\Delta P = P_{xx} + P_{yy} = \Delta Q = Q_{xx} + Q_{yy} = 0.$$

Atenție, însă, la formularea rezultatului: condiția de olomorfie este *necesară*, în niciun caz suficientă! Negarea corolarului de mai sus este:

Corola 1.2: *Dacă una dintre funcțiile P sau Q nu este armonică, atunci funcția $f = P + iQ$ nu poate fi olomorfă.*

1.4 Exerciții

1. Verificați dacă funcția de mai jos este olomorfă:

$$f : \mathbb{C} \longrightarrow \mathbb{C}, \quad f(z) = z^2 + \exp(iz).$$

Indicație: Se separă partea reală și partea imaginară a funcției și se verifică condițiile Cauchy-Riemann și faptul că cele două componente sînt armonice.

2. Fie $P(x, y) = e^{3x} \cos 2y + y^2 - x^2$. Determinați funcția olomorfă $f = P + iQ$ astfel încît $f(0) = 1$.

Indicație: Verificăm dacă $\Delta P = 0$ (condiție necesară!). Apoi, prin integrarea condițiilor Cauchy-Riemann, se obține componenta Q. În final, $f(z) = e^{2z} - z^2 + ki, k \in \mathbb{C}$ și folosind condiția din enunț, obținem $k = 0$.

3. Rezolvați ecuațiile:

(a) $\exp w = -2i$;

(b) $z^3 + 2 - 2i = 0$;

(c) $\sin z = 2$.

4. Calculați:

(a) $\sin(1 + i)$;

(b) $\sinh(1 - i)$;

(c) $\ln i$;

(d) $\ln(1 - i)$;

(e) $(1 + i)^{20}$;

(f) $\sqrt[5]{1 - i}$;

(g) $\arcsin(i\sqrt{3})$;

(h) $\arccos(i\sqrt{3})$;

(i) $\tan(1 - i)$.

5. Verificați dacă funcțiile date sînt armonice și, în caz afirmativ, determinați funcția complexă $f = P + iQ$, pentru:

(a) $P(x, y) = x^2 - y^2 - 2y$, știind că $f(0) = 1$;

(b) $P(x, y) = x^4 - 6x^2y^2 + y^4$, știind că $f(1) = 1$;

(c) $P(x, y) = (x \cos y - y \sin y)e^x$;

(d) $P(x, y) = xy + x + 2y$, știind că $f(2i) = -1 + 5i$;

(e) $P(x, y) = 4xy^3 - 4x^3y + x$, știind că $f(1 + i) = 5 + 4i$;

(f) $P(x, y) = 3x^2y + 2x^2 - y^3 - 2y^2$;

(g) $P(x, y) = x^4 - 6x^2y^2 + y^2$;

(h) $Q(x, y) = \dfrac{x}{x^2 + y^2}$.

6. Determinați $k \in \mathbb{C}$ astfel încît funcția:

$$P(x, y) = x^3 - kxy^2 + 12xy - 12x$$

să fie armonică și determinați f olomorfă pentru care $P = \mathrm{Re}(f)$.

7. Scrieți sub formă trigonometrică și polară numerele complexe:

(a) $z = 3 - i$;

(b) $z = 3 + i$;

(c) $z = i$;

(d) $z = 1$;

(e) $z = 1 + 2i$;

(f) $z = 2 + i$.

8. Găsiți forma canonică și ecuația complexă pentru:

(a) Cercul centrat în $(0, 1)$ și cu raza 2;

(b) Cercul centrat în $(1, 0)$ și cu raza 1;

(c) Cercul centrat în $(1, 2)$ și cu raza 1;

(d) Elipsa cu focarele în $(-1, 0)$ și $(3, 0)$ și cu axa mare de lungime 6;

(e) Elipsa cu focarele în $(0, 1)$ și $(0, -2)$ și cu axa mare de lungime 5.

Reprezentați grafic fiecare dintre cazurile de mai sus.

9. Fie punctele $A(1 + 2i)$ și $B(-1)$, iar M, mijlocul segmentului $[AB]$. Calculați distanța de la punctul M la punctul N, de afix $-2 + 3i$. Reprezentare grafică.

SEMINAR 2

INTEGRALE COMPLEXE

2.1 Teorema lui Cauchy

În multe situații, putem calcula integralele complexe direct, într-o manieră asemănătoare cu integralele curbilinii. Un exemplu simplu:

$$I_1 = \int_{|z|=1} z|dz|.$$

Folosind forma polară, $z = e^{it}$, deoarece integrala se face pe $|z| = 1$, iar $t \in [0, 2\pi]$. Rezultă $dz = ie^{it}dt$, deci $|dz| = dt$. Atunci:

$$I_1 = \int_0^{2\pi} e^{it}dt = \frac{1}{i}e^{it}\Big|_0^{2\pi} = 0.$$

Un alt exemplu:

$$I_2 = \int_S z|dz|,$$

unde S este segmentul care unește pe 0 și i. Putem parametriza acest segment: $S : z = ti, t \in [0, 1]$, deci $dz = idt$ și din nou $|dz| = dt$. Rezultă:

$$I_2 = \int_0^1 tidt = \frac{i}{2}.$$

Dar în unele situații, putem calcula chiar mai ușor:

Teoremă 2.1 (Cauchy)**:** *Fie $D \subseteq \mathbb{C}$ un domeniu simplu conex și $f : D \longrightarrow \mathbb{C}$ o funcție olomorfă pe D, cu $P = \mathrm{Re}f$ și $Q = \mathrm{Im}f$, funcții de clasă $\mathcal{C}^1(D)$.*

Fie $\gamma : [a, b] \longrightarrow D$ o curbă închisă și jordaniană (fără autointersecții) de clasă $\mathcal{C}^1$ pe porțiuni, astfel încît $\operatorname{Int}\gamma$ *să verifice condițiile formulei Green-Riemann.*

Atunci $\displaystyle\int_\gamma f(z)dz = 0.$

Acesta este un caz simplu în care calculul se termină imediat cu rezultat nul.

În exerciții, vom folosi adesea următoarea:

Teoremă 2.2 (Formula integrală Cauchy)**:** *Fie $D \subseteq \mathbb{C}$ un domeniu și $f : D \longrightarrow \mathbb{C}$ o funcție olomorfă pe D. Fie $\overline{\Delta} \subseteq D$, unde Δ este un domeniu simplu conex, mărginit, cu frontiera γ, care este o curbă închisă, jordaniană, de clasă $\mathcal{C}^1$ pe porțiuni, orientată pozitiv.*

Atunci pentru orice $a \in \Delta$ fixat are loc:

$$f(a) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z-a}dz.$$

Principala aplicație a acestei teoreme este să ne ajute să calculăm integrale pe domenii în interiorul cărora funcția pe care o integrăm are probleme. Un exemplu:

$$\int_{|z-2i|=1} \frac{1}{z^2+4}dz.$$

Observăm că $z = 2i$ este un punct cu probleme pentru funcția considerată și aplicăm formula integrală Cauchy.

Putem rescrie integrala astfel, izolînd punctul cu probleme:

$$\int_{|z-2i|=1} \frac{1}{z^2+4}dz = \int_{|z-2i|=1} \frac{\frac{1}{z+2i}}{z-2i}dz = \frac{f(z)}{z-2i}dz,$$

unde am introdus exact funcția cu probleme, adică $f(z) = \dfrac{1}{z+2i}$.

Aplicăm formula integrală Cauchy și obținem:

$$f(2i) = \frac{1}{2\pi i}\int_{|z-2i|=1} \frac{f(z)}{z-2i}dz \Longrightarrow \int_{|z-2i|=1} \frac{f(z)}{z-2i}dz = 2\pi i f(2i) = \frac{\pi}{2}.$$

Vor exista situații cînd punctul izolat nu poate fi eliminat atît de ușor (sau chiar deloc), cazuri în care vom aplica un rezultat fundamental, *teorema reziduurilor.*

Esențialmente, formula integrală Cauchy se aplică atunci cînd avem *un singur pol simplu*, aflat în interiorul domeniului de integrare. În afara teoremei reziduurilor, pe care o vom studia imediat, mai există un rezultat care ne permite să calculăm pentru cazul cînd avem un pol multiplu:

Teoremă 2.3 (Formula integrală Cauchy pentru derivate)**:** *În condițiile și ipoteza formulei integrale Cauchy, are loc și rezultatul:*

$$f^{(n)}(z_0) = \frac{n!}{2\pi i}\int_\gamma \frac{f(z)}{(z-z_0)^{n+1}}dz.$$

Această formulă permite calculul atunci cînd avem *un pol multiplu.* În cazul în care sînt mai mulți poli, se poate aplica descompunerea în fracții elementare sau teorema reziduurilor.

2.2 Exerciții

1. Calculați integrala $\int_\Gamma z^2 dz$, unde:

(a) $\Gamma = [-1, i] \cup [i, 1]$;

(b) $\Gamma = \{z(t) = 2 + it^2 \mid 0 \le t \le 1\}$;

(c) $\Gamma = \{z(t) = t + i\cos\frac{\pi t}{2} \mid -1 \le t \le 1\}$;

(d) $\Gamma = OA$, cu $O(0,0)$ și $A(2,1)$.

Indicații: Se parametrizează drumurile și se calculează ca în exemplele de mai sus.

2. Folosind teorema Cauchy sau formula integrală Cauchy, calculați:

(a) $\int_{|z-1|=3} z^4 dz$;

(b) $\int_{|z|=4} \frac{\cos z}{z^2 - 6z + 5} dz$;

(c) $\int_{|z|=1} \frac{\sin z}{z(z-2)} dz$;

(d) $\int_\Gamma \frac{\exp(z^2)}{z^2 - 6z} dz$, unde $\Gamma : |z-2| = r, r \in \{1, 3, 5\}$;

(e) $\int_{|z|=1} \frac{\exp(3z)}{z^4} dz$;

(f) $\int_{|z-i|=1} \frac{1}{(z^2+1)^2} dz$;

(g) $\int_{|z|=1} \frac{\sin z}{z^2(z-2)} dz$.

Indicații: Ideea de bază este să identificăm punctele cu probleme ale funcțiilor de integrat în interiorul domeniilor pe care integrăm, apoi să descompunem integrandul cu o funcție căreia i se poate aplica teorema Cauchy.

(a) funcția z^4 este olomorfă, deci integrala este nulă;

(b) avem $\frac{\cos z}{(z-1)(z-5)}$, dar singurul punct cu probleme din interiorul domeniului este $z_1 = 1$. Definim $f(z) = \frac{\cos z}{z-5}$, iar integrala devine $\int_{|z|=4} \frac{f(z)}{z-1} dz$, care se calculează cu formula Cauchy.

(c) pentru $r = 1$, funcția este olomorfă, deci integrala este nulă. Pentru $r = 3$, $z = 0$ este punct cu probleme, deci definim $f(z) = \dfrac{\exp(z^2)}{z-6}$.

3. Folosind formula integrală Cauchy pentru derivate, calculați:

(a) $I_1 = \displaystyle\oint_{|z|=2} \frac{z+2}{z^3(z-1)} dz$;

(b) $I_2 = \displaystyle\oint_{|z|=\frac{1}{2}} \frac{z+2}{z^3(z-1)} dz$.

Indicație: Pentru punctul (a), observăm că ambii poli ($z = 0$ și $z = 1$) se află în interiorul domeniului, astfel că, înainte de calcul, trebuie aplicată descompunerea în fracții simple.

Pentru punctul (b), singurul pol în interiorul domeniului este $z = 0$, care este pol triplu, deci se poate aplica direct formula.

2.3 Teorema reziduurilor

Similar cu orice funcție reală, și funcțiile complexe pot fi dezvoltate în serii de puteri. În cazul complex, seriile se numesc *serii Laurent* și pot conține și puteri negative.

Informal, punctele cu probleme care ne interesează se numesc *poli* sau *puncte singulare*. Ordinul unui pol $z = a$ este multiplicitatea algebrică a rădăcinii $z = a$ în dezvoltarea în serie Laurent a funcției f. În particular, avem *poli simpli, dubli* etc.

Definiţie 2.1: Fie $f : \mathbb{C} \longrightarrow \mathbb{C}$ o funcție complexă și dezvoltarea sa în serie Laurent în jurul unui punct $z_0 \in \mathbb{C}$:

$$f(z) = \sum_{n\in\mathbb{Z}} a_n(z-z_0)^n, \quad a_n \in \mathbb{R}.$$

Se numește *reziduul* funcției f în punctul singular z_0 coeficientul a_{-1} din dezvoltarea de mai sus, notat $\mathrm{Rez}(f, z_0)$.

Următoarea teoremă ne dă metode de calcul al reziduurilor, în funcție de multiplicitatea lor:

Teoremă 2.4 (Calculul reziduurilor)**:** *(1)* $\mathrm{Rez}(f, a) = c_{-1}$, *unde* c_{-1} *este coeficientul lui* $\dfrac{1}{z-a}$ *în dezvoltarea în serie Laurent a funcției* f *în vecinătatea singularității* $z = a$.

(2) Dacă $z = a$ *este pol de ordinul* $p \geq 2$ *pentru* f*, atunci:*

$$\mathrm{Rez}(f, a) = \frac{1}{(p-1)!} \lim_{z\to a} \Big[(z-a)^p f(z)\Big]^{(p-1)};$$

(3) Dacă $z = a$ *este pol simplu pentru* f*, atunci, particularizînd formula de mai sus, avem:*

$$\mathrm{Rez}(f, a) = \lim_{z\to a}(z-a)f(z);$$

(4) Dacă f se poate scrie ca un cît de funcții, $f = \frac{A}{B}$, olomorfe în jurul lui a și dacă $z = a$ este pol simplu pentru f, adică $B(a) = 0$, atunci:

$$\mathrm{Rez}(f, a) = \lim_{z \to a} \frac{A(z)}{B'(z)}.$$

Rezultatul esențial al acestei secțiuni ne arată că, dacă integrăm o funcție cu probleme, valoarea integralei este dată în mod esențial de reziduurile sale:

Teoremă 2.5 (Teorema reziduurilor): *Fie $D \subseteq \mathbb{C}$ un domeniu și $f : D - \{\alpha_1, \dots, \alpha_k\} \longrightarrow \mathbb{C}$ o funcție olomorfă pentru care α_i sînt poli.*

Fie $K \subseteq D$ un compact cu frontiera $\Gamma = \partial K$, o curbă de clasă $\mathcal{C}^1$, jordaniană, orientată pozitiv și care conține toate α_i în interior. Atunci:

$$\int_\Gamma f(z)dz = 2\pi i \sum_{j=1}^{k} \mathrm{Rez}(f, \alpha_j).$$

De exemplu, să calculăm integrala:

$$I = \int_{|z|=r} \frac{e^z}{(z-i)(z-2)} dz, r > 0, r \neq 1, 2.$$

Soluție: Dacă $0 < r < 1$, putem aplica teorema lui Cauchy (2.1) și găsim $I = 0$.

Dacă $1 < r < 2$, aplicăm formula integrală a lui Cauchy (2.2) și găsim:

$$\int_{|z|=r} \frac{e^z}{(z-i)(z-2)} dz = \int_{|z|=r} \frac{\frac{e^z}{z-2}}{z-i} dz = \int_{|z|=r} \frac{f(z)}{z-i} dz,$$

unde $f(z) = \frac{e^z}{z-2}$. Rezultă:

$$\int_{|z|=r} \frac{f(z)}{z-i} dz = 2\pi i f(i) = 2\pi i \frac{e^i}{i-2}.$$

Dacă $r > 2$, aplicăm teorema reziduurilor, cu i și 2 poli simpli. Avem:

$$\mathrm{Rez}(f, i) = \lim_{z \to i} (z-i) \frac{e^z}{(z-i)(z-2)} = \frac{e^i}{i-2}$$

$$\mathrm{Rez}(f, 2) = \lim_{z \to 2} \frac{e^z}{(z-i)(z-2)} = \frac{e^2}{2-i}.$$

Rezultă, din teorema reziduurilor:

$$\int_{|z|=r} \frac{e^z}{(z-i)(z-2)} dz = 2\pi i \left(\frac{e^i}{i-2} + \frac{e^2}{2-i} \right).$$

2.4 Exerciții

1. Calculați reziduurile funcțiilor în punctele a indicate:

(a) $f(z) = \dfrac{\exp\left(z^2\right)}{z-1}, a = 1;$

(b) $f(z) = \dfrac{\exp\left(z^2\right)}{(z-1)^2}, a = 1;$

(c) $f(z) = \dfrac{z+2}{z^2-2z}, a = 0;$

(d) $f(z) = \dfrac{1+e^z}{z^4}, a = 0;$

(e) $f(z) = \dfrac{\sin z}{4z^2}, a = 0;$

(f) $f(z) = \dfrac{z}{1-\cos z}, a = 0.$

Indicație: Verificăm multiplicitatea polului $z = a$ și aplicăm formula corespunzătoare din teorema 2.4.

2. Să se calculeze următoarele integrale:

(a) $I = \displaystyle\int_{|z|=2} \frac{dz}{z^2-1};$

(b) $I = \displaystyle\int_{\gamma} \frac{dz}{z^4+1}, \gamma : x^2 + y^2 - 2x = 0;$

(c) $I = \displaystyle\int_{|z|=3} \frac{z^2+1}{(z-1)^2(z+2)} dz.$

Soluție: (a) Punctele $z = \pm 1$ sînt poli de ordinul 1 pentru funcția $f(z) = \dfrac{1}{z^2-1}$. Ele sînt situate în interiorul discului pe care integrăm, cu $|z| = 2$, deci putem aplica teorema reziduurilor:

$$I = 2\pi i \cdot \Big(\mathrm{Rez}(f, z_1) + \mathrm{Rez}(f, z_2)\Big).$$

Calculăm separat reziduurile:

$$\mathrm{Rez}(f, z_1) = \lim_{z \to 1}(z-1) \cdot \frac{1}{z^2-1} = \frac{1}{2}$$
$$\mathrm{Rez}(f, z_2) = \lim_{z \to -1}(z+1) \cdot \frac{1}{z^2-1} = -\frac{1}{2}.$$

Rezultă:

$$I = \int_{|z|=2} \frac{dz}{z^2-1} = 2\pi i \cdot \left(\frac{1}{2} - \frac{1}{2}\right) = 0.$$

(b) Curba γ este un cerc centrat în $(1,0)$ și cu raza 1. Căutăm polii funcției $f(z) = \dfrac{1}{z^4+1}$ care se află în interiorul lui γ.

Avem succesiv:

$$\begin{aligned} z^4+1 = 0 &\Longrightarrow z^4 = -1 = \cos\pi + i\sin\pi \Longrightarrow \\ z &= \sqrt[4]{\cos\pi + i\sin\pi} \Longrightarrow \\ z_k &= \cos\frac{\pi+2k\pi}{4} + i\sin\frac{\pi+2k\pi}{4}, \, k = 0,1,2,3. \end{aligned}$$

Doar punctele z_0, z_3 se află în interiorul discului delimitat de γ și calculăm reziduurile în aceste puncte.

Putem aplica formula din Teorema 2.4 (4) și avem:

$$\begin{aligned} \operatorname{Rez}(f, z_0) &= \frac{A(z)}{B'(z)}\Big|_{z_0} = \frac{1}{4z^3}\Big|_{z_0} = -\frac{1}{4}e^{\frac{\pi i}{4}} \\ \operatorname{Rez}(f, z_3) &= \frac{A(z)}{B'(z)}\Big|_{z_3} = \frac{1}{4z^3}\Big|_{z_3} = -\frac{1}{4}e^{\frac{7\pi i}{4}}. \end{aligned}$$

Rezultă:

$$I = \int_\gamma \frac{dz}{z^4+1} = 2\pi i\left(-\frac{1}{4}e^{\frac{\pi i}{4}} - \frac{1}{4}e^{\frac{7\pi i}{4}}\right) = -\frac{\pi\sqrt{2}i}{2}.$$

(c) Avem doi poli, $z = 1, z = -2$ în interiorul conturului. Se vede că $z_1 = 1$ este pol de ordinul 2, iar $z_2 = -2$ este pol de ordinul 1. Calculăm reziduurile:

$$\begin{aligned} \operatorname{Rez}(f, z_1) &= \frac{1}{2!}\lim_{z\to 1}\left[(z-1)^2\frac{z^2+1}{(z-1)^2(z+2)}\right] \\ &= \frac{1}{2!}\lim_{z\to 1}\frac{z^2+4z-1}{(z+2)^2} \\ &= \frac{2}{9} \\ \operatorname{Rez}(f, z_2) &= \lim_{z\to -2}(z+2)\frac{z^2+1}{(z-1)^2(z+2)} = \frac{5}{9}. \end{aligned}$$

Rezultă:

$$I = \int_{|z|=3} \frac{z^2+1}{(z-1)^2(z+2)}\,dz = \frac{14}{9}\pi i.$$

3. Să se calculeze integralele:

(a) $\displaystyle\int_{|z|=1} \frac{dz}{\sin z}$;

(b) $\displaystyle\int_{|z|=1} \frac{e^z}{z^2}\,dz$;

(c) $\displaystyle\int_{|z|=5} z e^{\frac{3}{z}}\,dz$;

(d) $\displaystyle\int_{|z-1|=1} \frac{dz}{(z+1)(z-1)^3}$;

(e) $\displaystyle\int_{|z|=1} \frac{\sin z}{z^4}\,dz$.

Indicații: (a) $z = 0$ este singurul pol din interiorul domeniului;

(b), (c): Dezvoltăm în serie Laurent și identificăm reziduurile folosind definiția.

(d) Avem $z = 1$ pol de ordin 3 și $z = -1$ pol simplu. Doar $z = 1$ se află în interiorul domeniului și dezvoltăm în serie Laurent după puterile lui $z - 1$:

$$\begin{aligned}
\frac{1}{z+1} &= \frac{1}{2-(-(z-1))} \\
&= \frac{1}{2}\frac{1}{1-\left(-\frac{z-1}{2}\right)} \\
&= \frac{1}{2}\sum_{n\geq 0}(-1)^n\frac{(z-1)^n}{2^n} \\
&= \sum_{n\geq 0}(-1)^n\frac{(z-1)^n}{2^{n+1}},
\end{aligned}$$

pentru $|z-1| < 2$.

Rezultă:

$$\frac{1}{(z+1)(z-1)^3} = \sum_{n\geq 0}(-1)^n\frac{(z-1)^{n-3}}{2^{n+1}} = \frac{1}{2(z-1)^3} - \frac{1}{4(z-1)^2} + \frac{1}{8(z-1)} - \frac{1}{16} + \dots,$$

deci $\mathrm{Rez}(f, 1) = \dfrac{1}{8}$.

2.5 Aplicații ale teoremei reziduurilor

Putem folosi teorema reziduurilor pentru a calcula integrale trigonometrice de forma:

$$I = \int_0^{2\pi} R(\cos\theta, \sin\theta) d\theta,$$

unde R este o funcție rațională.

Facem schimbarea de variabilă $z = e^{i\theta}$ și atunci, pentru $\theta \in [0, 2\pi]$, z descrie cercul $|z| = 1$, o dată, în sens direct.

Folosim formulele lui Euler:

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} = \frac{1}{2}\left(z + \frac{1}{z}\right)$$
$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} = \frac{1}{2i}\left(z - \frac{1}{z}\right).$$

Atunci, dacă $z = e^{i\theta}$, rezultă $dz = ie^{i\theta}\, d\theta = izd\theta$, iar integrala devine:[1]

$$I = \int_0^{2\pi} R(\cos\theta, \sin\theta) d\theta = \oint_{|z|=1} R_1(z) dz,$$

unde:

$$R_1(z) = \frac{1}{iz} R\left(\frac{z^2+1}{2z}, \frac{z^2-1}{2iz}\right).$$

Această funcție poate avea poli și deci putem folosi teorema reziduurilor. Dacă $a_1, \ldots, a_n$ sînt polii din interiorul cercului unitate, avem:

$$I_1 = 2\pi i \sum_{k\geq 1} \mathrm{Rez}(R_1, a_k).$$

Să vedem cîteva exemple:

(a) $\displaystyle\int_0^{2\,\pi} \frac{d\theta}{2 + \cos\theta}$;

(b) $\displaystyle\int_0^{2\,\pi} \frac{d\theta}{1 + 3\cos^2\theta}$;

(c) $\displaystyle\int_0^{2\pi} \frac{\sin\theta}{1 + i\sin\theta}\, d\theta$;

[1] $\oint$ marchează o integrală pe un contur *închis*

(d) $\displaystyle\int_0^{2\pi} \frac{d\theta}{1 + a\sin\theta}, |a| < 1, a \in \mathbb{R};$

(e) $\displaystyle\int_{-\pi}^{\pi} \frac{\cos 3t}{5 - 4\cos t} dt.$

Soluție:

(a) Notăm $z = e^{i\theta}$, cu $\theta \in [0, 2\pi]$. Atunci avem succesiv:

$$\begin{aligned}
dz &= ie^{i\theta} d\theta = iz dz \Longrightarrow d\theta = \frac{dz}{iz} \\
\cos\theta &= \frac{e^{i\theta} + e^{-i\theta}}{2} = \frac{1}{2}\left(z + \frac{1}{z}\right) \\
\frac{1}{2 + \cos\theta} &= \frac{2z}{z^2 + 4z + 1} \\
\int_0^{2\pi} \frac{d\theta}{2 + \cos\theta} &= \oint_{|z|=1} \frac{2z}{z^2 + 4z + 1} \frac{dz}{iz} \\
&= -2i \oint_{|z|=1} \frac{dz}{z^2 + 4z + 1}.
\end{aligned}$$

Acum folosim teorema reziduurilor. Singularitățile funcției $f(z) = \dfrac{1}{z^2 + 4z + 1}$ sînt $z = -2 \pm \sqrt{3}$, care sînt poli simpli. Numai $z = -2 + \sqrt{3}$ se află în interiorul cercului $|z| = 1$ și calculăm reziduul folosind Teorema 2.4(2).

SEMINAR 3

SERII LAURENT ȘI REZIDUURI

3.1 Serii de puteri. Serii Laurent

Noțiunile privitoare la scrierea funcțiilor cu ajutorul seriilor de puteri, întîlnite în cazul real prin serii Taylor, se pot generaliza și în cazul complex. Situația este ceva mai problematică în acest caz, deoarece argumentele sînt obiecte bidimensionale. În orice caz, multe dintre noțiunile din cazul real pot fi preluate aproape fără modificare și în cazul complex.

Definiţie 3.1: O serie de forma $\sum_{n\geq 0} a_n(z-z_0)^n$, cu $z \in \mathbb{C}$ oarecare, $z_0 \in \mathbb{C}$ fixat și $a_n \in \mathbb{C}$ coeficienți, pentru orice $n \in \mathbb{N}$, se numește *serie de puteri* centrată în z_0.

Ca în cazul real, se poate calcula *raza de convergență* a seriilor de puteri cu una dintre formulele:

$$R = \frac{1}{\overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|}} = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right|.$$

Atunci, deoarece lucrăm în planul complex, se definește *discul de convergență* al seriei prin:

$$B(z_0, R) = \{z \in \mathbb{C} \mid |z - z_0| < R\}.$$

Știm, de asemenea, că:

- seria este absolut convergentă pe $|z - z_0| < R$;
- seria este divergentă pe $|z - z_0| > R$;
- seria este uniform convergentă pe $|z - z_0| \leq \rho$, pentru orice $\rho < R$,

iar pe frontiera discului trebuie testat separat.

În interiorul discului de convergență, suma seriei este o funcție olomorfă $S(z)$ și au sens rezultate de forma derivării sau integrării termen cu termen.

Definiție 3.2: O funcție $f : A \longrightarrow \mathbb{C}$ se numește *analitică* dacă poate fi dezvoltată într-o serie de puteri complexă cu discul de convergență inclus în A.

Dacă acesta este cazul, avem formula cunoscută:

$$f(z) = \sum_{n \geq 0} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n.$$

Are loc rezultatul important:

Teoremă 3.1 (Weierstrass-Riemann-Cauchy)**:** *O funcție complexă definită pe o mulțime deschisă este analitică dacă și numai dacă este olomorfă.*

Deducem de aici că, dacă domeniul de definiție A al funcției complexe studiate, $f : A \longrightarrow \mathbb{C}$, este o mulțime deschisă, atunci olomorfia (pe care o verificăm cu condițiile Cauchy-Riemann, de exemplu) este echivalentă cu analiticitatea, verificată cu ajutorul seriilor de puteri. Rezultă implicit că orice funcție olomorfă definită pe un deschis poate fi dezvoltată în serie de puteri.

Dar pentru cazul complex, avem și serii ceva mai complicate:

Definiție 3.3: O serie de puteri de forma $\sum_{n \in \mathbb{Z}} a_n(z - z_0)^n$, cu $a_n, z, z_0 \in \mathbb{C}$ se numește *serie Laurent* centrată în punctul $z_0 \in \mathbb{C}$.

O serie Laurent este convergentă dacă și numai dacă atît partea pozitivă $a_n > 0$, cît și partea negativă, $a_n \leq 0$ sînt simultan convergente.

Partea pozitivă se numește *partea Taylor*, iar partea negativă se numește *partea principală.*

Pentru cazul Taylor, seriile uzuale pentru funcțiile elementare, împreună cu domeniile de convergență, sînt date mai jos. În toate cazurile, $x \in \mathbb{R}$ se poate înlocui cu $z \in \mathbb{C}$.

$$
\begin{aligned}
e^x &= \sum_{n \geq 0} \frac{1}{n!} x^n, x \in \mathbb{R} \\
\frac{1}{1-x} &= \sum_{n \geq 0} x^n, |x| < 1 \\
\frac{1}{1+x} &= \sum_{n \geq 0} (-1)^n x^n, |x| < 1 \\
\cos x &= \sum_{n \geq 0} \frac{(-1)^n}{(2n)!} x^{2n}, x \in \mathbb{R} \\
\sin x &= \sum_{n \geq 0} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, x \in \mathbb{R} \\
(1+x)^\alpha &= \sum_{n \geq 0} \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!} x^n, |x| < 1, \alpha \in \mathbb{R} \\
\arctan x &= \sum_{n \geq 0} \frac{(-1)^n}{2n+1} x^{2n+1}, |x| \leq 1.
\end{aligned}
$$

Seriile pentru alte funcții se pot obține fie prin calcul direct, fie prin substituții, fie prin derivare sau integrare termen cu termen a seriilor de mai sus.

3.2 Singularități și reziduuri

În cazul funcțiilor reale, domeniul de definiție trebuie ales atent astfel încît să evităm „punctele cu probleme“. Exemplele tipice sînt cele în care se anulează numitorul unei fracții, cînd apar logaritmi sau radicali etc. Pentru funcții reale, cel mult putem calcula asimptote verticale în acele „puncte cu probleme“. Dar în cazul complex, distincția se face mult mai fin.

Definiție 3.4: Fie $A \subseteq \mathbb{C}$ o mulțime deschisă nevidă și $f : A \longrightarrow \mathbb{C}$ o funcție complexă, olomorfă pe A. Fie un punct $z_0 \in \mathbb{C}$.

Punctul z_0 se numește *punct singular izolat* (*singularitate izolată*) pentru f dacă există un disc centrat în z_0, de forma $B(z_0, r)$, cu $r \neq 0$, astfel încît, dacă eliminăm centrul discului, funcția este olomorfă pe $B(z_0, r) \setminus \{z_0\}$.

Amintim că, dacă funcția este olomorfă pe *discul punctat* (i.e. discul $B(z_0, r)$, din care am scos centrul), atunci ea are o dezvoltare în serie Laurent, conform teoremei Weierstrass-Riemann-Cauchy.

Definiție 3.5: În condițiile și cu notațiile din definiția anterioară, punctul $z_0 \in \mathbb{C}$ se numește *punct singular izolat* (*singularitate izolată*) dacă seria Laurent asociată funcției f are partea principală nulă, adică este o serie Taylor.

Definiţie 3.6: În condițiile și cu notațiile din definiția anterioară, punctul $z_0 \in \mathbb{C}$ se numește *pol* dacă în seria Laurent asociată funcției f există un număr finit de termeni nenuli în partea principală. Indicele ultimului termen nenul m (în sensul că $|m|$ este cel mai mare, iar $m < 0$) se numește *ordinul polului.*

Definiţie 3.7: În condițiile și cu notațiile din definiția anterioară, punctul $z_0 \in \mathbb{C}$ se numește *punct singular esențial* (*singularitate esențială*) dacă în partea principală a seriei Laurent asociată funcției f există o infinitate de termeni nenuli.

Idee intuitivă: Singularitățile, în general, sînt „puncte cu probleme". Chiar și în cazul real există această noțiune, întîlnită în analiză. De exemplu, un punct unghiular sau de întoarcere al unei funcții reale se numește singularitate. Alt exemplu la îndemînă: găurile negre sînt interpretate în cosmologie (ramura fizicii teoretice care utilizează geometria pentru studiul „formei" Universului) ca *singularități ale spațiu-timpului.* Dar, din fericire, în majoritatea cazurilor, singularitățile fie pot fi „ignorate" sau „eliminate" i.e. se poate extrage informația de bază și din domeniul din care ele au fost scoase. Amintiți-vă, de exemplu, că în analiza de liceu există o teoremă care afirmă că *dacă se scoate un număr numărabil de puncte din graficul unei funcții, integrala sa definită nu se schimbă.*

Similar este și cazul singularităților din analiza complexă: toate, cu excepția celor esențiale, pot fi „eliminate", adică se poate extrage informație foarte importantă despre comportamentul funcțiilor și în lipsa lor (sau, uneori, chiar *numai* din ele, ca în cazul reziduurilor, precum vom vedea).

Polii vor fi elementul central de studiu mai departe. Pentru ei, mai avem o caracterizare utilă:

Propoziţie 3.1: *În condițiile și cu notațiile de mai sus, punctul $z_0 \in \mathbb{C}$ este pol pentru funcția f dacă și numai dacă $\lim_{z \to z_0} f(z) = \infty$.*

Exemplu: Putem vedea foarte simplu acest lucru pentru funcția $f(z) = \dfrac{z}{z-1}$. Atunci $z = 1$ este pol simplu, deoarece $\lim_{z \to 1} f(z) = \infty$ și mai mult, putem scrie dezvoltarea în serie Laurent a funcției f în jurul lui $z = 1$, sub forma:

$$f(z) = \frac{1 + z - 1}{z - 1} = \frac{1}{z - 1} + 1,$$

care are partea principală $\dfrac{1}{z-1}$, cu un singur termen (care dă ordinul polului).

Propoziţie 3.2: *În condițiile și cu notațiile de mai sus, punctul $z_0 \in \mathbb{C}$ este singularitate esențială dacă și numai dacă nu există $\lim_{z \to z_0} f(z)$.*

Ajungem la o altă noțiune esențială:

Definiţie 3.8: În condițiile și cu notațiile de mai sus, fie z_0 o singularitate izolată a funcției f. Coeficientul a_{-1} din seria Laurent a funcției f în coroana $B(z_0; 0, r)$ se numește *reziduul* funcției în z_0 și se notează $\text{Rez}(f, z_0)$.

Observaţie 3.1: În multe situații, reziduurile se pot calcula simplu cu teorema 2.4, dar uneori sîntem obligați să folosim definiția, pentru că limitele complexe se calculează mult mai dificil decît în cazul real.

Un exemplu este următorul: considerăm funcția $f(z) = \dfrac{1}{z \sin z^2}$. Vrem să calculăm $\text{Rez}(f, 0)$. Evident, acest punct este pol, deoarece $\lim\limits_{z \to 0} f(z) = \infty$, dar orice formulă am folosi din teorema 2.4, nedeterminarea nu este eliminată. Astfel că sîntem obligați să folosim definiția, cu ajutorul seriei Laurent.

Ca în cazul seriei Taylor pentru funcția sinus, avem:

$$\begin{aligned} \sin z &= \frac{z}{1!} - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots \\ \Longrightarrow z \sin z^2 &= z \cdot \left(\frac{z^2}{1!} - \frac{z^6}{3!} + \frac{z^{10}}{5!} - \dots \right) \\ &= z^3 \left(1 - \frac{z^4}{3!} + \frac{z^8}{5!} - \dots \right). \end{aligned}$$

Atunci, inversînd, obținem funcția scrisă sub forma:

$$f(z) = z^{-3} \cdot \left(1 - \frac{z^4}{3!} + \frac{z^8}{5!} - \dots \right)^{-1}.$$

Vrem să inversăm seria din paranteză (admitem fără demonstrație că acest lucru este posibil), pentru ca în final, să putem scrie toată funcția ca o serie de puteri (în forma actuală nu este așa ceva!). Așadar, căutăm o serie de forma $\sum a_n z^n$, cu $n \in \mathbb{N}$, astfel încît:

$$\left(1 - \frac{z^4}{3!} + \frac{z^8}{5!} - \dots \right) \cdot \left(a_0 + a_1 z + a_2 z^2 + \dots \right) = 1.$$

Prin identificarea coeficienților, rezultă $a_0 = 1, a_2 = 0$. Făcînd acum înmulțirea, avem:

$$f(z) = \frac{1}{z^3} \sum_{n \geq 0} a_n z^n = \frac{a_0}{z^3} + \frac{a_1}{z^2} + \frac{a_2}{z} + a_3 + \dots.$$

Partea principală a seriei are 3 termeni nenuli, deci $z = 0$ este pol triplu (intuitiv, simplu pentru z și dublu pentru $\sin z^2$), deci rezultă $\text{Rez}(f, 0) = a_2 = 0$.

3.3 Exerciții

1. Determinați dezvoltarea în serie Taylor sau Laurent pentru funcțiile următoare, în jurul punctelor indicate z_0, precizînd și domeniile pe care sînt valabile:

(a) $f(z) = \cos^2 z$, $z_0 = 0$;

(b) $f(z) = \dfrac{z+3}{z^2 - 8z + 15}$, $z_0 = 4$;

(c) $f(z) = \sin \dfrac{1}{1-z}$, $z_0 = 1$;

(d) $f(z) = \dfrac{2z^2 + 9z + 5}{z^3 + z^2 - 8z - 12}$, $z_0 = 1$;

(e) $f(z) = \dfrac{z-1}{z-2}$, $z_0 = 1$;

(f) $f(z) = \dfrac{1}{z^2 - 3z + 2}$, pe domeniile $|z| < 1$, $1 < |z| < 2$ și $|z| > 2$.

Indicație: Se pot folosi seriile uzuale, cu anumite substituții, dar atenție la domeniile de convergență, precum și la eventualii poli ai funcțiilor. A se vedea exemplul rezolvat de mai jos.

2. Determinați punctele singulare și precizați natura lor pentru funcțiile:

(a) $f(z) = \dfrac{z^5}{(z^2+1)^2}$;

(b) $f(z) = z^3 \exp\left(-\dfrac{3}{z}\right)$;

(c) $f(z) = \sin \dfrac{1}{z}$;

(d) $f(z) = \dfrac{\sin 2z}{z^6}$;

(e) $f(z) = \dfrac{6z+1}{z-3}$;

(f) $f(z) = \dfrac{1}{z^2 \sin z}$;

Indicație: După identificarea „punctelor cu probleme“, se cercetează existentă limitei către punctele respective și/sau se utilizează dezvoltarea în serie Taylor/Laurent.

3. Calculați:

(a) $\displaystyle\int_{|z|=1} \frac{z^2 + \exp(z)}{z^3} dz;$

(b) $\displaystyle\int_{\gamma} (z+1)\exp\left(\frac{1}{z+1}\right) dz$, unde $\gamma \,:\, x^2 + y^2 + 2x = 0$;

(c) $\displaystyle\int_{|z|=r} \frac{\exp(z^{-1})}{1-z} dz$, pentru $r \in \left\{\frac{1}{2}, 1, 2\right\}$.

Indicație: Se folosește teorema reziduurilor.

Exemplu rezolvat pentru seria Laurent: Să se dezvolte funcția

$$f(z) = \frac{2z^2 + 3z - 1}{z^3 + z^2 - z - 1}$$

în jurul originii și în jurul punctelor $z = \pm 1$.

Soluție: Numitorul se descompune sub forma $(z-1)(z+1)^2$, deci avem un pol simplu $z = 1$ și un pol dublu $z = -1$.

Putem descompune funcția în fracții simple, sub forma:

$$f(z) = \frac{1}{z-1} + \frac{1}{z+1} + \frac{1}{(z+1)^2}.$$

Primii doi termeni sînt sume de serii geometrice, iar al treilea poate fi obținut prin derivarea seriei geometrice:

$$\frac{1}{z+1} = \sum (-1)^n z^n \Rightarrow -\frac{1}{(z+1)^2} = \sum (-1)^{n+1}(n+1) z^n,$$

după ce am făcut trecerea $n \mapsto n+1$, întrucît seria derivatelor ar fi pornit de la $n = 1$, dat fiind termenul z^{n-1}.

Pe cazuri acum:

Pentru $|z| < 1$: funcția este olomorfă, deoarece nu avem niciun pol în acest disc deschis. Atunci putem scrie direct seria ca sumă a celor trei serii corespunzătoare:

$$f(z) = -\sum_{n\geq 0} z^n + \sum_{n\geq 0} (-1)^n z^n + \sum_{n\geq 0} (-1)^n (n+1) z^n = \sum_{n\geq 0} z^n \left(-1 + (-1)^n + (-1)^n(n+1)\right).$$

În jurul punctului $z = -1$, unde avem pol, trebuie să rescriem funcția sub forma:

$$f(z) = \frac{1}{(z+1)^2} + \frac{1}{z+1} - \frac{1}{2} \cdot \frac{1}{1 - \frac{z+1}{2}},$$

astfel punînd în evidență puteri (negative!) ale lui $z+1$. Primii doi termeni nu necesită prelucrare, iar pentru al treilea, întrucît ne aflăm în domeniul de convergență a seriei geometrice, adică $|z+1| < 2$ putem scrie:

$$\frac{1}{1 - \frac{z+1}{2}} = \sum_{n\geq 0} \left(\frac{z+1}{2}\right)^n.$$

În jurul punctului $z = 1$, unde avem din nou pol, rescriem funcția sub forma:

$$f(z) = \frac{1}{z-1} + \frac{1}{2} \cdot \frac{1}{1 + \frac{z-1}{2}} + \frac{1}{4} \cdot \frac{1}{\left(1 + \frac{z-1}{2}\right)^2}.$$

Primul termen nu necesită prelucrare, iar pentru al doilea, remarcăm că ne aflăm în interiorul domeniului său de convergență, căci $|z-1| < 2$ și atunci avem direct suma seriei geometrice:

$$\frac{1}{1 + \frac{z-1}{2}} = \sum_{n \geq 0} (-1)^n \left(\frac{z-1}{2}\right)^n.$$

Al treilea termen se obține prin derivarea termen cu termen a seriei de mai sus (ca în cazul de la început) și avem:

$$\frac{1}{\left(1 + \frac{z-1}{2}\right)^2} = \sum_{n \geq 0} (-1)^n \frac{(n+1)(z-1)^n}{2^n}$$

și în fine, funcția se obține ca sumă a acestor trei serii, adică, după calcule:

$$f(z) = \frac{1}{z-1} + \sum_{n \geq 0} \frac{n+3}{2^{n+2}} (z-1)^n.$$

SEMINAR 4

RECAPITULARE ANALIZĂ COMPLEXĂ

1. Determinați funcția olomorfă $f = P + iQ$ în cazurile:

(a) $Q(x, y) = \ln(x^2 + y^2) + x - 2y$;

(b) $P(x, y) = x^2 - y^2$ și știind $f(0) = 0$;

(c) $P(x, y) = x^2 - y^2 - 2y$;

(d) $Q(x, y) = x^3 - 3xy^2$;

2. Calculați:

(a) $\exp\left(-\frac{5\pi i}{2}\right)$;

(b) $\mathrm{Ln}(-2)$;

(c) i^{3i};

(d) i^{1+i};

(e) $(1 + i\sqrt{3})^{42}$.

3. Să se dezvolte în serie de puteri ale lui z funcția $f(z) = \frac{1}{z^2 - 3z + 2}$ în domeniile $|z| < 1$, $1 < |z| < 2$ și $|z| > 2$.

4. Determinați singularitățile, precizați natura lor și calculați reziduurile pentru funcțiile:

(a) $f(z) = \dfrac{z^5}{(z^2+1)^2}$;

(b) $f(z) = \dfrac{\sin z}{z^2}$;

(c) $f(z) = z^2 \cdot \cos \dfrac{1}{z}$;

(d) $f(z) = \dfrac{\exp(z^{-1})}{z^4}$.

5. Calculați integralele complexe:

(a) $\displaystyle\int_0^{1+i} z^2 dz$;

(b) $\displaystyle\int_C (z^2 + \overline{z}^2) dz$, unde C este segmentul care unește punctele 0 și $2 + 2i$.

(c) $\displaystyle\int_{|z|=3} \overline{z}^2 dz$;

(d) $\displaystyle\int_0^{2\pi} \frac{\cos 2\theta}{2 + \cos^2\theta} d\theta$;

(e) $\displaystyle\int_0^{2\pi} \frac{1 + \sin\theta}{(3 - 2\cos\theta)^2} d\theta$;

(f) $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{(1+x^4)^2}$;

(g) $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{(1+x^2)^4}$.

Indicații: (f,g): Se definește un contur închis de forma $C = [-r, r] \cup \Gamma_r$, unde Γ_r este cercul centrat în origine și de rază r. Considerăm funcția complexă asociată $f(z) = \dfrac{1}{(1+z^4)^2}$ sau, respectiv, $f(z) = \dfrac{1}{(1+z^2)^4}$ și se calculează reziduurile pentru integrarea lui $f(z)$ pe C. Integrala reală revine la a calcula limita $r \longrightarrow \infty$ a integralei complexe, deoarece conform lemei Jordan, integrala pe semicercul Γ_r este nulă. Rămîne calculul cu reziduuri.

Pe scurt:

Luăm $\Gamma : (-r, r) \cup \gamma$, unde $\gamma : |z| = r, \mathrm{Im}(z) > 0$ și fie $f(z) = \dfrac{P(z)}{Q(z)}$.

Calculăm:

$$\begin{aligned} I_r &= \int_\Gamma f(z)dz = \int_{-r}^{r} f(x)dx + \int_\gamma f(z)dz \\ &= 2\pi i \sum_{k=1}^{n} \operatorname{Rez}(f, z_k), \end{aligned}$$

Apoi:

$$\lim_{r\to\infty} I_r = \int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx + \lim_{r\to\infty} \int_\gamma f(z)dz.$$

Și, deoarece $\lim\limits_{z\to\infty} |zf(z)| = 0$, rezultă, cu lema Jordan, că integrala pe semicerc e nulă, adică $\lim\limits_{r\to\infty} \int_\gamma f(z)dz = 0$, deci integrala reală căutată este egală cu rezultatul obținut cu teorema reziduurilor.

SEMINAR 5

TRANSFORMATA LAPLACE

5.1 Definiții și proprietăți

Transformata Laplace este o transformare integrală, care se poate aplica unor funcții speciale, numite *funcții original.*

Definiție 5.1: O funcție $f : \mathbb{R} \longrightarrow \mathbb{C}$ se numește *original* dacă:

(a) $f(t) = 0$ pentru orice $t < 0$;

(b) f este continuă (eventual pe porțiuni) pe intervalul $[0, \infty)$;

(c) Este mărginită de o exponențială, adică există $M > 0$ și $s_0 \geq 0$ astfel încît:

$$|f(t)| \leq Me^{s_0 t}, \quad \forall t \geq 0.$$

Numărul s_0 se mai numește *indicele de creștere* al funcției f.

Vom nota cu $\mathcal{O}$ mulțimea funcțiilor original.

Pornind cu o funcție original, definiția transformatei Laplace este:

Definiție 5.2: Păstrînd contextul și notațiile de mai sus, fie $f \in \mathcal{O}$ și mulțimea:

$$S(s_0) = \{s \in \mathbb{C} \mid \text{Re}(s) > s_0\}.$$

Funcția:

$$F : S(s_0) \longrightarrow \mathbb{C}, \quad F(s) = \int_0^\infty f(t)e^{-st}dt$$

se numește *transformata Laplace* a lui f sau *imaginea Laplace* a originalului f.

Vom mai folosi notația $F = \mathcal{L}f$ sau, explicit, $\mathcal{L}f(t) = F(s)$.

Proprietățile esențiale ale transformatei Laplace sînt date mai jos. Fiecare dintre ele va fi folosită pentru a calcula o transformată Laplace pentru o funcție care nu se regăsește direct într-un tabel de valori.

- **Liniaritate:** $\mathcal{L}(\alpha f + \beta g) = \alpha \mathcal{L} f + \beta \mathcal{L} g$, pentru orice $\alpha, \beta \in \mathbb{C}$, iar f, g funcții original;
- **Teorema asemănării:**
$$\mathcal{L} f(\alpha t) = \frac{1}{\alpha} F\left(\frac{s}{\alpha}\right);$$
- **Teorema deplasării:**
$$\mathcal{L}\left(f(t) e^{s_0 t}\right) = F(s - s_0);$$
- **Teorema întîrzierii:** Definim *întîrziata cu* τ a funcției $f \in \mathcal{O}$ prin:
$$f_\tau(t) = \begin{cases} 0, & t < \tau \\ f(t - \tau), & t \geq \tau \end{cases}.$$
Atunci, dacă $\mathcal{L} f(t) = F(s)$, $\mathcal{L} f_\tau(t) = e^{-s\tau} F(s)$;
- **Teorema derivării imaginii:**
$$\mathcal{L}\left(t^n f(t)\right) = (-1)^n F^{(n)}(s);$$
- **Teorema integrării originalului:** Fie $f \in \mathcal{O}$, $\mathcal{L} f(t) = F(s)$ și $g(t) = \int_0^t f(\tau) d\tau$. Atunci:
$$\mathcal{L} g(t) = \frac{1}{s} F(s);$$
- **Teorema integrării imaginii:** Fie $\mathcal{L} f(t) = F(s)$ și G o primitivă a lui F în $S(s_0)$, cu $G(\infty) = 0$. Atunci:
$$\mathcal{L} \frac{f(t)}{t} = -G(s).$$
- **Teorema de convoluție:** Definim *produsul de convoluție* al două funcții, $h = f * g$, prin formula:
$$h(t) = \int_0^t f(\tau) \cdot g(t - \tau) d\tau$$
și fie F, G transformatele Laplace ale funcțiilor f și respectiv g, presupuse original. Dacă H este imaginea Laplace a funcției h, atunci $H(s) = F(s) \cdot G(s)$. Cu alte cuvinte, transformata Laplace acționează ca înmulțire pe produsul de convoluție.

5.2 Tabel de transformate Laplace

În tabelul de mai jos, vom considera funcțiile $f(t)$ ca fiind funcții original, adică nule pentru argument negativ. Echivalent, putem gîndi $f(t)$ ca fiind, de fapt, înmulțite cu *funcția lui Heaviside*:

$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}.$$

$f(t) = \mathcal{L}^{-1}(F(s))$	$F(s) = \mathcal{L}(f(t))$
$u(t)$	$\frac{1}{s}$
$u(t-\tau)$	$\frac{1}{s}e^{-\tau s}$
t	$\frac{1}{s^2}$
t^n	$\frac{n!}{s^{n+1}}$
$t^n e^{-\alpha t}$	$\frac{n!}{(s+\alpha)^{n+1}}$
$e^{-\alpha t}$	$\frac{1}{s+\alpha}$
$\sin(\omega t)$	$\frac{\omega}{s^2+\omega^2}$
$\cos(\omega t)$	$\frac{s}{s^2+\omega^2}$
$\sinh(\alpha t)$	$\frac{\alpha}{s^2-\alpha^2}$
$\cosh(\alpha t)$	$\frac{s}{s^2-\alpha^2}$
$e^{-\alpha t}\sin(\omega t)$	$\frac{\omega}{(s+\alpha)^2+\omega^2}$
$e^{-\alpha t}\cos(\omega t)$	$\frac{s+\alpha}{(s+\alpha^2)+\omega^2}$
$\ln t$	$-\frac{1}{s}\left(\ln s + \gamma\right)$[1]

[1]Constanta Euler-Mascheroni, $\gamma \simeq 0,577\dots \in \mathbb{R} - \mathbb{Q}$

5.3 Exerciții

1. Calculați transformatele Laplace pentru funcțiile (presupuse original):

(a) $f(t) = 1, t \geq 0$;

(b) $f(t) = t, t \geq 0$;

(c) $f(t) = t^n, n \in \mathbb{N}$;

(d) $f(t) = e^{at}, t \geq 0, a \in \mathbb{R}$;

(e) $f(t) = \sin(at), t \geq 0, a \in \mathbb{R}$.

Soluție: (a) Avem direct din definiție:

$$F(s) = \int_0^\infty e^{-st} dt = \lim_{n \to \infty} \int_0^n e^{-st} dt = \frac{1}{s}, s > 0.$$

(b) Integrăm prin părți și obținem $F(s) = \frac{1}{s^2}$.

(c) Facem substituția $st = \tau$ și găsim:

$$F(s) = \int_0^\infty e^{-st} t^n dt = \frac{1}{s^{n+1}} \int_0^\infty e^{-\tau} \tau^n d\tau = \frac{n!}{s^{n+1}},$$

pentru $s > 0$, folosind funcția Gamma a lui Euler:

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, a > 0, \quad \Gamma(n+1) = n!, n \in \mathbb{N}.$$

(d) $F(s) = \displaystyle\int_0^\infty e^{at}\, e^{-st} dt = \int_0^\infty e^{-(s-a)t} dt = \frac{1}{s-a}, s > a.$

(e) Integrăm prin părți și ajungem la:

$$F(s) = \frac{1}{a} - \frac{s^2}{a^2} F(s) \Longrightarrow F(s) = \frac{a}{s^2 + a^2}, s > 0.$$

2. Folosind tabelul de valori și proprietățile, să se determine transformatele Laplace pentru funcțiile (presupuse original):

(a) $f(t) = 5$;

(b) $f(t) = 3t + 6t^2$;

(c) $f(t) = e^{-3t}$;

(d) $f(t) = 5e^{-3t}$;

(e) $f(t) = \cos(5t)$;

(f) $f(t) = \sin(3t)$;

(g) $f(t) = 3(t-1) + e^{-t-1}$;

(h) $f(t) = 3t^3(t-1) + e^{-5t}$;

(i) $f(t) = 5e^{-3t}\cos(5t)$;

(j) $f(t) = e^{2t}\sin(3t)$;

(k) $f(t) = te^{-t}\cos(4t)$;

(l) $f(t) = t^2\sin(3t)$;

(m) $f(t) = t^3\cos t$.

Indicații: În majoritatea cazurilor, se folosește tabelul și proprietatea de liniaritate. În plus:
(i, j) Folosim $\mathcal{L}(e^{at}f(t)) = F(s-a)$;
(k) Folosim $\mathcal{L}(tf(t)) = -\frac{d}{ds}\mathcal{L}(f(t))$;
(l) Folosim $\mathcal{L}(t^n f(t)) = (-1)^n \frac{d^n F(s)}{ds^n}$;

3. Folosind teorema derivării imaginii, să se determine transformatele Laplace pentru funcțiile (presupuse original):

(a) $f(t) = t$;

(b) $f(t) = t^2$;

(c) $f(t) = t\sin t$;

(d) $f(t) = te^t$.

Indicație: Conform proprietății de derivare a imaginii, avem:

$$\mathcal{L}(tf(t)) = -\frac{d}{ds}\mathcal{L}(f(t)).$$

4. Folosind teorema integrării originalului, să se determine transformatele Laplace pentru funcțiile (presupuse original):

(a) $f(t) = \int_0^t \cos(2\tau)d\tau$;

(b) $f(t) = \int_0^t e^{3\tau}\cos(2\tau)d\tau$;

(c) $f(t) = \int_0^t \tau e^{-3\tau}d\tau$.

Indicație: Conform proprietății de integrare a originalului, avem:

$$\mathcal{L}\int_0^t f(\tau)d\tau = \frac{F(s)}{s}.$$

5.4 Aplicații ale transformatei Laplace

Principala aplicație a transformatei Laplace este pentru rezolvarea ecuațiilor și sistemelor diferențiale de ordinul întîi sau superior.

Aceste aplicații se bazează pe calculele care se pot obține imediat din definiția transformatei Laplace și a proprietăților sale:

$$\begin{aligned}\mathcal{L}f' &= s\mathcal{L}f - f(0)\\ \mathcal{L}f'' &= s^2\mathcal{L}f - sf(0) - f'(0).\end{aligned}$$

De fapt, în general, avem:

$$\mathcal{L}f^{(n)} = s^n\mathcal{L}f - s^{n-1}f(0) - s^{n-2}f'(0) - \cdots - f^{(n-1)}(0).$$

De asemenea, pentru integrale, știm deja teorema integrării originalului:

$$\mathcal{L}\int_0^t f(\tau)d\tau = \frac{1}{s}F(s), \quad F(s) = \mathcal{L}f(t).$$

Rezultă, folosind transformarea inversă:

$$\int_0^t f(\tau)d\tau = \mathcal{L}^{-1}\Big(\frac{1}{s}F(s)\Big).$$

De exemplu, pentru a rezolva ecuația diferențială:

$$y'' + ay' + by = r(t), \quad y(0) = K_0, y'(0) = K_1,$$

aplicăm transformata Laplace și folosim proprietățile de mai sus. Fie $Y = \mathcal{L}y(t)$

Se obține ecuația algebrică:

$$(s^2Y - sy(0) - y'(0)) + a(sY - y(0)) + bY = R(s),$$

unde $R(s) = \mathcal{L}r$. Forma echivalentă este:

$$(s^2 + as + b)Y = (s + a)y(0) + y'(0) + R(s).$$

Împărțim prin $s^2 + as + b$ și folosim formula:

$$Q(s) = \frac{1}{s^2 + as + b} = \frac{1}{(s + \frac{1}{2}a)^2 + b - \frac{1}{4}a^2},$$

de unde rezultă:

$$Y(s) = \big((s + a)y(0) + y'(0)\big)Q(s) + R(s)Q(s).$$

În forma aceasta, descompunem $Y(s)$ în fracții simple, dacă este nevoie și folosim tabelul de transformate Laplace, pentru a afla $y = \mathcal{L}^{-1}(Y)$.

De exemplu:

$$y'' - y = t, \quad y(0) = 1, y'(0) = 1.$$

Soluție: Aplicăm transformata Laplace și ajungem la ecuația:

$$\begin{aligned} s^2Y - sy(0) - y'(0) - Y &= \frac{1}{s^2} \\ (s^2 - 1)Y &= s + 1 + \frac{1}{s^2}. \end{aligned}$$

Rezultă $Q = \dfrac{1}{s^2 - 1}$ și ecuația devine:

$$\begin{aligned} Y &= (s + 1)Q + \frac{1}{s^2}Q \\ &= \frac{s + 1}{s^2 - 1} + \frac{1}{s^2(s^2 - 1)} \\ &= \frac{1}{s - 1} + \left(\frac{1}{s^2 - 1} - \frac{1}{s^2}\right) \end{aligned}$$

Folosind tabelul și proprietățile transformatei Laplace, obținem soluția:

$$\begin{aligned} y(t) &= \mathcal{L}^{-1}Y \\ &= \mathcal{L}^{-1}\left(\frac{1}{s^{-1}}\right) + \mathcal{L}^{-1}\left(\frac{1}{s^2 - 1}\right) - \mathcal{L}^{-1}\left(\frac{1}{s^2}\right) \\ &= e^t + \sinh t - t. \end{aligned}$$

5.5 Exerciții

1. Să se rezolve următoarele probleme Cauchy, folosind transformata Laplace:

(a) $y'(t) + 2y(t) = 4t, y(0) = 1$;

(b) $y'(t) + y(t) = \sin 4t, y(0) = 0$;

(c) $y''(t) + 2y'(t) + 5y(t) = 0, y(0) = 0, y'(0) = 0$;

(d) $2y''(t) - 6y'(t) + 4y(t) = 3e^{3t}, y(0) = 1, y'(0) = -1$;

(e) $y''(t) - 2y'(t) + y(t) = e^t, y(0) = -2, y'(0) = -3$;

(f) $y''(t) + 4y(t) = 3\cos^2(t), y(0) = 1, y'(0) = 2$.

2. Să se rezolve următoarele sisteme diferențiale:

(a) $\begin{cases} x' + x + 4y & = 10 \\ x - y' - y & = 0 \\ x(0) = 4, & y(0) = 3 \end{cases}$

(b) $\begin{cases} x' + x - y & = e^t \\ y' + y - x & = e^t \\ x(0) = 1, & y(0) = 1 \end{cases}$

(c) $\begin{cases} x' + 2y' + x - y & = 5\sin t \\ 2x' + 3y' + x - y & = e^t \\ x(0) = 2, & y(0) = 1 \end{cases}$

Indicație: Aplicăm transformata Laplace fiecărei ecuații și notăm $\mathcal{L}x(t) = X(s)$ și $\mathcal{L}y(t) = Y(s)$. Apoi rezolvăm sistemul *algebric* obținut cu necunoscutele X și Y, cărora la final le aplicăm transformata Laplace inversă.

5.6 Formula de inversare Mellin-Fourier

În cazul simplu, dacă se dă o imagine Laplace și vrem să recuperăm funcția original, putem pur și simplu să citim tabelul de transformate de la stînga la dreapta, eventual după prelucrarea imaginii Laplace, precum am văzut în exemplele de mai sus. Dar avem nevoie și de o metodă care să funcționeze pe cazul general. Mai precis, avem nevoie de o metodă de a *inversa* transformatele Laplace. Această metodă este dată de formula de inversare Mellin-Fourier, pe care o prezentăm mai jos într-o variantă simplificată, mai potrivită pentru aplicații.

Teoremă 5.1 (Mellin-Fourier)**:** *Fie f o funcție original, $F = \mathcal{L}f$.*

Atunci are loc egalitatea:

$$f(t) = \sum_{k=1}^{n} \operatorname{Rez}\left(F(s)e^{st}, s_k\right).$$

Deoarece avem mai multe moduri particulare de a calcula reziduurile, distingem și aici două calcule speciale:

- Dacă $s_1, s_2, \dots, s_r$ sînt poli de ordin $n_1, n_2, \dots, n_r$, atunci formula de inversare se mai scrie:

$$f(t) = \sum_{k=1}^{r} \frac{1}{(n_k - 1)!} \cdot \lim_{s \to s_k} \frac{d^{n_k - 1}}{ds^{n_k - 1}} \left[F(s)(s - s_k)^{n_k} e^{st}\right];$$

- Dacă imaginea Laplace se poate scrie $F(s) = \dfrac{A(s)}{B(s)}$, unde A și B sînt polinoame cu coeficienți reali, astfel încît $\operatorname{grad}(A) < \operatorname{grad} B$, atunci formula de inversare devine:

$$f(t) = \sum_k \operatorname{Rez}\left(\frac{A(s)}{B(s)}e^{st}, s_k\right) + \sum_k 2\operatorname{Re}\left(\operatorname{Rez}\left(\frac{A(s)}{B(s)}e^{st}, s_k\right)\right),$$

  unde prima sumă se calculează pentru polii reali, iar cea de-a doua, pentru polii complecși cu partea imaginară pozitivă.

Exemplu: Determinați funcția original corespunzătoare imaginii Laplace:

$$F(s) = \frac{3s - 4}{s^2 - s - 6}.$$

Soluție: Observăm că avem $s_1 = 3$, $s_2 = -2$ poli simpli. Atunci rezultă:

$$f(t) = \operatorname{Rez}\left(\frac{3s - 4}{s^2 - s - s}e^{st}, 3\right) + \operatorname{Rez}\left(\frac{3s - 4}{s^2 - s - 6}e^{st}, -2\right).$$

Cum ambii sînt poli simpli, reziduurile se calculează folosind formula din teorma 2.4(3) și se obține în final $f(t) = e^{3t} + 2e^{-2t}$.

Observație 5.1: Această problemă, ca și altele, pot fi rezolvate și fără formula de inversare. Trebuie doar atenție și suficient exercițiu astfel încît, în afară de citirea „invers“ a tabelului de transformate Laplace, să putem folosi și teoreme de tip întîrziere, derivarea originalului etc. De exemplu, exercițiul de mai sus putea începe cu descompunerea în fracții simple:

$$\frac{3s - 4}{s^2 - s - 6} = \frac{A}{s - 3} + \frac{B}{s + 2}.$$

Se determină A și B, apoi se folosește tabelul de transformate.

5.7 Exerciții

Folosind, eventual, formula de inversare Mellin-Fourier, aflați funcțiile original care au transformatele Laplace date:

(1) $F(s) = \dfrac{s+3}{s^3+4s^2}$;

(2) $F(s) = \dfrac{s^2+3s+1}{(s+1)(s+2)(s+3)}$;

(3) $F(s) = \dfrac{2s+1}{s(s+1)}$;

(4) $F(s) = \dfrac{4s+10}{s^2-12s+32}$;

(5) $F(s) = \dfrac{5s+1}{s^2+1}$;

(6) $F(s) = \dfrac{s+2}{s^2(s+3)}$;

(7) $F(s) = \dfrac{1}{s^2(s-2)^2}$;

(8) $F(s) = \dfrac{2s-7}{s^2+2s+6}$.

SEMINAR 6

TRANSFORMATA Z

Această transformată se definește pe un caz discret, pornind de la:

Definiţie 6.1: Se numește *semnal discret* o funcție $x : \mathbb{Z} \longrightarrow \mathbb{C}$ dată de $n \mapsto x_n$ sau, echivalent, $x(n)$ ori $x[n]$.

Mulțimea semnalelor discrete se va nota cu S_d, iar cele cu suport pozitiv (nule pentru $n < 0$ se va nota S_d^+.

Un semnal particular este *impulsul unitar discret* la momentul k, definit prin:

$$\delta_k(n) = \begin{cases} 1, & n = k \\ 0 & n \neq k \end{cases},$$

definit pentru $k \in \mathbb{Z}$ fixat.

Pentru $k = 0$, vom nota $\delta_0 = \delta$.

Definiţie 6.2: Fie $x \in S_d$ și $k \in \mathbb{Z}$ fixat. Semnalul $y = (x_{n-k})$ se numește *întîrziatul lui x cu k momente.*

O operație foarte importantă, pe care se vor baza unele proprietăți esențiale ale transformatei Z este *convoluția:*

Definiţie 6.3: Fie $x, y \in S_d$. Dacă seria $\sum\limits_{k\in\mathbb{Z}} x_{n-k}y_k$ este convergentă pentru orice $n \in \mathbb{Z}$ și are suma z_n, atunci semnalul $z = (z_n)$ se numește *convoluția* semnalelor x și y și se notează $z = x * y$.

Trei proprietăți imediate sînt:

- $x * y = y * x$;

- $x * \delta = x$;
- $(x * \delta_k)(n) = x_{n-k}$.

Ajungem acum la definiția principală:

Definiţie 6.4: Fie $s \in S_d$, cu $s = (a_n)_n$. Se numește *transformata Z* sau *transformata Laplace discretă* a acestui semnal funcția definită prin:

$$L_s(z) = \sum_{n \in \mathbb{Z}} a_n z^{-n},$$

care se definește în domeniul de convergență al seriei Laurent din definiție.

Principalele proprietăți pe care le vom folosi în calcule sînt:

(1) **Liniaritatea:** Transformata Z este liniară în raport cu înmulțirea cu scalari și adunarea semnalelor discrete;

(2) **Inversarea transformării Z:** Fie $s \in S_d^+$, cu $s = (a_n)$. Presupunem că $L_s(z)$ este olomorfă în domeniul $|z| \in (r, R)$. Atunci putem recupera semnalul a_n prin formula:

$$a_n = \frac{1}{2\pi i} \int_\gamma z^{n-1} L_s(z) dz, \quad n \in \mathbb{Z},$$

unde γ este discul de rază $\rho \in (r, R)$.

(3) **Teorema de convoluție:** Fie $s, t \in S_d^+$. Atunci $s * t = S_d^+$ și are loc $L_{s*t} = L_s \cdot L_t$. În particular:

$$L_{s*\delta_k}(z) = z^{-k} L_s(z), \quad k \in \mathbb{Z};$$

Cîteva transformate uzuale sînt:

s	$L_s(z)$
$\begin{cases} h_n = 0, & n < 0 \\ h_n = 1, & n \geq 0 \end{cases}$	$\dfrac{z}{z-1}$
$\delta_k, k \in \mathbb{Z}$	$\dfrac{1}{z^k}$
$s = (n)_{n\in\mathbb{N}}$	$\dfrac{z}{(z-1)^2}$
$s = (n^2)_{n\in\mathbb{N}}$	$\dfrac{z(z+1)}{(z-1)^3}$
$s = (a^n)_{n\in\mathbb{N}}, a \in \mathbb{C}$	$\dfrac{z}{z-a}$
$s = (e^{an})_{n\in\mathbb{N}}, a \in \mathbb{R}$	$\dfrac{z}{z-e^a}$
$s = (\sin(\omega n))_{n\in\mathbb{N}}, \omega \in \mathbb{R}$	$\dfrac{z \sin \omega}{z^2 - 2z \cos \omega + 1}$
$s = (\cos(\omega n))_{n\in\mathbb{N}}, \omega \in \mathbb{R}$	$\dfrac{z(z - \cos \omega)}{z^2 - 2z \cos \omega + 1}$

6.1 Exerciții

1. Să se determine semnalul $x \in S_d^+$, a cărui transformată Z este dată de:

(a) $L_x(z) = \dfrac{z}{(z-3)^2}$;

(b) $L_x(z) = \dfrac{z}{(z-1)(z^2+1)}$;

(c) $L_x(z) = \dfrac{z}{(z-1)^2(z^2+z-6)}$;

(d) $L_x(z) = \dfrac{z}{z^2+2az+2a^2}$, $a > 0$ parametru.

Soluție:

(a) Avem:

$$\begin{aligned}
x_n &= \frac{1}{2\pi i}\int_{|z|=\rho} z^{n-1}L_s(z)dz \\
&= \mathrm{Rez}(z^{n-1}L_s(z),3) \\
&= \mathrm{Rez}\left(\frac{z^n}{(z-3)^2},3\right) \\
&= \lim_{z\longrightarrow 3}\left((z-3)^2\cdot\frac{z^n}{(z-3)^2}\right)' \\
&= \lim_{z\longrightarrow 3} nz^{n-1} \\
&= n3^{n-1}.
\end{aligned}$$

(b)

$$\begin{aligned}
x_n &= \frac{1}{2\pi i}\int_{|z|=\rho} z^{n-1}L_s(z)dz \\
&= \mathrm{Rez}(z^{n-1}L_s(z),1) + \mathrm{Rez}(z^{n-1}L_s(z),i) + \mathrm{Rez}(z^{n-1}L_s(z),-i)
\end{aligned}$$

$$\begin{aligned}
\mathrm{Rez}(z^{n-1}L_s(z),1) &= \lim_{z\longrightarrow 1} z^{n-1}\frac{z}{(z-1)(z^2+1)}\cdot(z-1) = \frac{1}{2} \\
\mathrm{Rez}(z^{n-1}L_s(z),i) &= \frac{i^n}{2i\cdot(i-1)} \\
\mathrm{Rez}(z^{n-1}L_s(z),-i) &= \frac{(-1)^n i^n}{2i(i+1)}.
\end{aligned}$$

Remarcăm că pentru $n = 4k$ și $n = 4k+1$, avem $x_n = 0$, iar în celelalte două cazuri, $x_n = 1$.

(c)

$$\begin{aligned}
\mathrm{Rez}(z^{n-1}L_s(z),1) &= \lim_{z\longrightarrow 1}\left[z^{n-1}\cdot(z-1)^2\cdot\frac{z^2}{(z-1)^2\cdot(z^2+z-6)}\right]' \\
&= -\frac{4n+3}{16}. \\
\mathrm{Rez}(z^{n-1}L_s(z),2) &= \frac{2^n}{5} \\
\mathrm{Rez}(z^{n-1}L_s(z),-3) &= -\frac{(-3)^n}{80}.
\end{aligned}$$

Obținem $x_n = -\dfrac{4n+3}{16} + \dfrac{2^n}{5} - \dfrac{(-3)^n}{80}$.

(d) $z_{1,2} = a(-1 \pm i)$ sînt poli simpli. Avem:

$$\begin{aligned} x_n &= \operatorname{Rez}\Big(\frac{z^n}{(z^2+2a+2a^2)}, z_1\Big) + \operatorname{Rez}\Big(\frac{z^n}{(z^2+2a+2a^2)}, z_2\Big) \\ &= \frac{a^n(-1+i)^n}{2z_1+2a} + \frac{a^n(-1-i)^n}{2z_1+2a} \\ &= -\frac{i}{2a}(z_1^n - z_2^n). \end{aligned}$$

Putem scrie trigonometric numerele z_1 și z_2:

$$\begin{aligned} z_1 &= a(-1+i) = a\sqrt{2}\Big(\cos\frac{3\pi}{4} + i\sin\frac{3\pi}{4}\Big) \\ z_2 &= a(-1-i) = a\sqrt{2}\Big(\cos\frac{3\pi}{4} - i\sin\frac{3\pi}{4}\Big). \end{aligned}$$

Deci: $x_n = 2^{\frac{n}{2}} a^{n-1} \sin\dfrac{3n\pi}{4}$.

2. Fie $x = (x_n) \in S_d^+$ și $y = (y_n)$, unde $y_n = x_0 + \cdots + x_n$. Să se arate că $Y(z) = \dfrac{z}{z-1}X(z)$, unde $Y(z) = L_x(z)$ și $X(z) = L_x(z)$ sînt transformatele Z pentru semnalele x și y.:

Soluție:

Avem $Y(z) = \sum_{n=0}^{\infty} y_n z^{-n}$. Dar:

$$X(z) = \sum_{n=0}^{\infty} x_n z^{-n} \text{ și } \sum_{n=0}^{\infty} x_{n-1} z^{n} = \frac{1}{z}X(z),$$

deoarece $x_{-1} = 0$. Putem continua și obținem $\sum_{n=0}^{\infty} x_{n-k} z^{-n} = \dfrac{1}{z^k}X(z)$. Așadar:

$$Y(z) = X(z)\cdot\Big(1 + \frac{1}{z} + \frac{1}{z^2} + \ldots\Big) = X(z)\cdot\frac{z}{z-1}.$$

3. Cu ajutorul transformării Z, să se determine șirurile (x_n) definite prin următoarele relații:

(a) $x_0 = 0, x_1 = 1, x_{n+2} = x_{n+1} + x_n, n \in \mathbb{N}$ (șirul lui Fibonacci);

(b) $x_0 = 0, x_1 = 1, x_{n+2} = x_{n+1} - x_n, n \in \mathbb{N}$;

(c) $x_0 = x_1 = 0, x_2 = -1, x_3 = 0, x_{n+4} + 2x_{n+3} + 3x_{n+2} + 2x_{n+1} + x_n = 0, n \in \mathbb{N}$;

(d) $x_0 = 2, x_{n+1} + 3x_n = 1, n \in \mathbb{Z}$;

(e) $x_0 = 0, x_1 = 1, x_{n+2} - 4x_{n+1} + 3x_n = (n+1)4^n, n \in \mathbb{N}$.

Soluție:

Abordarea generală este să considerăm șirul (x_n) ca fiind restricția unui semnal $x \in S_d^+$ la $\mathbb{N}$ și rescriem relațiile de recurență sub forma unor ecuații de convoluție $a * x = y$, pe care le rezolvăm în S_d^+.

(a) Fie $x \in S_d^+$, astfel încît restricția lui la $\mathbb{N}$ să fie șirul căutat. Deoarece avem:

$$x_{n+2} - x_{n+1} - x_n = y_n, n \in \mathbb{Z},$$

cu $y_n = 0$ pentru $n \neq -1$ și $y_{-1} = 1$, avem ecuația de convoluție:

$$a * x = y, \text{ unde } a = \delta_{-2} - \delta_{-1} - \delta, y = \delta_{-1}.$$

Aplicăm transformata Z și rezultă:

$$L_x(z) \cdot (z^2 - z - 1) = z \Longrightarrow x_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

(b) Ca în cazul anterior, avem $a * x = y$, cu $a = \delta_{-2} - \delta_{-1} + \delta$, unde $y = \delta_{-1}$. Aplicînd transformata Z, obținem:

$$L_x(z) \cdot (z^2 - z + 1) = z \Longrightarrow L_x(z) = \frac{z}{z^2 - z + 1}.$$

Obținem:

$$x_n = \operatorname{Rez}\left(z^{n-1}\frac{z}{z^2 - z + 1}, \frac{1 + i\sqrt{3}}{2}\right) + \operatorname{Rez}\left(z^{n-1}\frac{z}{z^2 - z + 1}, \frac{1 - i\sqrt{3}}{2}\right).$$

Calculăm reziduurile, cu notația $\varepsilon = \dfrac{1 + i\sqrt{3}}{2}$ și $\overline{\varepsilon} = \dfrac{1 - i\sqrt{3}}{2}$:

$$\begin{aligned}\operatorname{Rez}\left(\frac{z^n}{z^2 - z + 1}, \varepsilon\right) &= \lim_{z \to \varepsilon} \frac{z^n}{z^2 - z + 1}(z - \varepsilon) \\ &= \frac{\varepsilon^n}{i\sqrt{3}} \\ &= \frac{\cos\frac{2n\pi}{3} + i\sin\frac{2n\pi}{3}}{i\sqrt{3}}.\end{aligned}$$

Similar:

$$\operatorname{Rez}\left(\frac{z^n}{z^2 - z + 1}, \overline{\varepsilon}\right) = \frac{\cos\frac{2n\pi}{3} - i\sin\frac{2n\pi}{3}}{-i\sqrt{3}}.$$

Rezultă:

$$x_n = \frac{2}{\sqrt{3}}\sin\frac{2n\pi}{3}, n \in \mathbb{N}.$$

(c) Ecuația $a * x = y$ este valabilă pentru:

$$a = \delta_{-4} + 2\delta_{-3} + 3\delta_{-2} + 2\delta_{-1} + \delta, \quad y = -\delta_{-2} - 2\delta_{-1}.$$

Aplicăm transformata Z și obținem: $L_x(z) = -\dfrac{z(z+2)}{(z^2+z+1)^2}$. Descompunem în fracții simple, calculăm reziduurile și ținem cont de faptul că rădăcinile numitorului, $\varepsilon_1, \varepsilon_2$ sînt poli de ordinul 2, obținem:

$$x_n = \frac{(2n-4)(\varepsilon_1^n - \varepsilon_2^n) - (n+1)(\varepsilon_1^{n-1} + \varepsilon_1^{n-2} - \varepsilon_2^{n-1} - \varepsilon_2^{n-2})}{(\varepsilon_2 - \varepsilon_1)^3} = \frac{2(n-1)}{\sqrt{3}} \sin \frac{2n\pi}{3}, n \in \mathbb{N}.$$

(d) Ecuația corespunzătoare este $a * x = y$, cu $a = \delta_{-1} + 3\delta$ și $y_n = 1, \forall n \geq 1$, iar $y_{-1} = x_0 + 3x_{-1} = 2$, cu $y_n = 0, \forall n \leq -2$, adică $y = 1 + 2\delta_{-1}$.

Așadar:

$$\delta_{-1} * x + 3\delta * x = 1 + 2\delta_{-1}.$$

Aplicăm transformata Z și obținem:

$$\begin{aligned}
zL_x(z) + 3L_x(z) &= \frac{z}{z-1} + 2z \\
&= \frac{2z^2 + 3z}{z-1} \\
\Longrightarrow L_x(z) &= \frac{2z^2 + 3z}{(z-1)(z+3)} \\
\Longrightarrow x_n &= \mathrm{Rez}(z^{n-1} \cdot \frac{2z^2+3z}{(z-1)(z+3)}, 1) + \mathrm{Rez}(z^{n-1} \cdot \frac{2z^2+3z}{(z-1)(z+3)}, -3) \\
\mathrm{Rez}(z^{n-1} \cdot \frac{2z^2+3z}{(z-1)(z+3)}, 1) &= \lim_{z \to 1} (z-1) \cdot z^{n-1} \cdot \frac{2z^2+3z}{(z-1)(z+3)} = \frac{5}{4} \\
\mathrm{Rez}(z^{n-1} \cdot \frac{2z^2+3z}{(z-1)(z+3)}, -3) &= \lim_{z \to -3} (z+3) z^{n-1} \cdot \frac{2z^2+3z}{(z-1)(z+3)} \\
&= (-3)^{n-1} \cdot \frac{12}{-4} = -3 \cdot (-3)^{n-1}.
\end{aligned}$$

Rezultă: $x_n = \dfrac{5}{4} - 3 \cdot (-3)^{n-1}$.

(e) Avem ecuația: $a * x = y$, unde $a = \delta_{-2} - 4\delta_{-1} + 3\delta$, cu $y_n = 0, \forall n \leq -2, y_{-1} = 1$ și $y_n = (n+1)4^n, \forall n \in \mathbb{N}$.

Fie $s_1 = (n4^n)_n$, $s_2 = (4^n)_n$. Atunci:

$$
\begin{aligned}
L_{s_1}(z) &= -zL_{s_2'}(z) = -z\Big(\frac{z}{z-4}\Big)' = \frac{4z}{(z-4)^2} \\
\Longrightarrow L_x(z)(z^2-4z+3) &= \frac{4z}{(z-4)^2} + \frac{z}{z-4} + z \\
&= \frac{z^2}{(z-4)^2} + z \\
\Longrightarrow L_x(z) &= \frac{z(z^2-7z+16)}{(z-4)^2(z-1)(z-3)}.
\end{aligned}
$$

Descompunem în fracții simple și obținem, în fine:

$$x_n = \frac{1}{9}\big[18 \cdot 3^n + (3n-13)4^n - 5\big], n \in \mathbb{N}.$$

OBSERVAȚIE: Toate exercițiile cu recurențe se mai pot rezolva în alte două moduri:
(1) Se poate aplica teorema de convoluție relației de recurență. De exemplu, din recurența:

$$x_{n+2} - 2x_{n+1} + x_n = 2$$

putem obține:

$$L_{x_{n+2}}(z) - 2L_{x_{n+1}}(z) + L_{x_n}(z) = 2L_1(z),$$

iar $L_{x_{n+2}}(z) = L_{x_n * \delta_{-2}}(z) = L_{x_n}(z) \cdot L_{\delta_{-2}}(z)$ etc.

(2) Se poate aplica teorema de deplasare. În aceeași recurență de mai sus, de exemplu, avem:

$$L_{x_{n+2}}(z) = z^n \left(L_{x_n}(z) - x_0 - x_1 z^{-1}\right)$$

și la fel pentru celelalte.

6.2 Exerciții suplimentare (L. Costache)

1. Să se determine semnalul $x \in S_d^+$ a cărui transformată Z este:

(a) $L_x(z) = \dfrac{2z+3}{z^2-5z+6}$;

(b) $L_x(z) = \dfrac{z^2+1}{z^2-z+1}$;

(c) $L_x(z) = \dfrac{z}{(z-5)^2}$;

(d) $L_x(z) = \dfrac{z+2}{z^2+z+1}$.

2. Cu ajutorul transformatei Z, să se determine șirurile $(x_n)_{n\in\mathbb{N}}$ definite prin următoarele relații de recurență:

(a) $x_0 = 4, x_1 = 6$ și $x_{n+2} - 3x_{n+1} + 2x_n = 0, \forall n \in \mathbb{N}$;

(b) $x_0 = 0, x_1 = 11, x_2 = -8, x_3 = 6$ și $x_{n+4} - \dfrac{5}{2}x_{n+3} + \dfrac{5}{2}x_{n+1} - x_n = 1, \forall n \in \mathbb{N}$;

(c) $x_0 = 0, x_1 = 3$ și $x_{n+2} - 4x_{n+1} + 3x_n = 2, \forall n \in \mathbb{N}$;

(d) $x_0 = 0, x_1 = -1$ și $x_{n+2} + x_{n+1} - 6x_n = n, \forall n \in \mathbb{N}$;

(e) $x_0 = 0, x_1 = 1$ și $x_{n+2} - 5x_{n+1} + 6x_n = 4 \cdot 5^n, \forall n \in \mathbb{N}$.

SEMINAR 7

TRANSFORMAREA FOURIER

Transformarea Fourier, cu diversele sale variante (discretă, rapidă etc.) este foarte utilă pentru studiul undelor sinusoidale. Dacă o *serie* Fourier dezvoltă o funcție într-o serie în care termenii sînt sinusuri și cosinusuri, *transformata* Fourier este o transformare integrală, în care aspectele periodice corespunzătoare funcțiilor trigonometrice se păstrează.

De aceea, veți întîlni foarte des transformatele Fourier în cazul analizelor semnalelor sonore, fie pentru digitalizare sau pentru diverse descompuneri. Un exemplu concret este în inteligența artificială și învățarea automată, unde se lucrează la programe care să facă recunoaștere vocală, transcrieri și traduceri. În toate aceste aplicații, o analiză a semnalului audio cu ajutorul transformatei Fourier este indispensabil.

7.1 Elemente teoretice

Foarte pe scurt, vom da definițiile și principalele proprietăți care vor fi de folos în exerciții.

Avem nevoie de următoarea noțiune:

Definiție 7.1: Fie $f : \mathbb{R} \to \mathbb{R}$ o funcție reală. Spunem că ea este L^1 (formal, *aparține mulțimii* $L^1(\mathbb{R})$), dacă are proprietatea:

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty.$$

Intuitiv, mutînd tot graficul funcției deasupra axei Ox (prin oglindire pe intervalele negative), aria de sun grafic este finită. Deci funcția nu are nicio zonă în care „explodează“ spre $\pm\infty$ asimptotic.

Acestea sînt funcțiile cărora le vom defini transformările Fourier. Deci, dacă nu se precizează altfel, toate funcțiile care se transformă vor fi presupuse L^1.

Definiția principală urmează.

Definiţie 7.2: Fie f o funcție L^1. Se definește *transformata Fourier* a lui f, notată $\mathcal{F}$ (alternativ, $\mathcal{F}[f]$ sau, și mai explicit, $\mathcal{F}[f(t)](\omega)$)[1]funcția complexă definită prin:

$$\mathcal{F}[f] : \mathbb{R} \longrightarrow \mathbb{C}, \quad \mathcal{F}[f](\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt.$$

Se vede aici importanța faptului că f este L^1.

Folosind formula lui Euler pentru exponențiala complexă care apare în integrală, precum și paritatea funcțiilor trigonometrice, facem următoarele observații imediate:

- Dacă f este funcție pară, atunci transformata Fourier are partea imaginară nulă (integrala sinsului pe un interval simetric față de origine — în acest caz special, $(-\infty, \infty)$ — este nulă, iar integrala funcției pare cosinus este dublul integralei calculate pe jumătate din interval) și rămînem cu:
$$\mathcal{F}[f](\omega) = 2\int_{0}^{\infty} f(t)\cos(\omega t)dt;$$
- Dacă f este funcție impară, folosind argumentul din cazul anterior, transformarea Fourier are partea reală nulă și rămînem cu:
$$\mathcal{F}[f](\omega) = -2i\int_{0}^{\infty} f(t)\sin(\omega t)dt.$$

Cele două cazuri speciale se numesc, respectiv, *transformarea (Fourier) prin sinus* și *transformarea (Fourier) prin cosinus* ale funcției f.

Transformarea Fourier se poate inversa relativ simplu:

Teoremă 7.1 (Transformarea Fourier inversă)**:** *Fie f o funcție L^1 și $\mathcal{F}[f]$ transformata sa Fourier. Presupunem că și $\mathcal{F}[f]$ este funcție L^1 și atunci se poate recupera funcția f prin formula:*

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{F}[f]e^{i\omega t}d\omega.$$

Dacă funcția f care se transformă are „puncte cu probleme“ (care sigur aparțin domeniului de integrare, deoarece acesta este întreg $\mathbb{R}$), atunci transformarea Fourier se calculează cu ajutorul reziduurilor. Fie, așadar, $p_1, \dots, p_n$ poli ai funcției f și presupunem că $\lim\limits_{|z|\longrightarrow\infty} f(z) = 0$. Atunci:

- Dacă $\mathrm{Im} p_k > 0$, atunci:
$$\mathcal{F}[f](\omega) = 2\pi i \sum_{k=1}^{n} \mathrm{Rez}\left(f(z)e^{-i\omega z}, p_k\right), \quad \text{pentru } \omega < 0;$$

[1]Pentru a nu încărca inutil notația, vom folosi această variantă explicită doar aici. Regula care se va păstra în continuare este următoarea: argumentul funcției f care se transformă este t, iar argumentul transformatei $\mathcal{F}[f]$ este ω. De asemenea, argumentul t al funcției f se mai numește *timp*, iar argumentul ω al transformatei se mai numește *frecvență*.

- Dacă $\mathrm{Im} p_k < 0$, atunci:

$$\mathcal{F}[f](\omega) = -2\pi i \sum_{k=1}^{n} \mathrm{Rez}\left(f(z)e^{-i\omega z}, p_k\right), \qquad \text{pentru } \omega > 0;$$

În exerciții, ne vor mai fi de folos și următoarele proprietăți ale transformatei Fourier:

(1) **Liniaritatea:** $\mathcal{F}[\alpha f + \beta g] = \alpha\mathcal{F}[f] + \beta\mathcal{F}[g]$, pentru orice f, g funcții L^1 și $\alpha, \beta \in \mathbb{C}$ scalari;

(2) **Simetria:** $\mathcal{F}\left[\mathcal{F}[f](\omega)\right] = 2\pi f(-\omega)$;

(3) **Rescalarea:** Fie $\alpha \in \mathbb{C}$. Atunci:

$$\mathcal{F}\left[f(\alpha t)\right](\omega) = \frac{1}{|\alpha|}\mathcal{F}[f]\left(\frac{\omega}{\alpha}\right).$$

(4) **Translația după t:**

$$\mathcal{F}[f(t - t_0)] = \mathcal{F}[f(t)] \cdot e^{-i\omega t_0}, \quad \forall t_0 \in \mathbb{R};$$

(5) **Translația după ω:**

$$\mathcal{F}\left[e^{i\omega_0 t} f(t)\right] = \mathcal{F}[f(t)](\omega - \omega_0), \quad \forall \omega_0 \in \mathbb{R};$$

(6) **Derivarea după t:**

$$\mathcal{F}[f^{(n)}(t)] = (i\omega)^n \mathcal{F}[f](\omega),$$

în ipoteza că f este de n ori derivabilă, pentru un anume n;

(7) **Derivarea după ω:**

$$\mathcal{F}\left[(-it)^n f(t)\right] = \mathcal{F}^{(n)}[f](\omega),$$

în ipoteza că derivatele de pînă la n ori au sens;

(8) **Transformata conjugatei complexe:** Notăm $f^*(t)$ funcția conjugată complexă asociată lui $f(t)$. Atunci:

$$\mathcal{F}[f^*(t)] = \left(\mathcal{F}[f(t)]\right)^*(-\omega).$$

Altfel spus, se transformă funcția inițială, apoi se ia conjugatul rezultatului și se schimbă semnul argumentului;

(9) **Convoluția în timp:** Definim:

$$h(t) = (f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

produsul de convoluție al funcțiilor f și g. Atunci:

$$\mathcal{F}[f * g] = \mathcal{F}[f] \cdot \mathcal{F}[g];$$

(10) **Convoluția în frecvență:**

$$\mathcal{F}\left[f(t) \cdot g(t)\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{F}[f](y) \cdot \mathcal{F}[g](\omega - y)dy.$$

7.2 Aplicații la ecuații integrale

Folosind formula de transformare Fourier, precum și inversarea acesteia, putem rezolva ecuații integrale. Vom prezenta acest lucru pe un exemplu.

Exemplu: Rezolvăm ecuația integrală:

$$\int_0^\infty y(t)\cos tx dt = \frac{1}{x^2+1}.$$

Soluție: Pentru a face legătura cu transformările Fourier, avem nevoie să prelungim funcția y (astfel încît integrala să fie pe $(-\infty, \infty)$). Deoarece integrandul este par, acest lucru conduce la dublarea rezultatului. Deci avem:

$$\int_{-\infty}^\infty y(t)\cos tx dt = \frac{2}{x^2+1}, \quad \int_{-\infty}^\infty y(t)\sin tx dt = 0.$$

Acest lucru conduce la:

$$\frac{1}{2\pi}\int_{-\infty}^\infty y(t)e^{-itx}dt = \frac{1}{\pi(x^2+1)}.$$

Recunoaștem în membrul stîng formula de trasnformare Fourier, deci rezultă:

$$\mathcal{F}[y(t)](x) = \frac{1}{\pi(x^2+1)}.$$

Folosim acum formula de inversare și rezultă:

$$y(t) = \frac{1}{\pi}\int_{-\infty}^\infty \frac{1}{x^2+1}e^{itx}dx.$$

Acest calcul se poate face cu teorema reziduurilor, deoarece integrandul are doi poli, $p_{1,2} = \pm i$, ambii simpli. Obținem:

- Pentru $t < 0$ și polul $p_1 = i$:

$$\begin{aligned} y(t) &= \frac{1}{\pi}\cdot 2\pi i \text{Rez}\left(\frac{1}{x^2+1}e^{itx}, i\right) \\ &= 2i \cdot \lim_{z\to i}(z-i)\frac{1}{z^2+1}e^{itz} \\ &= 2i\cdot e^{-t}; \end{aligned}$$

- Pentru $t > 0$ și polul $p_2 = -i$:

$$\begin{aligned} y(t) &= \frac{1}{pi}\cdot -2\pi i \text{Rez}\left(\frac{1}{x^2+1}e^{itx}, -i\right) \\ &= -2i\cdot \lim_{z\to -i}(z+i)\frac{1}{z^2+1}e^{itz} \\ &= 2ie^t. \end{aligned}$$

Concluzia este că soluția arată astfel:

$$y(t) = \begin{cases} 2i \cdot \exp(-t), & t < 0 \\ 2i \cdot \exp(t), & t > 0 \end{cases} = 2i \cdot \exp(|t|).$$

7.3 Exerciții

1. Rezolvați ecuațiile integrale:

(a) $\displaystyle\int_0^\infty f(t) \sin txdt = e^{-x}, \quad x > 0;$

(b) $\displaystyle\int_0^\infty f(t) \cos txdt = \frac{1}{(1+x^2)^2};$

2. Calculați transformatele Fourier pentru funcțiile:

(a) $f(t) = \begin{cases} 1 - |t|, & t \in [-1, 1] \\ 0, & \text{în rest} \end{cases};$

(b) $f(t) = \dfrac{t}{t^2 + a^2}, a > 0;$

(c) $f(t) = \dfrac{1}{(t^2+1)^2};$

(d) $f(t) = \begin{cases} \exp(2t), & t \le 0 \\ 0, & t > 0 \end{cases};$

(e) $f(t) = \begin{cases} t^2 - t, & 0 < t < 1 \\ 0, & \text{în rest} \end{cases};$

(f) $f(t) = \begin{cases} t, & |t| < 1 \\ 0, & |t| > 1 \end{cases}.$

3. Calculați transformatele Fourier pentru funcțiile de mai jos, folosind teorema reziduurilor:

(a) $f(t) = \dfrac{2}{t^2+9};$

(b) $f(t) = \dfrac{t}{(t^2+9)(t^2+1)};$

(c) $f(t) = \dfrac{1}{(t^2+4)^2}.$

SEMINAR 8

RECAPITULARE TEST (L. COSTACHE)

Integrale complexe

Calculați:

(a) $\displaystyle\int_{|z|=4} \frac{z\,dz}{\sin z(1-\cos z)}$;

(b) $\displaystyle\int_{|z|=4} \frac{dz}{3\sin z - \sin 3z}$;

(c) $\displaystyle\int_{|z|=2} \frac{z^n \cdot e^{\frac{1}{z}}}{z^2-1}\,dz$;

(d) $\displaystyle\int_0^{2\pi} \frac{\cos 4\theta}{1-2a\cos\theta + a^2}\,d\theta,\ a > 1$;

(e) $\displaystyle\int_{-\infty}^{\infty} \frac{x\cos x}{x^2-6x+13}\,dx$;

(f) $\displaystyle\int_0^{\infty} \frac{x\sin ax}{(x^2+b^2)(x^2+c^2)}\,dx,\ b, c > 0.$

Transformata Laplace

1. Să se determine transformata Laplace a funcției:

$$f(t) = \begin{cases} 2, & 0 < t < 1 \\ \dfrac{t^2}{2}, & 1 < t < \dfrac{\pi}{2} \\ \cos t, & t > \dfrac{\pi}{2} \end{cases}$$

2. Folosind transformata Laplace, rezolvați ecuațiile diferențiale:

(a) $y'' + 3y' + 2y = u(t-1) - u(t-2)$, știind $y(0) = 0, y'(0) = 0$;

(b) $y'' + 4y = \begin{cases} 8t^2, & 0 < t < 5 \\ 200, & t > 5 \end{cases}$, știind $y(1) = 1 + \cos 2$ și $y'(1) = 4 - 2\sin 2$;

(c) $x''(t) - 2x'(t-1) = t$, știind $x(0) = x'(0) = 0$;

(d) $y'' + 2y' - 2y = \begin{cases} 3\sin t - \cos t, & 0 < t < 2\pi \\ 3\sin 2t - \cos 2t, & t > 2\pi \end{cases}$, știind $y(0) = 1, y'(0) = 0$.

3. Folosind transformata Laplace, rezolvați ecuația integrală:

$$x(t) = t + 2\int_0^t [t - \tau - \sin(t-\tau)] \cdot x(\tau)d\tau$$

Transformata Fourier

1. Determinați transformata Fourier a semnalului:

$$f(t) = (t+2)e^{-(t+2)(3i+5)}, \quad t > 0.$$

2. Rezolvați ecuația integrală, folosind transformata Fourier:

$$\int_0^\infty f(t)\sin(\omega t)dt = \begin{cases} \frac{\pi}{2}\sin\omega, & \omega \in (0,\pi) \\ 0, & \omega \geq \pi \end{cases}$$

SEMINAR 9

SPAȚII DE PROBABILITATE

9.1 Noțiuni teoretice

Trecem acum la studiul probabilităților într-un context mai abstract decît în liceu, unde se defineau și studiau evenimentele aleatoare concret.

Definiţie 9.1: Se numește *spațiu (cîmp) discret de probabilitate* o mulțime finită sau numărabilă $\Omega = (\omega_n)_n$, împreună cu un șir $(p_n)_n$, cu $0 \leq p_n \leq 1$, care satisface condiția $\sum_n p_n = 1$.

Orice submulțime $A \subseteq \Omega$ este un *eveniment*, căruia i se atașează o probabilitate

$$P(A) = \sum_{\omega_n \in A} p_n.$$

Intuitiv, putem privi elementele mulțimii Ω drept *experiențe*, pe care le putem colecta în *evenimente*, iar p_n este probabilitatea ca o experiență să se realizeze.

Alte definiții elementare sînt următoarele:

Definiţie 9.2: *Evenimentul sigur* este un eveniment care se realizează cu certitudine la fiecare repetare a experienței.

Evenimentul imposibil nu se produce la nicio efectuare a experienței.

Dat un eveniment $A \subseteq \Omega$, lui i se asociază *evenimentul contrar*, notat $\overline{A}$, CA sau A^c, care constă în nerealizarea lui A.

Două evenimente $A, B \subseteq \Omega$ se numesc *compatibile* dacă se pot produce simultan și *incompatibile* în caz contrar.

Pentru evenimente incompatibile, avem o definiție alternativă, calculatorie:

Definiţie 9.3: Dacă A și B sînt evenimente incompatibile, atunci:

$$P(A \cup B) = P(A) + P(B).$$

Mai general, pentru orice șir (A_n) de evenimente două cîte două incompatibile, avem:

$$P\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} P(A_n).$$

Definiția abstractă a contextului în care vom studia teoria probabilităților urmează.

Definiţie 9.4: Se numește *spațiu de probabilitate* un triplet $(\Omega, \mathcal{K}, P)$, unde Ω este o mulțime de evenimente elementare (individuale), $\mathcal{K}$ este o submulțime a $\mathcal{P}(\Omega)$, iar $P : \mathcal{K} \longrightarrow [0,1]$ este o funcție de probabilitate, care satisface:

- $P(\Omega) = 1$;
- $P\Big(\bigcup_{n\geq 0} A_n\Big) = \sum_{n\geq 0} P(A_n)$, pentru orice șir (A_n) de evenimente două cîte două incompatibile.

Pornind de la această definiție generală, putem obține cazuri particulare cunoscute:

Definiția clasică a probabilității: Dacă Ω este o mulțime cu N elemente, putem defini un spațiu discret de probabilitate, definind $p_n = \frac{1}{N}$, pentru orice $n \in \{1, \dots, N\}$. În acest caz, probabilitățile sînt egale pentru toate evenimentele, deci spunem că *evenimentele sînt echiprobabile.* Pentru orice eveniment $A \subseteq \Omega$, avem:

$$P(A) = \frac{\operatorname{card}(A)}{\operatorname{card}(\Omega)},$$

care coincide cu definiția clasică a raportului dintre numărul cazurilor favorabile unui eveniment și numărul cazurilor posibile.

Probabilități geometrice: Dacă luăm $\Omega \subseteq \mathbb{R}^n$ o submulțime oarecare, pe care considerăm *măsura Lebesgue* (cea cu ajutorul căreia calculăm integralele), atunci obținem un spațiu de probabilitate definind pentru orice eveniment A probabilitatea prin:

$$P(A) = \frac{\mu(A)}{\mu(\Omega)},$$

unde μ este măsura Lebesgue din $\mathbb{R}^n$ (în particular, e vorba de lungimea dl pe $\mathbb{R}$, aria $dA = dxdy$ din $\mathbb{R}^2$ etc.).

Următoarele sînt proprietăți elementare ale probabilităților:

Fie $(\Omega, \mathcal{K}, P)$ un spațiu de probabilitate.

(1) Dacă $A, B \in \mathcal{K}$ și $A \subseteq B$, atunci $P(B - A) = P(B) - P(A)$;

(2) **Poincaré**: Fie n evenimente arbitrare $A_1, \dots, A_n \in \mathcal{K}$. Atunci are loc:

$$P\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} P(A_i) - \sum_{i \neq j} P(A_i \cap A_j) + \cdots + (-1)^{n-1} P(A_1, \dots, A_n).$$

(3) Pentru orice șir crescător de evenimente $A_0 \subseteq A_1 \subseteq \dots \subseteq A_n \subseteq \dots$, avem:

$$P\Big(\bigcup_{n \geq 0}\Big) = \lim_{n \longrightarrow \infty} P(A_n).$$

Definiţie 9.5: Evenimentele A și B se numesc *independente* dacă $P(A \cap B) = P(A) \cdot P(B)$.

În general, evenimentele $A_1, \dots, A_n$ se numesc *independente în ansamblu* dacă pentru orice $m \geq n$ și $1 \leq j_1 \leq \cdots \leq j_m \leq n$, avem:

$$P(A_{j_1} \cap \cdots \cap A_{j_m}) = P(A_{j_1}) \cdots P(A_{j_m}).$$

Observaţie 9.1: Dacă avem n evenimente independente două cîte două (cf. primei părți a definiției de mai sus), nu rezultă neapărat că sînt evenimente în ansamblul lor (cf. părții a doua a definiției).

Un contraexemplu, datorat lui S. N. Bernstein este următorul. Considerăm un tetraedru omogen cu fețele colorate alb, negru, roșu, iar a patra, cu toate cele trei culori. Aruncăm acest corp și notăm cu A_i probabilitatea ca tetraedrul să se așeze (cu baza) pe fața cu numărul $i \in \{1, 2, 3, 4\}$. Aceste evenimente sînt elementare și $P(A_i) = \frac{1}{4}, \forall i$. Acum fie $A = A_1 \cup A_2$, $B = A_1 \cup A_3$ și $C = A_1 \cup A_4$. Atunci $P(A) = P(B) = P(C) = \frac{1}{2}$, deoarece pentru fiecare culoare sînt 4 cazuri posibile și 2 favorabile (fața cu culoarea respectivă și fața cu toate culorile).

În plus, avem:

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

deci evenimentele sînt independente două cîte două.

În plus, avem:

$$P(A \cap B \cap C) = P(A_1) = \frac{1}{4}$$
$$P(A)P(B)P(C) = \frac{1}{8},$$

deci evenimentele nu sînt independente în ansamblul lor.

Mai avem nevoie de următorul concept:

Definiţie 9.6: Fie A și B două evenimente, cu $P(B) \neq 0$.

Probabilitatea lui A condiționată de B, notată $P(A/B)$ sau $P_B(A)$ se definește prin:

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

De asemenea, în calcule, vom mai folosi:
Formula de înmulțire a probabilităților: Fie $A_1, \dots A_n$ evenimente. Atunci are loc:

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \cdots P(A_n/A_1 \cap \dots \cap A_{n-1}).$$

Formula probabilității totale: Dacă avem descompunerea evenimentului sigur Ω în reuniunea de n evenimente incompatibile $H_1, \dots H_n$, atunci pentru orice eveniment $A \in \mathcal{K}$, are loc:

$$P(A) = \sum_{i=1}^{n} P(A/H_i) \cdot P(H_i).$$

Formula lui Bayes: În contextul și cu notațiile de mai sus:

$$P(H_j/A) = \frac{P(A/H_j) \cdot P(H_j)}{\sum_{i=1}^{n} P(A/H_i) \cdot P(H_i)}.$$

În particular, pentru două evenimente A, B, avem:

$$P(A) = P(A/B) \cdot P(B) + P(A/B^c) \cdot P(B^c)$$
$$P(B/A) = \frac{P(A/B) \cdot P(B)}{P(A/B) \cdot P(B) + P(A/B^c) \cdot P(B^c)}.$$

9.2 Exerciții

9.2.1 Calcul „clasic“

1. Într-un spațiu de probabilitate se considerăm evenimentele A, B, cu:

$$P(A) = \frac{1}{3}, P(B) = \frac{1}{4}, P(A \cap B) = \frac{1}{6}.$$

Determinați $P(A^c), P(A^c \cup B), P(A \cup B^c), P(A^c \cup B^c), P(A^c \cap B^c)$.

2. Care dintre următoarele evenimente rezultate în urma aruncării cu zarul este mai probabil:

(a) obținerea numărului 6 în cel puțin una din 4 aruncări;

(b) obținerea cel puțin a unei duble de 6 în 24 aruncări cu 2 zaruri?

3. O urnă conține 12 bile numerotate de la 1 la 12. Să se determine probabilitatea ca bilele numerotate cu $5, 7, 11$ să iasă la extragerile cu numărul de ordine $5, 7, 11$.
Soluție: Cazurile posibile sînt $12!$ la număr.

Cazurile favorabile sînt 9!, deoarece dacă fixăm de fiecare dată bilele cu numerele $5, 7, 11$, rămîn 9! libere.

Rezultă $P = \dfrac{9!}{12!} = \dfrac{1}{1320}$.

4. O urnă conține 50 bile, dintre care 10 sînt negre, iar restul albe. Se scot la întîmplare 5 bile. Care e probabilitatea ca între cele 5 bile să fie bile negre?

Indicație: Se calculează mai ușor probabilitatea evenimentului complementar.

5. Coeficienții întregi ai ecuației $ax^2 + bx + c = 0$ sînt obținuți prin aruncarea unui zar de 3 ori. Să se determine probabilitatea ca rădăcinile ecuației să fie reale.

Indicație: Calculăm clasic probabilitatea ca $\dfrac{b^2}{4} \geq ac$, unde $a, b, c \in \{1, 2, \dots, 6\}$.

6. Într-o cameră întunecoasă se găsesc 5 perechi de pantofi. Se aleg la întîmplare 5 pantofi.

(a) Care este probabilitatea ca între cei 5 pantofi aleși să fie cel puțin o pereche, în ipoteza că cele 5 perechi sînt identice (deci este suficient să avem un pantof stîng și unul drept)?

(b) Care e probabilitatea ca între cei 5 pantofi aleși să fie cel puțin o pereche, dacă toate cele 5 perechi sînt distincte (culoare, număr)?

7. **Problema aniversării:** Într-o cameră sînt k persoane. Care este probabilitatea ca cel puțin două din aceste persoane să aibă aceeași zi de naștere, adică aceeași zi și lună a anului?

8. Se aruncă 3 zaruri. Calculați probabilitatea ca suma punctelor obținute să fie:

(a) mai mică decît 8;

(b) mai mare decît 7;

(c) egală cu 12.

9. Un scafandru are 2 sisteme de oxigen independente. Presupunem că probabilitatea ca primul sistem să funcționeze este 0,9, iar probabilitatea ca al doilea sistem să funcționeze este 0,8.

(a) Găsiți probabilitatea ca niciun sistem să nu se defecteze;

(b) Găsiți probabilitatea ca cel puțin unul din sisteme să funcționeze.

Soluție: Fie $S_{1,2}$ evenimentul ca sistemul $1, 2$ să funcționeze. (a) Cum evenimentele sînt independente, avem:

$$P(S_1 \cap S_2) = P(S_1) \cdot P(S_2) = 0,72.$$

(b)

$$P(S_1 \cup S_2) = P(S_1) + P(S_2) - P(S_1 \cap S_2) = 0,98.$$

9.2.2 Probabilități geometrice

10. Care este probabilitatea ca suma a 3 numere din intervalul $[0, a]$ alese la întîmplare să fie mai mare decît a?

Soluție: Spațiul de probabilitate este $\Omega = [0, a]^3$. Evenimentul cerut este dat de punctele mulțimii:

$$E = \{(x, y, z) \in \Omega \mid x + y + z \geq a\}.$$

Alegem un sistem ortogonal de axe și Ω devine un cub de latură a situat în primul octant. În acest caz, E este una din regiunile lui Ω separate de planul $x + y + z = a$. Rezultă:

$$P(E) = \frac{a^3 - \frac{a^3}{2}}{a^3} = \frac{1}{2}.$$

11. Pe un plan orizontal considerăm un sistem de axe XOY și mulțimea E a punctelor cu coordonate întregi. O monedă cu diametrul $\frac{1}{2}$ este aruncată la întîmplare pe acest plan. Care e probabilitatea ca moneda să acopere un punct din E?

Soluție: Fie $C(x_0, y_0)$ cel mai apropiat punct din E de centrul M al monedei, deci coordonatele lui M sînt de forma:

$$(x_0 + x, y_0 + y), \quad -\frac{1}{2} < x, y < \frac{1}{2}.$$

Spațiul de probabilitate este:

$$\Omega = \left\{(x, y) \in \mathbb{R}^2 \mid -\frac{1}{2} < x, y < \frac{1}{2}\right\}.$$

Mulțimea evenimentelor favorabile este dată de:

$$A = \left\{(x, y) \in \Omega \mid (x - x_0)^2 + (y - y_0)^2 < \frac{1}{16}\right\},$$

deoarece raza monedei este $\frac{1}{4}$.

Rezultă:

$$p = \frac{\text{aria}(A)}{\text{aria}(\Omega)} = \frac{\pi}{16}.$$

12. Se alege aleatoriu un număr real x în intervalul $[0, 3]$. Care este probabilitatea ca x să fie mai apropiat de 0 decît de 1?

Soluție: Interpretînd problema geometric, ne putem gîndi la x ca fiind un punct pe axa reală. Atunci cazurile favorabile se află în segmentul $[0; 0,5)$, iar cazurile posibile sînt pe tot segmentul.

Probabilitatea, calculată cu formula geometrică, este dată de raportul lungimilor celor două segmente:

$$P = \frac{0,5}{3} = \frac{1}{6} \simeq 17\%.$$

13. Se trage cu o săgeată la o țintă circulară de rază r. Care este probabilitatea ca săgeata să ajungă mai aproape de centru decît de margine?

Soluție: Aria cercului, care ne dă și „cazurile posibile“, este πr^2.

Cazurile favorabile se află în cercul concentric de rază $< \frac{r}{2}$, care are aria $\frac{\pi r^2}{4}$. Atunci:

$$P = \frac{1}{4}.$$

14. **Problema autobuzului (cu așteptare):** Să presupunem că ajungeți în stația de autobuz la o oră aleatorie din intervalul 12 și 1. Așteptați 20 minute în stație, iar dacă autobuzul nu vine, plecați. Același „program“ îl are și autobuzul, dar cu condiția că așteaptă 5 minute în stație, apoi pleacă.

Care este probabilitatea să prindeți autobuzul?[1]

Soluție: Avem două variabile independente: momentul în care ajungeți în stație și momentul în care autobuzul ajunge. Deci putem modela problema bidimensional. Mai precis, să luăm un pătrat cu latura de 60 de unități (1/minut) și-i notăm laturile cu c și a („călător“ și „autobuz“).

Remarcăm că avem cazurile (cf. Figura 9.1):

- Pentru a prinde autobuzul la fix, ar trebui să ne situăm pe segmentul $a = c$;
- Cum autobuzul așteaptă 5 minute, de fapt putem să ne situăm în zona $c \leq a + 5$, *netinînd seama de faptul că și călătorul așteaptă (deocamdată)!*;
- Cum călătorul așteaptă 20 minute, ne situăm în zona $c \geq a - 20$ (*netinînd seama că și autobuzul așteaptă!*);
- Luînd toate restricțiile în considerare, ne interesează, de fapt, zona dintre cele două puncte anterioare, deci sub segmentul $c = a + 5$ și deasupra lui $c = a - 20$.

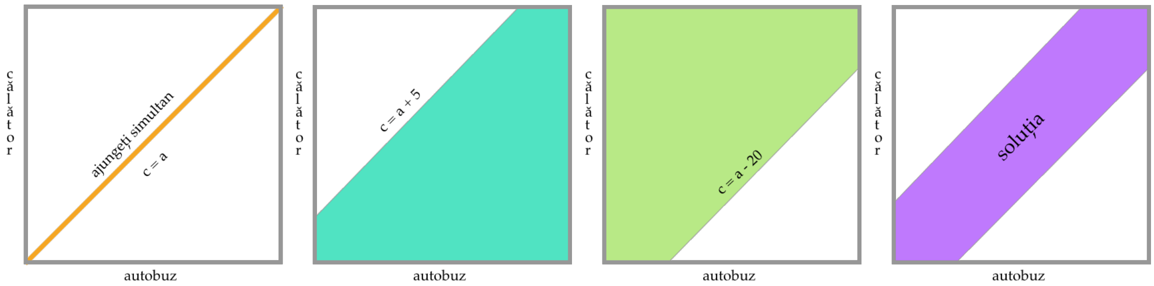
Acum trebuie doar să calculăm aria zonei corespunzătoare și să o împărțim la aria pătratului:

$$P = \frac{60^2 - \frac{55^2}{2} - \frac{40^2}{2}}{60^2} \simeq 36\%.$$

15. Trei persoane aleg la întîmplare un număr real între 0 și 1. Care este probabilitatea ca suma pătratelor numerelor alese de ei să nu depășească 1?

Soluție: Pentru a modela problema geometric, e clar că o putem gîndi ca pe alegerea unui punct din cubul unitate, $(x, y, z) \in [0, 1]^3$. Cazurile posibile sînt, atunci, date de volumul cubului, care este 1.

[1] Mai multe exemple și explicații puteți găsi, de exemplu, aici.



(a) $c = a$ (b) $c \leq a + 5$ (c) $c \leq a - 20$ (d) Soluția

Figura 9.1: *Problema autobuzului (cu așteptare)*

Pentru cazurile favorabile, avem nevoie de $x^2 + y^2 + z^2 \leq 1$, care este sfera unitate. Dar, cum ne interesează doar alegerile de numere pozitive, i.e. partea din sferă care se găsește în interiorul cubului, adică $\frac{1}{8}$ din ea.

Avem, deci:

$$P = \frac{\frac{1}{8} \cdot \frac{4\pi \cdot 1^3}{3}}{1} \simeq 52\%.$$

Mult mai multe exemple sofisticate de probabilități geometrice se pot găsi la Wolfram MathWorld.

SEMINAR 10

SCHEME CLASICE DE PROBABILITATE

Schemele clasice de probabilitate sînt modele matematice care ne permit calculul probabilității de realizare a unui eveniment în cazul unor distribuții anume.

10.1 Schema lui Poisson

Schema lui Poisson se formulează astfel. Presupunem că avem n urne U_i, $1 \leq i \leq n$, care conțin bile albe și bile negre în proporții cunoscute. Fie p_i probabilitatea de extragere a unei bile albe din urna U_i, respectiv q_i, probabilitatea de extragere a unei bile negre din urna U_i.

Extragem cîte o bilă din fiecare urnă. Probabilitatea de a obține k bile albe este coeficientul lui X^k din polinomul:

$$\prod_{i=1}^{n}(p_iX + q_i) = (p_1X + q_1)(p_2X + q_2)\cdots(p_nX + q_n).$$

Observație: În cazul schemei lui Poisson, extragerea se face *fără revenire*. O bilă, după ce a fost extrasă, nu este repusă în urnă, astfel că, după fiecare extragere, numărul de bile din urnă scade.

Exemplu: Avem 3 sisteme de siguranță, primul funcționează cu probabilitatea de 80%, al doilea, de 70%, iar al treilea, în 90% din cazuri. Să se determine probabilitatea ca oricare două dintre sisteme să funcționeze simultan.

Soluție: Ne aflăm în cazul schemei lui Poisson, „bilele“ fiind sistemele de siguranță, iar „extragerea“ înseamnă funcționarea lor. Așadar, căutăm coeficientul lui X^2 din polinomul:

$$\pi = (0,8 \cdot X + 0,2)(0,7 \cdot X + 0,3)(0,9 \cdot X + 0,1),$$

coeficient care este $p = 0,398$.

10.2 Schema lui Bernoulli (binomială)

Situația este similară cu cea din cazul Poisson, doar că acum extragerile se fac *cu repunere*. Adică, după ce fiecare bilă este extrasă și se înregistrează culoarea sa, este repusă în urnă, pentru a putea fi extrasă din nou, eventual.

Așadar, schema lui Bernoulli se formulează astfel. Avem o urnă cu a bile albe și b bile negre. Extragem cu repunere n bile. Probabilitatea să extragem k bile albe, cu $0 \leq k \leq n$ este dată de:

$$p(n,k) = C_n^k p^k q^{n-k}, \quad \text{unde } p = \frac{a}{a+b}, q = \frac{b}{a+b}.$$

De remarcat că în acest caz, p corespunde evenimentului favorabil, iar $q = 1 - p$.

Exemplu: Se aruncă două zaruri de 10 ori. Care este probabilitatea să se obțină de exact 4 ori suma 7?

Soluție: Ne aflăm în cazul schemei lui Bernoulli, unde „bilele" sînt sumele de pe zaruri, iar „extragerea" este aruncarea zarului. Evident, modelul potrivit este acela al extragerii cu revenire, deoarece orice sumă poate fi obținută la fiecare aruncare.

În acest caz, avem $p = \frac{1}{6}$, deoarece suma 7 poate fi obținută în 6 moduri din punctele de pe două zaruri, iar cazurile posibile sînt 36. Rezultă $q = \frac{5}{6}$.

În plus, avem $n = 10, k = 4$, deci:

$$p(10,4) = C_{10}^4 \frac{1}{6^4} \cdot \frac{5^6}{6^6}.$$

10.3 Schema multinomială

Schema lui Bernoulli se mai numește *schema binomială*, deoarece $p(n,k)$ este, de fapt, coeficientul lui X^k din dezvoltarea binomului $(pX + q)^n$.

O generalizare a acestei situații este *schema multinomială*, în care presupunem că avem o urnă cu bile de $s \in \mathbb{N}$ culori și se extrag cu repunere n bile. Probabilitatea de a extrage k_i bile de culoare i, cu $1 \leq i \leq s$ este dată de formula:

$$p(n; k_1, \ldots, k_s) = \frac{n!}{k_1! k_2! \cdots k_s!} \cdot p_1^{k_1} \cdot p_2^{k_2} \cdots p_s^{k_s},$$

unde $n = k_1 + \cdots + k_s$ și $p_1 + \cdots + p_s = 1$, iar p_i este probabilitatea de a extrage o bilă de culoare i.

Exemplu: Se aruncă cu un zar de 10 ori. Care este probabilitatea ca exact de 2 ori să apară fața cu 1 punct și de 3 ori să apară fața cu 2 puncte?

Soluție: Evident, sîntem în cazul schemei multinomiale. „Culorile" sînt punctele de pe zaruri și vrem să „extragem 2 bile de culoarea 1 și 3 bile de culoarea 2".

Atunci $n = 10$, deoarece facem 10 „extrageri“, $k_1 = 2$ și $k_2 = 3$. În plus, avem $p_1 = p_2 = \frac{1}{6}$, deoarece orice număr de pe zar are aceeași probabilitate de a apărea. Rezultă, de asemenea, că avem nevoie și de $k_3 = n - k_1 - k_2 = 5$ și $p_3 = 1 - p_1 - p_2 = \frac{2}{3}$.

Probabilitatea cerută este:

$$p(10; 2, 3, 5) = \frac{10!}{2! \cdot 3! \cdot 5!} \cdot \frac{1}{6^2} \cdot \frac{1}{6^3} \cdot \frac{2^5}{3^5}.$$

10.4 Schema geometrică

Dintr-o urnă cu a bile albe și b bile negre, extragem fără repunere n bile, cu $n \leq a + b$. Probabilitatea de a obține $k \leq a$ bile albe, *fără repunere*, este dată de:

$$p_{a,b}^{k,n-k} = \frac{C_a^k \cdot C_b^{n-k}}{C_{a+b}^n}.$$

Mai general, dacă avem a_i bile de culoarea i, $1 \leq i \leq s$ și extragem n bile fără repunere, atunci probabilitatea de a extrage k_i, cu $k_1 + \cdots + k_s = n$ bile de culoarea i este dată de formula:

$$p_{a_1,\ldots,a_s}^{k_1,\ldots,k_s} = \frac{C_{a_1}^{k_1} \cdots C_{a_s}^{k_s}}{C_{k_1+\cdots+k_s}^n}.$$

Exemplu: Avem un lot de 100 articole, dintre care 80 sînt corespunzătoare, 15 sînt cu defecțiuni remediabile și 5 rebuturi. Alegem 6 articole. Care este probabilitatea ca dintre acestea, 3 să fie bune, 2 cu defecțiuni remediabile și un rebut?

Soluție: Presupunem că extragerile se fac cu repunere. Atunci sîntem în cazul schemei multinomiale, deci:

$$p(6; 3, 2, 1) = \frac{6!}{3! \cdot 2! \cdot 1!} \left(\frac{80}{100}\right)^3 \cdot \left(\frac{15}{100}\right)^2 \cdot \left(\frac{5}{100}\right)^1.$$

Dacă extragerile se fac fără repunere, ne aflăm în cazul schemei geometrice și avem:

$$p_{80,15,5}^{3,2,1} = \frac{C_{80}^3 \cdot C_{15}^2 \cdot C_5^1}{C_{100}^6}.$$

10.5 Exerciții

1. Un lot de 50 de procesoare conține 3 defecte. Alegem la întîmplare 10, simultan, și le verificăm. Care este probabilitatea ca 2 să fie defecte?

Indicație: Schema geometrică (fără repunere).

2. (a) Un semnal este transmis pe 3 canale diferite, iar probabilitățile de recepționare corectă sînt 0,9; 0,8 și 0,7. Care este probabilitatea să se recepționeze corect un semnal?

(b) Dar dacă semnalul este transmis pe un canal ales la întîmplare și presupunem că orice canal poate fi ales cu aceeași probabilitate?

Indicație: (a) Schema lui Poisson.

(b) Formula probabilității totale: $\frac{1}{3} \cdot (0,9 + 0,8 + 0,7)$.

3. O urnă conține 2 bile albe și 3 bile negre. Alegem la întîmplare 2 bile, fără repunere. Care este probabilitatea ca acestea să fie ambele negre?

Dar dacă extragerea se face cu repunere?

4. Se testează 5 dispozitive care funcționează în condiții identice, independent și cu randament de 0,9 fiecare. Care este probabilitatea ca exact două să funcționeze?

Indicație: Schema Bernoulli.

5. Se consideră 3 urne: U_1, care conține 5 bile albe și 5 negre, U_2, care conține 4 bile albe și 6 negre și U_3, care conține 4 bile albe și 5 negre. Din fiecare urnă se extrag cu repunere cîte 5 bile.

Care este probabilitatea ca din 2 urne să obținem cîte 2 bile albe și 3 negre, iar din a treia urnă să obținem o altă combinație?

Indicație: Schema Poisson sau Bernoulli, după cazuri (cu/fără repunere).

6. Se consideră urnele din problema precedentă. Din fiecare urnă se extrage cîte o bilă. Dacă se repetă experiența de 5 ori, care este probabilitatea ca de 3 ori să se obțină o bilă albă și 2 negre?

Indicație: Schema Poisson.

7. Fie 8 canale de transmisie a informației care funcționează independent. Presupunem că un canal este activ cu probabilitatea 1/3.

Să se calculeze probabilitatea ca la un moment dat să fie mai mult de 6 canale active.

Indicație: Schema Bernoulli.

8. Șase vînători au zărit o vulpe și au tras simultan. Presupunem că de la distanța de la care au tras, fiecare vînător nimerește în mod obișnuit vulpea cu o probabilitate de $\frac{1}{3}$. Aflați probabilitatea ca vulpea să fie nimerită.

9. Se aruncă o monedă de 8 ori. Aflați probabilitatea ca stema să apară de 6 ori. Aflați probabilitatea ca stema să apară de *cel puțin* 6 ori.

10. Un muncitor produce piese astfel încît 99% sînt bune, 7% au defecte remediabile și 3% sînt rebuturi.

Se aleg aleatoriu 3 piese produse de muncitor. Care este probabilitatea ca dintre aceste 3 piese, cel puțin una să fie bună și cel puțin una să fie rebut?

SEMINAR 11

VARIABILE ALEATOARE

11.1 Cazul discret

Variabilele aleatoare sînt moduri de reprezentare a informației în calculul probabilităților, precum matricele ne pot ajuta să „codificăm“ informații geometrice sau algebrice. De obicei, ele se reprezintă asemenea permutărilor, anume ca tablouri bidimensionale.

Începem cu definițiile fundamentale.

Definiţie 11.1: O variabile a cărei valoare se determină de un eveniment rezultat în urma unei experiențe se numește *variabilă aleatoare.*

Definiţie 11.2: Fie X o variabilă aleatoare care poate lua valorile $x_1, \dots, x_n$, cu probabilitățile $f(x_1)$, $\dots, f(x_n)$. Mulțimea de perechi ordonate $\{(x_i, f(x_i))\}_i$ definește *repartiția* variabilei aleatoare.

Mai general, dacă X este o variabilă aleatoare reală (adică $\mathrm{Im} f \subseteq \mathbb{R}$), atunci funcția $F : \mathbb{R} \longrightarrow \mathbb{R}$ definită prin:

$$F_X(x) = P(X < x), \forall x \in \mathbb{R}$$

se numește *funcția de repartiție* a variabilei aleatoare X.

Putem interpreta funcția de repartiție în sensul următor: calculată în x, ea ne dă probabilitatea ca valorile variabilei aleatoare să fie pînă în x.

Se poate vedea din definiție că au loc următoarele *proprietăți*:

(1) $\lim\limits_{x \longrightarrow -\infty} F_X(x) = 0$, iar $\lim\limits_{x \longrightarrow \infty} F_X(x) = 1$;

(2) F_X este crescătoare (deoarece acumulează) și este continuă la dreapta în orice punct $x \in \mathbb{R}$;

(3) $P(X = x) = F(x) - F(x - 0)$, unde $F(x - 0)$ notează limita la stînga

$$\lim_{x_0 \nearrow x} F(x_0).$$

(4) $P(a < X \leq b) = F(b) - F(a)$.

Ne vor interesa două cazuri, anume acelea ale variabilelor aleatoare continue și discrete, pe care le definim mai jos.

Definiție 11.3: Variabila aleatoare X se numește *discretă* dacă mulțimea valorilor ei este o mulțime cel mult numărabile de numere reale sau complexe (în bijecție cu o submulțime [nestrictă] a lui $\mathbb{N}$), $\{a_n\}_n$.

În acest caz, dacă $P(X = a_n) = p_n$, atunci condiția de normare este $\sum p_n = 1$, iar funcția de repartiție F_X este o funcție în scară, cu:

$$F_X(x) = \sum_{a_n < x} p_n.$$

În acest caz, putem reprezenta variabila aleatoare într-o formă matriceală:

$$X \sim \begin{pmatrix} a_0 & a_1 & \dots & a_n & \dots \\ p_0 & p_1 & \dots & p_n & \dots \end{pmatrix},$$

matrice care se numește *matricea de repartiție* a lui X sau *distribuția* sa.

Cazul continuu este definit mai jos.

Definiție 11.4: Dacă $P(X = x) = 0$, pentru orice $x \in \mathbb{R}$, atunci variabila aleatoare X se numește *continuă*. Echivalent, funcția ei de repartiție este o funcție reală continuă (v. proprietatea (3)).

În multe situații, vom lucra cu variabile aleatoare discrete, reprezentate matriceal. Va fi util să definim cîteva operații elementare între aceste obiecte.

Fie, așadar, două variabile aleatoare discrete:

$$X : \begin{pmatrix} x_i \\ p_i \end{pmatrix}, \quad Y : \begin{pmatrix} y_j \\ q_j \end{pmatrix}$$

Definim $p_{ij} = P(X = x_i, Y = y_j), \forall i, j$. Atunci au loc formulele de calcul:

$$\begin{aligned} p_i &= \sum_j p_{ij} \\ q_j &= \sum_i p_{ij} \\ 1 &= \sum_{i,j} p_{ij}. \end{aligned}$$

Variabilele X și Y se numesc **independente** dacă $p_{ij} = p_i \cdot q_j$.

Suma variabilelor are repartiția:

$$X + Y : \begin{pmatrix} x_i + y_j \\ p_{ij} \end{pmatrix}.$$

Dacă $\alpha \in \mathbb{R}$ este o constantă reală, atunci variabila $X + \alpha$ are repartiția:

$$X + \alpha \; : \; \begin{pmatrix} x_i + \alpha \\ p_i \end{pmatrix}.$$

Produsul variabilelor X și Y are repartiția:

$$XY \; : \; \begin{pmatrix} x_i y_j \\ p_{ij} \end{pmatrix}$$

Produsul αX, pentru $\alpha \in \mathbb{R}$ are repartiția:

$$\alpha X \; : \; \begin{pmatrix} \alpha x_i \\ p_i \end{pmatrix}.$$

Să vedem un exemplu de calcul:
Exemplu: Se dau variabilele independente:

$$X \; : \; \begin{pmatrix} 0 & 1 & 2 \\ 0,2 & 0,4 & 0,4 \end{pmatrix}, \quad Y \; : \; \begin{pmatrix} 1 & 2 \\ 0,6 & 0,4 \end{pmatrix}.$$

Determinați repartițiile variabilelor $X + Y, XY, 2 + X, X^2, 3Y$.

Soluție: Vom prezenta repartițiile punînd valorile pentru evenimente în ordine crescătoare. De exemplu, $X + Y$ poate lua valorile 1, 2, 3, 4, iar X^2 poate lua valorile 0, 1, 4.

Folosim independența și formulele de calcul pentru probabilități și obținem, de exemplu:

$$\begin{aligned}
P(X + Y = 1) &= P(X = 0 \wedge Y = 1) \\
&= P(X = 0) \cdot P(Y = 1) \\
&= 0,12 \\
P(X + Y = 2) &= P\Big((X = 0 \wedge Y = 2) \vee (X = 1 \wedge Y = 1)\Big) \\
&= P(X = 0 \wedge Y = 2) + P(X = 1 \wedge Y = 1) \\
&= 0,32
\end{aligned}$$

Similar se calculează și celelalte valori și, folosind definițiile operațiilor, obținem:

$$X+Y : \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0,12 & 0,32 & 0,4 & 0,16 \end{pmatrix}$$
$$XY : \begin{pmatrix} 0 & 1 & 2 & 4 \\ 0,2 & 0,24 & 0,4 & 0,6 \end{pmatrix}$$
$$2+X : \begin{pmatrix} 2 & 3 & 4 \\ 0,2 & 0,4 & 0,4 \end{pmatrix}$$
$$X^2 : \begin{pmatrix} 0 & 1 & 4 \\ 0,2 & 0,4 & 0,4 \end{pmatrix}$$
$$3Y : \begin{pmatrix} 3 & 6 \\ 0,6 & 0,4 \end{pmatrix}$$

Pentru studiul variabilelor aleatoare se folosesc mai multe mărimi, pe care le introducem, pe rînd.

Definiţie 11.5: Fie X o variabilă aleatoare discretă.

(1) Dacă seria $\sum_{n\geq 0} x_n p_n$ este absolut convergentă, atunci spunem că X admite medie, iar numărul:

$$E(X) = \sum_{n\geq 0} x_n p_n$$

se numește *media* lui X;

(2) Pentru $n \neq 0$, media $E(X^n)$ a variabilei X^n se numește *momentul de ordinul n* al lui X;

(3) Media variabilei $(X - E(X))^2$ se numește *varianța* sau *dispersia* lui X, notată $\mathrm{Var}(X)$ sau $D^2(X)$.

Funcția de medie are următoarele proprietăți esențiale:

(a) **Liniaritatea:** $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$, pentru orice $\alpha, \beta \in \mathbb{C}$;

(b) **Monotonia:** Dacă $X \leq Y$, atunci $E(X) \leq E(Y)$;

(c) Dacă X este o constantă c, atunci $E(X) = c$;

(d) Dacă X și Y sînt variabile aleatoare independente, adică descriu evenimente independente, atunci $E(XY) = E(X) \cdot E(Y)$.

Funcția de dispersie are proprietățile:

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

(b) $D^2(X) \geq 0$;

(c) $D^2(X) = 0 \iff P(X = E(X)) = 1$, adică X este o constantă cu probabilitatea 1 (X descrie doar evenimentul sigur);

(d) **Cebîșev:** Pentru orice $\varepsilon > 0$, are loc:

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

Echivalent, se mai poate scrie:

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

(e) $D^2(\alpha X) = \alpha^2 D^2(X)$, pentru orice $\alpha \in \mathbb{C}$;

(f) Dacă X și Y sînt independente, atunci:

$$D^2(X + Y) = D^2(X) + D^2(Y).$$

Mai definim două mărimi care se asociază variabilelor aleatorii.

Definiție 11.6: Fie X și Y două variabile aleatoare. Numim *covarianță* a variabilelor X și Y, notată $\operatorname{cov}(X, Y)$ numărul:

$$\operatorname{cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y).$$

Proprietățile esențiale ale covarianței sînt date de:

(a) Dacă X și Y sînt variabile aleatoare independente, atunci $\operatorname{cov}(X, Y) = 0$;

(b) Dacă $X_1, \dots, X_n$ sînt variabile aleatoare cu dispersiile D_i^2 și covarianțele $\operatorname{cov}(X_i, X_j)$, $i \neq j$, atunci are loc:

$$D^2\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n D_i^2 + 2\sum_{i<j} \operatorname{cov}(X_i, X_j).$$

Definiție 11.7: Dacă X și Y sînt variabile aleatoare, se numește *coeficient de corelație* al lor, notat $\rho(X, Y)$, numărul calculat prin:

$$\rho(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sqrt{D^2(X) \cdot D^2(Y)}}$$

Se poate vedea din definiție că, atît covarianța, cît și corelația arată „gradul de independență“ a variabilelor aleatoare X și Y. Într-adevăr, dacă X și Y sînt independente, atunci avem $\rho(X, Y) = 0$. Afirmația reciprocă este, însă, falsă!

Să luăm un exemplu: Fie $A = \{a_1, a_2, a_3, a_4\}$, cu $P(a_i) = \frac{1}{4}$.

Presupunem că variabila aleatoare X realizează corespondența:

$$X(a_1) = 1, X(a_2) = -1, X(a_3) = 2, X(a_4) = -2.$$

Atunci, ea are matricea de repartiție:

$$X \sim \begin{pmatrix} -2 & -1 & 1 & 2 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Presupunem că o altă variabilă aleatoare Y realizează corespondența:

$$Y(a_1) = 1, Y(a_2) = 1, Y(a_3) = -1, Y(a_4) = -1.$$

Atunci matricea ei se poate reduce la:

$$Y \sim \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Produsul celor două variabile aleatoare are matricea de repartiție:

$$XY : \begin{pmatrix} -2 & -1 & 1 & 2 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Rezultă că $E(X) = E(Y) = E(XY)$, deci $\text{cov}(X, Y) = 0$. Însă putem observa că X și Y nu sînt independente. Într-adevăr, avem:

$$P(X = 1 \wedge Y = 1) = P(a_1) = \frac{1}{4} \neq P(X = 1) \cdot P(Y = 1) = \frac{1}{4} \cdot \frac{1}{2}.$$

În fine, o altă noțiune de care avem nevoie este *funcția generatoare.*

Definiție 11.8: Fie X o variabilă aleatoare discretă care ia valori naturale.

Funcția $G_X(t)$, definită prin:

$$G_X(t) = E(t^X) = \sum_{n \geq 0} p_n t^n,$$

cu $t \leq 1$ și $p_n = P(X = n)$ se numește *funcția generatoare* a variabilei aleatoare X.

Funcția generatoare codifică, de fapt, în coeficienții unei serii formale, repartiția unei variabile aleatoare. Ea are următoarele proprietăți:

(a) Dacă avem variabilele aleatoare independente $X_1, \dots, X_n$, iar $X = X_1 + \cdots + X_n$, atunci:

$$G_X(t) = G_{X_1}(t) \cdots G_{X_n}(t);$$

(b) $E(X) = G'_X(1)$;

(c) $D^2(X) = G''_X(1) + G'_X(1) - [G'_X(1)]^2$;

(d) Două variabile aleatoare cu aceeași funcție generatoare au aceeași repartiție.

11.2 Exemple de repartiții discrete

Vom regăsi acum, în contextul teoriei variabilelor aleatoare, scheme clasice de probabilitate.

11.2.1 Repartiția binomială (Bernoulli)

Presupunem că se fac n experiențe independente, în fiecare din experiențe, probabilitatea de realizare a unui eveniment A fiind constantă și egală cu p (rezultă că probabilitatea ca A să nu se realizeze este $q = 1 - p$).

Numărul de realizări ale evenimentului A în cele n experiențe se constituie ca o variabilă aleatoare X, ale cărei valori posibile sînt $k = 0, 1, \dots, n$, în funcție de cîte din experiențe realizează A. Așadar, avem $P(X = k) = p_k$.

Cînd considerăm fiecare k, cu probabilitatea sa de realizare, mulțimea perechilor (k, p_k), $k \in \{0, 1, \dots, n\}$ se numește *repartiție binomială* sau *repartiție Bernoulli*.

Pentru acestea, valorile principalilor parametri sînt:

- $P(X = k) = C_n^k p^k q^{n-k}, k \in \{0, 1, \dots, n\}$;
- $E(X) = np$;
- $D^2(X) = npq$;
- $G_X(t) = (pt + q)^n$.

Observație: Pentru $n \longrightarrow \infty, p \longrightarrow 0$, astfel încît $np = \lambda$, probabilitățile $P(X = k)$ pot fi aproximate prin valorile repartiției Poisson (vezi mai jos).

11.2.2 Repartiția geometrică

Fie X numărul de experimente de tip Bernoulli, cu probabilitatea p de succes, care trebuie repetate pînă la apariția primului succes. Atunci:

- $P(X = k) = p(1 - p)^{k-1}, k \in \mathbb{N}^*$;

- $E(X) = \frac{1}{p}$;
- $D^2(X) = \frac{1-p}{p^2}$;
- $G_X(t) = \frac{pt}{1-qt}$.

11.2.3 Repartiția binomial negativă

Fie X numărul de experimente de tip Bernoulli cu probabilitatea de succes p care trebuie efectuate pentru a obține m succese. Atunci:

- $P(X = n) = C_{n-1}^{m-1} p^m q^{n-m}$ (pentru $m = 1$, găsim repartiția geometrică);
- $E(X) = \frac{m}{p}$;
- $D^2(X) = \frac{mq}{p^2}$;
- $G_X(t) = \frac{p^m t^m}{(1-qt)^m}$.

11.2.4 Repartiția hipergeometrică

Fie X o variabilă aleatoare care reprezintă numărul de bile albe obținute după n extrageri fără repunere dintr-o urnă ce conține a bile albe și b bile negre. Atunci avem:

- $P(X = x) = \frac{C_a^x C_b^{n-x}}{C_{a+b}^n}, 0 \leq x \leq n$;
- $E(X) = np$;
- $D^2(X) = xpq \cdot \frac{a+b-x}{a+b-1}$, unde $p = \frac{a}{a+b}$, iar $q = 1 - p = \frac{b}{a+b}$.

11.2.5 Repartiția Poisson

Repartiția Poisson se caracterizează printr-un parametru $\lambda > 0$ (v. repartiția binomială [Bernoulli] mai sus). Avem:

- $P(X = n) = \frac{\lambda^n}{n!} \cdot e^{-\lambda}, n \in \mathbb{N}$;
- $E(X) = \lambda$;

- $D^2(X) = \lambda$;
- $G_X(t) = e^{\lambda(t-1)}$.

11.3 Cazul continuu

Definiţie 11.9: Variabila aleatoare X se numește *absolut continuă* dacă există o funcție integrabilă $f : \mathbb{R} \longrightarrow \mathbb{R}_+$ astfel încît funcția de repartiție $F(x)$ a lui X să fie dată de:

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

În acest caz, f se numește *densitatea de probabilitate* a lui X.

Ca în cazul discret, avem următoarele proprietăți esențiale:

(a) Funcția f este pozitivă și are loc condiția de normare

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

(b) În orice punct de continuitate a lui f, avem $F'(x) = f(x)$;

(c) $P(a \leq X \leq b) = \displaystyle\int_{a}^{b} f(x)dx$;

(d) $E(X) = \displaystyle\int_{-\infty}^{\infty} xf(x)dx$;

(e) Dacă X are densitatea X, iar $Y = aX + b$ este o altă variabilă aleatoare, cu $a, b \in \mathbb{R}, a \neq 0$, atunci Y are densitatea:

$$f_Y(x) = \frac{1}{|a|} f\left(\frac{x-b}{a}\right).$$

Definiţie 11.10: Pentru orice variabilă aleatoare X, funcția:

$$\varphi_X(t) = E(e^{itX})$$

se numește *funcția caracteristică* a lui X.

Funcția caracteristică joacă un rol similar funcției generatoare din cazul discret. Astfel, ea are următoarele proprietăți:

(a) Dacă $X_1, \ldots, X_n$ sînt variabile aleatoare independente, iar $X = X_1 + \cdots + X_n$, atunci:

$$\varphi_X(t) = \prod_{k=1}^{n} \varphi_{X_k}(t);$$

(b) Dacă X admite moment de ordinul n, atunci:

$$E(X^k) = \frac{\varphi_X^{(k)}(0)}{i^k}, \quad 1 \le k \le n;$$

(c) Două variabile aleatoare cu aceeași funcție caracteristică au aceeași repartiție;

(d) Dacă X este discretă, atunci obținem:

$$\varphi_X(t) = \sum_n e^{ita_n} p_n;$$

(e) Dacă X are densitatea de repartiție $f(x)$, atunci:

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx,$$

iar în punctele de continuitate ale lui f, are loc relația:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t)dt.$$

În continuare, prezentăm cele mai cunoscute repartiții absolut continue.

11.4 Repartiții absolut continue

11.4.1 Repartiția normală (gaussiană)

Dată media m și abaterea pătratică σ, repartiția variabilei aleatoare X are densitatea:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

În acest caz, se mai notează $X \sim N(m, \sigma)$.

Pentru cazul particular $m = 0$ și $\sigma = 1$, X se numește *variabilă aleatoare standard.*

Funcția caracteristică este dată de $\varphi_X(t) = e^{imt-\frac{\sigma^2 t^2}{2}}$.

Pentru cazul variabilelor normale standard, funcția de repartiție are o formă particulară și se numește *funcția lui Laplace* (eng. CDF, *Cumulative Distribution Function*):

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Observație: Valorile funcției Laplace sînt tabelate și, de obicei, se dau în probleme.

În general, dacă avem o variabilă aleatoare normală de tip $N(m, \sigma)$, atunci funcția ei de repartiție F se poate obține din funcția lui Laplace în forma:

$$F(x) = \Phi\left(\frac{x - m}{\sigma}\right).$$

Așadar, avem:

$$P(a \leq X \leq b) = \Phi\left(\frac{b - m}{\sigma}\right) - \Phi\left(\frac{a - m}{\sigma}\right), \tag{11.1}$$

de unde putem obține mai departe:

$$P(|X - m| < \varepsilon\sigma) = \Phi(\varepsilon) - \Phi(-\varepsilon) = 2\Phi(\varepsilon) - 1.$$

11.4.2 Repartiția uniformă

Pentru un interval $[a, b]$, densitatea de repartiție este:

$$f(x) = \begin{cases} \dfrac{1}{b - a}, & a \leq x \leq b \\ 0, & \text{în rest} \end{cases}.$$

În acest caz, avem media $E(X) = \dfrac{a + b}{2}$ și $D^2(X) = \dfrac{(b - a)^2}{12}$.

11.4.3 Repartiția exponențială

Dat un parametru $\lambda > 0$, repartiția exponențială are densitatea:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}.$$

Pentru aceasta, avem:

- $E(X) = \dfrac{1}{\lambda}$;
- $D^2(X) = \dfrac{1}{\lambda^2}$;
- $\varphi_X(t) = \left(1 - \dfrac{it}{\lambda}\right)^{-1}$.

11.4.4 Repartiția Gamma

Dați parametrii $\lambda, p > 0$, densitatea variabilei aleatoare de repartiție Gamma este:

$$f(x) = \begin{cases} \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases},$$

unde Γ este funcția lui Euler. În acest caz, avem:

- $E(X) = \frac{p}{\lambda}$;
- $D^2(X) = \frac{p}{\lambda^2}$;
- $\varphi_X(t) = \left(1 - \frac{it}{\lambda}\right)^{-p}$.

Din repartiția Gamma, putem obține alte cîteva cazuri particulare:

- pentru $p = 1$, se obține repartiția exponențială;
- pentru $\lambda = \frac{1}{2}, p = \frac{n}{2}$, se obține repartiția numită *hi pătrat*, cu n grade de libertate, notată $\chi^2(n)$.

11.4.5 Repartiția Cauchy

Aceasta are densitatea $f(x) = \frac{1}{\pi(1 + x^2)}$.

În acest caz, X nu admite valoare medie.

11.5 Exerciții rezolvate

1. Doi parteneri egal cotați boxează 12 runde. Probabilitatea ca oricare dintre ei să cîștige o rundă este 50%. Să se calculeze valoarea medie, dispersia și abaterea medie pătratică a variabilei aleatoare care reprezintă numărul de runde cîștigate de unul din parteneri.

Soluție: Variabila aleatoare are o distribuție binomială. Astfel, avem:

- $P(X = k) = C_{12}^k \frac{1}{2^k} \cdot \frac{1}{2^{12-k}}$;
- $E(X) = np = 12 \cdot \frac{1}{2}$;

- $D^2(X) = npq = 12 \cdot \dfrac{1}{2} \cdot \frac{1}{2} = 3;$
- $D(X) = \sqrt{D^2(X)} = \sqrt{3}.$

2. La o agenție de turism s-a constatat că 5% dintre persoanele care au făcut rezervări renunță. Se presupune că s-au făcut 100 de rezervări pentru un hotel cu 95 locuri. Care este probabilitatea ca toate persoanele ce se prezintă la hotel să aibă loc?

Soluție: Fie X variabila aleatoare corespunzătoare numărului de persoane care se prezintă la hotel. X urmează o repartiție binomială, cu $n = 100$ și $p = 0,95$. Atunci avem:

$$P(X \leq 95) = 1 - P(X > 95) = 1 - \sum_{k=96}^{100} C_{100}^{k}(0,05)^{100-k} \cdot (0,95)^{k}.$$

3. La examenul de matematică, probabilitatea ca o teză să fie notată cu notă de trecere este 75%. Se aleg la întîmplare 10 lucrări. Fie X variabila aleatoare ce reprezintă numărul tezelor ce vor fi notate cu notă de trecere. Calculați:

(a) repartiția lui X ;

(b) $E(X), D^2(X)$;

(c) $P(X \geq 5), P(7 \leq X \leq 10/X \geq 8), P(X = 10)$;

(d) funcția caracteristică a lui X.

Soluție: (a) Avem o variabilă aleatoare cu repartiție binomială și $n = 10$, $p = 0,75$. Atunci X ia valorile $x \in \{0, 1, \ldots, 10\}$ cu probabilitățile corespunzătoare distribuției binomiale, deci:

$$p(x) = C_{10}^{x}(0,75)^{x} \cdot (0,25)^{10-x}.$$

(b) Conform formulelor de calcul, avem:

$$\begin{aligned} E(X) &= np = 10 \cdot 0,75 = 7,5 \\ D^2(X) &= npq = 10 \cdot 0,75 \cdot 0,25 = 1,875. \end{aligned}$$

(c)

$$P(X \geq 5) = \sum_{x=5}^{10} C_{10}^{x}(0,75)^{x} \cdot (0,25)^{10-x}$$

Pentru $P(7 \leq X \leq 10/X \geq 8)$ avem un *eveniment condiționat.* Dacă A și B sînt două evenimente, amintim că probabilitatea $P(A/B)$, adică probabilitatea evenimentului A condiționat de evenimentul prealabil B este:

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

Atunci, obținem:

$$\begin{aligned}
P(7 \le X \le 10/X \ge 8) &= \frac{P(8 \le X \le 10)}{P(X \ge 8)} \\
&= \frac{\sum_{x=8}^{10} C_{10}^{x}(0,75)^{x}(0,25)^{10-x}}{\sum_{x=8}^{10} C_{10}^{x}(0,75)^{x}(0,25)^{10-x}} \\
&= 1
\end{aligned}$$

În fine:

$$P(X = 10) = C_{10}^{10}(0,75)^{x}(0,25)^{0}.$$

(d) Calculăm succesiv:

$$\begin{aligned}
\varphi_X(t) &= E(e^{itX}) \\
&= \sum_{x=0}^{10} e^{itx} C_{10}^{x}(0,75)^{x}(0,25)^{10-x} \\
&= (0,75 \cdot e^{it} + 0,25)^{10}.
\end{aligned}$$

4. Presupunem că la 100 de convorbiri telefonice au loc 1000 de bruiaje. Care este probabilitatea de a avea o convorbire fără bruiaje? Dar una cu cel puțin 2 bruiaje?

Variabila aleatoare ce reprezintă numărul de bruiaje se consideră a fi repartizată Poisson, cu $\lambda = E(X)$.

Soluție: Avem $E(X) = \dfrac{1000}{100} = 10 = \lambda$, care este și parametrul distribuției Poisson. Atunci:

$$\begin{aligned}
P(X = 0) &= e^{-10} \cdot \frac{10^0}{0!} = e^{-10} \\
P(X \ge 2) &= 1 - P(X = 0) - P(X = 1) \\
&= 1 - e^{-10} - 10 \cdot e^{-10}.
\end{aligned}$$

5. Într-o mină au loc în medie 2 accidente pe săptămînă, distribuite după legea Poisson. Să se calculeze probabilitatea de a avea cel mult 2 accidente:

(a) într-o săptămînă;

(b) în 2 săptămîni;

(c) în fiecare săptămînă dintr-un interval de 2 săptămîni.

Soluție: (a) Fie X variabila aleatoare ce desemnează numărul de accidente dintr-o săptămînă. Distribuția este de tip Poisson, cu $\lambda = E(X) = 2$. Atunci:

$$P(X \leq 2) = e^{-2}\left(1 + \frac{2}{1!} + \frac{4}{2!}\right) = 5e^{-2}.$$

(b) Fie Y variabila aleatoare ce desemnează numărul de accidente în 2 săptămîni. Distribuția este de asemenea de tip Poisson, cu $\lambda = 2 + 2 = 4$. Atunci:

$$P(Y \leq 2) = e^{-5}\left(1 + \frac{4}{1!} + \frac{16}{2!}\right) = 13e^{-4}.$$

(c) Probabilitatea cerută este:

$$P(X \leq 2) \cdot P(X \leq 2) = 25e^{-4}.$$

6. La un anumit cîntar erorile de măsurare sînt normal distribuite cu $m = 0$ și $\sigma = 0,1g$. Dacă se cîntărește un obiect la acest aparat, care este probabilitatea ca eroarea de măsurare să fie mai mică decît 0,15g?

Soluție: Fie X variabila aleatoare care reprezintă eroarea de măsurare dată cînd un obiect este cîntărit. Această variabilă este continuă și distribuită normal cu parametrii dați. Notăm acest lucru cu $X \sim N(0; 0,1)$. Vrem să calculăm probabilitatea $P(-0,15 \leq X \leq 0,15)$.

Avem succesiv (cf. (11.1)):

$$\begin{aligned} P(-0,15 \leq X \leq 0,15) &= P\left(\frac{-0,15-0}{0,1} \leq X \leq \frac{-0,15+0}{0,1}\right) \\ &= P(-1,5 \leq X \leq 1,5) \\ &= \Phi(1,5) - \Phi(-1,5) \\ &= \Phi(1,5) - 1 + \Phi(1,5) \\ &= 0,8664. \end{aligned}$$

11.6 Exerciții propuse

1. Dacă efectuăm o măsurătoare, iar rezultatul este cuprins între valorile k și $k+1$, îl aproximăm cu $k + \frac{1}{2}$. Astfel, eroarea comisă poate fi presupusă uniform distribuită pe intervalul $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

Care este probabilitatea ca eroarea să fie mai mare ca $\frac{1}{4}$?

2. Fie variabila aleatoare X, distribuită uniform pe $(-1,1)$. Determinați densitățile de probabilitate, mediile și dispersiile pentru variabilele $Y = 2X + 1$ și $Z = 2X^2 + 1$.

3. Durata medie de funcționare a unei baterii este o variabilă aleatoare, pe care o notăm X și care este repartizată normal, $N(m, \sigma^2)$, unde $m = 120$ zile este timpul mediu de funcționare, iar $\sigma = 10$ zile este abaterea față de medie. Determinați:

(a) probabilitatea ca bateria să funcționeze cel puțin 100 zile;

(b) probabilitatea ca bateria să funcționeze între 100 și 150 zile;

(c) intervalul de timp în care se poate spune că bateria funcționează cu o probabilitate de cel puțin 75% („aproape sigur").

4. Fie $a, b > 0$ și $n \in \mathbb{N}$. Determinați a, b astfel încît funcția $f_n(x) = ax^n e^{-b^2x^2}$ pentru $x > 0$ să fie o densitate de probabilitate pentru o variabilă aleatoare cu media m.

5. Fie densitatea de probabilitate a unei variabile aleatoare X de forma:

$$f(x) = \begin{cases} c \cdot \ln\left(\dfrac{a}{x}\right), & 0 < x < a \\ 0, & \text{în rest} \end{cases}$$

Să se determine constanta c astfel încît f să fie corect definită și să se calculeze valoarea medie și dispersia variabilei X.

6. Determinați funcția caracteristică a variabilei aleatoare X cu densitatea de probabilitate dată de:

$$f(x) = \begin{cases} \dfrac{1}{2}\left(1 - \dfrac{|x|}{2}\right), & |x| \leq 2 \\ 0, & |x| > 2 \end{cases}$$

Aflați funcția caracteristică a acestei variabile aleatoare.

7. Fie X, Y două variabile aleatoare independente, cu $X : N(1, 2)$ și $Y : U(0, 2)$. Determinați $M(X+Y), M(XY), M(X^2), M(X-Y^2), D^2(X+Y), D^2(X-Y)$.

8. Presupunem că emisia de particule dintr-un experiment este modelată de legea Poisson, cu $\lambda = 30$ particule pe oră. Care este probabilitatea ca în 10 minute să fie emise cel mult 20 particule?

9. Fie variabilele aleatoare X, Y, în relația $Y = 2 - 3X$. Calculați media și dispersia pentru Y, precum și corelația și coeficientul de corelație, știind că $m_X = -1, D_X^2 = 4$.

10. Presupunem că modulul vectorului viteză a unei particule este o variabile aleatoare care este repartizată după legea Maxwell, adică are:

$$f(v) = \frac{4h^2}{\sqrt{\pi}} \cdot v^2 \cdot e^{-h^2v^2}, \quad v > 0.$$

Aflați media și dispersia vitezei.

SEMINAR 12

VECTORI ALEATORI BIDIMENSIONALI

Putem gîndi vectorii aleatori bidimensionali ca pe niște *matrice aleatoare.* Însă noțiunea este chiar ceva mai generală. Astfel, dacă X și Y sînt două variabile aleatoare, atunci pur și simplu formăm perechea (X, Y), care se numește *vector aleator bidimensional.* În majoritatea cazurilor, vom lucra cu X și Y definite pe același spațiu de evenimente, deci dacă $\text{card}(X) = \text{card}(Y) = n$, atunci perechea lor devine o matrice aleatoare $2 \times n$.

Repartiția unei astfel de matrice aleatoare se calculează intuitiv folosind repartițiile individuale. Avem, deci:

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

unde se pot considera $1 \leq i \leq n$ și $1 \leq j \leq m$. Mai notăm probabilitățile individuale cu p_i, respectiv q_j, adică $p_i = P(X = x_i)$ și $q_j = P(Y = y_j)$. Aceste mărimi se mai numesc *probabilități marginale.*

Evident, dacă fixăm oricare dintre cele două variabile, obținem, de exemplu, pentru X fixat:

$$\cup_j (X = x_i, Y = y_j) = \sum_j P(X = x_i, Y = y_j) = \sum_j p_{ij} = p_i,$$

deoarece ambele variabile aleatoare X și Y sînt, individual, normate. Adică $\sum_i p_i = \sum_i P(X = x_i) = \sum_j q_j = \sum_j P(Y = y_j) = 1$.

Analog, desigur, $\sum_i p_{ij} = q_j$.

Luate împreună, obținem automat condiția de normare pentru perechea de variabile aleatoare, adică $\sum_{i,j} p_{ij} = 1$, în ipoteza că ambele componente ale perechii sînt normate.

Notăm cu $F_{(X,Y)}(x, y) = P(X < x, Y < y)$ funcția de repartiție a perechii de variabile aleatoare (vectorului aleator bidimensional) (X, Y). Separat, $F_X(x)$ și $F_Y(y)$ se numesc *funcții de repartiție marginale* ale vectorului aleator (X, Y).

Pentru *cazul continuu*, întîlnim noțiunea cunoscută din cazul vectorilor aleatori 1-dimensionali. Mai precis, avem:

$$F_{(X,Y)}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) du dv,$$

unde f este *densitatea de probabilitate* și este o funcție cu valori pozitive, dublu-normată, adică:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u,v)dudv = 1.$$

Totodată, putem recupera din această ecuație și densitatea de probabilitate știind funcția de repartiție, prin:

$$f(x,y) = \frac{\partial^2 F_{(X,Y)}(x,y)}{\partial x \partial y},$$

evident, pentru punctele (x, y) unde egalitatea are sens.

12.1 Repartiții condiționate

Pornim cu un vector aleator bidimensional (X, Y), discret. Ne punem problema din cazul probabilităților condiționate, i.e. vom nota cu $P(x_i/y_j)$ probabilitatea să aibă loc evenimentul $X = x_i$, știind că a avut loc (i.e. condiționat de) evenimentul $Y = y_j$.

Aceasta se poate calcula simplu, ca în cazul probabilităților condiționate:

$$P(x_i/y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{q_j}.$$

Invers, obținem desigur $P(y_j/x_i) = \frac{p_{ij}}{p_i}$.

Dacă fixăm un eveniment $Y = y$, putem defini o *funcție de repartiție condiționată*, $F(x/y) = P(X < x/Y = y)$, asociată vectorului X (deoarece Y este fixat). Aceasta se calculează generalizînd formulele de mai sus pe cazul continuu, anume:

$$F(x/y) = \frac{1}{f_Y(y)} \cdot \int_{-\infty}^{x} f(u,y)du.$$

Prin derivare în raport cu x, obținem densitatea de probabilitate a lui X, în ipoteza că $Y = y$:

$$f_X(x/y) = \frac{f(x,y)}{f_Y(y)}.$$

12.2 Variabile aleatoare independente

Știm deja că două variabile aleatoare X și Y sînt independente dacă evenimentele $X = x_i$ și $Y = y_j$ sînt independente, adică dacă are loc:

$$P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j).$$

Egalitatea poate fi formulată și cun funcțiile de repartiție:

$$F_{(X,Y)}(x,y) = F_X(x) \cdot F_Y(y).$$

Mai mult decît atît, rezultatul de mai jos ne arată că egalitatea are loc și pentru densitățile de probabilitate:

Teoremă 12.1: *Fie X, Y variabile aleatoare cu densitățile de probabilitate f_X, respectiv f_Y. Atunci X și Y sînt independente dacă și numai dacă are loc:*

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

unde $f(x,y)$ este densitatea de probabilitate a vectorului (X,Y), format de cele două variabile aleatoare.

Atenție: Dacă lucrăm cu alte variabile aleatoare U, V, obținute din X și Y prin transformări funcționale de forma generală:

$$U = \varphi(X,Y), \quad V = \psi(X,Y),$$

atunci calculele care folosesc densitățile de probabilitate și funcțiile de repartiție se fac folosind schimbări de variabile de forma:

$$u = \varphi(x,y), v = \psi(x,y) \iff x = \varphi_1(u,v), y = \psi_1(u,v),$$

după care noua densitate de probabilitate se schimbă cu ajutorul jacobianului transformărilor de mai sus:

$$f_{(U,V)}(u,v) = \left|\frac{D(x,Y)}{D(u,v)}\right| \cdot f_{(X,Y)}(\varphi_1(u,v), \psi_1(u,v)),$$

unde:

$$\left|\frac{D(x,y)}{D(u,v)}\right| = \det \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix},$$

determinantul matricei jacobiene a transformării ($x_u = \dfrac{\partial x}{\partial u}$ etc.).

Dintre toate caracteristicile numerice asociate variabilelor aleatoare (medie, dispersie, coeficient de corelație, covarianță), pentru cazul bidimensional avem în plus *media condiționată*. Astfel, fie X, Y variabile aleatoare *discrete*. Atunci putem calcula mediile condiționate prin:

- $E(X/Y = y_j) = \sum_i x_i \cdot P(x_i/y_j)$;
- $E(Y/X = x_i) = \sum_j y_j P(y_j/x_i)$.

Pentru *cazul continuu*, sumele devin integrale, iar probabilitățile devin densități:

- $E(X/Y = y) = \int_{-\infty}^{\infty} x \cdot f_X(x/y) dx$;
- $E(Y/X = x) = \int_{-\infty}^{\infty} y \cdot f_Y(y/x) dy$.

12.3 Caracteristici numerice

Media și dispersia unei variabile aleatoare au fost discutate anterior. De asemenea, atunci cînd lucrăm cu două sau mai multe variabile aleatoare, devin relevante și alte mărimi, precum *covarianța* și *coeficientul de corelație*. Le redăm mai jos, împreună cu proprietăți importante ale lor și ale mediei și dispersiei.

Funcția de medie are următoarele proprietăți esențiale:

(a) **Liniaritatea:** $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$, pentru orice $\alpha, \beta \in \mathbb{C}$;

(b) **Monotonia:** Dacă $X \leq Y$, atunci $E(X) \leq E(Y)$;

(c) Dacă X este o constantă c, atunci $E(X) = c$;

(d) Dacă X și Y sînt variabile aleatoare independente, adică descriu evenimente independente, atunci $E(XY) = E(X) \cdot E(Y)$.

Funcția de dispersie are proprietățile:

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

(b) $D^2(X) \geq 0$;

(c) $D^2(X) = 0 \iff P(X = E(X)) = 1$, adică X este o constantă cu probabilitatea 1 (X descrie doar evenimentul sigur);

(d) **Cebîșev:** Pentru orice $\varepsilon > 0$, are loc:

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

Echivalent, se mai poate scrie:

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

(e) $D^2(\alpha X) = \alpha^2 D^2(X)$, pentru orice $\alpha \in \mathbb{C}$;

(f) Dacă X și Y sînt independente, atunci:

$$D^2(X + Y) = D^2(X) + D^2(Y).$$

Mai definim două mărimi care se asociază variabilelor aleatorii.

Definiţie 12.1: Fie X și Y două variabile aleatoare. Numim *covarianță* a variabilelor X și Y, notată $\text{cov}(X, Y)$ numărul:

$$\text{cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y).$$

Proprietățile esențiale ale covarianței sînt date de:

(a) Dacă X și Y sînt variabile aleatoare independente, atunci $\operatorname{cov}(X, Y) = 0$;

(b) Dacă $X_1, \dots, X_n$ sînt variabile aleatoare cu dispersiile D_i^2 și covarianțele $\operatorname{cov}(X_i, X_j)$, $i \neq j$, atunci are loc:

$$D^2\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} D_i^2 + 2\sum_{i<j} \operatorname{cov}(X_i, X_j).$$

Definiție 12.2: Dacă X și Y sînt variabile aleatoare, se numește *coeficient de corelație* al lor, notat $\rho(X, Y)$, numărul calculat prin:

$$\rho(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sqrt{D^2(X) \cdot D^2(Y)}}$$

Se poate vedea din definiție că, atît covarianța, cît și corelația arată „gradul de independență“ a variabilelor aleatoare X și Y. Într-adevăr, dacă X și Y sînt independente, atunci avem $\rho(X, Y) = 0$. Afirmația reciprocă este, însă, falsă!

Să luăm un exemplu: Fie $A = \{a_1, a_2, a_3, a_4\}$, cu $P(a_i) = \dfrac{1}{4}$.

Presupunem că variabila aleatoare X realizează corespondența:

$$X(a_1) = 1, X(a_2) = -1, X(a_3) = 2, X(a_4) = -2.$$

Atunci, ea are matricea de repartiție:

$$X \sim \begin{pmatrix} -2 & -1 & 1 & 2 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Presupunem că o altă variabilă aleatoare Y realizează corespondența:

$$Y(a_1) = 1, Y(a_2) = 1, Y(a_3) = -1, Y(a_4) = -1.$$

Atunci matricea ei se poate reduce la:

$$Y \sim \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Produsul celor două variabile aleatoare are matricea de repartiție:

$$XY : \begin{pmatrix} -2 & -1 & 1 & 2 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Rezultă că $E(X) = E(Y) = E(XY)$, deci $\operatorname{cov}(X, Y) = 0$. Însă putem observa că X și Y nu sînt independente. Într-adevăr, avem:

$$P(X = 1 \wedge Y = 1) = P(a_1) = \frac{1}{4} \neq P(X = 1) \cdot P(Y = 1) = \frac{1}{4} \cdot \frac{1}{2}.$$

În fine, o altă noțiune de care avem nevoie este *funcția generatoare*.

Definiţie 12.3: Fie X o variabilă aleatoare discretă care ia valori naturale.

Funcția $G_X(t)$, definită prin:

$$G_X(t) = E(t^X) = \sum_{n \geq 0} p_n t^n,$$

cu $t \leq 1$ și $p_n = P(X = n)$ se numește *funcția generatoare* a variabilei aleatoare X.

Funcția generatoare codifică, de fapt, în coeficienții unei serii formale, repartiția unei variabile aleatoare. Ea are următoarele proprietăți:

(a) Dacă avem variabilele aleatoare independente $X_1, \ldots, X_n$, iar $X = X_1 + \cdots + X_n$, atunci:

$$G_X(t) = G_{X_1}(t) \cdots G_{X_n}(t);$$

(b) $E(X) = G'_X(1)$;

(c) $D^2(X) = G''_X(1) + G'_X(1) - [G'_X(1)]^2$;

(d) Două variabile aleatoare cu aceeași funcție generatoare au aceeași repartiție.

12.4 Exerciții

1. Fie variabilele aleatoare:

$$X = \begin{pmatrix} -1 & 2 \\ 0.5 & 0.5 \end{pmatrix}, \quad Y = \begin{pmatrix} -2 & 0 & 1 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}.$$

(a) Calculați $E(X/Y = 1)$;

(b) Calculați $D^2(X)$ și $D^2(Y)$;

(c) Calculați XY și apoi $E(XY)$;

(d) Determinați coeficientul de corelația $\rho(X, Y)$.

2. Fie vectorul aleator (X, Y), cu densitatea de repartiție:

$$f(x, y) = \begin{cases} e^{-y}, & 0 < x < y \\ 0, & \text{în rest} \end{cases}.$$

Determinați:

(a) Densitățile marginale $f_X(x)$ și $f_Y(y)$;

(b) Densitatea variabilei aleatoare Y condiționată de X;

(c) Probabilitatea ca $(x, y) \in (X, Y)$ să aparțină pătratului unitate $[0, 1] \times [0, 1]$.

3. Fie vectorul aleator (X, Y), cu densitatea de repartiție:

$$f(x, y) = \begin{cases} k, & 0 < y \leq x < 1 \\ 0, & \text{în rest} \end{cases},$$

unde $k \in \mathbb{R}$ este o constantă fixată.

(a) Determinați k, astfel încît funcția de mai sus să fie corect definită, ca densitate de repartiție;

(b) Determinați densitățile de repartiție marginale $f_X(x), f_Y(y)$;

(c) Calculați $P(0 < X < 0.5, 0 < Y < 0.5)$;

(d) Determinați densitățile condiționate $f(y/x)$ și $f(x/y)$.

4. Fie funcția:

$$f(x, y) = \begin{cases} x^2 + \dfrac{xy}{c}, & x \in (0, 1), y \in (0, 2) \\ 0, & \text{în rest} \end{cases}.$$

(a) Determinați $c \in \mathbb{R}$ astfel încît $f(x, y)$ să fie o densitate de repartiție;

(b) Determinați funcția de repartiție $F_{(X,Y)}(x, y)$ a vectorului aleator bidimensional (X, Y) care are densitatea de repartiție calculată mai sus;

(c) Determinați densitățile marginale $f_X(x), f_Y(y)$ și funcțiile de repartiție $F_X(x), F_Y(y)$;

(d) Calculați $P(X > 0.5)$ și $P(Y < 0.5/X < 0.5)$;

(e) Calculați densitatea de repartiție condiționată $F(Y|X = x)$.

5. Vectorul aleator (X, Y) se dă prin tabloul de repartiție comună:

$X \mid Y$	$y_1 = -2$	$y_2 = -1$	$y_3 = 4$	$y_4 = 5$
$x_1 = 1$	$p_{11} = 0.1$	$p_{12} = 0.2$	$p_{13} = 0$	$p_{14} = 0.3$
$x_2 = 2$	$p_{21} = 0.2$	$p_{22} = 0.1$	$p_{23} = 0.1$	$p_{24} = 0$

(a) Determinați repartițiile marginale p_X și p_Y ale celor două variabile aleatoare X și Y;

(b) Aflați dacă variabilele aleatoare sînt independente;

(c) Calculați $E(X)$ și $E(Y)$;

(d) Calculați $\mathrm{cov}(X, Y)$ si determinați dacă variabilele aleatoare sînt corelate sau nu;

(e) Calculați $\rho(X, Y)$.

6. Vectorul aleator (X, Y) se dă prin tabloul de repartiție comună:

$X \mid Y$	$y_1 = 0$	$y_2 = 1$
$x_1 = -1$	$p_{11} = \lambda$	$p_{12} = 3/20$
$x_2 = 0$	$p_{21} = 3/20$	$p_{22} = 1/4$
$x_3 = 1$	$p_{31} = 1/4$	$p_{32} = 3/20$

.

(a) Determinați parametrul λ;

(b) Determinați $D^2(X)$ și $D^2(Y)$;

(c) Verificați dacă cele două variabile aleatoare sînt independente.

7. Vectorul aleator (X, Y) se dă prin tabloul de repartiție comună:

$X \mid Y$	$y_1 = 0$	$y_2 = 1$	$y_3 = 2$
$x_1 = 2$	$p_{11} = 0.2$	$p_{12} = 0.3$	$p_{13} = 0.1$
$x_2 = 3$	$p_{21} = 0.1$	$p_{22} = 0.1$	$p_{23} = \lambda$

.

(a) Determinați parametrul λ;

(b) Determinați repartițiile marginale ale vectorului aleator (X, Y);

(c) Calculați $E(Y/X = 2)$ și $E(X/Y = 2)$.

INDEX

MEDIA SI DISPERSIA UNEI VARIABILE ALEATOARE. PROBLEME

Lect. Dr. Costache Luminita

1. Variabila aleatoare X are repartiţia $X \sim \begin{pmatrix} -1 & 0 & 1 \\ 0,2 & 0,3 & 0,5 \end{pmatrix}$. Să se calculeze valorile medii $E(X), E(2X), E(X+1), E(2X+1), E(X^2)$, $E((X-0,3)^2)$.

2. Doi jucători de tenis joacă în 4 meciuri 12 seturi. Considerând că probabilităţile celor 2 jucători de a câştiga un set sunt egale, să se calculeze valoarea medie, dispersia şi abaterea medie pătratică a variabilei aleatoare ce reprezintă numărul de seturi căştigate de unul dintre jucători.

3. Doi parteneri cu forţă egală boxează 10 runde (probabilitatea ca oricare din ei să câştige o rundă este $\frac{1}{2}$). Să se calculeze valoarea medie, dispersia şi abaterea medie pătratică a v. a. care reprezintă numărul de runde câştigate de unul din parteneri.

4. O firmă se aprovizionează de la 3 furnizori. Din datele statistice privind furnizorii, firma estimează că probabilitatea cu care furnizorii nu pot onora contractul sunt $p_1 = 0,1, p_2 = 0,3, p_3 = 0,2$. Fie X variabila aleatoare ce indică numărul furnizorilor ce nu-şi pot onora contractul . Să se afle:

   a) repartiţia v.a. X;

   b) $E(X), D^2(X)$.

5. Se fac experimente asupra alegerii filamentului unui girofar până când acesta este aprins. La fiecare experiment probabilitatea de succes este $\frac{1}{5}$. Se cer media şi dispersia numărului de experimente.

6. Persoanele A şi B joacă în următoarele condiţii : se aruncă 2 zaruri şi când suma punctelor obţinute e mai mică decât 10, atunci B plăteşte lui A suma de 5 dolari. In caz contrar, A plăteşte lui B o sumă fixă a. Să se determine această sumă astfel încât jocul să fie echitabil.

7. Se dau densităţile de probabilitate

$$a) f_X(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, \text{în rest} & \end{cases}$$

$$b) f_X(x) = \begin{cases} |x|, & |x| < 1 \\ 0, & \text{în rest} \end{cases}$$

   Să se calculeze valoarea medie şi dispersia variabilei X.

8. Se dau funcţiile de repartiţie

$$a) F_X(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

$$b) F_X(x) = \begin{cases} 0, & x < 0 \\ x^{\frac{1}{2}}, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

Să se calculeze valorile medii şi dispersiile corespunzătoare.

9. Fie X o variabilă aleatoare a cărei densitate de repartiţie este

$$f_X(x) = \begin{cases} c \ln\left(\frac{a}{x}\right), & 0 \leq x < a \\ 0, & \text{în rest} \end{cases}$$

Să se determine constanta c şi să se calculeze valoarea medie,momentul de ordinul doi şi dispersia variabilei X.

10. Se dă densitatea de repartiţie

$$f(x) = \begin{cases} 1 - |1 - x|, & 0 < x < 2 \\ 0, & \text{în rest} \end{cases}$$

Să se calculeze media şi dispersia.

11. Variabila a;eatoare X are densitatea de probabilitate

$$f(x) = \begin{cases} cx(1 - x), & \text{dacă } 0 \leq x \leq 1 \\ 0, & \text{în rest} \end{cases}$$

Să se determine: a) constanta c astfel încât f să fie densitate de probabilitate; b) $P(X \leq \frac{1}{2})$; c) funcţia de repartiţie $F_X(x)$ a variabilei aleatoare X; d) $D^2(X)$.

12. Cunoscând funcţia de repartiţie a depărtării stelelor de la steaua cea mai apropiată este $F_R(r) = 1 - \mathrm{e}^{-\frac{4}{3}\pi\lambda r^3}$ şi numărul mediu de stele în cubul parsec din vecinătatea soarelui $\lambda = 0,00163$, să se afle valoarea medie şi dispersia variabilelor R.

13. Fiind dată densitatea de probabilitate a vectorului aleator (X, Y)

$$f(x, y) = \begin{cases} 2xy\mathrm{e}^{-(x^2+y^2)}, & x > 0, y > 0 \\ 0, & \text{în rest} \end{cases}$$

să se determine $E(X), E(Y), D^2(X), D^2(Y)$.

14. Fie (X, Y) un vector aleator cu densitatea

$$f(x,y) = \begin{cases} \frac{1}{8}(6-x-y), & (x,y) \in (0,2) \times (2,4) \\ 0, & \text{în rest} \end{cases}$$

Să se afle:

a) densitățile de repartiție condiționate $f(x/y), f(y/x)$;

b) $E(X/Y = y), E(Y/X = x)$

SIRURI DE VARIABILE ALEATOARE. TEOREME LIMITA
Lect. Dr. Costache Luminita

1 Inegalitatea lui Cebâşev

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

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

2 Tipuri de convergenţă

Definiţia 1. Fie $(X_n)_{n\in\mathbb{N}^*}$ un şir de v. a. reale sau complexe definite pe un spaţiu de probabilitate $(\Omega, \mathcal{K}, P)$, iar X o altă v.a. definită pe acelaşi spaţiu de probabilitate.

1) Sirul $(X_n)_{n\in\mathbb{N}^*}$ **converge în probabilitate către** X ($X_n \xrightarrow{P} X$, când $n \longrightarrow \infty$), dacă $\forall \varepsilon > 0, P(|X_n - X| > \varepsilon) \longrightarrow 0, n \longrightarrow \infty$ sau $\forall \varepsilon > 0, P(|X_n - X| < \varepsilon) \longrightarrow 1, n \longrightarrow \infty$.

2) Sirul $(X_n)_{n\in\mathbb{N}^*}$ **converge aproape sigur (tare) către** X ($X_n \xrightarrow{a.s.} X$, când $n \longrightarrow \infty$), dacă $P(\{\omega / \lim\limits_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$.

3) Sirul $(X_n)_{n\in\mathbb{N}^*}$ **converge în medie de ordinul** r **către** X ($X_n \xrightarrow{r} X$, când $n \longrightarrow \infty$), dacă există momentele absolute $E(|X_n|^r), n \in \mathbb{N}^*$ şi $E(|X|)$ şi dacă $E(|X_n - X|^r) \longrightarrow 0$, când $n \longrightarrow \infty$.

4) Sirul $(X_n)_{n\in\mathbb{N}^*}$ **converge în repartiţie sau slab către** X ($X_n \xrightarrow{w} X$, când $n \longrightarrow \infty$), dacă $F_{X_n}(x) \longrightarrow F_X(x), n \longrightarrow \infty, \forall x$ punct de continuitate al funcţiei F_X, unde $F_{X_n}, n \in \mathbb{N}^*$ şi F_X sunt funcţii de repartiţie ale v. a. $X_n, n \in \mathbb{N}^*$ şi respectiv X.

Relaţii între tipurile de convergenţă

1) Dacă $X_n \xrightarrow{a.s.} X, n \longrightarrow \infty$, atunci $X_n \xrightarrow{P} X, n \longrightarrow \infty$.

2) Dacă $X_n \xrightarrow{P} X, n \longrightarrow \infty$, atunci există un subşir $(X_{n_k})_{k\in\mathbb{N}^*}$ astfel încât $X_{n_k} \xrightarrow{a.s.} X, k \longrightarrow \infty$.

3) Dacă $X_n \xrightarrow{r} X, n \longrightarrow \infty$, atunci $X_n \xrightarrow{P} X, n \longrightarrow \infty$. Implicaţia reciprocă nu are loc, deoarece $E(|X_n - X|^r)$ s-ar putea să nu existe.

4) Dacă $X_n \xrightarrow{r} X, n \longrightarrow \infty$, atunci $X_n \xrightarrow{r'} X, n \longrightarrow \infty$, pentru $r' < r$.

5) Dacă $X_n \xrightarrow{P} X, n \longrightarrow \infty$, atunci $X_n \xrightarrow{w} X, n \longrightarrow \infty$.

3 Legea numerelor mari

Am văzut că nu putem şti înainte de efectuarea experienţei ce valoare va lua variabila aleatoare pe care o studiem. S-ar părea că , întrucât despre fiecare variabilă aleatoare dispunem de informaţii reduse, cu greu am putea determina comportarea mediei aritmetice a unui număr suficient de mare de variabile aleatoare. In realitate, în condiţii puţin restrictive media aritmetică a unui număr suficient de mare de variabile aleatoare îşi pierde caracterul întâmplător. Pentru practică este foarte important să cunoaştem condiţiile în care acţiunea combinată a mai multţi factori întâmplători conduce la un rezultat care să nu depindă de întâmplare, deci care să ne permită să prevedem mersul fenomenului studiat. Astfel de condiţii se dau în teoremele cunoscute în calculul probabilităţilor sub denumirea comună de **legea numerelor mari**. Termenul de lege a numerelor mari a fost folosit pentru prima oară de Poisson, deşi, cu aproximativ un secol înainte, Jacob Bernoulli a pus în evidenţă acţiunea legii numerelor mari cu referire la repartiţia binomială . In 1867, Cebâşev precizează riguros din punct de vedere matematic legea numerelor mari în condiţii generale.

Fie $(\Omega, \mathcal{K}, P)$ un spaţiu de probabilitate şi $(X_n)_{n\in\mathbb{N}^*}$ un şir de v. a. reale definite pe acest spaţiu.

Ne interesează cazul în care există un şir de numere reale $(a_n)_{n\in\mathbb{N}^*}$ astfel încât:

1) $\frac{1}{n}\sum\limits_{k=1}^{n} X_k - a_n \xrightarrow{P} 0, n \longrightarrow \infty$

sau

2) $\frac{1}{n}\sum\limits_{k=1}^{n} X_k - a_n \xrightarrow{a.s.} 0, n \longrightarrow \infty$

De obicei se consideră cazul în care $a_n = \frac{1}{n}\sum\limits_{k=1}^{n} E(X_k)$.

In cazul 1) (resp. 2)) se spune că şirul $(X_n)_{n\in\mathbb{N}^*}$ satisface **legea slabă a numerelor mari** (resp. **legea tare a numerelor mari**) sau că $(X_n)_{n\in\mathbb{N}^*}$ este **slab stabil** (resp. **tare stabil**).

Teorema 2. ***(Cebâşev)*** *Fie $(X_n)_{n\in\mathbb{N}^*}$ un şir de v. a. independente două câte două ale căror dispersii verifică $D^2(X_i) < c, \forall i \in \mathbb{N}^*$, $c \in \mathbb{R}$. Atunci*

$$\lim_{n\to\infty} P(|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)| < \varepsilon) = 1$$

sau echivalent

$$\lim_{n\to\infty} P(|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)| \geq \varepsilon) = 0$$

Deşi v.a. independente pot lua valori depărtate de mediile lor, media aritmetică a unui număr suficient de mare de astfel de v.a. ia, cu o probabilitate mare, valori apropiate de un număr constant $\frac{1}{n}\sum\limits_{i=1}^{n}E(X_i)$.

In acest fel, între comportarea fiecărei v.a. şi a mediei lor aritmetice există o mare deosebire: nu este posibil să prevedem ce valoare va lua fiecare dintre v.a., însă cu o probabilitate mare putem prevedea ce valoare va lua media aritmetică a acestor v.a. De aici tragem concluzia că media aritmetică a unui număr suficient de mare de v.a. (cu dispersii mărginite) îşi pierde caracterul de v.a.

Exemplul 1. 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 probabilitate

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

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

Demonstraţie. Teorema este aplicabilă dacă media este finită şi dispersia este mărginită .

$E(X_n) = (-5n)\cdot\frac{1}{3n^2} + 0\cdot(1-\frac{2}{3n^2}) + 5n\cdot\frac{1}{3n^2} = 0, E(X_n^2) = (-5n)^2\cdot$ $\cdot\frac{1}{3n^2} + 0^2\cdot(1-\frac{2}{3n^2}) + (5n)^2\cdot\frac{1}{3n^2} = \frac{50}{3}, D^2(X_n) = \frac{50}{3} \Longrightarrow$ teorema se aplică

$E(Y_n) = (-n^2)\cdot\alpha^{-n} + 0\cdot(1-2\alpha^{-n}) + n^2\cdot\alpha^{-n} = 0, E(Y_n^2) =$ $= (-n^2)^2\cdot\cdot\alpha^{-n} + 0^2\cdot(1-2\alpha^{-n}) + (n^2)^2\cdot\alpha^{-n} = \frac{2n^4}{\alpha^n}, D^2(Y_n) = \frac{2n^4}{\alpha^n} \Longrightarrow$ $\Longrightarrow \lim\limits_{n\to\infty} D^2(Y_n) = 0 \Longrightarrow \exists c \in (0,\infty)$ astfel încât $\forall n \geq 1, D^2(Y_n) \leq c \Longrightarrow$ teorema se aplică

$E(Z_n) = \int_0^\infty x\lambda^{-n}\mathrm{e}^{-\frac{x}{\lambda^n}}dx = \lambda^n, E(Z_n^2) = \int_0^\infty x^2\lambda^{-n}\mathrm{e}^{-\frac{x}{\lambda^n}}dx = 2\lambda^{2n} \Longrightarrow$ $\Longrightarrow D^2(Z_n) = \lambda^{2n} \Longrightarrow$

$$\lim_{n\to\infty} D^2(Z_n) = \begin{cases} 0, & \lambda \in (0,1) \\ 1, & \lambda = 1 \\ \infty, & \lambda > 1 \end{cases}$$

deci teorema se aplică numai dacă $\lambda \in (0,1]$ □

Teorema 3. ***(Teorema lui Bernoulli)*** *Fie ν_n numărul de realizări ale unui eveniment A în n experimente independente şi p probabilitatea de realizare a lui A în fiecare experiment. Atunci pentru*

$$\forall\varepsilon > 0, \lim_{n\to\infty} P(|\frac{\nu_n}{n} - p| < \varepsilon) = 1$$

sau echivalent

$$\lim_{n\to\infty} P(|\frac{\nu_n}{n} - p| \geq \varepsilon) = 0$$

Observaţia 1. Rezultatul poate fi formulat şi astfel: şirul de v.a. $(\frac{\nu_n}{n})_n$ converge în probabilitate către p.

In condiţiile teoremei lui Bernoulli, Borel a arătat în 1909 un rezultat mai profund : $\frac{\nu_n}{n} \xrightarrow{a.s.} p$.

Observaţia 2. In cazul unei populaţii de volum mare, dacă se efectuează o selecţie de volum n şi se obţin ν_n rezultate favorabile, atunci cu o probabilitate apropiată de unitate, putem afirma că probabilitatea evenimentului cercetat este dată de frecvenţa relativă . Prin urmare, în studiul populaţiilor pentru care nu putem determina apriori probabilitatea de realizare a unui eveniment, probabilitatea teoretică p se poate exprima pe cale experimentală prin frecvenţa relativă $\frac{\nu_n}{n}$ a evenimentului considerat, fapt ce constituie justificarea teoretică a folosirii frecvenţei în loc de probabilitate.

Corolarul 1. *Pentru* $\forall\varepsilon > 0$ *avem* $\lim_{n\to\infty} P(|f_n - p| < \varepsilon) \geq 1 - \dfrac{pq}{n\varepsilon^2}$, *unde* $f_n = \frac{\nu_n}{n}$, ν_n*=de câte ori s-a realizat evenimentul* A *în* n *probe independente.*

Exemplul 2. Pentru un studiu biologic sunt cercetate 1000 de probe independente, dacă au sau nu o anumită caracteristică biologică . Să se determine o margine inferioară a probabilităţii ca diferenţa, în valoare absolută , dintre frecvenţa relativă şi probabilitatea p de a apărea caracteristica într-o probă , să fie mai mică decât 0,03.

Demonstraţie. Cf. consecinţei teoremei lui Bernoulli $\Longrightarrow p = q = \frac{1}{2}, \varepsilon =$
$= 0,03 \Longrightarrow P(|\frac{x}{n} - p| < 0,03) \geq 1 - \frac{pq}{n\varepsilon^2} = 1 - \frac{\frac{1}{2}\cdot\frac{1}{2}}{1000\cdot 0,03^2} = 0,72$ □

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* $Z_n = \frac{Y_n - E(Y_n)}{D^2(Y_n)}$, *unde* $Y_n = \sum_{i=1}^{n} X_i$. *Atunci* $\forall x \in \mathbb{R}, \lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P(Z_n < x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} \mathrm{e}^{-\frac{u^2}{2}} du$, *adică şirul* $(Z_n)_n$ *converge în repartiţie către o v.a.* $Z \sim N(0,1)$.

Când v.a. au repartiţii diferite avem o altă variantă a celebrei teoreme limită centrală :

Teorema 5. ***(Teorema lui A. M. Leapunov)****. Fie* $(X_n)_n$ *un şir de v.a. independente. Presupunem că* $E(X_k) = m_k, D^2(X_k) = \sigma_k^2$,
$E(|X_k - m_k|^3) = \rho_k^3, 1 \leq k \leq n$ *există şi notăm* $Y_n = \sum_{k=1}^{n} X_k, \sigma^2(n) = \sum_{k=1}^{n} \sigma_k^2, \rho^3(n) = \sum_{k=1}^{n} \rho_k^3, Z_n = \dfrac{Y_n - E(Y_n)}{\sigma(n)}$ *şi* F_n *funcţia de repartiţie a v.a.*

Z_n. *Dacă* $\lim\limits_{n\to\infty}\dfrac{\rho(n)}{\sigma(n)}=0$, *atunci* $\lim\limits_{n\to\infty}F_n(x)=\lim\limits_{n\to\infty}P(Z_n<x)=$
$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{u^2}{2}}du$, *adică* $(Z_n)_n$ *converge în repartiţie către v. a.* $Z\sim N(0,1)$.

Exemplul 3. Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente astfel încât $P(X_n=\frac{1}{n^\beta})=P(X_n=-\frac{1}{n^\beta})=p$, cu $\frac{1}{3}<\beta\leq\frac{1}{2}, n\in\mathbb{N}^*$ şi $P(X_n=0)=1-2p$. Să se arate că şirului $(X_n)_{n\geq 1}$ i se poate aplica teorema lui Leapunov.

Demonstraţie. Obţinem $E(X_k)=\frac{1}{k^\beta}\cdot p+(-\frac{1}{k^\beta})\cdot p=0$

$E(X_k^2)=(\frac{1}{k^\beta})^2\cdot p+(-\frac{1}{k^\beta})^2\cdot p=\frac{2p}{k^{2\beta}}=\mu_k^2$

$E(|X_k|^3)=\frac{2p}{k^{3\beta}}=\rho_k^3$

Atunci $\sigma^2(n)=\displaystyle\sum_{k=1}^{n}\sigma_k^2=2p\sum_{k=1}^{n}\frac{1}{k^{2\beta}}, \rho^3(n)=\sum_{k=1}^{n}\rho_k^3=2p\sum_{k=1}^{n}\frac{1}{k^{3\beta}}$

Cum $\frac{1}{3}<\beta\leq\frac{1}{2}\Longrightarrow 1<3\beta\leq\frac{3}{2}\Longrightarrow \displaystyle\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{k^{3\beta}}=\sum_{k=1}^{\infty}\frac{1}{k^{3\beta}}$ este o serie convergentă (serie Riemann)

$\displaystyle\sum_{k=1}^{n}\frac{1}{k^{2\beta}}>\sum_{k=1}^{n}\frac{1}{k}$ ce tinde la ∞ când $n\longrightarrow\infty$

Condiţia lui Leapunov $\displaystyle\lim_{n\to\infty}\frac{\sqrt[3]{\rho(n)^3}}{\sqrt{\sigma(n)^2}}=0$ este îndeplinită , deci şirului $(X_n)_{n\geq 1}$ i se poate aplica teorema lui Leapunov $\square$

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*

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

sau

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

SEMINAR 10

ȘIRURI DE VARIABILE ALEATOARE

Pentru aspectele teoretice privitoare la aceste noțiuni, puteți consulta *Lecții de matematici speciale*, L. Costache.

În continuare, vom insista pe exerciții.

10.1 Exerciții rezolvate

1. Fie variabila aleatoare:

$$X = \begin{pmatrix} 0.2 & 0.3 & 0.4 & 0.5 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{pmatrix}.$$

Estimați probabilitatea $P(\{|X - m| < 0.2\})$, unde $m = M[X] = E[X]$ este media lui X.

Soluție: Se folosește inegalitatea lui Cebîșev, sub forma:

$$P(\{|X - m| < \varepsilon\}) \geq 1 - \frac{\sigma^2}{\varepsilon^2},$$

unde σ este dispersia, iar $\varepsilon = 0.2$ în acest caz (alegem).

Avem succesiv: $M[X] = 0.4$, deci $(M[X])^2 = 0.16$, iar variabila la pătrat este:

$$X^2 = \begin{pmatrix} 0.04 & 0.09 & 0.16 & 0.25 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{pmatrix}.$$

Rezultă $M[X^2] = 0.17$, de unde obținem $\sigma^2 = D^2[X] = M[X^2] - (M[X])^2 = 0.01$.

Din egalitatea lui Cebîșev, rezultă în fine:

$$P(\{|X - m| < 0.02\}) \geq 0.9975.$$

2. Se aruncă 2 monede, simultan, de 800 de ori. Determinați probabilitatea de apariție a stemei pe ambele monede de un număr de ori cuprins între 150 și 250.

Soluție: Fie evenimentul A, care arată că apare stema pe ambele monede. Pentru o singură aruncare, avem $p = P(A) = \frac{1}{4}$.

Fie acum o variabilă aleatoare cu repartiție binomială (potrivită pentru studiul aruncării monedelor), deci de forma:

$$X = \begin{pmatrix} k \\ C_n^k p^k q^{n-k} \end{pmatrix},$$

unde $n = 800, 1 \leq k \leq 800$ și $p = \frac{1}{4}$.

Media variabilei (folosim direct formulele cunoscute din seminariile anterioare) este $M[X] = np = 800 \cdot \frac{1}{4} = 200$, iar dispersia se calculează cu:

$$\sigma^2 = np(1-p) = 150.$$

Acum, din egalitatea lui Cebîșev:

$$P\left(\{|X-m| < \varepsilon\}\right) \geq 1 - \frac{\sigma^2}{\varepsilon^2},$$

unde $m = M[X]$, trebuie să determinăm ε convenabil.

Evenimentul căutat $|X - m| < \varepsilon$ poate fi scris în mod echivalent:

$$m - \varepsilon < X < \varepsilon + m \Longrightarrow 200 - \varepsilon < X < \varepsilon + 200.$$

Dar în ipoteză, avem cerința $150 < X < 250$, deci se ajunge la sistemul:

$$\begin{cases} 200 - \varepsilon & = 150 \\ 200 + \varepsilon & = 250 \end{cases},$$

sistem care are soluția $\varepsilon = 50$. Acum revenim în inegalitatea lui Cebîșev și obținem:

$$P(150 < X < 250) \geq 1 - \frac{150}{50^2} \Longrightarrow P(150 < X < 250) \geq 0.94.$$

3. Un zar se aruncă de 1200 de ori. Să se determine probabilitatea minimă ca numărul de apariții ale feței cu i puncte, pentru un i fixat, să fie între 150 și 250.

Soluție: Fie evenimentul A care arată că apare fața cu i puncte, pentru un i fixat, la o anume aruncare. Considerăm $n = 1200$ variabile aleatoare identic repartizate, cu forma Bernoulli (potrivită pentru aruncarea zarului), deci:

$$X_k = \begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix},$$

unde $p + q = 1$ și $P(X_k = 1) = P(A) = p = \dfrac{1}{6}$.

Media unei asemenea variabile (din formula corespunzătoare cazului Bernoulli) este $M[X_k] = p$, pentru orice k.

Vom folosi legea slabă a numerelor mari, în forma generală, unde frecvența relativă de apariție a evenimentului A este $f_{1200} = \dfrac{k}{1200}$. Rezultă:

$$P\left(\left|\frac{k}{1200} - \frac{1}{6}\right| < \varepsilon\right) > 1 - \delta, \qquad \textbf{Cor. 1, p. 102}$$

unde avem $n = \dfrac{1}{4\delta\varepsilon^2}$, care conduce la $\delta = \dfrac{1}{4 \cdot 1200 \cdot \varepsilon^2}$.

Pentru a determina valoarea lui ε, calculăm:

$$\left|\frac{k}{1200} - \frac{1}{6}\right| < \varepsilon \Longleftrightarrow \frac{1}{6} - \varepsilon < \frac{k}{1200} < \frac{1}{6} + \varepsilon,$$

de unde, în fine: $200 - 1200\varepsilon < k < 200 + 1200\varepsilon$. Deoarece cerința problemei este $150 < k < 250$, se ajunge la sistemul:

$$\begin{cases} 200 - 1200\varepsilon & = 150 \\ 200 + 1200\varepsilon & = 250 \end{cases} \Longrightarrow \varepsilon = \frac{1}{24}.$$

Pentru această valoare, avem $\delta = 0.12$, deci $1 - \delta = 0.88$. Concluzia este că probabilitatea cerută este de 80%.

4. Folosind teorema Moivre-Laplace, estimați probabilitatea ca din 100 de aruncări ale unei monede, să apară banul de un număr de ori cuprins între 40 și 60.

Soluție: Fie $n = 100$ variabile aleatoare repartizate identic, binomial (model potrivit pentru aruncarea monedelor). Putem considera simplu:

$$X_k = \begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix},$$

unde $P(X_k = 1) = p = 0.5$. Media fiecăreia dintre aceste variabile este $M[X_k] = p$. Avem de calculat probabilitatea:

$$P\left(a < \sum_{k=1}^{n} X_k < b\right), \qquad a = 40, b = 60.$$

Aplicăm teorema Moivre-Laplace:

$$\ell = \lim_{n\to\infty} P\left(a < \sum_{k=1}^{n} X_k < b\right) = \phi\left(\frac{b - np}{\sqrt{npq}}\right) - \phi\left(\frac{a - np}{\sqrt{npq}}\right). \qquad \textit{(vezi p. 78)}$$

Rezultă $\ell = \phi(2) - \phi(-2) = \phi(2) - (1 - \phi(2))$. Acum, cu valorile din tabel pentru funcția lui Laplace, avem rezultatul cerut ca fiind 0.9544.

5. Fie șirul de variabile aleatoare, independente $(X_n)_{n\in\mathbb{N}}$, cu repartiția:

$$X_n = \begin{pmatrix} -n & 0 & n \\ \frac{1}{2n^2} & 1 - \frac{1}{n^2} & \frac{1}{2n^2} \end{pmatrix}.$$

Arătați că acest șir de variabile aleatoare îndeplinește condițiile legii numerelor mari în forma Cebîșev.

Soluție: Avem de arătat că dispersiile șirului de variabile aleatoare sînt finite și uniform mărginite, adică există $c \in \mathbb{R}$ constantă, astfel încît $\sigma^2(X_k) \le c$, pentru toți k.

Dar calcule imediate arată:

$$\begin{aligned} M[X_n] &= -\frac{n}{2n^2} + \frac{n}{2n^2} = 0 \\ M[X_n^2] &= \frac{n^2}{2n^2} + \frac{n^2}{2n^2} = 1 \\ \sigma^2 &= M[X_n^2] - M^2[X_n] = 1 - 0 = 1, \end{aligned}$$

care adeveresc condiția.

10.2 Exerciții propuse

1. Timpul mediu de răspuns al unui calculator este de 15 secunde pentru o anumită operație, cu abaterea standard de 3 secunde.

Estimați probabilitatea ca timpul de răspuns să se abată cu mai puțin de 5 secunde față de medie.

Indicație: Folosiți inegalitatea Cebîșev. (seminarTP, ex. 22, p. 85)

2.* Arătați că șirul de variabile aleatore $(X_n)_n$, independente, cu repartiția:

$$X_n = \begin{pmatrix} -\sqrt{\ln n} & \sqrt{\ln n} \\ 0.5 & 0.5 \end{pmatrix}$$ (Curs & Seminar RP)

urmează legea slabă a numerelor mari.

Indicație: Verificați mai întîi dacă variabilele din șir au dispersii finite, iar în caz negativ, verificați formularea Markov.

3. Într-o școală se află 100 de unități electrice care funcționează independent. Probabilitatea deconectării uneia dintre ele într-un interval de timp dat este de 5%. Determinați probabilitatea ca în acest interval anume să fie deconectate:

(a) nu mai puțin de 5 unități;

(b) nu mai mult de 5 unități;

(c) între 5 și 10 unități.

Indicație: Teorema Moivre-Laplace, pentru variabile repartizate Bernoulli.

4. Se știe că 5% dintre produsele unui anumit lot sînt defecte. Care este probabilitatea ca dintr-o selecție de 200 de produse, cel puțin 10% să fie defecte?
Indicație: Teorema Moivre-Laplace, pentru variabile repartizate Bernoulli.

5. O fabrică folosește 2 tipuri de aparate, A și B. După un timp, tipul A se strică și trebuie reparate cu probabilitatea de 20%, iar tipul B, cu probabilitatea de 40%.
Numărul total de aparate din tip A disponibile în fabrică este de 1000, iar cele de tip B sînt în număr de 1500.

(a) Care este probabilitatea ca la un moment dat, numărul de aparate care trebuie reparate (de oricare dintre tipuri) să fie între 750 și 850?

(b) Determinați, cu o probabilitate de 0.99, numărul maxim de aparate care ar trebui reparate la un moment dat.

Indicație: Teorema Moivre-Laplace, pentru variabile repartizate Bernoulli (considerați două șiruri, corespunzătoare celor două tipuri de aparate).

INDEX

MATEMATICI SPECIALE
Culegere de probleme

TANIA-LUMINIŢA COSTACHE

Capitolul 4

Şiruri de variabile aleatoare

4.1 Noţiuni teoretice

Tipuri de convergenţă

Definiţia 4.1. Fie $(X_n)_{n\in\mathbb{N}^*}$ un şir de v. a. reale sau complexe definite pe un spaţiu de probabilitate $(\Omega, \mathcal{K}, P)$, iar X o altă v.a. definită pe acelaşi spaţiu de probabilitate.

1) Sirul $(X_n)_{n\in\mathbb{N}^*}$ **converge în probabilitate către** X ($X_n \xrightarrow{P} X$, când $n \longrightarrow \infty$), dacă $\forall\varepsilon > 0, P(|X_n - X| > \varepsilon) \longrightarrow 0, n \longrightarrow \infty$ sau $\forall\varepsilon > 0, P(|X_n - X| < \varepsilon) \longrightarrow 1, n \longrightarrow \infty$.

2) Sirul $(X_n)_{n\in\mathbb{N}^*}$ **converge aproape sigur către** X ($X_n \xrightarrow{a.s.} X$, când $n \longrightarrow \infty$), dacă $P(\{\omega / \lim\limits_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$.

3) Sirul $(X_n)_{n\in\mathbb{N}^*}$ **converge în medie de ordinul** r **către** X ($X_n \xrightarrow{r} X$, când $n \longrightarrow \infty$), dacă există momentele absolute $E(|X_n|^r), n \in \mathbb{N}^*$ şi $E(|X|)$ şi dacă $E(|X_n - X|^r) \longrightarrow 0$, când $n \longrightarrow \infty$.

4) Sirul $(X_n)_{n\in\mathbb{N}^*}$ **converge în repartiţie sau slab către** X ($X_n \xrightarrow{w} X$, când $n \longrightarrow \infty$), dacă $F_{X_n}(x) \longrightarrow F_X(x), n \longrightarrow \infty, \forall x \in C(F_X)$, unde $F_{X_n}, n \in \mathbb{N}^*$ şi F_X sunt funcţii de repartiţie ale v. a. $X_n, n \in \mathbb{N}^*$ şi respectiv X, iar $C(F_X) = \{x / F_X(x)\text{este continuă în} x\}$ reprezintă mulţimea de continuitate a lui F_X.

Relaţii între tipurile de convergenţă

1)Dacă $X_n \xrightarrow{a.s.} X, n \longrightarrow \infty$, atunci $X_n \xrightarrow{P} X, n \longrightarrow \infty$.

2) Dacă $X_n \xrightarrow{P} X, n \longrightarrow \infty$, atunci există un subşir $(X_{n_k})_{k\in\mathbb{N}^*}$ astfel încât $X_{n_k} \xrightarrow{a.s.} X, k \longrightarrow \infty$.

3) Dacă $X_n \xrightarrow{r} X, n \longrightarrow \infty$, atunci $X_n \xrightarrow{P} X, n \longrightarrow \infty$. Implicaţia reciprocă nu are loc, deoarece $E(|X_n - X|^r)$ s-ar putea să nu existe.

4) Dacă $X_n \xrightarrow{r} X, n \longrightarrow \infty$, atunci $X_n \xrightarrow{r'} X, n \longrightarrow \infty$, pentru $r' < r$.

5) Dacă $X_n \xrightarrow{P} X, n \longrightarrow \infty$, atunci $X_n \xrightarrow{w} X, n \longrightarrow \infty$.

Teorema 4.1. *Fie $X, X_1, X_2, \ldots$ v. a. discrete cu valori întregi nenegative.*

Dacă $G_{X_n}(t) \longrightarrow G_X(t), n \longrightarrow \infty$, *atunci* $X_n \xrightarrow{w} X, n \longrightarrow \infty$.

Teorema 4.2. ***(Helly)*** *Dacă* $X_n \xrightarrow{w} X, n \longrightarrow \infty$, *atunci şirul* $(\varphi_{X_n})_{n\geq 0}$ *converge uniform în orice interval mărginit către* φ_X. *Reciproc, dacă şirul* $(\varphi_{X_n})_{n\geq 0}$ *converge punctual pe* $\mathbb{R}$ *către o funcţie* φ *continuă în origine, atunci există o v. a.* X *cu funcţia caracteristică* $\varphi_X = \varphi$, *astfel încât* $X_n \xrightarrow{w} X, n \longrightarrow \infty$.

Legea numerelor mari

Am văzut că nu putem şti înainte de efectuarea experienţei ce valoare va lua variabila aleatoare pe care o studiem. S-ar părea că , întrucât despre fiecare variabilă aleatoare dispunem de informaţii reduse, cu greu am putea determina comportarea mediei aritmetice a unui număr suficient de mare de variabile aleatoare. In realitate, în condiţii puţin restrictive media aritmetică a unui număr suficient de mare de variabile aleatoare îşi pierde caracterul întâmplător. Pentru practică este foarte important să cunoaştem condiţiile în care acţiunea combinată a mai mulţi factori întâmplători conduce la un rezultat care să nu depindă de întâmplare, deci care să ne permită să prevedem mersul fenomenului studiat. Astfel de condiţii se dau în teoremele cunoscute în calculul probabilităţilor sub denumirea comună de **legea numerelor mari**. Termenul de lege a numerelor mari a fost folosit pentru prima oară de Poisson, deşi, cu aproximativ un secol înainte, Jacob Bernoulli a pus în evidenţă acţiunea legii numerelor mari cu referire la repartiţia binomială . In 1867, Cebîşev precizează riguros din punct de vedere matematic legea numerelor mari în condiţii generale.

Fie $(\Omega, \mathcal{K}, P)$ un spaţiu de probabilitate şi $(X_n)_{n\in\mathbb{N}^*}$ un şir de v. a. reale definite pe acest spaţiu.

Ne interesează cazul în care există un şir de numere reale $(a_n)_{n\in\mathbb{N}^*}$ astfel încât:

1) $\frac{1}{n}\sum\limits_{k=1}^{n} X_k - a_n \xrightarrow{P} 0, n \longrightarrow \infty$

sau

2) $\frac{1}{n}\sum\limits_{k=1}^{n} X_k - a_n \xrightarrow{a.s.} 0, n \longrightarrow \infty$

De obicei se consideră cazul în care $a_n = \frac{1}{n}\sum\limits_{k=1}^{n} E(X_k)$.

In cazul 1) (resp. 2)) se spune că şirul $(X_n)_{n\in\mathbb{N}^*}$ satisface **legea slabă a numerelor mari** (resp. **legea tare a numerelor mari**) sau că $(X_n)_{n\in\mathbb{N}^*}$ este **slab stabil** (resp. **tare stabil**).

Teorema 4.3. ***(Hincin)*** *Fie* $(X_n)_{n\in\mathbb{N}^*}$ *un şir de v. a. independente, identic repartizate, având valoarea medie* m $(m < \infty)$. *Atunci şirul* $(X_n)_{n\in\mathbb{N}^*}$ *verifică legea slabă a numerelor mari.*

Teorema 4.4. ***(Teorema lui Bernoulli)*** *Să presupunem că se fac n experienţe independente, în fiecare experienţă probabilitatea evenimentului A fiind p, şi fie ν numărul de realizări ale evenimentului A, în cele n experienţe. Dacă ε este un număr pozitiv arbitrar suficient de mic, atunci* $\lim\limits_{n\to\infty} P(|\frac{\nu}{n} - p| < \varepsilon) = 1$.

Observaţia 4.1. In cazul unei populaţii de volum mare, dacă se efectuează o selecţie de volum n şi se obţin ν rezultate favorabile, atunci cu o probabilitate apropiată de unitate, putem afirma că probabilitatea evenimentului cercetat este dată de frecvenţa relativă . Prin urmare, dacă în studiul populaţiilor pentru care nu putem determina apriori probabilitatea de realizare a unui eveniment, probabilitatea teoretică p se poate exprima pe cale experimentală prin frecvenţa relativă $\frac{\nu}{n}$ a evenimentului considerat, fapt ce constituie justificarea teoretică a folosirii frecvenţei în loc de probabilitate.

Corolarul 4.1. *Pentru* $\forall\varepsilon > 0$ *avem* $\lim\limits_{n\to\infty} P(|f_n - p| < \varepsilon) \geq 1 - \dfrac{pq}{n\varepsilon^2}$, *unde* $f_n = \frac{\nu}{n}$,*ν=de câte ori s-a realizat evenimentul A în n probe independente.*

Teorema 4.5. ***(Teorema lui Poisson)*** *Fie şirul de evenimente $A_1, A_2, \ldots\ldots, A_n, \ldots$ ale căror probabilităţi de verificare au valorile succesive $p_1, p_2, \ldots, p_n, \ldots$. Dacă notăm cu f_n frecvenţa relativă a numărului care indică de câte ori s-au realizat evenimentele $A_1, A_2, \ldots\ldots, A_n, \ldots$ şi cu p expresia* $p = \lim\limits_{n\to\infty} \dfrac{p_1 + p_2 + \ldots + p_n}{n}$, *atunci*

$$\lim_{n\to\infty} P(|f_n - \frac{p_1 + p_2 + \ldots + p_n}{n}| < \varepsilon) = 1$$

Teorema 4.6. ***(Cebîşev)*** *Fie $(X_n)_{n\in\mathbb{N}^*}$ un şir de v. a. independente astfel încât $E(X_i) = m_i, D^2(X_i) = \sigma_i^2, \forall i \in \mathbb{N}^*$. Dacă există o constantă $M < \infty$ astfel încât $\sigma_i^2 \leq M, \forall i \in \mathbb{N}^*$, atunci şirul $(X_n)_{n\in\mathbb{N}^*}$ verifică legea slabă a numerelor mari.*

Problema limită centrală

Teorema 4.7. ***(Teorema Moivre-Laplace)*** *Se consideră v. a. bernoulliene $X_1, X_2, \ldots, X_n, n \in \mathbb{N}^*$, adică*

$$X_i = \begin{cases} 1, & \textit{cu probabilitatea } p \\ 0, & \textit{cu probabilitatea } q = 1 - p \end{cases}$$

$\forall 1 \leq i \leq n$ şi v. a. $S_n = X_1 + X_2 + \ldots + X_n, n \in \mathbb{N}^*$. Atunci şirul de v. a. $(Y_n)_{n\in\mathbb{N}^*}, Y_n = \frac{S_n - np}{\sqrt{npq}}, n \in \mathbb{N}^*$ converge în repartiţie către o v. a. repartizată normal $N(0,1)$.

Teorema 4.8. ***(Teorema lui A. M. Leapunov)*** *Fie $X_1, X_2, \ldots, X_n$ v. a. independente. Să notăm $m_k = E(X_k), \mu_k^2 = D^2(X_k)$,*

$\rho_k^3 = E(|X_k - m_k|^3), 1 \leq k \leq n, \mu^2(n) = \sum_{k=1}^{n} \mu_k^2, \rho^3(n) = \sum_{k=1}^{n} \rho_k^3$. *Dacă* $\lim_{n\to\infty} \frac{\rho(n)}{\mu(n)} = 0$, *atunci* $(Y_n)_n$ *converge în repartiţie către v. a.* Y, *unde*
$Y_n = \frac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} m_k}{\mu(n)}$ *şi* $Y \sim N(0,1)$.

4.2 Probleme rezolvate

1. Fie X o v. a. a cărei densitate de repartiţie este

$$f(x) = \begin{cases} \frac{x^m}{m!} \cdot \mathrm{e}^{-x}, & x > 0 \\ 0, & \text{în rest} \end{cases}$$

Să se arate că $P(0 < X < 2(m+1)) > \frac{m}{m+1}$.

Soluţie. Folosim inegalitatea lui Cebîşev:
$P(|X - E(X)| < \varepsilon) > 1 - \frac{D^2(X)}{\varepsilon^2}$
Avem $E(X) = \int_0^\infty x \cdot f(x)dx = \frac{1}{m!} \int_0^\infty x^{m+1} \cdot \mathrm{e}^{-x} dx = \frac{\Gamma(m+2)}{m!} =$
$= \frac{(m+1)!}{m!} = m + 1$
$E(X^2) = \int_0^\infty x^2 \cdot f(x)dx = \frac{1}{m!} \int_0^\infty x^{m+2} \cdot \mathrm{e}^{-x} dx = \frac{\Gamma(m+3)}{m!} =$
$= \frac{(m+2)!}{m!} = (m+1)(m+2)$
$D^2(X) = E(X^2) - [E(X)]^2 = m + 1$
Luăm $\varepsilon = m + 1 \Longrightarrow P(|X - (m+1)| < m + 1) > 1 - \frac{m+1}{(m+1)^2} =$
$= \frac{m}{m+1} \Longrightarrow P(0 < X < 2(m+1)) > \frac{m}{m+1}$ □

2. Se aruncă o monedă de n ori. Cât de mare trebuie să fie n pentru ca $P\left(\left|\frac{\alpha}{n} - \frac{1}{2}\right| < \frac{1}{100}\right) > 0,99$, ştiind că α reprezintă numărul de apariţii ale unei feţe alese de mai înainte.

Soluţie. Se ştie că $D^2(X) = E((X - E(X))^2) = E_2(|X - E(X)|)$
Folosim inegalitatea lui Cebîşev şi obţinem
$P\left(\left|\frac{\alpha}{n} - \frac{1}{2}\right| < \frac{1}{100}\right) > 1 - \frac{E_2\left(\left|\frac{\alpha}{n} - \frac{1}{2}\right|\right)}{10^{-4}}$
$E\left(\left|\frac{\alpha}{n} - \frac{1}{2}\right|\right) = E\left(\left|\frac{\alpha}{n} - \frac{1}{2}\right|^2\right) = E\left(\frac{\alpha^2}{n^2} - \frac{\alpha}{n} + \frac{1}{4}\right) = \frac{E(\alpha^2)}{n^2} - \frac{E(\alpha)}{n} + \frac{1}{4} =$
$= \frac{n^2p^2}{n^2} + \frac{npq}{n^2} - \frac{np}{n} + \frac{1}{4} = \frac{1}{4n}$, deoarece $p = q = \frac{1}{2}$
Deci $P\left(\left|\frac{\alpha}{n} - \frac{1}{2}\right| < \frac{1}{100}\right) > 1 - \frac{10^4}{4n}$
Aflăm n din inegalitatea $1 - \frac{10^4}{4n} > 0,99 \Longrightarrow n > 5^2 \cdot 10^4$ □

3. O variabilă aleatoare X are $E(X) = 80, E_2(X) = 6416$. Să se determine o limită inferioară a probabilităţii $P(40 < X < 120)$.

   Soluţie. $40 < X < 120 \Longleftrightarrow 80 - 40 < X < 80 + 40 \Longleftrightarrow -40 < X - 80 < 40 \Longleftrightarrow |X - 80| < 40 \Longrightarrow P(40 < X < 120) = P(|X - 80| < 40)$

   Folosim inegalitatea lui Cebîşev şi luăm $\varepsilon = 40$

   Calculăm $D^2(X) = E_2(X) = (E(X))^2 = 16$

   Avem $P(|X - 80| < 40) > 1 - \frac{16}{1600} = 0,99$ □

4. Limita superioară a probabilităţii ca abaterile, în modul, ale valorilor v. a. X faţă de medie, să fie mai mare decât 3, este 0,96. Să se afle $D^2(X)$.

   Soluţie. Folosim inegalitatea lui Cebîşev :

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

   $P(|X - E(X)| \geq 3) < \frac{D^2(X)}{9} = 0,96 \Longrightarrow D^2(X) = 8,64$ □

5. Cu ce probabilitate putem afirma că din 100 de aruncări ale unei monede, stema apare de un număr de ori cuprins între 40 şi 60?

   Soluţie. Aplicăm formula lui Moivre-Laplace:
   $P(\alpha \leq \frac{X_n - np}{\sqrt{npq}} \leq \beta) \simeq \Phi(\beta) - \Phi(\alpha) \Longleftrightarrow P(\alpha \leq \sqrt{\frac{n}{pq}}(X_n - p) \leq \beta) \simeq$
   $\simeq \Phi(\beta) - \Phi(\alpha)$, unde X_n e v. a. ce ne dă numărul de apariţii ale stemei când aruncăm moneda de n ori

   Avem $p = q = \frac{1}{2}, n = 100 \Longrightarrow \alpha = \sqrt{\frac{100}{\frac{1}{2} \cdot \frac{1}{2}}}(0,4 - 0,5) = -2, \beta =$
   $= \sqrt{\frac{100}{\frac{1}{2} \cdot \frac{1}{2}}}(0,6 - 0,5) = 2$

   Atunci $P(0,4 \leq X_n \leq 0,6) \simeq \Phi(2) - \Phi(-2) = 2\Phi(2) - 1 \simeq 2 \cdot 0,9772-$
   $-1 = 0,9544$ □

6. De câte ori trebuie aruncată o monedă pentru ca să putem spune cu o probabilitate de 0,6 că abaterea frecvenţei de apariţie a stemei de la probabilitatea $p = \frac{1}{2}$ să fie mai mică decât 0,01?

   Soluţie. Cf. teoremei Moivre-Laplace avem $P(\sqrt{\frac{n}{pq}}|X_n - p| \leq \beta) \simeq$
   $\simeq 2\Phi(\beta) - 1$

   Ştim $2\Phi(\beta) - 1 = 0,6 \Longrightarrow \Phi(\beta) = 0,8 \Longrightarrow \beta = 0,84$

   Dar $0,84 = \sqrt{\frac{n}{\frac{1}{2} \cdot \frac{1}{2}}} \cdot 0,01 \Longrightarrow n = 1764$ □

7. Se aruncă de 360 de ori o pereche de zaruri. Cu ce probabilitate ne putem aştepta să apară 12 puncte (dubla 6) de un număr de ori cuprins între 8 şi 12?

   Soluţie. Probabilitatea ca să apară 12 puncte într-o aruncare cu 2 zaruri este $p = \frac{1}{36}$, iar probabilitatea ca să nu apară 12 puncte este $q = 1 - p = \frac{35}{36}$

   Cf. teoremei Moivre-Laplace $\Longrightarrow P(\alpha \leq \sqrt{\frac{n}{pq}}(X_n - p) \leq \beta) \simeq \Phi(\beta) -$ $-\Phi(\alpha)$

   $n = 360, \alpha = \sqrt{\frac{360}{\frac{1}{36}\cdot\frac{35}{36}}} \cdot \left(\frac{8}{360} - \frac{1}{36}\right) = -\frac{6}{5\sqrt{3,5}} \simeq -0,641, \beta = \sqrt{\frac{360}{\frac{1}{36}\cdot\frac{35}{36}}} \cdot$ $\cdot \left(\frac{12}{360} - \frac{1}{36}\right) = \frac{6}{5\sqrt{3,5}} \simeq 0,641$

   $P(-\frac{6}{5\sqrt{3,5}} \leq X_n - \frac{1}{36} \leq \frac{6}{5\sqrt{3,5}}) \simeq 2\Phi(\beta) - 1 = 2 \cdot 0,7389 - 1 = 0,48$ □

8. Probabilitatea ca o anumită operaţie chirurgicală să fie un succes este 0,8. Dacă operaţia este făcută la 120 de persoane, găsiţi probabilitatea ca mai mult de 90 de operaţii să aibă succces?

   Soluţie. Dacă X este v. a. binomială ce reprezintă numărul de operaţii cu succes, atunci vrem să aflăm probabilitatea $P(X \geq 90)$.

   Cum n e mare, vom folosi distribuţia normală cu $m = 120 \cdot 0,8 = 96$ şi $\sigma = \sqrt{120 \cdot 0,8 \cdot 0,2} = 4,38$ pentru a aproxima distribuţia binomială a lui X.

   Obţinem $P(X \geq 90) = P(\frac{X-96}{4,38} \geq \frac{90-96}{4,38}) = P(\frac{X-96}{4,38} \geq -1,37) =$ $1 - P(\frac{X-96}{4,38} < -1,37) = 1 - \Phi(-1,37) = 0,9147.$ □

9. Probabilitatea obţinerii unei piese rebut din producţia unei maşini automate este $p = 0,005$. Să se determine probabilitatea ca din 10000 de piese fabricate la această maşină , numărul pieselor rebut să fie:

   a) între 60 şi 70;

   b) cel mult 70.

   Soluţie. Fie $X_k \sim \begin{pmatrix} 0 & 1 \\ 0,995 & 0,005 \end{pmatrix}, k \in \mathbb{N}^*$ şi $S_n = \sum_{k=1}^{n} X_k$

   S_n reprezintă numărul de piese rebut; este o v. a. cu o repartiţie binomială de parametrii $n = 10000, p = 0,005$. Deci $E(S_n) = np =$ $= 50, D^2(S_n) = npq = 49,75, \sigma = \sqrt{D^2(S_n)} \approx 7$

   Cf. teoremei Moivre-Laplace, repartiţia limită a repartiţiei binomiale este repartiţia normală de parametrii np şi $\sqrt{npq}$ şi rezultă că

   $P(a \leq S_n \leq b) \simeq \Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right)$

a) $P(60 \leq S_n \leq 70) \simeq \Phi\left(\frac{70-50}{7}\right) - \Phi\left(\frac{60-50}{7}\right) = \Phi\left(\frac{20}{7}\right) - \Phi\left(\frac{10}{7}\right) =$
$= 0,49788 - 0,42364 \simeq 0,07$

b) $P(0 \leq S_n \leq 70) \simeq \Phi\left(\frac{70-50}{7}\right) - \Phi\left(-\frac{50}{7}\right) \simeq 0,997$ □

10. 500 de aparate de tipul I şi 1000 de aparate de tipul II cu aceeaşi destinaţie sunt în serviciu. Ele trebuie reparate după exploatare cu probabilitatea de 0,3 pentru tipul I şi de 0,4 pentru tipul II.

    1) Să se determine probabilitatea ca numărul de aparate ce trebuie reparate să fie cuprins între 250 şi 350;

    2) Să se determine numărul de reparaţii n ce trebuie efectuate în atelier pentru ca probabilitatea p a numărului de aparate în reparaţie să fie $p = 0,9$.

    Soluţie. 1) Fie evenimentul A= aparatul trebuie reparat şi X= numărul de apariţii ale evenimentului A; X este o v. a. repartizată normal

    Atunci $E(X) = 500 \cdot 0,3 + 1000 \cdot 0,4 = 550$, $D^2(X) = 500 \cdot 0,3 \cdot 0,7 +$
    $+1000 \cdot 0,4 \cdot 0,6 = 345$

    $P(250 \leq X \leq 350) \simeq P(250 < X < 350) = P(|X - 300| < 50) =$
    $= 2\Phi\left(\frac{50}{\sqrt{345}}\right) - 1$

    2) $P(X < n) = 0,9 = \Phi\left(\frac{n-300}{\sqrt{345}}\right)$

    Din tabele $\Phi(1,27) \simeq 0,9 \Longrightarrow \frac{n-300}{\sqrt{345}} = 1,27 \Longrightarrow n \simeq 323$ □

11. O întreprindere fabrică un anumit produs cu $5^0/_0$ rebuturi. Ce comandă trebuie să facă un beneficiar pentru ca, cu probabilitatea 0,8, să nu achiziţioneze mai mult de 4 produse defecte?

    Soluţie. Fie X= numărul de produse fără defecte achiziţionate dintr-o comandă de n produse

    Cf. teoremei Moivre-Laplace $\Longrightarrow 0,8 = P(X \geq 5) \simeq 1 - \Phi\left(\frac{5-0,05n}{\sqrt{0,05 \cdot 0,95 \cdot n}}\right)$

    Din tabel $\Phi(0,84) \simeq 0,8$

    Deci $\frac{5-0,05n}{\sqrt{0,05 \cdot 0,95 \cdot n}} = -0,84 \Longrightarrow n \simeq 144$ (deoarece $\Phi(0,84) = 1-$
    $-\Phi(-0,84)$) □

12. Presupunem că 120 de persoane stau la coada unei casierii pentru a-şi primi drepturile băneşti. Sumele care trebuie primite de fiecare sunt v. a. $X_1, \ldots, X_{120}$ independente şi identic repartizate cu media $m = 50$ şi abaterea standard $\sigma = 30$. Casieria dispune de 6500 unităţi monetare. Calculaţi probabilitatea ca toate persoanele să-şi primească drepturile.

Soluţie. Notăm $S = X_1 + \ldots + X_{120}$

S poate fi aproximată cu o v. a. normală cu $E(S) = E(\sum_{k=1}^{120} X_k) =$
$= \sum_{k=1}^{120} E(X_k) = 120 \cdot 50 = 6000$ şi $D^2(S) = D^2(\sum_{k=1}^{120} X_k) = \sum_{k=1}^{120} D^2(X_k) =$
$= 120 \cdot 30^2 = 108000$

Atunci $P(S \leq 6500) \simeq \Phi\left(\frac{6500-6000}{\sqrt{108000}}\right) \simeq 0,936$ □

13. La un restaurant pot servi prânzul 75 de clienţi. Din practică , proprietarul ştie ca $20^0/_0$ dintre clienţii care au rezervat loc nu se mai prezintă .

    a) Proprietarul acceptă 90 de rezervări. Care este probabilitatea ca să se prezinte mai mult de 50 de clienţi?

    b) Câte rezervări trebuie să accepte proprietarul pentru ca, cu o probabilitate mai mare sau egală cu 0,9, să poată servi toţi clienţii care se prezintă ? $(\Phi(1,281) = 0,9)$

    Soluţie. a) Fie X v. a. ce reprezintă numărul de clienţi care se prezintă la restaurant ; X este repartizată binomial cu $p = 0,8$.

    Avem $n = 90, E(X) = np = 72, D^2(X) = npq = 14,4 \Longrightarrow \sigma =$
    $= \sqrt{14,4} \simeq 3,795$

    Cf. th. Moivre -Laplace obţinem $P(X > 50) = 1 - P(X \leq 50) = 1-$
    $-P\left(\frac{X-np}{\sigma} \leq \frac{50-72}{3,795}\right) = 1-\Phi(-5,797) = 1-1+\Phi(5,975) = \Phi(5,975) \simeq$
    $\simeq 1$

    b) Vom determina n astfel încât $P(X \leq 75) \geq 0,9$

    $P(X \leq 75) = P\left(\frac{X-np}{\sqrt{npq}} \leq \frac{75-0,8n}{\sqrt{0,16n}}\right) \geq 0,9 = \Phi(1,281) \Longrightarrow \frac{75-0,8n}{\sqrt{0,16n}} \geq$
    $\geq 1,281 \Longrightarrow n < 88$ □

14. Probabilitatea de câştig la ruletă este de $\frac{1}{35}$. Presupunem că la fiecare joc, un jucător poate câştiga sau pierde 1 dolar.

    a) Câte jocuri trebuie jucate zilnic în cazino astfel încât cu probabilitatea 0,5, câştigul cazinoului să fie cel puţin 1000 dolari zilnic. $(\Phi(0) = 0,5)$

    b) Să se determine apoi procentul zilelor în care cazinoul pierde.

    Soluţie. a) Fie X_i v. a. ce reprezintă profitul cazinoului într-o zi; $X_i \sim \begin{pmatrix} 0 & 1 \\ \frac{34}{35} & \frac{1}{35} \end{pmatrix}$

$P(\sum_{i=1}^{n} X_i > 1000) = 1 - P(\sum_{i=1}^{n} X_i \leq 1000) = 1-$

$$-P\left(\frac{\sum_{i=1}^{n} X_i - n\cdot\dfrac{1}{35}}{\sqrt{n\cdot\frac{1}{35}\cdot\frac{34}{35}}} \leq \frac{1000-n\cdot\frac{1}{35}}{\sqrt{n\cdot\frac{1}{35}\cdot\frac{34}{35}}}\right) = 1 - \Phi\left(\frac{1000-n\cdot\frac{1}{35}}{\sqrt{n\cdot\frac{1}{35}\cdot\frac{34}{35}}}\right) = 0,5 \Longrightarrow$$

$$\Longrightarrow \Phi\left(\frac{1000-n\cdot\frac{1}{35}}{\sqrt{n\cdot\frac{1}{35}\cdot\frac{34}{35}}}\right) = 0,5 \Longrightarrow \frac{1000-n\cdot\frac{1}{35}}{\sqrt{n\cdot\frac{1}{35}\cdot\frac{34}{35}}} = 0 \Longrightarrow n = 35000$$

b) $P\left(\sum_{i=1}^{35000} X_i < 0\right) = \Phi\left(\frac{-1000}{\sqrt{35000\cdot\frac{1}{35}\cdot\frac{34}{35}}}\right) \simeq 0$ □

15. Calitatea unor piese este apreciată prin caracteristica X care este repartizată normal cu media 10 cm şi abaterea medie pătratică 0,15 cm. Piesele sunt acceptate numai dacă valorile caracteristicii X sunt cuprinse în intervalul (9,8 cm, 10,2 cm). Firma producătoare are o comandă de 3000 de piese. Să se determine numărul de piese care trebuie să fie produse de firmă astfel încât să se poată onora contractul.

Soluţie. Fie n numărul de piese care trebuie să fie produse de firmă astfel încât contractul să fie onorat, x numărul pieselor contractate şi p probabilitatea realizării evenimentului (9,8 cm< X <10,2 cm)

Cf. teoremei lui Bernoulli, pentru n suficient de mare se poate considera $n \approx x \cdot \frac{1}{p}$

$P(9,8 < X < 10,2) = \Phi(\frac{10,2-10}{0,15}) - \Phi(\frac{9,8-10}{0,15}) = 0,81$

Deci $n = 3000 \cdot \frac{1}{0,81} = 3690$ □

16. Pentru un studiu biologic sunt cercetate 1000 de probe independente, dacă au sau nu o anumită caracteristică biologică . Să se determine o margine inferioară a probabilităţii ca diferenţa, în valoare absolută , dintre frecvenţa relativă şi probabilitatea p de a apărea caracteristica într-o probă , să fie mai mică decât 0,03.

Soluţie. Cf. consecinţei teoremei lui Bernoulli $\Longrightarrow p = q = \frac{1}{2}, \varepsilon =$ $= 0,03 \Longrightarrow P(|\frac{x}{n} - p| < 0,03) \geq 1 - \frac{pq}{n\varepsilon^2} = 1 - \frac{\frac{1}{2}\cdot\frac{1}{2}}{1000\cdot 0,03^2} = 0,72$ □

17. 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.

Soluţie. Teorema este aplicabilă dacă media e finită şi dispersia este egal mărginită .

$E(X_n) = (-5n) \cdot \frac{1}{3n^2} + 0 \cdot (1 - \frac{2}{3n^2}) + 5n \cdot \frac{1}{3n^2} = 0, E(X_n^2) = (-5n)^2 \cdot$ $\cdot \frac{1}{3n^2} + 0^2 \cdot (1 - \frac{2}{3n^2}) + (5n)^2 \cdot \frac{1}{3n^2} = \frac{50}{3}, D^2(X_n) = \frac{50}{3} \Longrightarrow$ teorema se aplică

$E(Y_n) = (-n^2) \cdot \alpha^{-n} + 0 \cdot (1 - 2\alpha^{-n}) + n^2 \cdot \alpha^{-n} = 0, E(Y_n^2) =$ $= (-n^2)^2 \cdot \cdot \alpha^{-n} + 0^2 \cdot (1 - 2\alpha^{-n}) + (n^2)^2 \cdot \alpha^{-n} = \frac{2n^4}{\alpha^n}, D^2(Y_n) = \frac{2n^4}{\alpha^n} \Longrightarrow$ $\Longrightarrow \lim\limits_{n\to\infty} D^2(Y_n) = 0 \Longrightarrow \exists c \in (0, \infty)$ astfel încât $\forall n \geq 1, D^2(Y_n) \leq$ $\leq c \Longrightarrow$ teorema se aplică

$E(Z_n) = \int_0^\infty x\lambda^{-n} e^{-\frac{x}{\lambda^n}} dx = \lambda^n, E(Z_n^2) = \int_0^\infty x^2 \lambda^{-n} e^{-\frac{x}{\lambda^n}} dx = 2\lambda^{2n} \Longrightarrow$ $\Longrightarrow D^2(Z_n) = \lambda^{2n} \Longrightarrow$

$$\lim_{n\to\infty} D^2(Z_n) = \begin{cases} 0, & \lambda \in (0,1) \\ 1, & \lambda = 1 \\ \infty, & \lambda > 1 \end{cases}$$

deci teorema se aplică numai dacă $\lambda \in (0, 1]$ □

18. Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente ce pot lua valorile $\pm\sqrt{\lg n}$ cu probabilităţile $P(X_k = \sqrt{\lg k}) = P(X_k = -\sqrt{\lg k}) =$ $= \frac{1}{2}, k = 2, 3, \ldots, P(X_1 = 0) = 1$. Să se arate că şirul dat se supune legii numerelor mari în formularea lui Cebîşev.

Soluţie. Se constată că $E(X_k) = \sqrt{\lg k} \cdot \frac{1}{2} + (-\sqrt{\lg k}) \cdot \frac{1}{2} = 0, E(X_k^2) =$ $= (\sqrt{\lg k})^2 \cdot \frac{1}{2} + (-\sqrt{\lg k})^2 \cdot \frac{1}{2} = \lg k, D^2(X_k) = \lg k, k = 2, 3, \ldots$

Luăm $Y_n = \frac{1}{n}\sum\limits_{k=1}^{n} X_k \Longrightarrow E(Y_n) = \dfrac{1}{n}\sum\limits_{k=1}^{n} E(X_k) = 0$

$D^2(Y_n) = \frac{1}{n^2}\sum\limits_{k=1}^{n} D^2(X_k) = \dfrac{1}{n^2}\sum\limits_{k=1}^{n} \lg k$

$D^2(Y_n)$ poate fi majorată astfel:

$\sum\limits_{k=1}^{n} \lg k < \displaystyle\int_1^{n+1} \lg x dx = x \lg x /_1^{n+1} - \int_1^{n+1} dx = (n+1)\lg(n+1) -$ $-n \Longrightarrow D^2(Y_n) < \frac{(n+1)\lg(n+1)}{n^2} \Longrightarrow \lim\limits_{n\to\infty} D^2(Y_n) = 0$

Sunt îndeplinite condiţiile teoremei Markov şi rezultă
$\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\sum\limits_{k=1}^{n} X_k - \frac{1}{n}\sum\limits_{k=1}^{n} E(X_k)\right| < \varepsilon\right) = 1 \Longrightarrow$ şirului dat i se poate aplica legea numerelor mari $\square$

19. Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente astfel încât $P(X_n = \frac{1}{n^\beta}) = P(X_n = -\frac{1}{n^\beta}) = p$, cu $\frac{1}{3} < \beta \leq \frac{1}{2}, n \in \mathbb{N}^*$ şi $P(X_n = 0) = 1 - 2p$. Să se arate că şirului $(X_n)_{n\geq 1}$ i se poate aplica teorema lui Leapunov.

Soluţie. Obţinem $E(X_k) = \frac{1}{k^\beta} \cdot p + (-\frac{1}{k^\beta}) \cdot p = 0, E(X_k^2) = (\frac{1}{k^\beta})^2 \cdot p +$
$+ (-\frac{1}{k^\beta})^2 \cdot p = \frac{2p}{k^{2\beta}} = \mu_k^2, E(|X_k|^3) = \frac{2p}{k^{3\beta}} = \rho_k^3 \Longrightarrow \mu^2(n) = \sum\limits_{k=1}^{n} \mu_k^2 =$
$= 2p\sum\limits_{k=1}^{n}\frac{1}{k^{2\beta}}, \rho^3(n) = \sum\limits_{k=1}^{n}\rho_k^3 = 2p\sum\limits_{k=1}^{n}\frac{1}{k^{3\beta}}$

Cum $\frac{1}{3} < \beta \leq \frac{1}{2} \Longrightarrow 1 < 3\beta \leq \frac{3}{2} \Longrightarrow \lim\limits_{n\to\infty}\sum\limits_{k=1}^{n}\frac{1}{k^{3\beta}} = \sum\limits_{k=1}^{\infty}\frac{1}{k^{3\beta}}$ e o serie convergentă

$\sum\limits_{k=1}^{n}\frac{1}{k^{2\beta}} > \sum\limits_{k=1}^{n}\frac{1}{k}$ ce tinde la ∞ când $n \longrightarrow \infty$

Condiţia lui Leapunov $\lim\limits_{n\to\infty}\frac{\sqrt[3]{\rho(n)^3}}{\sqrt{\mu(n)^2}} = 0$ e îndeplinită $\Longrightarrow$ şirului $(X_n)_{n\geq 1}$ i se poate aplica teorema lui Leapunov $\square$

20. Fie $(X_n)_{n\geq 1}$ un şir de variabile aleatoare independente pentru care $P(X_n = 3^{n-1}) = P(X_n = -3^{n-1}) = \frac{1}{2}, n = 1, 2, 3, \ldots$ Dacă notăm cu $Y_n = \frac{1}{S_n}\sum\limits_{k=1}^{n} X_k$, unde $S_n = D^2(\sum\limits_{k=1}^{n} X_k)$, să se arate că Y_n nu converge în probabilitate către 0.

Soluţie. Avem $E(X_k) = 3^{k-1} \cdot \frac{1}{2} + (-3^{k-1}) \cdot \frac{1}{2} = 0, E(X_k^2) = (3^{k-1})^2 \cdot$
$\cdot \frac{1}{2} + (-3^{k-1})^2 \cdot \frac{1}{2} = 3^{2(k-1)}, D^2(X_k) = 3^{2(k-1)}$

$D^2(\sum\limits_{k=1}^{n} X_k) = \sum\limits_{k=1}^{n} D^2(X_k) = \sum\limits_{k=1}^{n} 3^{2(k-1)} = \frac{3^{2n} - 1}{3^2 - 1} = \frac{9^n - 1}{8}$

$|X_n| = |X_n + X_{n-1} - X_{n-1}| \leq |X_n + X_{n-1}| + |X_{n-1}| \leq |X_n + X_{n-1} +$
$+ \ldots + X_1| + |X_{n-1}| + \ldots + |X_1| \Longrightarrow |Y_n| \geq \frac{|X_n| - |X_1| - |X_2| - \ldots - |X_{n-1}|}{S_n} =$
$= \frac{3^{n-1} - (1 + 3 + \ldots + 3^{n-2})}{\sqrt{\frac{9^n - 1}{8}}} > \frac{\sqrt{2}}{3} \Longrightarrow P(|Y_n| > \frac{\sqrt{2}}{3}) = 1 \Longrightarrow P(|Y_n| > \varepsilon)$ nu

tinde la 0, dacă $\varepsilon \leq \frac{\sqrt{2}}{3} \Longrightarrow (Y_n)_n$ nu converge în probabilitate către 0 $\square$

21. Dacă funcţiile de repartiţie corespunzătoare şirului de variabile $(X_n)_n$ tind către o repartiţie limită şi dacă şirul $(Y_n)_n$ converge în probabilitate la 0, atunci şirul $(X_nY_n)_n$ converge în probabilitate către 0.

Soluţie. Fie $a \in \mathbb{R}, a > 0 \Longrightarrow \big\{\omega/|X_n(\omega)Y_n(\omega)| > \varepsilon\big\} \subset$
$\subset \big\{\omega/|X_n(\omega)| > a\big\} \cup \big\{\omega/|Y_n(\omega)| > \frac{\varepsilon}{a}\big\} \Longrightarrow$
$\Longrightarrow P(\big\{\omega/|X_n(\omega)Y_n(\omega)| > \varepsilon\big\}) \leq P(\big\{\omega/|X_n(\omega)| > a\big\})+$
$+P(\big\{\omega/|Y_n(\omega)| > \frac{\varepsilon}{a}\big\}) \leq P(X_n(\omega) < -a) + 1 - P(X_n(\omega) \leq a)+$
$+P(|Y_n(\omega)| > \frac{\varepsilon}{a}) \leq F_n(-a) + 1 - F_n(a) + P(|Y_n(\omega)| > \frac{\varepsilon}{a})$

Dacă a e luat astfel încât $-a$ şi a să fie puncte de continuitate pentru funcţia de repartiţie limită F, atunci din ipoteză obţinem
$\limsup\limits_{n\to\infty} P(\big\{\omega/|X_n(\omega)Y_n(\omega)| > \varepsilon\big\}) \leq F(-a) + 1 - F(a)$

Alegem a astfel încât $-a$ şi a să fie puncte de continuitate pentru F şi, în plus, să avem $F(-a) < \frac{\delta}{2}, 1 - F(a) < \frac{\delta}{2}$ cu $\delta > 0 \Longrightarrow$
$\Longrightarrow \limsup\limits_{n\to\infty} P(\big\{\omega/|X_n(\omega)Y_n(\omega)| > \varepsilon\big\}) \leq \delta$

Cum δ e arbitrar $\Longrightarrow (X_nY_n)_n$ converge în probabilitate către 0 $\square$

22. Fie $(X_n)_n$ un şir de variabile aleatoare Poisson, independente cu $E(X_k) = \lambda_k$ şi $Y_n = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$. Să se arate că dacă există $\lim\limits_{n\to\infty} \frac{1}{n}\sum\limits_{k=1}^{n}\lambda_k = \lambda$, atunci şirul de variabile aleatoare $(Y_n)_n$ converge în probabilitate către λ.

Soluţie. Cum $(X_n)_n$ un şir de variabile aleatoare Poisson $\Longrightarrow \lambda_k = E(X_k) = D^2(X_k)$

Cum variabilele X_n sunt independente $\Longrightarrow D^2(Y_n) = \frac{1}{n^2}\sum\limits_{k=1}^{n} D^2(X_k) =$

$$= \frac{1}{n^2}\sum_{k=1}^{n}\lambda_k = \frac{1}{n} \cdot \frac{\sum\limits_{k=1}^{n}\lambda_k}{n} \longrightarrow 0, \text{ când } n \longrightarrow \infty$$

$$E(Y_n) = \frac{1}{n}\sum_{k=1}^{n} E(X_k) = \frac{1}{n}\sum_{k=1}^{n}\lambda_k \longrightarrow \lambda$$

Folosind inegalitatea lui Cebîşev $\Longrightarrow 0 \leq P(|Y_n - \lambda| \geq \varepsilon) < \frac{D^2(Y_n)}{\varepsilon^2} \longrightarrow 0$, deci $P(|Y_n - \lambda| \geq \varepsilon) \longrightarrow 0 \Longrightarrow (Y_n)_n$ converge în probabilitate către 0 $\square$

23. Fie $(X_n)_n$ un şir de v. a. independente astfel încât $P(X_n = 1) = = 1 - \frac{1}{n}, P(X_n = n) = \frac{1}{n}, n > 1$. Să se arate că $(X_n)_n$ converge în probabilitate la 1, când $n \longrightarrow \infty$, dar $(X_n)_n$ nu converge aproape sigur la 1, când $n \longrightarrow \infty$.

Soluţie. $\forall \varepsilon > 0, P(|X_n - 1| > \varepsilon) = P(X_n = n) = \frac{1}{n} \longrightarrow 0 \Longrightarrow X_n \xrightarrow{P} 1$

$X_n \xrightarrow{a.s} X \iff \forall \varepsilon > 0, \forall \delta \in (0,1) \exists n_0$ astfel încât $\forall n > n_0$ avem $P(\bigcap\limits_{m>n} \{|X_m - X| < \varepsilon\}) > 1 - \delta$ (1)

Pentru $\forall \varepsilon > 0, \delta \in (0,1), N > n$ obţinem $P(\bigcap\limits_{m>n} \{|X_m - 1| < \varepsilon\}) \leq$

$\leq P(\bigcap\limits_{m=n+1}^{N} \{|X_m - 1| < \varepsilon\}) = \prod\limits_{m=n+1}^{N} P(|X_m - 1| < \varepsilon) =$

$= \prod\limits_{m=n+1}^{N} P(X_m = 1) = \prod\limits_{m=n+1}^{N} \left(1 - \frac{1}{m}\right) = \frac{n}{N} < 1 - \delta$ cu condiţia să alegem N astfel încât $N > \frac{n}{1-\delta} \Longrightarrow$ nu există n_0 astfel încât să aibă loc (1) $\Longrightarrow (X_n)_n$ nu converge aproape sigur la 1, când $n \longrightarrow \infty$ □

24. Fie $\alpha > 0$ şi $(X_n)_n$ un şir de v. a. astfel încât $P(X_n = 1) = = 1 - \frac{1}{n^\alpha}, P(X_n = n) = \frac{1}{n^\alpha}, n > 1$. Să se arate că $(X_n)_n$ converge în probabilitate la 1 şi $(X_n)_n$ converge în medie de ordinul r la 1, când $r < \alpha$, dar $(X_n)_n$ nu converge în medie de ordinul r la 1, când $r \geq \alpha$.

Soluţie. $P(|X_n - 1| > \varepsilon) = P(X_n = n) = \frac{1}{n^\alpha} \longrightarrow 0 \Longrightarrow X_n \xrightarrow{P} 1$

Cum $E(|X_n - 1|^r) = 0 \cdot \left(1 - \frac{1}{n^\alpha}\right) + |n-1|^r \cdot \frac{1}{n^\alpha} = \frac{(n-1)^r}{n^\alpha} \Longrightarrow$

$$E(|X_n - 1|^r) \longrightarrow \begin{cases} 0, & r < \alpha \\ 1, & r = \alpha \\ \infty, & r > \alpha \end{cases}$$

□

25. Fie $(X_n)_n$ un şir de v. a. având repartiţia $X_n \sim \begin{pmatrix} -n^{\frac{2}{r}} & 0 & n^{\frac{2}{r}} \\ \frac{1}{2n^2} & 1 - \frac{1}{n^2} & \frac{1}{2n^2} \end{pmatrix}$. Să se arate că $(X_n)_n$ converge aproape sigur către 0.

Soluţie. $A_{j,\delta} = \{\omega / |X_j(\omega)| \leq \delta\}$

$B_{n,\delta} = \bigcap\limits_{j=n}^{\infty} A_{j,\delta}$

$$B^c_{n,\delta} = \bigcup_{j=n}^{\infty} A^c_{j,\delta}$$

$$P(B^c_{n,\delta}) = P(\bigcup_{j=n}^{\infty} A^c_{j,\delta}) \leq \sum_{j=n}^{\infty} P(A^c_{j,\delta}) = \sum_{j=n}^{\infty} P(|X_j(\omega)| > \delta)$$

Din definiţia şirului $(X_n)_n \Longrightarrow$ pentru $n \geq 1, \delta < 1$ avem $P(|X_j(\omega)| > \delta) = P(X_j(\omega) = -j^{\frac{2}{r}}) + P(X_j(\omega) = j^{\frac{2}{r}}) = \frac{1}{j} \Longrightarrow P(B_{n,\delta}) \leq \sum_{j=n}^{\infty} \frac{1}{j^2} \longrightarrow$
$\longrightarrow 0$ (când $n \longrightarrow \infty$)

$$\lim_{n\to\infty} P(B_{n,\delta}) = \lim_{n\to\infty} P(\bigcap_{j=n}^{\infty} \{\omega / |X_j(\omega)| \leq \delta\}) = 1 \qquad \square$$

26. Fie $(X_n)_n$ un şir de v. a. pozitive cu densităţile de repartiţie date de

$$f_n(x) = \begin{cases} \frac{x^{n-1} e^{-x}}{(n-1)!}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Să se arate că şirul $\left(\frac{X_n - E(X_n)}{\sqrt{D^2(X_n)}}\right)_n$ urmează la limită o lege normală $N(0,1)$.

Soluţie. $E(X_n) = \frac{1}{(n-1)!}\int_0^{\infty} x^n e^{-x} dx = \frac{\Gamma(n+1)}{(n-1)!} = n$

$E(X_n^2) = \frac{1}{(n-1)!}\int_0^{\infty} x^{n+1} e^{-x} dx = \frac{\Gamma(n+2)}{(n-1)!} = n(n+1)$

$D^2(X_n) = n(n+1) - n^2 = n$

Şirul $Y_n = \frac{X_n - E(X_n)}{\sqrt{D^2(X_n)}}$ devine $Y_n = \frac{X_n - n}{\sqrt{n}} = \frac{1}{\sqrt{n}} X_n - \sqrt{n}$

Dacă notăm $\varphi_n(t)$ funcţia caracteristică a v. a. Y_n şi reuşim să arătăm că $\lim_{n\to\infty} \varphi_n(t) = e^{-\frac{t^2}{2}}$, demonstraţia s-a încheiat, deoarece folosim teorema ce leagă între ele şirurile de funcţii de repartiţie şi cele caracteristice corespunzătoare unui şir de v. a.

Folosind definiţia funcţiei caracteristice avem
$\varphi_n(t) = \int_0^{\infty} e^{it\left(\frac{x}{\sqrt{n}} - \sqrt{n}\right)} f_n(x) dx = e^{-it\sqrt{n}} \cdot \frac{1}{(n-1)!}\int_0^{\infty} x^{n-1} e^{-\left(1-\frac{it}{\sqrt{n}}\right)x} dx =$
$= e^{-it\sqrt{n}}\left(1 - \frac{it}{\sqrt{n}}\right)^{-n}$ (cu ajutorul funcţiei Γ)

$\ln \varphi_n(t) = -it\sqrt{n} - n \ln\left(1 - \frac{it}{\sqrt{n}}\right) = -it\sqrt{n} + n \cdot \left(\frac{it}{\sqrt{n}} - \frac{t^2}{2n} + o(n)\right) =$
$= -\frac{t^2}{2} + o'(n)$, dacă $\left|\frac{t}{\sqrt{n}}\right| < 1$

Cum $\lim\limits_{n\to\infty} \ln \varphi_n(t) = -\dfrac{t^2}{2} \Longrightarrow \lim\limits_{n\to\infty} \varphi_n(t) = \mathrm{e}^{-\frac{t^2}{2}} \Longrightarrow$

$$\Longrightarrow \lim_{n\to\infty} P\left(\frac{X_n(\omega)-E(X_n)}{\sqrt{D^2(X_n)}} < x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\,dt \qquad \square$$

27. Se ştie că dacă $(X_n)_{n\in\mathbf{N}^*}$ este un şir de v. a. convergent în medie pătratică , atunci $\lim\limits_{n\to\infty} E(X_n) = E(X)$ şi $\lim\limits_{n\to\infty} E(X_n^2) = E(X^2)$, unde $X = \lim\limits_{n\to\infty} X_n$. Să se arate că dacă se înlocuieşte condiţia de convergenţă în medie pătratică cu convergenţa în probabilitate, afirmaţia nu mai este adevărată .

    Soluţie. Fie $(X_n)_{n\in\mathbf{N}^*}$ un şir de v. a. care pot lua valorile $-(n+4), -1, n+4$ cu probabilităţile $P(X_n = -n-4) = \frac{1}{n+4}, P(X_n = -1) = 1 - \frac{4}{n+4}, P(X_n = n+4) = \frac{3}{n+4}$

    Avem $\lim\limits_{n\to\infty} P(|X_n+1| > \varepsilon) = 0$, ceea ce arată că şirul $(X_n)_{n\in\mathbf{N}^*}$ converge în probabilitate către -1.

    Pe de altă parte $E(X_n) = 1 + \frac{4}{n+4}$, deci $\lim\limits_{n\to\infty} E(X_n) = 1$, de unde rezultă concluzia că $\lim\limits_{n\to\infty} E(X_n) = 1 \neq -1 = E(\lim\limits_{n\to\infty} X_n)$. $\square$

28. Să se arate că există şiruri de v. a. care converg atât în medie de ordinul r cât şi aproape sigur.

    Soluţie. Fie şirul de v. a. $(X_n)_{n\in\mathbf{N}^*}$, unde $X_n \sim \begin{pmatrix} -\frac{1}{n} & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ Deoarece $E(|X_n|^r) = \frac{1}{n^r}$ rezultă că $\lim\limits_{n\to\infty} E(|X_n|^r) = 0$, ceea ce arată că şirul $(X_n)_{n\in\mathbf{N}^*}$ converge în medie de ordinul r.

    Fie $T_{j,\delta} = \left\{\omega/|X_j(\omega)| \le \delta\right\}$

    Din felul cum am definit şirul $(X_n)_{n\in\mathbf{N}^*}$ rezultă că pentru orice $j < k$ avem $|X_j| > |X_k|$.

    Deci $\left\{\omega/|X_j(\omega)| \le \delta\right\} \subset \left\{\omega/|X_k(\omega)| \le \delta\right\} \Longrightarrow T_{1,\delta} \subset T_{2,\delta} \subset \ldots \subset \subset T_{n,\delta} \subset T_{n+1,\delta} \subset \ldots$

    In acest caz $S_{n,\delta} = \bigcap\limits_{j=n}^{\infty} T_{j,\delta} = T_{n,\delta}$

    Fie $\delta > 0$ dat. Dacă $n > \frac{1}{\delta}$ din definiţia şirului şi din relaţiile găsite avem $P(S_{n,\delta}) = P(T_{n,\delta}) = P(\left\{\omega/|X_n(\omega)| \le \delta\right\}) = 1$, ceea ce înseamnă că $(X_n)_{n\in\mathbf{N}^*}$ converge aproape sigur către 0. $\square$

29. Să se arate că poate exista un şir de v. a. convergent în medie pătratică şi care să nu fie convergent aproape sigur.

Soluţie. Fie $(Y_n)_{n\in\mathbf{N}^*}$ un şir de v. a. independente care iau doar valorile -1,0,1 cu probabilităţile $P(Y_n=-1)=P(Y_n=1)=\frac{1}{2\sqrt[4]{n}}$, $P(Y_n=0)=1-\frac{1}{\sqrt[4]{n}}, n=1,2,\ldots$

Pentru $n\geq 2$ definim evenimentul E_n ca fiind evenimentul ce constă în faptul că toţi $Y_i=0$ cu $n-\sqrt{n}\leq i<n$.

Probabilitatea acestui eveniment este, datorită independenţei v. a. $(Y_n)_{n\in\mathbf{N}^*}$:

$$P(E_n)=\prod_{n-\sqrt{n}\leq i<n}P(Y_i=0)=\prod_{n-\sqrt{n}\leq i<n}(1-\frac{1}{\sqrt[4]{i}})$$

Se ştie că pentru orice numır real, $0\leq b<1$ este adevărată inegalitatea $1-b\leq \mathrm{e}^{-b}$.

Utilizând acest fapt, putem scrie $P(E_n)\leq \mathrm{e}^{-\sum\limits_{n-\sqrt{n}\leq i<n}\frac{1}{\sqrt[4]{i}}}$, de unde rezultă $\lim\limits_{n\to\infty}P(E_n)=0$.

Definim, cu ajutorul şirului $(Y_n)_{n\in\mathbf{N}^*}$ considerat mai înainte şirul de v. a. $(X_n)_{n\in\mathbf{N}^*}$ astfel :

$X_1=Y_1$

$$X_n=\begin{cases} Y_n, & \text{dacă se realizează } E_n^c \\ \pm 1, & \text{cu probabilităţi egale dacă se realizează evenimentul } E_n \end{cases}$$

pentru orice $n\geq 2$

Avem $E(X_n)=0$ şi $E(X_n^2)=P(E_n)+(1-P(E_n))\frac{1}{\sqrt[4]{n}}$ de unde $\lim\limits_{n\to\infty}E(X_n^2)=0$, adică $(X_n)_{n\in\mathbf{N}^*}$ converge în medie de ordinul 2 către 0. Să arătăm că şirul $(X_n)_{n\in\mathbf{N}^*}$ nu converge aproape sigur către 0.

Să presupunem că şirul $(X_n)_{n\in\mathbf{N}^*}$ converge aproape sigur către 0.

Atunci, deoarece v. a. limită $X=0$, rezultă că pentru orice $\delta<1$ avem

$T_{j,\delta}=\{\omega/|X_j(\omega)|\leq\delta\}=\{\omega/X_j(\omega)=0\}=E_j^c\cap\{\omega/Y_j(\omega)=0\}$

Prin urmare $T_{j,\delta}\subset\{\omega/Y_j(\omega)=0\}$ şi $T_{j,\delta}\subset E_j^c$

Deoarece $S_{n,\delta}=\bigcap\limits_{j=n}^{\infty}T_{j,\delta}$ va rezulta că putem scrie

$S_{[m-\sqrt{m}],\delta}=\bigcap\limits_{j=[m-\sqrt{m}]}^{\infty}T_{j,\delta}\subset\bigcap\limits_{j=[m-\sqrt{m}]}^{\infty}\{\omega/Y_j(\omega)=0\}\subset E_m$, unde prin $[m-\sqrt{m}]$ înţelegem partea întreagă a numărului $m-\sqrt{m}$.

Mai departe putem scrie $S_{n,\delta}=\bigcap\limits_{j=n}^{\infty}T_{j,\delta}\subset\bigcap\limits_{j=n}^{\infty}E_j^c\subset E_n^c$.

Deoarece am presupus că şirul $(X_n)_{n\in\mathbb{N}^*}$ converge aproape sigur, rezultă că pentru orice $\varepsilon > 0$ există un rang $N = N(\varepsilon)$ astfel încât $P(S_{n,\delta}) >$ $> 1 - \varepsilon$ pentru toţi $n \geq N$.

Alegând pe m astfel încât $m - \sqrt{m} \geq N$ obţinem
$P(E_m) \geq P(S_{[m-\sqrt{m}],\delta}) > 1 - \varepsilon$, rezultat în contradicţie cu
$P(E_m^c) \geq P(S_{m,\delta}) > 1 - \varepsilon$ □

30. Convergenţa aproape sigură nu implică convergenţa în medie de ordinul r.

Soluţie. Fie $(X_n)_{n\in\mathbb{N}^*}$ un şir de v. a. cu repartiţiile

$$X_n \sim \begin{pmatrix} -n^{\frac{2}{r}} & 0 & n^{\frac{2}{r}} \\ \frac{1}{2n^2} & 1-\frac{1}{n^2} & \frac{1}{2n^2} \end{pmatrix}, \ r > 0, n \in \mathbb{N}^*$$

Fie $T_{j,\delta} = \{\omega/|X_j(\omega)| \leq \delta\}, S = \bigcap_{j=n}^{\infty} T_{j,\delta}$ şi $S_{n,\delta}^c = \bigcup_{j=n}^{\infty} T_{j,\delta}^c$

Atunci $P(S_{n,\delta}^c) = P(\bigcup_{j=n}^{\infty} T_{j,\delta}^c) \leq \sum_{j=n}^{\infty} P(T_{j,\delta}^c)) =$

$$= \sum_{j=n}^{\infty} P(\{\omega/|X_j(\omega)| > \delta\})$$

Din definiţia şirului de v. a. $(X_n)_{n\in\mathbb{N}^*}$ urmează că pentru $n \geq 1$ şi $\delta < 1$ avem $P(\{\omega/|X_j(\omega)| > \delta\}) = P(\{\omega/X_j(\omega) = -j^{\frac{2}{r}}\})+$
$+P(\{\omega/X_j(\omega) = j^{\frac{2}{r}}\}) = \frac{1}{j^2}$, aşa că $P(S_{n,\delta}) \leq \sum_{j=n}^{\infty} \frac{1}{j^2}$

Deci $\lim_{n\to\infty} P(S_{n,\delta}) = \lim_{n\to\infty} P(\bigcap_{j=n}^{\infty} \{\omega/|X_j(\omega)| \leq \delta\}) = 1$, adică $(X_n)_{n\in\mathbb{N}^*}$ converge aproape sigur către 0.

Pe de altă parte, $E(|X_n|^r) = P(|X_n|^r \neq 0) + P(|X_n|^r = 0) = 1$, deci şirul $(X_n)_{n\in\mathbb{N}^*}$ nu converge în medie de ordinul r către 0. □

31. Să se arate că dacă un şir de v. a. converge în probabilitate, nu rezultă că şirul dat converge aproape sigur.

Soluţie. Considerăm câmpul de probabilitate (Ω, K, P), unde $\Omega =$ $= [0, 1), K = \mathcal{B}_{[0,1)}, P$ măsura Lebesgue pe dreaptă .

Pentru fiecare număr natural m vom considera şirul de v. a.
$X_1^{(m)}, X_2^{(m)}, \ldots, X_m^{(m)}$ definite astfel:

$$X_j^{(m)}(\omega) = \begin{cases} 1, & \text{dacă } \omega \in \left[\frac{j-1}{m}, \frac{j}{m}\right) \\ 0, & \text{în rest} \end{cases}$$

Punem $X_1^{(1)}(\omega) = 1, \omega \in [0,1)$ şi considerăm următorul şir de v. a. $X_1^{(1)}, X_1^{(2)}, X_2^{(2)}, X_1^{(3)}, X_2^{(3)}, X_3^{(3)}, \ldots, X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)}, \ldots$.

Oricare ar fi $\varepsilon > 0$ avem
$P(\{\omega/|X_k^{(n)}(\omega)| \geq \varepsilon\}) = P(\{\omega/\omega \in \left[\frac{k-1}{n}, \frac{k}{n}\right) \}) = \frac{1}{n}$.

Dacă $n > N(\varepsilon, \eta)$ atunci $P(\{\omega/|X_k^{(n)}(\omega)| \geq \varepsilon\}) < \eta$, deci şirul de v. a. considerat converge în probabilitate către 0.

Se observă că şirul nu converge aproape sigur către 0. Intr-adevăr, dacă $\omega_0 \in [0,1)$, atunci pentru orice m există un j astfel încât $\omega_0 \in \left[\frac{j-1}{m}, \frac{j}{m}\right)$, deci $X_j^{(m)}(\omega_0) = 1$ şi de aici rezultă că în şirul $X_1^{(1)}(\omega_0), X_1^{(2)}(\omega_0), X_2^{(2)}(\omega_0), X_1^{(3)}(\omega_0), X_2^{(3)}(\omega_0), X_3^{(3)}(\omega_0), \ldots$ oricare ar fi rangul termenului din şir găsim după el termeni ai şirului egali cu 1, care demonstrează afirmaţia. □

32. Convergenţa în probabilitate nu implică întotdeauna convergenţa în medie de ordinul r.

Soluţie. Fie şirul de v. a. $(X_n)_{n\in\mathbf{N}^*}$, unde $P(X_n = 0) = 1 - \frac{1}{n^2}$, $P(X_n = -n) = P(X_n = n) = \frac{1}{2n^2}$.

Atunci $\forall\varepsilon > 0$ şi $\forall\eta > 0 \exists n > N(\varepsilon, \eta)$ astfel încât $P(|X_n| > \varepsilon) = \frac{1}{2n^2} <$ $< \eta$ îndată ce $n > N(\varepsilon, \eta)$, ceea ce arată că şirul $(X_n)_{n\in\mathbf{N}^*}$, converge în probabilitate către 0.

Pe de altă parte, $E(|X_n - 0|^2) = E(X_n^2) = 1$, deci şirul nu converge în medie de ordinul 2. □

33. Să se arate că dacă $(X_n)_{n\in\mathbf{N}^*}$ converge în repartiţie, nu rezultă că el converge şi în probabilitate.

Soluţie. Considerăm câmpul de probabilitate (Ω, K, P), unde $\Omega =$ $= [0,1], K = \mathcal{B}_{[0,1]}$ şi P măsura Lebesgue pe dreaptă .

Definim şirul de v. a. $(X_n)_{n\in\mathbf{N}^*}$ astfel:

$$X_n(\omega) = \begin{cases} 0, & \text{dacă } \omega \in \left[0, \frac{1}{2}\right) \\ 1, & \text{dacă } \omega \in \left[\frac{1}{2}, 1\right] \end{cases}$$

şi v. a. X astfel

$$X(\omega) = \begin{cases} 1, & \text{dacă } \omega \in \left[0, \frac{1}{2}\right) \\ 0, & \text{dacă } \omega \in \left[\frac{1}{2}, 1\right] \end{cases}$$

Atunci $P(\{\omega/|X_n(\omega) - X(\omega)| = 1\}) = 1$, deci şirul $(X_n)_{n\in\mathbf{N}^*}$ nu converge în probabilitate către X.

Funcţiile de repartiţie ale v. a X_n şi X sunt:

$$F_n(x) = P(\{\omega/X_n(\omega) < x\}) = \begin{cases} 0, & \text{dacă } x \leq 0 \\ \frac{1}{2}, & \text{dacă } 0 < x \leq 1 \\ 1, & \text{dacă } x > 1 \end{cases}$$

$$F(x) = P(\{\omega/X(\omega) < x\}) = \begin{cases} 0, & \text{dacă } x \leq 0 \\ \frac{1}{2}, & \text{dacă } 0 < x \leq 1 \\ 1, & \text{dacă } x > 1 \end{cases}$$

deci $F_n(x) = F(x), x \in \mathbb{R}, n = 1, 2, \ldots \Longrightarrow \lim\limits_{n\to\infty} F_n(x) = F(x)$ ceea ce dovedeşte că şirul $(X_n)_{n\in\mathbb{N}^*}$ converge în repartiţie către X. □

34. Se ştie că funcţia caracteristică este continuă pentru orice $t \in \mathbb{R}$, precum şi teorema lui Helly. Să se arate că este esenţial ca limita $\varphi(t)$ să fie continuă în $t = 0$.

Soluţie. Fie

$$F_n(x) = \begin{cases} 0, & \text{dacă } x \leq -n \\ \frac{x+n}{2n}, & \text{dacă } -n < x < n \\ 1, & \text{dacă } x \geq 1 \end{cases}$$

Densitatea de repartiţie corespunzătoare este

$$f_n(x) = \begin{cases} \frac{1}{2n}, & \text{dacă } -n < x < n \\ 0, & \text{ın rest} \end{cases}$$

Sirul funcţiilor caracteristice este dat de $\varphi_n(t) = \frac{1}{2n}\int_{-n}^{n} \mathrm{e}^{itx}dx = \frac{\sin nt}{nt}$

$$\lim_{n\to\infty} \varphi_n(t) = \varphi(t) = \begin{cases} 1, & \text{dacă } t = 0 \\ 0, & \text{în rest} \end{cases}$$

Se vede că funcţia limită nu este continuă în $t = 0$.

Corespunzător acestui fapt avem pentru orice x fixat $\lim\limits_{n\to\infty} F_n(x) = \dfrac{1}{2}$, adică limita şirului $(F_n(x))_{n\in\mathbb{N}^*}$ nu este o funcţie de repartiţie. □

35. Să se arate că şirul funcţiilor de repartiţie $(F_n(x))_{n\in\mathbb{N}^*}$ corespunzător şirului de funcţii caracteristice $(\varphi_n)_{n\in\mathbb{N}^*}$ date prin relaţiile $\varphi_n(t) = \mathrm{e}^{int}, n \in \mathbb{N}^*$ nu converge către o funcţie de repartiţie.

Soluţie. Se observă că şirul $(\varphi_n)_{n\in\mathbb{N}^*}$ nu converge în afară de $t = 2k\pi$. Nefiind îndeplinite condiţiile din teorema lui Helly, rezultă că $F_n(x))_{n\in\mathbb{N}^*}$ nu converge către o funcţie de repartiţie.

Acest lucru se observă şi direct, şi anume $\varphi_n(t) = \mathrm{e}^{int}$ reprezintă funcţia caracteristică a v. a. X_n cu masa concentrată în punctul n.

Deci

$$F_n(x) = P(\{\omega/X_n(\omega) \le x\}) = \varepsilon(x-n) == \begin{cases} 0, & \text{dacă } x < n \\ 1, & \text{în rest} \end{cases}$$

Sub această formă se vede că $\lim\limits_{n\to\infty} F_n(x) = 0$ pentru orice x fixat. $\square$

4.3 Probleme propuse

1. Aplicând inegalitatea lui Cebîşev, să se găsească limita inferioară a probabilităţii inegalităţii $\left|\frac{\alpha}{10^5} - \frac{1}{6}\right| < \frac{1}{100}$, unde α reprezintă numărul de apariţii ale feţei 5 în 100000 aruncări de zar.

   R: $\frac{71}{72}$

2. Să se determine numărul n al probelor independente începând cu care are loc $P\left(\left|\frac{x}{n} - p\right| < 0,1\right) \ge 0,97$, dacă într-o singură probă evenimentul se realizează cu o probabilitate $p = 0,8$.

   R: Cf. teoremei lui Bernoulli $\Longrightarrow n \ge 534$

3. Fie $(X_n)_n$ un şir de v. a. independente ale căror valori şi probabilităţi sunt
   $X_k \sim \begin{pmatrix} -k & -(k-1) & \dots & -1 & 0 & 1 & \dots & k-1 & k \\ \frac{1}{3k^3} & \frac{1}{3(k-1)^3} & \dots & \frac{1}{3} & 1-\frac{2}{3}(1+\ldots+\frac{1}{k^3}) & \frac{1}{3} & \dots & \frac{1}{3(k-1)^3} & \frac{1}{3k^3} \end{pmatrix}$
   $k = 1,2,\ldots$ Să se arate că şirul dat se supune legii numerelor mari în formularea lui Cebîşev.

4. Fie $(X_n)_n$ un şir de v. a. independente care pot lua valorile $\pm\sqrt{n}$ şi 0 cu probabilităţile $P(X_1 = 0) = 1, P(X_k = \sqrt{k}) = P(X_k = -\sqrt{k}) = = \frac{1}{k}, P(X_k = 0) = 1 - \frac{2}{k}, k = 2,3,\ldots$ Să se arate că şirul dat se supune legii numerelor mari.

5. Fie $(X_n)_n$ un şir de v. a. independente de valori medii $E(X_n) = = 0, \forall n \ge 1$ şi dispersii $D^2(X_n) = n^\lambda, \forall n \ge 1$, unde $0 < \lambda < 1$. Să se arate că şirul dat se supune legii numerelor mari.

6. Fie $(X_n)_n$ un şir de v. a. independente astfel încât $P(X_k = -\sqrt{k}) = = P(X_k = \sqrt{k}) = \frac{1}{2}$. Să se arate că şirului i se poate aplica teorema lui Leapunov.

7. Fie $(X_n)_n$ un şir de v. a. cu densităţile de repartiţie

$$f_{X_k}(x) = \begin{cases} \frac{1}{ck^\alpha} \cdot \mathrm{e}^{-\frac{x}{ck^\alpha}}, & x > 0, c > 0, 0 < \alpha < \frac{1}{2} \\ 0, & x \le 0 \end{cases}$$

Notăm $Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k$, să se arate că şirul de v. a. $\left(Y_n - \frac{cn^\alpha}{\alpha+1}\right)_n$ converge în probabilitate la 0.

R: Calculăm dispersia variabilei $Y_n - \frac{cn^\alpha}{\alpha+1}$ şi apoi folosim inegalitatea lui Cebîşev

Chebyshev's inequality

In probability theory, **Chebyshev's inequality** (also called the **Bienaymé–Chebyshev inequality**) guarantees that, for a wide class of probability distributions, no more than a certain fraction of values can be more than a certain distance from the mean. Specifically, no more than $1/k^2$ of the distribution's values can be k or more standard deviations away from the mean (or equivalently, over $1 - 1/k^2$ of the distribution's values are less than k standard deviations away from the mean). The rule is often called Chebyshev's theorem, about the range of standard deviations around the mean, in statistics. The inequality has great utility because it can be applied to any probability distribution in which the mean and variance are defined. For example, it can be used to prove the weak law of large numbers.

Its practical usage is similar to the 68–95–99.7 rule, which applies only to normal distributions. Chebyshev's inequality is more general, stating that a minimum of just 75% of values must lie within two standard deviations of the mean and 88.89% within three standard deviations for a broad range of different probability distributions.[1][2]

The term *Chebyshev's inequality* may also refer to Markov's inequality, especially in the context of analysis. They are closely related, and some authors refer to Markov's inequality as "Chebyshev's First Inequality," and the similar one referred to on this page as "Chebyshev's Second Inequality."

Contents

History

The theorem is named after Russian mathematician Pafnuty Chebyshev, although it was first formulated by his friend and colleague Irénée-Jules Bienaymé.[3]:98 The theorem was first stated without proof by Bienaymé in 1853[4] and later proved by Chebyshev in 1867.[5] His student Andrey Markov provided another proof in his 1884 Ph.D. thesis.[6]

Statement

Chebyshev's inequality is usually stated for random variables, but can be generalized to a statement about measure spaces.

Probabilistic statement

Let X (integrable) be a random variable with finite expected value μ and finite non-zero variance σ^2. Then for any real number $k > 0$,

$$\Pr(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Only the case $k > 1$ is useful. When $k \leq 1$ the right-hand side $\frac{1}{k^2} \geq 1$ and the inequality is trivial as all probabilities are ≤ 1.

As an example, using $k = \sqrt{2}$ shows that the probability that values lie outside the interval $(\mu - \sqrt{2}\sigma, \mu + \sqrt{2}\sigma)$ does not exceed $\frac{1}{2}$.

Because it can be applied to completely arbitrary distributions provided they have a known finite mean and variance, the inequality generally gives a poor bound compared to what might be deduced if more aspects are known about the distribution involved.

k	Min. % within *k* standard deviations of mean	Max. % beyond *k* standard deviations from mean
1	0%	100%
√2	50%	50%
1.5	55.56%	44.44%
2	75%	25%
2√2	87.5%	12.5%
3	88.8889%	11.1111%
4	93.75%	6.25%
5	96%	4%
6	97.2222%	2.7778%
7	97.9592%	2.0408%
8	98.4375%	1.5625%
9	98.7654%	1.2346%
10	99%	1%

Measure-theoretic statement

Let (X, Σ, μ) be a measure space, and let f be an extended real-valued measurable function defined on X. Then for any real number $t > 0$ and $0 < p < \infty$,[7]

$$\mu(\{x \in X \,:\, |f(x)| \geq t\}) \leq \frac{1}{t^p} \int_{|f| \geq t} |f|^p \, d\mu.$$

More generally, if g is an extended real-valued measurable function, nonnegative and nondecreasing, with $g(t) \neq 0$ then:

$$\mu(\{x \in X \,:\, f(x) \geq t\}) \leq \frac{1}{g(t)} \int_X g \circ f \, d\mu.$$

The previous statement then follows by defining $g(x)$ as $|x|^p$ if $x \geq t$ and 0 otherwise.

Example

Suppose we randomly select a journal article from a source with an average of 1000 words per article, with a standard deviation of 200 words. We can then infer that the probability that it has between 600 and 1400 words (i.e. within $k = 2$ standard deviations of the mean) must be at least 75%, because there is no more than $1/k^2 = \frac{1}{4}$ chance to be outside that range, by Chebyshev's inequality. But if we additionally know that the distribution is normal, we can say there is a 75% chance the word count is between 770 and 1230 (which is an even tighter bound).

Sharpness of bounds

As shown in the example above, the theorem typically provides rather loose bounds. However, these bounds cannot in general (remaining true for arbitrary distributions) be improved upon. The bounds are sharp for the following example: for any $k \geq 1$,

$$X = \begin{cases} -1, & \text{with probability } \frac{1}{2k^2} \\ 0, & \text{with probability } 1 - \frac{1}{k^2} \\ 1, & \text{with probability } \frac{1}{2k^2} \end{cases}$$

For this distribution, the mean $\mu = 0$ and the standard deviation $\sigma = \frac{1}{k}$, so

$$\Pr(|X - \mu| \geq k\sigma) = \Pr(|X| \geq 1) = \frac{1}{k^2}.$$

Chebyshev's inequality is an equality for precisely those distributions that are a linear transformation of this example.

Proof (of the two-sided version)

Probabilistic proof

Markov's inequality states that for any real-valued random variable Y and any positive number a, we have $\Pr(|Y| > a) \leq \mathrm{E}(|Y|)/a$. One way to prove Chebyshev's inequality is to apply Markov's inequality to the random variable $Y = (X - \mu)^2$ with $a = (k\sigma)^2$.

It can also be proved directly using conditional expectation:

$$\begin{aligned} \sigma^2 &= \mathbb{E}[(X-\mu)^2] \\ &= \mathbb{E}[(X-\mu)^2 \mid k\sigma \leq |X-\mu|] \Pr[k\sigma \leq |X-\mu|] + \mathbb{E}[(X-\mu)^2 \mid k\sigma > |X-\mu|] \Pr[k\sigma > |X-\mu|] \\ &\geq (k\sigma)^2 \Pr[k\sigma \leq |X-\mu|] + 0 \cdot \Pr[k\sigma > |X-\mu|] \\ &= k^2\sigma^2 \Pr[k\sigma \leq |X-\mu|] \end{aligned}$$

Chebyshev's inequality then follows by dividing by $k^2\sigma^2$.

This proof also shows why the bounds are quite loose in typical cases: the conditional expectation on the event where $|X - \mu| < k\sigma$ is thrown away, and the lower bound of $k^2\sigma^2$ on the event $|X - \mu| \geq k\sigma$ can be quite poor.

Measure-theoretic proof

Fix t and let A_t be defined as $A_t = \{x \in X \mid f(x) \geq t\}$, and let 1_{A_t} be the indicator function of the set A_t. Then, it is easy to check that, for any x,

$$g(t)1_{A_t}(x) \le g(f(x))\,1_{A_t}(x),$$

since g is nondecreasing, and therefore,

$$\begin{aligned} g(t)\mu(A_t) &= \int_X g(t)1_{A_t}\,d\mu \\ &\le \int_{A_t} g\circ f\,d\mu \\ &\le \int_X g\circ f\,d\mu, \end{aligned}$$

where the last inequality is justified by the non-negativity of g. The desired inequality follows from dividing the above inequality by $g(t)$.

Proof assuming random variable *X* is continuous

Using the definition of probability density function $f(x)$, and a standard characterization of variance $\mathrm{Var}(X)$:

$$\Pr(a \le X \le b) = \int_a^b f_X(x)\,dx,$$

$$\mathrm{Var}(X) = \sigma^2 = \int_{\mathbb{R}} (x-\mu)^2 f(x)\,dx,$$

we have:

$$\begin{aligned} \Pr(|X-\mu| \ge k\sigma) &= \int_{|x-\mu|\ge k\sigma} f(x)\,dx \\ &\le \int_{|x-\mu|\ge k\sigma} \frac{|x-\mu|}{k\sigma} f(x)\,dx \\ &\le \int_{|x-\mu|\ge k\sigma} \frac{(x-\mu)^2}{k^2\sigma^2} f(x)\,dx \\ &= \int_{|x-\mu|\ge k\sigma} \frac{1}{k^2\sigma^2}(x-\mu)^2 f(x)\,dx \\ &= \frac{1}{k^2\sigma^2}\int_{|x-\mu|\ge k\sigma} (x-\mu)^2 f(x)\,dx \\ &\le \frac{1}{k^2\sigma^2}\int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx \\ &= \frac{1}{k^2\sigma^2}\sigma^2 \\ &= \frac{1}{k^2}. \end{aligned}$$

Replacing $k\sigma$ with ε, where $k = \varepsilon/\sigma$, we have another form of the Chebyshev's inequality:

$$\Pr(|X-\mu| \ge \varepsilon) \le \frac{\sigma^2}{\varepsilon^2},$$

or, the equivalent

$$\Pr(|X-\mu| < \varepsilon) > 1 - \frac{\sigma^2}{\varepsilon^2},$$

where $\boldsymbol{\varepsilon}$ is defined the same way as $\mathbf{k}$; any positive real number.

Extensions

Several extensions of Chebyshev's inequality have been developed.

Asymmetric two-sided

If X has mean μ and variance σ^2, then

$$\Pr(\ell < X < u) \geq \frac{4[(\mu-\ell)(u-\mu)-\sigma^2]}{(\ell-u)^2},$$ [8]

if $(\mu-\ell)(h-\mu) \geq \sigma^2$ and $(\mu-\ell)(h-\mu)-k^2 \leq 2\sigma^2$, where $k = \min(\mu-\ell, h-\mu)$ and $\ell < \mu < h$.[9]

This reduces to Chebyshev's inequality in the symmetric case (ℓ and u equidistant from the mean).

Bivariate generalization

Let X_1, X_2 be two random variables with means μ_1, μ_2 and finite variances σ_1, σ_2 respectively. Then a union bound shows that

$$\Pr\left(\ell_1 \leq \frac{X_1-\mu_1}{\sigma_1} \leq u_1, \ell_2 \leq \frac{X_2-\mu_2}{\sigma_2} \leq u_2\right) \geq 1 - \frac{4+(u_1+\ell_1)^2}{(u_1-\ell_1)^2} - \frac{4+(u_2+\ell_2)^2}{(u_2-\ell_2)^2}$$

This bound does not require X_1 and X_2 independent.[9]

Bivariate, known correlation

Berge derived an inequality for two correlated variables X_1, X_2.[10] Let ρ be the correlation coefficient between X_1 and X_2 and let σ_i^2 be the variance of X_i. Then

$$\Pr\left(\bigcap_{i=1}^{2}\left[\frac{|X_i-\mu_i|}{\sigma_i} < k\right]\right) \geq 1 - \frac{1+\sqrt{1-\rho^2}}{k^2}.$$

Lal later obtained an alternative bound[11]

$$\Pr\left(\bigcap_{i=1}^{2}\left[\frac{|X_i-\mu_i|}{\sigma_i} \leq k_i\right]\right) \geq 1 - \frac{k_1^2+k_2^2+\sqrt{(k_1^2+k_2^2)^2-4k_1^2k_2^2\rho}}{2(k_1k_2)^2}$$

Isii derived a further generalisation.[12] Let

$$Z = \Pr\left((-k_1 < X_1 < k_2) \cap (-k_1 < X_2 < k_2)\right), \qquad 0 < k_1 \leq k_2.$$

and define:

$$\lambda = \frac{k_1(1+\rho)+\sqrt{(1-\rho^2)(k_1^2+\rho)}}{2k_1-1+\rho}$$

There are now three cases.

- **Case A:** If $2k_1^2 > 1-\rho$ and $k_2 - k_1 \geq 2\lambda$ then

  $$Z \leq \frac{2\lambda^2}{2\lambda^2+1+\rho}.$$

- **Case B:** If the conditions in case A are not met but $k_1k_2 \geq 1$ and

  $$2(k_1k_2-1)^2 \geq 2(1-\rho^2)+(1-\rho)(k_2-k_1)^2$$

  then

  $$Z \leq \frac{(k_2-k_1)^2+4+\sqrt{16(1-\rho^2)+8(1-\rho)(k_2-k_1)}}{(k_1+k_2)^2}.$$

- **Case C:** If none of the conditions in cases A or B are satisfied then there is no universal bound other than 1.

Multivariate

The general case is known as the Birnbaum–Raymond–Zuckerman inequality after the authors who proved it for two dimensions.[13]

$$\Pr\left(\sum_{i=1}^{n}\frac{(X_i-\mu_i)^2}{\sigma_i^2 t_i^2}\geq k^2\right)\leq\frac{1}{k^2}\sum_{i=1}^{n}\frac{1}{t_i^2}$$

where X_i is the i-th random variable, μ_i is the i-th mean and σ_i^2 is the i-th variance.

If the variables are independent this inequality can be sharpened.[14]

$$\Pr\left(\bigcap_{i=1}^{n}\frac{|X_i-\mu_i|}{\sigma_i}\leq k_i\right)\geq\prod_{i=1}^{n}\left(1-\frac{1}{k_i^2}\right)$$

Olkin and Pratt derived an inequality for n correlated variables.[15]

$$\Pr\left(\bigcap_{i=1}^{n}\frac{|X_i-\mu_i|}{\sigma_i}<k_i\right)\geq 1-\frac{1}{n^2}\left(\sqrt{u}+\sqrt{n-1}\sqrt{n\sum_i\frac{1}{k_i^2}-u}\right)^2$$

where the sum is taken over the n variables and

$$u=\sum_{i=1}^{n}\frac{1}{k_i^2}+2\sum_{i=1}^{n}\sum_{j<i}\frac{\rho_{ij}}{k_i k_j}$$

where ρ_{ij} is the correlation between X_i and X_j.

Olkin and Pratt's inequality was subsequently generalised by Godwin.[16]

Finite-dimensional vector

Ferentinos[9] has shown that for a vector $X = (x_1, x_2, ...)$ with mean $\mu = (\mu_1, \mu_2, ...)$, standard deviation $\sigma = (\sigma_1, \sigma_2, ...)$ and the Euclidean norm $\|\cdot\|$ that

$$\Pr(\|X-\mu\|\geq k\|\sigma\|)\leq\frac{1}{k^2}.$$

A second related inequality has also been derived by Chen.[17] Let n be the dimension of the stochastic vector X and let $\mathrm{E}(X)$ be the mean of X. Let S be the covariance matrix and $k > 0$. Then

$$\Pr\left((X-\mathrm{E}(X))^T S^{-1}(X-\mathrm{E}(X))<k\right)\geq 1-\frac{n}{k}$$

where Y^{T} is the transpose of Y. A simple proof was obtained in Navarro[18] as follows:

$$Z=(X-\mathrm{E}(X))^T S^{-1}(X-\mathrm{E}(X))=(X-\mathrm{E}(X))^T S^{-1/2}S^{-1/2}(X-\mathrm{E}(X))=Y^T Y\geq 0$$

where

$$Y=(Y_1,\ldots,Y_n)^T=S^{-1/2}(X-\mathrm{E}(X))$$

and $S^{-1/2}$ is a symmetric invertible matrix such that: $S^{-1/2}S^{-1/2}=S^{-1}$. Hence $\mathrm{E}(Y)=(0,\ldots,0)^T$ and $\mathbf{Cov}(Y)=I_n$ where I_n represents the identity matrix of dimension n. Then $\mathrm{E}(Y_i^2)=\mathbf{Var}(Y_i)=1$ and

$$\mathrm{E}(Z)=\mathrm{E}(Y^T Y)=\sum_{i=1}^{n}\mathrm{E}(Y_i^2)=n$$

Finally, by applying Markov's inequality to Z we get

$$\Pr(Z\geq k)=\Pr\left((X-\mathrm{E}(X))^T S^{-1}(X-\mathrm{E}(X))\geq k\right)\leq\frac{\mathrm{E}(Z)}{k}=\frac{n}{k}$$

and so the desired inequality holds.

The inequality can be written in terms of the Mahalanobis distance as

$$\Pr\left(d_S^2(X,\mathrm{E}(X))<k\right)\geq 1-\frac{n}{k}$$

where the Mahalanobis distance based on S is defined by

$$d_S(x,y)=\sqrt{(x-y)^T S^{-1}(x-y)}$$

Navarro[19] proved that these bounds are sharp, that is, they are the best possible bounds for that regions when we just know the mean and the covariance matrix of X.

Stellato et al.[20] showed that this multivariate version of the Chebyshev inequality can be easily derived analytically as a special case of Vandenberghe et al.[21] where the bound is computed by solving a semidefinite program (SDP).

Infinite dimensions

There is a straightforward extension of the vector version of Chebyshev's inequality to infinite dimensional settings. Let X be a random variable which takes values in a Fréchet space $\mathcal{X}$ (equipped with seminorms $\|\cdot\|_\alpha$). This includes most common settings of vector-valued random variables, e.g., when $\mathcal{X}$ is a Banach space (equipped with a single norm), a Hilbert space, or the finite-dimensional setting as described above.

Suppose that X is of "strong order two", meaning that

$$\mathrm{E}\left(\|X\|_\alpha^2\right)<\infty$$

for every seminorm $\|\cdot\|_\alpha$. This is a generalization of the requirement that X have finite variance, and is necessary for this strong form of Chebyshev's inequality in infinite dimensions. The terminology "strong order two" is due to Vakhania.[22]

Let $\mu\in\mathcal{X}$ be the Pettis integral of X (i.e., the vector generalization of the mean), and let

$$\sigma_a:=\sqrt{\mathrm{E}\,\|X-\mu\|_\alpha^2}$$

be the standard deviation with respect to the seminorm $\|\cdot\|_\alpha$. In this setting we can state the following:

General version of Chebyshev's inequality. $\forall k>0: \quad \Pr\left(\|X-\mu\|_\alpha\geq k\sigma_\alpha\right)\leq\frac{1}{k^2}.$

Proof. The proof is straightforward, and essentially the same as the finitary version. If $\sigma_\alpha=0$, then X is constant (and equal to μ) almost surely, so the inequality is trivial.

If

$$\|X-\mu\|_\alpha\geq k\sigma_\alpha^2$$

then $\|X-\mu\|_\alpha>0$, so we may safely divide by $\|X-\mu\|_\alpha$. The crucial trick in Chebyshev's inequality is to recognize that $1=\frac{\|X-\mu\|_\alpha^2}{\|X-\mu\|_\alpha^2}$.

The following calculations complete the proof:

$$\begin{aligned}
\Pr\left(\|X-\mu\|_\alpha\geq k\sigma_\alpha\right) &= \int_\Omega \mathbf{1}_{\|X-\mu\|_\alpha\geq k\sigma_\alpha}\,\mathrm{d}\Pr \\
&= \int_\Omega\left(\frac{\|X-\mu\|_\alpha^2}{\|X-\mu\|_\alpha^2}\right)\cdot\mathbf{1}_{\|X-\mu\|_\alpha\geq k\sigma_\alpha}\,\mathrm{d}\Pr \\
&\leq \int_\Omega\left(\frac{\|X-\mu\|_\alpha^2}{(k\sigma_\alpha)^2}\right)\cdot\mathbf{1}_{\|X-\mu\|_\alpha\geq k\sigma_\alpha}\,\mathrm{d}\Pr \\
&\leq \frac{1}{k^2\sigma_\alpha^2}\int_\Omega\|X-\mu\|_\alpha^2\,\mathrm{d}\Pr && \mathbf{1}_{\|X-\mu\|_\alpha\geq k\sigma_\alpha}\leq 1 \\
&= \frac{1}{k^2\sigma_\alpha^2}\left(\mathrm{E}\,\|X-\mu\|_\alpha^2\right) \\
&= \frac{1}{k^2\sigma_\alpha^2}\left(\sigma_\alpha^2\right) \\
&= \frac{1}{k^2}
\end{aligned}$$

Higher moments

An extension to higher moments is also possible:

$$\Pr\left(|X-\mathrm{E}(X)|\geq k\,\mathrm{E}(|X-\mathrm{E}(X)|^n)^{\frac{1}{n}}\right)\leq\frac{1}{k^n},\qquad k>0, n\geq 2.$$

Exponential moment

A related inequality sometimes known as the exponential Chebyshev's inequality[23] is the inequality

$$\Pr(X\geq\varepsilon)\leq e^{-t\varepsilon}\,\mathrm{E}\left(e^{tX}\right),\qquad t>0.$$

Let $K(t)$ be the cumulant generating function,

$$K(t)=\log\left(\mathrm{E}\left(e^{tx}\right)\right).$$

Taking the Legendre–Fenchel transformation of $K(t)$ and using the exponential Chebyshev's inequality we have

$$-\log(\Pr(X\geq\varepsilon))\geq\sup_t(t\varepsilon-K(t)).$$

This inequality may be used to obtain exponential inequalities for unbounded variables.[24]

Bounded variables

If $P(x)$ has finite support based on the interval $[a, b]$, let $M = \max(|a|, |b|)$ where $|x|$ is the absolute value of x. If the mean of $P(x)$ is zero then for all $k > 0$[25]

$$\frac{\mathrm{E}(|X|^r)-k^r}{M^r}\leq\Pr(|X|\geq k)\leq\frac{\mathrm{E}(|X|^r)}{k^r}.$$

The second of these inequalities with $r = 2$ is the Chebyshev bound. The first provides a lower bound for the value of $P(x)$.

Sharp bounds for a bounded variate have been proposed by Niemitalo, but without a proof though[26]

Let $0 \leq X \leq M$ where $M > 0$. Then

- **Case 1:**

$$\Pr(X<k)=0\qquad\text{if}\qquad \mathrm{E}(X)>k\quad\text{and}\quad \mathrm{E}(X^2)<k\,\mathrm{E}(X)+M\,\mathrm{E}(X)-kM$$

- **Case 2:**

$$\Pr(X<k)\geq 1-\frac{k\,\mathrm{E}(X)+M\,\mathrm{E}(X)-\mathrm{E}(X^2)}{kM}\qquad\text{if}\qquad\begin{cases}\mathrm{E}(X)>k\quad\text{and}\quad \mathrm{E}(X^2)\geq k\,\mathrm{E}(X)+M\,\mathrm{E}(X)-kM\\ \text{or}\\ \mathrm{E}(X)\leq k\quad\text{and}\quad \mathrm{E}(X^2)\geq k\,\mathrm{E}(X)\end{cases}$$

- **Case 3:**

$$\Pr(X<k)\geq\frac{\mathrm{E}(X)^2-2k\,\mathrm{E}(X)+k^2}{\mathrm{E}(X^2)-2k\,\mathrm{E}(X)+k^2}\qquad\text{if}\qquad \mathrm{E}(X)\leq k\quad\text{and}\quad \mathrm{E}(X^2)<k\,\mathrm{E}(X)$$

Finite samples

Univariate case

Saw *et al* extended Chebyshev's inequality to cases where the population mean and variance are not known and may not exist, but the sample mean and sample standard deviation from N samples are to be employed to bound the expected value of a new drawing from the same distribution.[27]

$$P(|X-m|\geq ks)\leq\frac{g_{N+1}\left(\frac{Nk^2}{N-1+k^2}\right)}{N+1}\left(\frac{N}{N+1}\right)^{1/2}$$

where X is a random variable which we have sampled N times, m is the sample mean, k is a constant and s is the sample standard deviation. $g(x)$ is defined as follows:

Let $x \geq 1$, $Q = N + 1$, and R be the greatest integer less than Q/x. Let

$$a^2=\frac{Q(Q-R)}{1+R(Q-R)}.$$

Now

$$g_Q(x) = \begin{cases} R & \text{if } R \text{ is even,} \\ R & \text{if } R \text{ is odd and } x < a^2, \\ R-1 & \text{if } R \text{ is odd and } x \geq a^2. \end{cases}$$

This inequality holds even when the population moments do not exist, and when the sample is only weakly exchangeably distributed; this criterion is met for randomised sampling. A table of values for the Saw–Yang–Mo inequality for finite sample sizes (N < 100) has been determined by Konijn.[28] The table allows the calculation of various confidence intervals for the mean, based on multiples, C, of the standard error of the mean as calculated from the sample. For example, Konijn shows that for N = 59, the 95 percent confidence interval for the mean m is (m − Cs, m + Cs) where C = 4.447 × 1.006 = 4.47 (this is 2.28 times larger than the value found on the assumption of normality showing the loss on precision resulting from ignorance of the precise nature of the distribution).

Kabán gives a somewhat less complex version of this inequality.[29]

$$P(|X-m| \geq ks) \leq \frac{1}{N+1}\left\lfloor \frac{N+1}{N}\left(\frac{N-1}{k^2}+1\right)\right\rfloor$$

If the standard deviation is a multiple of the mean then a further inequality can be derived,[29]

$$P(|X-m| \geq ks) \leq \frac{N-1}{N}\frac{1}{k^2}\frac{s^2}{m^2}+\frac{1}{N}.$$

A table of values for the Saw–Yang–Mo inequality for finite sample sizes (N < 100) has been determined by Konijn.[28]

For fixed N and large m the Saw–Yang–Mo inequality is approximately[30]

$$P(|X-m| \geq ks) \leq \frac{1}{N+1}.$$

Beasley *et al* have suggested a modification of this inequality[30]

$$P(|X-m| \geq ks) \leq \frac{1}{k^2(N+1)}.$$

In empirical testing this modification is conservative but appears to have low statistical power. Its theoretical basis currently remains unexplored.

Dependence on sample size

The bounds these inequalities give on a finite sample are less tight than those the Chebyshev inequality gives for a distribution. To illustrate this let the sample size N = 100 and let k = 3. Chebyshev's inequality states that at most approximately 11.11% of the distribution will lie at least three standard deviations away from the mean. Kabán's version of the inequality for a finite sample states that at most approximately 12.05% of the sample lies outside these limits. The dependence of the confidence intervals on sample size is further illustrated below.

For N = 10, the 95% confidence interval is approximately ±13.5789 standard deviations.

For N = 100 the 95% confidence interval is approximately ±4.9595 standard deviations; the 99% confidence interval is approximately ±140.0 standard deviations.

For N = 500 the 95% confidence interval is approximately ±4.5574 standard deviations; the 99% confidence interval is approximately ±11.1620 standard deviations.

For N = 1000 the 95% and 99% confidence intervals are approximately ±4.5141 and approximately ±10.5330 standard deviations respectively.

The Chebyshev inequality for the distribution gives 95% and 99% confidence intervals of approximately ±4.472 standard deviations and ±10 standard deviations respectively.

Samuelson's inequality

Although Chebyshev's inequality is the best possible bound for an arbitrary distribution, this is not necessarily true for finite samples. Samuelson's inequality states that all values of a sample will lie within $\sqrt{N-1}$ standard deviations of the mean. Chebyshev's bound improves as the sample size increases.

When N = 10, Samuelson's inequality states that all members of the sample lie within 3 standard deviations of the mean: in contrast Chebyshev's states that 99.5% of the sample lies within 13.5789 standard deviations of the mean.

When N = 100, Samuelson's inequality states that all members of the sample lie within approximately 9.9499 standard deviations of the mean: Chebyshev's states that 99% of the sample lies within 10 standard deviations of the mean.

When N = 500, Samuelson's inequality states that all members of the sample lie within approximately 22.3383 standard deviations of the mean: Chebyshev's states that 99% of the sample lies within 10 standard deviations of the mean.

Multivariate case

Stellato et al.[20] simplified the notation and extended the empirical Chebyshev inequality from Saw et al.[27] to the multivariate case. Let $\xi \in \mathbb{R}^{n_\xi}$ be a random variable and let $N \in \mathbb{Z}_{\geq n_\xi}$. We draw $N+1$ iid samples of ξ denoted as $\xi^{(1)}, \dots, \xi^{(N)}, \xi^{(N+1)} \in \mathbb{R}^{n_\xi}$. Based on the first N samples, we define the empirical mean as $\mu_N = \frac{1}{N}\sum_{i=1}^N \xi^{(i)}$ and the unbiased empirical covariance as $\Sigma_N = \frac{1}{N}\sum_{i=1}^N (\xi^{(i)} - \mu_N)(\xi^{(i)} - \mu_N)^\top$. If Σ_N is nonsingular, then for all $\lambda \in \mathbb{R}_{\geq 0}$ then

$$P^{N+1}\left((\xi^{(N+1)} - \mu_N)^\top \Sigma_N^{-1}(\xi^{(N+1)} - \mu_N) \geq \lambda^2\right)$$

$$\leq \min\left\{1, \frac{1}{N+1}\left\lfloor \frac{n_\xi(N+1)(N^2-1+N\lambda^2)}{N^2\lambda^2} \right\rfloor\right\}.$$

Remarks

In the univariate case, i.e. $n_\xi = 1$, this inequality corresponds to the one from Saw et al.[27] Moreover, the right-hand side can be simplified by upper bounding the floor function by its argument

$$P^{N+1}\left((\xi^{(N+1)} - \mu_N)^\top \Sigma_N^{-1}(\xi^{(N+1)} - \mu_N) \geq \lambda^2\right) \leq \min\left\{1, \frac{n_\xi(N^2-1+N\lambda^2)}{N^2\lambda^2}\right\}.$$

As $N \to \infty$, the right-hand side tends to $\min\left\{1, \frac{n_\xi}{\lambda^2}\right\}$ which corresponds to the multivariate Chebyshev inequality over ellipsoids shaped according to Σ and centered in μ.

Sharpened bounds

Chebyshev's inequality is important because of its applicability to any distribution. As a result of its generality it may not (and usually does not) provide as sharp a bound as alternative methods that can be used if the distribution of the random variable is known. To improve the sharpness of the bounds provided by Chebyshev's inequality a number of methods have been developed; for a review see eg.[31]

Standardised variables

Sharpened bounds can be derived by first standardising the random variable.[32]

Let X be a random variable with finite variance Var(X). Let Z be the standardised form defined as

$$Z = \frac{X - \mathrm{E}(X)}{\mathrm{Var}(X)^{1/2}}.$$

Cantelli's lemma is then

$$P(Z \geq k) \leq \frac{1}{1+k^2}.$$

This inequality is sharp and is attained by k and $-1/k$ with probability $1/(1 + k^2)$ and $k^2/(1 + k^2)$ respectively.

If $k > 1$ and the distribution of X is symmetric then we have

$$P(Z \geq k) \leq \frac{1}{2k^2}.$$

Equality holds if and only if $Z = -k$, 0 or k with probabilities $1 / 2 k^2$, $1 - 1 / k^2$ and $1 / 2 k^2$ respectively.[32] An extension to a two-sided inequality is also possible.

Let u, $v > 0$. Then we have[32]

$$P(Z \leq -u \text{ or } Z \geq v) \leq \frac{4 + (u-v)^2}{(u+v)^2}.$$

Semivariances

An alternative method of obtaining sharper bounds is through the use of semivariances (partial variances). The upper (σ_+^2) and lower (σ_-^2) semivariances are defined as

$$\sigma_+^2 = \frac{\sum_{x>m}(x-m)^2}{n-1},$$

$$\sigma_-^2 = \frac{\sum_{x<m}(m-x)^2}{n-1},$$

where m is the arithmetic mean of the sample and n is the number of elements in the sample.

The variance of the sample is the sum of the two semivariances:

$$\sigma^2 = \sigma_+^2 + \sigma_-^2.$$

In terms of the lower semivariance Chebyshev's inequality can be written[33]

$$\Pr(x \le m - a\sigma_-) \le \frac{1}{a^2}.$$

Putting

$$a = \frac{k\sigma}{\sigma_-}.$$

Chebyshev's inequality can now be written

$$\Pr(x \le m - k\sigma) \le \frac{1}{k^2}\frac{\sigma_-^2}{\sigma^2}.$$

A similar result can also be derived for the upper semivariance.

If we put

$$\sigma_u^2 = \max(\sigma_-^2, \sigma_+^2),$$

Chebyshev's inequality can be written

$$\Pr(|x \le m - k\sigma|) \le \frac{1}{k^2}\frac{\sigma_u^2}{\sigma^2}.$$

Because $\sigma_u^2 \le \sigma^2$, use of the semivariance sharpens the original inequality.

If the distribution is known to be symmetric, then

$$\sigma_+^2 = \sigma_-^2 = \frac{1}{2}\sigma^2$$

and

$$\Pr(x \le m - k\sigma) \le \frac{1}{2k^2}.$$

This result agrees with that derived using standardised variables.

Note

The inequality with the lower semivariance has been found to be of use in estimating downside risk in finance and agriculture.[33][34][35]

Selberg's inequality

Selberg derived an inequality for $P(x)$ when $a \le x \le b$.[36] To simplify the notation let

$$Y = \alpha X + \beta$$

where

$$\alpha = \frac{2k}{b-a}$$

and

$$\beta = \frac{-(b+a)k}{b-a}.$$

The result of this linear transformation is to make $P(a \le X \le b)$ equal to $P(|Y| \le k)$.

The mean (μ_X) and variance (σ_X) of X are related to the mean (μ_Y) and variance (σ_Y) of Y:

$$\mu_Y = \alpha\mu_X + \beta$$

$$\sigma_Y^2 = \alpha^2\sigma_X^2.$$

With this notation Selberg's inequality states that

$$\Pr(|Y| < k) \ge \frac{(k-\mu_Y)^2}{(k-\mu_Y)^2 + \sigma_Y^2} \quad \text{if} \quad \sigma_Y^2 \le \mu_Y(k-\mu_Y)$$

$$\Pr(|Y| < k) \ge 1 - \frac{\sigma_Y^2 + \mu_Y^2}{k^2} \quad \text{if} \quad \mu_Y(k-\mu_Y) \le \sigma_Y^2 \le k^2 - \mu_Y^2$$

$$P(|Y| < k) \ge 0 \quad \text{if} \quad k^2 - \mu_Y^2 \le \sigma_Y^2.$$

These are known to be the best possible bounds.[37]

Cantelli's inequality

Cantelli's inequality[38] due to Francesco Paolo Cantelli states that for a real random variable (X) with mean (μ) and variance (σ^2)

$$P(X - \mu \ge a) \le \frac{\sigma^2}{\sigma^2 + a^2}$$

where $a \ge 0$.

This inequality can be used to prove a one tailed variant of Chebyshev's inequality with $k > 0$[39]

$$\Pr(X - \mu \ge k\sigma) \le \frac{1}{1+k^2}.$$

The bound on the one tailed variant is known to be sharp. To see this consider the random variable X that takes the values

$$X = 1 \text{ with probability } \frac{\sigma^2}{1+\sigma^2}$$
$$X = -\sigma^2 \text{ with probability } \frac{1}{1+\sigma^2}.$$

Then $E(X) = 0$ and $E(X^2) = \sigma^2$ and $P(X < 1) = 1 / (1 + \sigma^2)$.

An application - distance between the mean and the median

The one-sided variant can be used to prove the proposition that for probability distributions having an expected value and a median, the mean and the median can never differ from each other by more than one standard deviation. To express this in symbols let μ, ν, and σ be respectively the mean, the median, and the standard deviation. Then

$$|\mu - \nu| \le \sigma.$$

There is no need to assume that the variance is finite because this inequality is trivially true if the variance is infinite.

The proof is as follows. Setting $k = 1$ in the statement for the one-sided inequality gives:

$$\Pr(X - \mu \ge \sigma) \le \frac{1}{2} \implies \Pr(X \ge \mu + \sigma) \le \frac{1}{2}.$$

Changing the sign of X and of μ, we get

$$\Pr(X \le \mu - \sigma) \le \frac{1}{2}.$$

As the median is by definition any real number m that satisfies the inequalities

$$\mathrm{P}(X \leq m) \geq \frac{1}{2} \text{ and } \mathrm{P}(X \geq m) \geq \frac{1}{2}$$

this implies that the median lies within one standard deviation of the mean. A proof using Jensen's inequality also exists.

Bhattacharyya's inequality

Bhattacharyya[40] extended Cantelli's inequality using the third and fourth moments of the distribution.

Let $\mu = 0$ and σ^2 be the variance. Let $\gamma = \mathrm{E}(X^3)/\sigma^3$ and $\kappa = \mathrm{E}(X^4)/\sigma^4$.

If $k^2 - k\gamma - 1 > 0$ then

$$P(X > k\sigma) \leq \frac{\kappa - \gamma^2 - 1}{(\kappa - \gamma^2 - 1)(1 + k^2) + (k^2 - k\gamma - 1)}.$$

The necessity of $k^2 - k\gamma - 1 > 0$ requires that k be reasonably large.

Mitzenmacher and Upfal's inequality

Mitzenmacher and Upfal[41] note that

$$(X - \mathrm{E}[X])^{2k} > 0$$

for any integer $k > 0$ and that

$$\mathrm{E}[(X - \mathrm{E}(X))^{2k}]$$

is the $2k^{\text{th}}$ central moment. They then show that for $t > 0$

$$\Pr\left(|X - \mathrm{E}[X]| > t\,\mathrm{E}[(X - \mathrm{E}[X])^{2k}]^{1/2k}\right) \leq \frac{1}{t^{2k}}.$$

For $k = 1$ we obtain Chebyshev's inequality. For $t \geq 1$, $k > 2$ and assuming that the k^{th} moment exists, this bound is tighter than Chebyshev's inequality.

Related inequalities

Several other related inequalities are also known.

Zelen's inequality

Zelen has shown that[42]

$$\Pr(X - \mu \geq k\sigma) \leq \left[1 + k^2 + \frac{\left(k^2 - k\theta_3 - 1\right)^2}{\theta_4 - \theta_3^2 - 1}\right]^{-1}$$

with

$$k \geq \frac{\theta_3 + \sqrt{\theta_3^2 + 4}}{2}, \qquad \theta_m = \frac{M_m}{\sigma}$$

where M_m is the m-th moment and σ is the standard deviation.

He, Zhang and Zhang's inequality

For any collection of n non-negative independent random variables X_i with expectation 1 [43]

$$\Pr\left(\frac{\sum_{i=1}^{n} X_i}{n} - 1 \geq \frac{1}{n}\right) \leq \frac{7}{8}.$$

Hoeffding's lemma

Let X be a random variable with $a \leq X \leq b$ and $\mathrm{E}[X] = 0$, then for any $s > 0$, we have

$$E\left[e^{sX}\right] \leq e^{\frac{1}{8}s^2(b-a)^2}.$$

Van Zuijlen's bound

Let X_i be a set of independent Rademacher random variables: $\Pr(X_i = 1) = \Pr(X_i = -1) = 0.5$. Then[44]

$$\Pr\left(\left|\frac{\sum_{i=1}^n X_i}{\sqrt{n}}\right| \leq 1\right) \geq 0.5.$$

The bound is sharp and better than that which can be derived from the normal distribution (approximately $\Pr > 0.31$).

Unimodal distributions

A distribution function F is unimodal at ν if its cumulative distribution function is convex on $(-\infty, \nu)$ and concave on (ν,∞)[45] An empirical distribution can be tested for unimodality with the dip test.[46]

In 1823 Gauss showed that for a unimodal distribution with a mode of zero[47]

$$P(|X| \geq k) \leq \frac{4\,\mathrm{E}(X^2)}{9k^2} \quad \text{if} \quad k^2 \geq \frac{4}{3}\,\mathrm{E}(X^2),$$

$$P(|X| \geq k) \leq 1 - \frac{k}{\sqrt{3}\,\mathrm{E}(X^2)} \quad \text{if} \quad k^2 \leq \frac{4}{3}\,\mathrm{E}(X^2).$$

If the mode is not zero and the mean (μ) and standard deviation (σ) are both finite, then denoting the median as ν and the root mean square deviation from the mode by ω, we have

$$\sigma \leq \omega \leq 2\sigma$$

and

$$|\nu - \mu| \leq \sqrt{\frac{3}{4}}\omega.$$

Winkler in 1866 extended Gauss' inequality to r^{th} moments [48] where $r > 0$ and the distribution is unimodal with a mode of zero:

$$P(|X| \geq k) \leq \left(\frac{r}{r+1}\right)^r \frac{\mathrm{E}(|X|)^r}{k^r} \quad \text{if} \quad k^r \geq \frac{r^r}{(r+1)^{r+1}}\,\mathrm{E}(|X|^r),$$

$$P(|X| \geq k) \leq \left(1 - \left[\frac{k^r}{(r+1)\,\mathrm{E}(|X|)^r}\right]^{1/r}\right) \quad \text{if} \quad k^r \leq \frac{r^r}{(r+1)^{r+1}}\,\mathrm{E}(|X|^r).$$

Gauss' bound has been subsequently sharpened and extended to apply to departures from the mean rather than the mode due to the Vysochanskiï–Petunin inequality. The latter has been extended by Dharmadhikari and Joag-Dev[49]

$$P(|X| > k) \leq \max\left(\left[\frac{r}{(r+1)k}\right]^r E|X^r|, \frac{s}{(s-1)k^r} E|X^r| - \frac{1}{s-1}\right)$$

where s is a constant satisfying both $s > r + 1$ and $s(s - r - 1) = r^r$ and $r > 0$.

It can be shown that these inequalities are the best possible and that further sharpening of the bounds requires that additional restrictions be placed on the distributions.

Unimodal symmetrical distributions

The bounds on this inequality can also be sharpened if the distribution is both unimodal and symmetrical.[50] An empirical distribution can be tested for symmetry with a number of tests including McWilliam's R*.[51] It is known that the variance of a unimodal symmetrical distribution with finite support $[a, b]$ is less than or equal to $(b - a)^2 / 12$.[52]

Let the distribution be supported on the finite interval $[-N, N]$ and the variance be finite. Let the mode of the distribution be zero and rescale the variance to 1. Let $k > 0$ and assume $k < 2N/3$. Then[50]

$$P(X \geq k) \leq \frac{1}{2} - \frac{k}{2\sqrt{3}} \quad \text{if} \quad 0 \leq k \leq \frac{2}{\sqrt{3}},$$

$$P(X \geq k) \leq \frac{2}{9k^2} \quad \text{if} \quad \frac{2}{\sqrt{3}} \leq k \leq \frac{2N}{3}.$$

If $0 < k \leq 2 / \sqrt{3}$ the bounds are reached with the density[50]

$$f(x) = \frac{1}{2\sqrt{3}} \quad \text{if} \quad |x| < \sqrt{3}$$

$$f(x) = 0 \quad \text{if} \quad |x| \geq \sqrt{3}.$$

If $2 / \sqrt{3} < k \leq 2N / 3$ the bounds are attained by the distribution

$$(1 - \beta_k)\delta_0(x) + \beta_k f_k(x),$$

where $\beta_k = 4 / 3k^2$, δ_0 is the Dirac delta function and where

$$f_k(x) = \frac{1}{3k} \quad \text{if} \quad |x| < \frac{3k}{2},$$

$$f_k(x) = 0 \quad \text{if} \quad |x| \geq \frac{3k}{2}.$$

The existence of these densities shows that the bounds are optimal. Since N is arbitrary these bounds apply to any value of N.

The Camp–Meidell's inequality is a related inequality.[53] For an absolutely continuous unimodal and symmetrical distribution

$$P(|X - \mu| \geq k\sigma) \leq 1 - \frac{k}{\sqrt{3}} \quad \text{if} \quad k \leq \frac{2}{\sqrt{3}},$$

$$P(|X - \mu| \geq k\sigma) \leq \frac{4}{9k^2} \quad \text{if} \quad k > \frac{2}{\sqrt{3}}.$$

DasGupta has shown that if the distribution is known to be normal[54]

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{3k^2}.$$

Notes

Effects of symmetry and unimodality

Symmetry of the distribution decreases the inequality's bounds by a factor of 2 while unimodality sharpens the bounds by a factor of 4/9.

Because the mean and the mode in a unimodal distribution differ by at most $\sqrt{3}$ standard deviations[55] at most 5% of a symmetrical unimodal distribution lies outside $(2\sqrt{10} + 3\sqrt{3})/3$ standard deviations of the mean (approximately 3.840 standard deviations). This is sharper than the bounds provided by the Chebyshev inequality (approximately 4.472 standard deviations).

These bounds on the mean are less sharp than those that can be derived from symmetry of the distribution alone which shows that at most 5% of the distribution lies outside approximately 3.162 standard deviations of the mean. The Vysochanskiï–Petunin inequality further sharpens this bound by showing that for such a distribution that at most 5% of the distribution lies outside $4\sqrt{5}/3$ (approximately 2.981) standard deviations of the mean.

Symmetrical unimodal distributions

For any symmetrical unimodal distribution

- at most approximately 5.784% of the distribution lies outside 1.96 standard deviations of the mode
- at most 5% of the distribution lies outside $2\sqrt{10}/3$ (approximately 2.11) standard deviations of the mode

Normal distributions

DasGupta's inequality states that for a normal distribution at least 95% lies within approximately 2.582 standard deviations of the mean. This is less sharp than the true figure (approximately 1.96 standard deviations of the mean).

Bounds for specific distributions

- DasGupta has determined a set of best possible bounds for a normal distribution for this inequality.[54]

- Steliga and Szynal have extended these bounds to the Pareto distribution.[8]
- Grechuk et al. developed a general method for deriving the best possible bounds in Chebyshev's inequality for any family of distributions, and any deviation risk measure in place of standard deviation. In particular, they derived Chebyshev inequality for distributions with log-concave densities.[56]

Zero means

When the mean (μ) is zero Chebyshev's inequality takes a simple form. Let σ^2 be the variance. Then

$$P(|X| \geq 1) \leq \sigma^2.$$

With the same conditions Cantelli's inequality takes the form

$$P(X \geq 1) \leq \frac{\sigma^2}{1+\sigma^2}.$$

Unit variance

If in addition E(X^2) = 1 and E(X^4) = ψ then for any $0 \leq \varepsilon \leq 1$[57]

$$\Pr(|X| > \varepsilon) \geq \frac{(1-\epsilon^2)^2}{\psi - 1 + (1-\varepsilon^2)^2} \geq \frac{(1-\varepsilon^2)^2}{\psi}.$$

The first inequality is sharp. This is known as the Paley–Zygmund inequality.

It is also known that for a random variable obeying the above conditions that[58]

$$P(X \geq \varepsilon) \geq \frac{C_0}{\psi} - \frac{C_1}{\sqrt{\psi}}\varepsilon + \frac{C_2}{\psi\sqrt{\psi}}\varepsilon$$

where

$$C_0 = 2\sqrt{3} - 3 \quad (\approxeq 0.464),$$

$$C_1 = 1.397,$$

$$C_2 = 0.0231.$$

It is also known that[58]

$$\Pr(X > 0) \geq \frac{C_0}{\psi}.$$

The value of C_0 is optimal and the bounds are sharp if

$$\psi \geq \frac{3}{\sqrt{3}+1} \quad (\approxeq 1.098).$$

If

$$\psi \leq \frac{3}{\sqrt{3}+1}$$

then the sharp bound is

$$P(X > 0) \geq \frac{2}{3 + \psi + \sqrt{(1+\psi)^2 - 4}}.$$

Integral Chebyshev inequality

There is a second (less well known) inequality also named after Chebyshev[59]

If $f, g : [a, b] \to \mathbf{R}$ are two monotonic functions of the same monotonicity, then

$$\frac{1}{b-a}\int_a^b f(x)g(x)\,dx \geq \left[\frac{1}{b-a}\int_a^b f(x)\,dx\right]\left[\frac{1}{b-a}\int_a^b g(x)\,dx\right].$$

If f and g are of opposite monotonicity, then the above inequality works in the reverse way.

This inequality is related to Jensen's inequality,[60] Kantorovich's inequality,[61] the Hermite–Hadamard inequality[61] and Walter's conjecture.[62]

Other inequalities

There are also a number of other inequalities associated with Chebyshev:

- Chebyshev's sum inequality
- Chebyshev–Markov–Stieltjes inequalities

Haldane's transformation

One use of Chebyshev's inequality in applications is to create confidence intervals for variates with an unknown distribution. Haldane noted,[63] using an equation derived by Kendall,[64] that if a variate (x) has a zero mean, unit variance and both finite skewness (γ) and kurtosis (κ) then the variate can be converted to a normally distributed standard score (z):

$$z = x - \frac{\gamma}{6}(x^2 - 1) + \frac{x}{72}[2\gamma^2(4x^2 - 7) - 3\kappa(x^2 - 3)] + \cdots$$

This transformation may be useful as an alternative to Chebyshev's inequality or as an adjunct to it for deriving confidence intervals for variates with unknown distributions.

While this transformation may be useful for moderately skewed and/or kurtotic distributions, it performs poorly when the distribution is markedly skewed and/or kurtotic.

Notes

The Environmental Protection Agency has suggested best practices for the use of Chebyshev's inequality for estimating confidence intervals.<ref>Calculating Upper Confidence Limits for Exposure Point Concentrations at hazardous Waste Sites (http://nepis.epa.gov/Exe/ZyNET.exe/P100CYCE.TXT?ZyActionD=ZyDocument&Client=EPA&Index=2000+Thru+2005&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C00thru05%5CTxt%5C00000029%5CP100CYCE.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=p%7Cf&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL#) (Report). Office of Emergency and Remedial Response of the U.S. Environmental Protection Agency. December 2002. Retrieved 5 August 2016.

See also

- Multidimensional Chebyshev's inequality
- Concentration inequality – a summary of tail-bounds on random variables.
- Cornish–Fisher expansion
- Eaton's inequality
- Kolmogorov's inequality
- Proof of the weak law of large numbers using Chebyshev's inequality
- Le Cam's theorem
- Paley–Zygmund inequality
- Vysochanskiï–Petunin inequality — a stronger result applicable to unimodal probability distributions

References

1. Kvanli, Alan H.; Pavur, Robert J.; Keeling, Kellie B. (2006). *Concise Managerial Statistics* (https://books.google.com/books?id=h6CQ1J0gwNgC&pg=PT95). cEngage Learning. pp. 81–82. ISBN 9780324223880.
2. Chernick, Michael R. (2011). *The Essentials of Biostatistics for Physicians, Nurses, and Clinicians* (https://books.google.com/books?id=JP4azqd8ONEC&pg=PA50). John Wiley & Sons. pp. 49–50. ISBN 9780470641859.
3. Knuth, Donald (1997). *The Art of Computer Programming: Fundamental Algorithms, Volume 1* (http://www-cs-faculty.stanford.edu/~uno/taocp.html) (3rd ed.). Reading, Massachusetts: Addison–Wesley. ISBN 978-0-201-89683-1. Retrieved 1 October 2012.
4. Bienaymé, I.-J. (1853). "Considérations àl'appui de la découverte de Laplace". *Comptes Rendus de l'Académie des Sciences*. **37**: 309–324.
5. Tchebichef, P. (1867). "Des valeurs moyennes". *Journal de Mathématiques Pures et Appliquées*. 2. **12**: 177–184.
6. Markov A. (1884) On certain applications of algebraic continued fractions, Ph.D. thesis, St. Petersburg
7. Grafakos, Lukas (2004). *Classical and Modern Fourier Analysis*. Pearson Education Inc. p. 5.
8. Steliga, Katarzyna; Szynal, Dominik (2010). "On Markov-Type Inequalities" (http://ijpam.eu/contents/2010-58-2/2/2.pdf) (PDF). *International Journal of Pure and Applied Mathematics*. **58** (2): 137–152. ISSN 1311-8080 (https://www.worldcat.org/issn/1311-8080). Retrieved 10 October 2012.

9. Ferentinos, K (1982). "On Tchebycheff type inequalities". *Trabajos Estadıst Investigacion Oper*. **33**: 125–132. doi:10.1007/BF02888707 (https://doi.org/10.1007%2FBF02888707). S2CID 123762564 (https://api.semanticscholar.org/CorpusID:123762564).
10. Berge, P. O. (1938). "A note on a form of Tchebycheff's theorem for two variables". *Biometrika*. **29** (3/4): 405–406. doi:10.2307/2332015 (https://doi.org/10.2307%2F2332015). JSTOR 2332015 (https://www.jstor.org/stable/2332015).
11. Lal D. N. (1955) A note on a form of Tchebycheff's inequality for two or more variables. Sankhya 15(3):317–320
12. Isii K. (1959) On a method for generalizations of Tchebycheff's inequality. Ann Inst Stat Math 10: 65–88
13. Birnbaum, Z. W.; Raymond, J.; Zuckerman, H. S. (1947). "A Generalization of Tshebyshev's Inequality to Two Dimensions" (http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.aoms/1177730493). *The Annals of Mathematical Statistics*. **18** (1): 70–79. doi:10.1214/aoms/1177730493 (https://doi.org/10.1214%2Faoms%2F1177730493). ISSN 0003-4851 (https://www.worldcat.org/issn/0003-4851). MR 0019849 (https://www.ams.org/mathscinet-getitem?mr=0019849). Zbl 0032.03402 (https://zbmath.org/?format=complete&q=an:0032.03402). Retrieved 7 October 2012.
14. Kotz, Samuel; Balakrishnan, N.; Johnson, Norman L. (2000). *Continuous Multivariate Distributions, Volume 1, Models and Applications* (http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471183873.html) (2nd ed.). Boston [u.a.]: Houghton Mifflin. ISBN 978-0-471-18387-7. Retrieved 7 October 2012.
15. Olkin, Ingram; Pratt, John W. (1958). "A Multivariate Tchebycheff Inequality" (https://doi.org/10.1214%2Faoms%2F1177706720). *The Annals of Mathematical Statistics*. **29** (1): 226–234. doi:10.1214/aoms/1177706720 (https://doi.org/10.1214%2Faoms%2F1177706720). MR 0093865 (https://www.ams.org/mathscinet-getitem?mr=0093865). Zbl 0085.35204 (https://zbmath.org/?format=complete&q=an:0085.35204).
16. Godwin H. J. (1964) Inequalities on distribution functions. New York, Hafner Pub. Co.
17. Xinjia Chen (2007). "A New Generalization of Chebyshev Inequality for Random Vectors". arXiv:0707.0805v2 (https://arxiv.org/abs/0707.0805v2) [math.ST (https://arxiv.org/archive/math.ST)].
18. Jorge Navarro (2016). "A very simple proof of the multivariate Chebyshev's inequality". *Communications in Statistics – Theory and Methods*. **45** (12): 3458–3463. doi:10.1080/03610926.2013.873135 (https://doi.org/10.1080%2F03610926.2013.873135). S2CID 121107480 (https://api.semanticscholar.org/CorpusID:121107480).
19. Jorge Navarro (2014). "Can the bounds in the multivariate Chebyshev inequality be attained?". *Statistics and Probability Letters*. **91**: 1–5. doi:10.1016/j.spl.2014.03.028 (https://doi.org/10.1016%2Fj.spl.2014.03.028).
20. Stellato, Bartolomeo; Parys, Bart P. G. Van; Goulart, Paul J. (2016-05-31). "Multivariate Chebyshev Inequality with Estimated Mean and Variance". *The American Statistician*. **71** (2): 123–127. arXiv:1509.08398 (https://arxiv.org/abs/1509.08398). doi:10.1080/00031305.2016.1186559 (https://doi.org/10.1080%2F00031305.2016.1186559). ISSN 0003-1305 (https://www.worldcat.org/issn/0003-1305). S2CID 53407286 (https://api.semanticscholar.org/CorpusID:53407286).
21. Vandenberghe, L.; Boyd, S.; Comanor, K. (2007-01-01). "Generalized Chebyshev Bounds via Semidefinite Programming". *SIAM Review*. **49** (1): 52–64. Bibcode:2007SIAMR..49...52V (https://ui.adsabs.harvard.edu/abs/2007SIAMR..49...52V). CiteSeerX 10.1.1.126.9105 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.126.9105). doi:10.1137/S0036144504440543 (https://doi.org/10.1137%2FS0036144504440543). ISSN 0036-1445 (https://www.worldcat.org/issn/0036-1445).
22. Vakhania, Nikolai Nikolaevich. Probability distributions on linear spaces. New York: North Holland, 1981.
23. Section 2.1 (http://www.math.utah.edu/~firas/Papers/rassoul-seppalainen-ldp.pdf) Archived (https://web.archive.org/web/20150430075226/http://www.math.utah.edu/~firas/Papers/rassoul-seppalainen-ldp.pdf) April 30, 2015, at the Wayback Machine
24. Baranoski, Gladimir V. G.; Rokne, Jon G.; Xu, Guangwu (15 May 2001). "Applying the exponential Chebyshev inequality to the nondeterministic computation of form factors". *Journal of Quantitative Spectroscopy and Radiative Transfer*. **69** (4): 199–200. Bibcode:2001JQSRT..69..447B (https://ui.adsabs.harvard.edu/abs/2001JQSRT..69..447B). doi:10.1016/S0022-4073(00)00095-9 (https://doi.org/10.1016%2FS0022-4073%2800%2900095-9). (the references for this article are corrected by Baranoski, Gladimir V. G.; Rokne, Jon G.; Guangwu Xu (15 January 2002). "Corrigendum to: 'Applying the exponential Chebyshev inequality to the nondeterministic computation of form factors' ". *Journal of Quantitative Spectroscopy and Radiative Transfer*. **72** (2): 199–200. Bibcode:2002JQSRT..72..199B (https://ui.adsabs.harvard.edu/abs/2002JQSRT..72..199B). doi:10.1016/S0022-4073(01)00171-6 (https://doi.org/10.1016%2FS0022-4073%2801%2900171-6).)
25. Dufour (2003) Properties of moments of random variables (http://www2.cirano.qc.ca/~dufourj/Web_Site/ResE/Dufour_1999_C_TS_Moments.pdf)
26. Niemitalo O. (2012) One-sided Chebyshev-type inequalities for bounded probability distributions. (http://yehar.com/blog/?p=1225)
27. Saw, John G.; Yang, Mark C. K.; Mo, Tse Chin (1984). "Chebyshev Inequality with Estimated Mean and Variance". *The American Statistician*. **38** (2): 130–2. doi:10.2307/2683249 (https://doi.org/10.2307%2F2683249). ISSN 0003-1305 (https://www.worldcat.org/issn/0003-1305). JSTOR 2683249 (https://www.jstor.org/stable/2683249).
28. Konijn, Hendrik S. (February 1987). "Distribution-Free and Other Prediction Intervals". *The American Statistician*. **41** (1): 11–15. doi:10.2307/2684311 (https://doi.org/10.2307%2F2684311). JSTOR 2684311 (https://www.jstor.org/stable/2684311).
29. Kabán, Ata (2012). "Non-parametric detection of meaningless distances in high dimensional data". *Statistics and Computing*. **22** (2): 375–85. doi:10.1007/s11222-011-9229-0 (https://doi.org/10.1007%2Fs11222-011-9229-0). S2CID 6018114 (https://api.semanticscholar.org/CorpusID:6018114).
30. Beasley, T. Mark; Page, Grier P.; Brand, Jaap P. L.; Gadbury, Gary L.; Mountz, John D.; Allison, David B. (January 2004). "Chebyshev's inequality for nonparametric testing with small *N* and α in microarray research". *Journal of the Royal Statistical Society*. C (Applied Statistics). **53** (1): 95–108. doi:10.1111/j.1467-9876.2004.00428.x (https://doi.org/10.1111%2Fj.1467-9876.2004.00428.x). ISSN 1467-9876 (https://www.worldcat.org/issn/1467-9876).
31. Savage, I. Richard. "Probability inequalities of the Tchebycheff type." Journal of Research of the National Bureau of Standards-B. Mathematics and Mathematical Physics B 65 (1961): 211-222 (http://nvlpubs.nist.gov/nistpubs/jres/65B/jresv65Bn3p211_A1b.pdf)
32. Ion, Roxana Alice (2001). "Chapter 4: Sharp Chebyshev-type inequalities" (http://dare.uva.nl/document/60326). *Nonparametric Statistical Process Control*. Universiteit van Amsterdam. ISBN 978-9057760761. Retrieved 1 October 2012.
33. Berck, Peter; Hihn, Jairus M. (May 1982). "Using the Semivariance to Estimate Safety-First Rules". *American Journal of Agricultural Economics*. **64** (2): 298–300. doi:10.2307/1241139 (https://doi.org/10.2307%2F1241139). ISSN 0002-9092 (https://www.worldcat.org/issn/0002-9092). JSTOR 1241139 (https://www.jstor.org/stable/1241139).

34. Nantell, Timothy J.; Price, Barbara (June 1979). "An Analytical Comparison of Variance and Semivariance Capital Market Theories". *The Journal of Financial and Quantitative Analysis*. **14** (2): 221–42. doi:10.2307/2330500 (https://doi.org/10.2307%2F2330500). JSTOR 2330500 (https://www.jstor.org/stable/2330500).
35. Neave, Edwin H.; Ross, Michael N.; Yang, Jun (2009). "Distinguishing upside potential from downside risk". *Management Research News*. **32** (1): 26–36. doi:10.1108/01409170910922005 (https://doi.org/10.1108%2F01409170910922005). ISSN 0140-9174 (https://www.worldcat.org/issn/0140-9174).
36. Selberg, Henrik L. (1940). "Zwei Ungleichungen zur Ergänzung des Tchebycheffschen Lemmas" [Two Inequalities Supplementing the Tchebycheff Lemma]. *Skandinavisk Aktuarietidskrift (Scandinavian Actuarial Journal)* (in German). **1940** (3–4): 121–125. doi:10.1080/03461238.1940.10404804 (https://doi.org/10.1080%2F03461238.1940.10404804). ISSN 0346-1238 (https://www.worldcat.org/issn/0346-1238). OCLC 610399869 (https://www.worldcat.org/oclc/610399869).
37. Conlon, J.; Dulá, J. H. "A geometric derivation and interpretation of Tchebyscheff's Inequality" (http://www.people.vcu.edu/~jdula/WORKINGPAPERS/tcheby.pdf) (PDF). Retrieved 2 October 2012.
38. Cantelli F. (1910) Intorno ad un teorema fondamentale della teoria del rischio. Bolletino dell Associazione degli Attuari Italiani
39. Grimmett and Stirzaker, problem 7.11.9. Several proofs of this result can be found in Chebyshev's Inequalities (http://www.mcdowella.demon.co.uk/Chebyshev.html) by A. G. McDowell.
40. Bhattacharyya, B. B. (1987). "One-sided chebyshev inequality when the first four moments are known". *Communications in Statistics – Theory and Methods*. **16** (9): 2789–91. doi:10.1080/03610928708829540 (https://doi.org/10.1080%2F03610928708829540). ISSN 0361-0926 (https://www.worldcat.org/issn/0361-0926).
41. Mitzenmacher, Michael; Upfal, Eli (January 2005). *Probability and Computing: Randomized Algorithms and Probabilistic Analysis* (http://www.cambridge.org/us/knowledge/isbn/item1171566/?site_locale=en_US) (Repr. ed.). Cambridge [u.a.]: Cambridge Univ. Press. ISBN 9780521835404. Retrieved 6 October 2012.
42. Zelen M. (1954) Bounds on a distribution function that are functions of moments to order four. J Res Nat Bur Stand 53:377–381
43. He, S.; Zhang, J.; Zhang, S. (2010). "Bounding probability of small deviation: A fourth moment approach". *Mathematics of Operations Research*. **35** (1): 208–232. doi:10.1287/moor.1090.0438 (https://doi.org/10.1287%2Fmoor.1090.0438). S2CID 11298475 (https://api.semanticscholar.org/CorpusID:11298475).
44. Martien C. A. van Zuijlen (2011) On a conjecture concerning the sum of independent Rademacher random variables (https://arxiv.org/abs/1112.4988)
45. Feller, William (1966). *An Introduction to Probability Theory and Its Applications, Volume 2* (https://books.google.com/books?id=LhrvAAAAMAAJ&q=%22An+introduction+to+probability+theory+and+its+applications%22+%22volume+2%22+feller) (2 ed.). Wiley. p. 155. ISBN 9789994311071. Retrieved 6 October 2012.
46. Hartigan, J. A.; Hartigan, P. M. (1985). "The Dip Test of Unimodality" (https://doi.org/10.1214%2Faos%2F1176346577). *The Annals of Statistics*. **13**: 70–84. doi:10.1214/aos/1176346577 (https://doi.org/10.1214%2Faos%2F1176346577). MR 0773153 (https://www.ams.org/mathscinet-getitem?mr=0773153).
47. Gauss C. F. Theoria Combinationis Observationum Erroribus Minimis Obnoxiae. Pars Prior. Pars Posterior. Supplementum. Theory of the Combination of Observations Least Subject to Errors. Part One. Part Two. Supplement. 1995. Translated by G. W. Stewart. Classics in Applied Mathematics Series, Society for Industrial and Applied Mathematics, Philadelphia
48. Winkler A. (1886) Math-Natur theorie Kl. Akad. Wiss Wien Zweite Abt 53, 6–41
49. Dharmadhikari, S. W.; Joag-Dev, K. (1985). "The Gauss–Tchebyshev inequality for unimodal distributions" (http://www.mathnet.ru/links/1043b08ee307d37ab64be90bd9332b88/tvp1919.pdf) (PDF). *Teoriya Veroyatnostei i ee Primeneniya*. **30** (4): 817–820.
50. Clarkson, Eric; Denny, J. L.; Shepp, Larry (2009). "ROC and the bounds on tail probabilities via theorems of Dubins and F. Riesz" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2828638). *The Annals of Applied Probability*. **19** (1): 467–76. arXiv:0903.0518 (https://arxiv.org/abs/0903.0518). Bibcode:2009arXiv0903.0518C (https://ui.adsabs.harvard.edu/abs/2009arXiv0903.0518C). doi:10.1214/08-AAP536 (https://doi.org/10.1214%2F08-AAP536). PMC 2828638 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2828638). PMID 20191100 (https://pubmed.ncbi.nlm.nih.gov/20191100).
51. McWilliams, Thomas P. (1990). "A Distribution-Free Test for Symmetry Based on a Runs Statistic". *Journal of the American Statistical Association*. **85** (412): 1130–3. doi:10.2307/2289611 (https://doi.org/10.2307%2F2289611). ISSN 0162-1459 (https://www.worldcat.org/issn/0162-1459). JSTOR 2289611 (https://www.jstor.org/stable/2289611).
52. Seaman, John W., Jr.; Young, Dean M.; Odell, Patrick L. (1987). "Improving small sample variance estimators for bounded random variables". *Industrial Mathematics*. **37**: 65–75. ISSN 0019-8528 (https://www.worldcat.org/issn/0019-8528). Zbl 0637.62024 (https://zbmath.org/?format=complete&q=an:0637.62024).
53. Bickel, Peter J.; Krieger, Abba M. (1992). "Extensions of Chebyshev's Inequality with Applications" (http://www.math.uni.wroc.pl/~pms/files/13.2/Article/13.2.11.pdf) (PDF). *Probability and Mathematical Statistics*. **13** (2): 293–310. ISSN 0208-4147 (https://www.worldcat.org/issn/0208-4147). Retrieved 6 October 2012.
54. DasGupta, A (2000). "Best constants in Chebychev inequalities with various applications". *Metrika*. **5** (1): 185–200. doi:10.1007/s184-000-8316-9 (https://doi.org/10.1007%2Fs184-000-8316-9). S2CID 121436601 (https://api.semanticscholar.org/CorpusID:121436601).
55. "More thoughts on a one tailed version of Chebyshev's inequality – by Henry Bottomley" (http://www.se16.info/~se16/hgb/cheb2.htm#3unimodalinequalities). se16.info. Retrieved 2012-06-12.
56. Grechuk, B., Molyboha, A., Zabarankin, M. (2010). Chebyshev Inequalities with Law Invariant Deviation Measures (https://www.researchgate.net/publication/231939730_Chebyshev_inequalities_with_law-invariant_deviation_measures), Probability in the Engineering and Informational Sciences, 24(1), 145-170.
57. Godwin H. J. (1964) Inequalities on distribution functions. (Chapter 3) New York, Hafner Pub. Co.
58. Lesley F. D., Rotar V. I. (2003) Some remarks on lower bounds of Chebyshev's type for half-lines. J Inequalities Pure Appl Math 4(5) Art 96
59. Fink, A. M.; Jodeit, Max, Jr. (1984). "On Chebyshev's other inequality" (http://projecteuclid.org/euclid.lnms/1215465617). In Tong, Y. L.; Gupta, Shanti S. (eds.). *Inequalities in Statistics and Probability*. Institute of Mathematical Statistics Lecture Notes - Monograph Series. **5**. pp. 115–120. doi:10.1214/lnms/1215465637 (https://doi.org/10.1214%2Flnms%2F1215465637). ISBN 978-0-940600-04-1. MR 0789242 (https://www.ams.org/mathscinet-getitem?mr=0789242). Retrieved 7 October 2012.

60. Niculescu, Constantin P. (2001). "An extension of Chebyshev's inequality and its connection with Jensen's inequality" (http://emis.matem.unam.mx/journals/HOA/JIA/Volume6_4/462.html). *Journal of Inequalities and Applications*. **6** (4): 451–462. CiteSeerX 10.1.1.612.7056 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.612.7056). doi:10.1155/S1025583401000273 (https://doi.org/10.1155%2FS1025583401000273). ISSN 1025-5834 (https://www.worldcat.org/issn/1025-5834). Retrieved 6 October 2012.
61. Niculescu, Constantin P.; Pečarić, Josip (2010). "The Equivalence of Chebyshev's Inequality to the Hermite–Hadamard Inequality" (http://www.csm.ro/reviste/Mathematical_Reports/Pdfs/2010/2/Niculescu.pdf) (PDF). *Mathematical Reports*. **12** (62): 145–156. ISSN 1582-3067 (https://www.worldcat.org/issn/1582-3067). Retrieved 6 October 2012.
62. Malamud, S. M. (15 February 2001). "Some complements to the Jensen and Chebyshev inequalities and a problem of W. Walter" (https://www.ams.org/journals/proc/2001-129-09/S0002-9939-01-05849-X/). *Proceedings of the American Mathematical Society*. **129** (9): 2671–2678. doi:10.1090/S0002-9939-01-05849-X (https://doi.org/10.1090%2FS0002-9939-01-05849-X). ISSN 0002-9939 (https://www.worldcat.org/issn/0002-9939). MR 1838791 (https://www.ams.org/mathscinet-getitem?mr=1838791). Retrieved 7 October 2012.
63. Haldane, J. B. (1952). "Simple tests for bimodality and bitangentiality". *Annals of Eugenics*. **16** (4): 359–364. doi:10.1111/j.1469-1809.1951.tb02488.x (https://doi.org/10.1111%2Fj.1469-1809.1951.tb02488.x). PMID 14953132 (https://pubmed.ncbi.nlm.nih.gov/14953132).
64. Kendall M. G. (1943) The Advanced Theory of Statistics, 1. London

Further reading

- A. Papoulis (1991), *Probability, Random Variables, and Stochastic Processes*, 3rd ed. McGraw–Hill. ISBN 0-07-100870-5. pp. 113–114.
- G. Grimmett and D. Stirzaker (2001), *Probability and Random Processes*, 3rd ed. Oxford. ISBN 0-19-857222-0. Section 7.3.

External links

- "Chebyshev inequality in probability theory" (https://www.encyclopediaofmath.org/index.php?title=Chebyshev_inequality_in_probability_theory), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]
- Formal proof (https://web.archive.org/web/20131204193123/http://mws.cs.ru.nl/mwiki/random_2.html#T7) in the Mizar system.

WIKIPEDIA

Convergence of random variables

In probability theory, there exist several different notions of **convergence of random variables**. The convergence of sequences of random variables to some limit random variable is an important concept in probability theory, and its applications to statistics and stochastic processes. The same concepts are known in more general mathematics as **stochastic convergence** and they formalize the idea that a sequence of essentially random or unpredictable events can sometimes be expected to settle down into a behavior that is essentially unchanging when items far enough into the sequence are studied. The different possible notions of convergence relate to how such a behavior can be characterized: two readily understood behaviors are that the sequence eventually takes a constant value, and that values in the sequence continue to change but can be described by an unchanging probability distribution.

Contents

Background

"Stochastic convergence" formalizes the idea that a sequence of essentially random or unpredictable events can sometimes be expected to settle into a pattern. The pattern may for instance be

- Convergence in the classical sense to a fixed value, perhaps itself coming from a random event
- An increasing similarity of outcomes to what a purely deterministic function would produce
- An increasing preference towards a certain outcome
- An increasing "aversion" against straying far away from a certain outcome
- That the probability distribution describing the next outcome may grow increasingly similar to a certain distribution

Some less obvious, more theoretical patterns could be

- That the series formed by calculating the expected value of the outcome's distance from a particular value may converge to 0
- That the variance of the random variable describing the next event grows smaller and smaller.

These other types of patterns that may arise are reflected in the different types of stochastic convergence that have been studied.

While the above discussion has related to the convergence of a single series to a limiting value, the notion of the convergence of two series towards each other is also important, but this is easily handled by studying the sequence defined as either the difference or the ratio of the two series.

For example, if the average of n independent random variables Y_i, $i = 1, ..., n$, all having the same finite mean and variance, is given by

$$X_n = \frac{1}{n}\sum_{i=1}^{n} Y_i\,,$$

then as n tends to infinity, X_n converges *in probability* (see below) to the common mean, μ, of the random variables Y_i. This result is known as the weak law of large numbers. Other forms of convergence are important in other useful theorems, including the central limit theorem.

Throughout the following, we assume that (X_n) is a sequence of random variables, and X is a random variable, and all of them are defined on the same probability space $(\Omega, \mathcal{F}, \Pr)$.

Convergence in distribution

With this mode of convergence, we increasingly expect to see the next outcome in a sequence of random experiments becoming better and better modeled by a given probability distribution.

Examples of convergence in distribution

Dice factory

Convergence in distribution is the weakest form of convergence typically discussed, since it is implied by all other types of convergence mentioned in this article. However, convergence in distribution is very frequently used in practice; most often it arises from application of the central limit theorem.

Suppose a new dice factory has just been built. The first few dice come out quite biased, due to imperfections in the production process. The outcome from tossing any of them will follow a distribution markedly different from the desired uniform distribution.

As the factory is improved, the dice become less and less loaded, and the outcomes from tossing a newly produced die will follow the uniform distribution more and more closely.

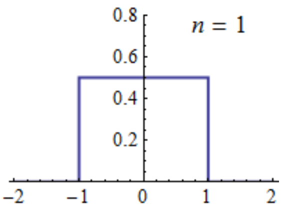
Tossing coins

Let X_n be the fraction of heads after tossing up an unbiased coin n times. Then X_1 has the Bernoulli distribution with expected value $\mu = 0.5$ and variance $\sigma^2 = 0.25$. The subsequent random variables $X_2, X_3, \ldots$ will all be distributed binomially.

As n grows larger, this distribution will gradually start to take shape more and more similar to the bell curve of the normal distribution. If we shift and rescale X_n appropriately, then $Z_n = \frac{\sqrt{n}}{\sigma}(X_n - \mu)$ will be **converging in distribution** to the standard normal, the result that follows from the celebrated central limit theorem.

Graphic example

Suppose $\{X_i\}$ is an iid sequence of uniform $U(-1, 1)$ random variables. Let $Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$ be their (normalized) sums. Then according to the central limit theorem, the distribution of Z_n approaches the normal $N(0, \frac{1}{3})$ distribution. This convergence is shown in the picture: as n grows larger, the shape of the probability density function gets closer and closer to the Gaussian curve.



Definition

A sequence $X_1, X_2, \ldots$ of real-valued random variables is said to **converge in distribution**, or **converge weakly**, or **converge in law** to a random variable X if

$$\lim_{n\to\infty} F_n(x) = F(x),$$

for every number $x \in \mathbb{R}$ at which F is continuous. Here F_n and F are the cumulative distribution functions of random variables X_n and X, respectively.

The requirement that only the continuity points of F should be considered is essential. For example, if X_n are distributed uniformly on intervals $(0, \frac{1}{n})$, then this sequence converges in distribution to a degenerate random variable $X = 0$. Indeed, $F_n(x) = 0$ for all n when $x \leq 0$, and $F_n(x) = 1$ for all $x \geq \frac{1}{n}$ when $n > 0$. However, for this limiting random variable $F(0) = 1$, even though $F_n(0) = 0$ for all n. Thus the convergence of cdfs fails at the point $x = 0$ where F is discontinuous.

Convergence in distribution may be denoted as

$$\begin{aligned} &X_n \xrightarrow{d} X,\ \ X_n \xrightarrow{\mathcal{D}} X,\ \ X_n \xrightarrow{\mathcal{L}} X,\ \ X_n \xrightarrow{d} \mathcal{L}_X, \\ &X_n \rightsquigarrow X,\ \ X_n \Rightarrow X,\ \ \mathcal{L}(X_n) \to \mathcal{L}(X), \end{aligned} \qquad (1)$$

where $\mathcal{L}_X$ is the law (probability distribution) of X. For example, if X is standard normal we can write $X_n \xrightarrow{d} \mathcal{N}(0, 1)$.

For random vectors $\{X_1, X_2, \ldots\} \subset \mathbf{R}^k$ the convergence in distribution is defined similarly. We say that this sequence **converges in distribution** to a random k-vector X if

$$\lim_{n\to\infty} \Pr(X_n \in A) = \Pr(X \in A)$$

for every $A \subset \mathbf{R}^k$ which is a continuity set of X.

The definition of convergence in distribution may be extended from random vectors to more general random elements in arbitrary metric spaces, and even to the "random variables" which are not measurable — a situation which occurs for example in the study of empirical processes. This is the "weak convergence of laws without laws being defined" — except asymptotically.[1]

In this case the term **weak convergence** is preferable (see weak convergence of measures), and we say that a sequence of random elements $\{X_n\}$ converges weakly to X (denoted as $X_n \Rightarrow X$) if

$$\mathrm{E}^* h(X_n) \to \mathrm{E}\, h(X)$$

for all continuous bounded functions h.[2] Here E^* denotes the *outer expectation*, that is the expectation of a "smallest measurable function g that dominates $h(X_n)$".

Properties

- Since $F(a) = \Pr(X \le a)$, the convergence in distribution means that the probability for X_n to be in a given range is approximately equal to the probability that the value of X is in that range, provided n is sufficiently large.
- In general, convergence in distribution does not imply that the sequence of corresponding probability density functions will also converge. As an example one may consider random variables with densities $f_n(x) = (1 - \cos(2\pi nx))\mathbf{1}_{(0,1)}$. These random variables converge in distribution to a uniform $U(0, 1)$, whereas their densities do not converge at all.[3]
  - However, according to *Scheffé's theorem*, convergence of the probability density functions implies convergence in distribution.[4]
- The **portmanteau lemma** provides several equivalent definitions of convergence in distribution. Although these definitions are less intuitive, they are used to prove a number of statistical theorems. The lemma states that $\{X_n\}$ converges in distribution to X if and only if any of the following statements are true:[5]
  - $\Pr(X_n \le x) \to \Pr(X \le x)$ for all continuity points of $x \mapsto \Pr(X \le x)$;
  - $\mathrm{E} f(X_n) \to \mathrm{E} f(X)$ for all bounded, continuous functions f (where E denotes the expected value operator);
  - $\mathrm{E} f(X_n) \to \mathrm{E} f(X)$ for all bounded, Lipschitz functions f;
  - $\liminf \mathrm{E} f(X_n) \ge \mathrm{E} f(X)$ for all nonnegative, continuous functions f;
  - $\liminf \Pr(X_n \in G) \ge \Pr(X \in G)$ for every open set G;
  - $\limsup \Pr(X_n \in F) \le \Pr(X \in F)$ for every closed set F;
  - $\Pr(X_n \in B) \to \Pr(X \in B)$ for all continuity sets B of random variable X;
  - $\limsup \mathrm{E} f(X_n) \le \mathrm{E} f(X)$ for every upper semi-continuous function f bounded above;
  - $\liminf \mathrm{E} f(X_n) \ge \mathrm{E} f(X)$ for every lower semi-continuous function f bounded below.
- The **continuous mapping theorem** states that for a continuous function g, if the sequence $\{X_n\}$ converges in distribution to X, then $\{g(X_n)\}$ converges in distribution to $g(X)$.
  - Note however that convergence in distribution of $\{X_n\}$ to X and $\{Y_n\}$ to Y does in general *not* imply convergence in distribution of $\{X_n + Y_n\}$ to $X + Y$ or of $\{X_nY_n\}$ to XY.
- **Lévy's continuity theorem:** the sequence $\{X_n\}$ converges in distribution to X if and only if the sequence of corresponding characteristic functions $\{\varphi_n\}$ converges pointwise to the characteristic function φ of X.
- Convergence in distribution is metrizable by the Lévy–Prokhorov metric.
- A natural link to convergence in distribution is the Skorokhod's representation theorem.

Convergence in probability

The basic idea behind this type of convergence is that the probability of an "unusual" outcome becomes smaller and smaller as the sequence progresses.

The concept of convergence in probability is used very often in statistics. For example, an estimator is called consistent if it converges in probability to the quantity being estimated. Convergence in probability is also the type of convergence established by the weak law of large numbers.

Definition

A sequence $\{X_n\}$ of random variables **converges in probability** towards the random variable X if for all $\varepsilon > 0$

$$\lim_{n\to\infty} \Pr\big(|X_n - X| > \varepsilon\big) = 0.$$

More explicitly, let $P_n(\varepsilon)$ be the probability that X_n is outside the ball of radius ε centered at X. Then X_n is said to converge in probability to X if for any $\varepsilon > 0$ and any $\delta > 0$ there exists a number N (which may depend on ε and δ) such that for all $n \ge N$, $P_n(\varepsilon) < \delta$ (the definition of limit).

Notice that for the condition to be satisfied, it is not possible that for each n the random variables X and X_n are independent (and thus convergence in probability is a condition on the joint cdf's, as opposed to convergence in distribution, which is a condition on the individual cdf's), unless X is deterministic like for the weak law of large numbers. At the same time, the case of a deterministic X cannot, whenever the deterministic value is a discontinuity point (not isolated), be handled by convergence in distribution, where discontinuity points have to be explicitly excluded.

Examples of convergence in probability

Height of a person

Consider the following experiment. First, pick a random person in the street. Let X be his/her height, which is *ex ante* a random variable. Then ask other people to estimate this height by eye. Let X_n be the average of the first n responses. Then (provided there is no systematic error) by the law of large numbers, the sequence X_n will converge in probability to the random variable X.

Predicting random number generation

Suppose that a random number generator generates a pseudorandom floating point number between 0 and 1. Let random variable X represent the distribution of possible outputs by

Convergence in probability is denoted by adding the letter p over an arrow indicating convergence, or using the "plim" probability limit operator:

the algorithm. Because the pseudorandom number is generated deterministically, its next value is not truly random. Suppose that as you observe a sequence of randomly generated numbers, you can deduce a pattern and make increasingly accurate predictions as to what the next randomly generated number will be. Let X_n be your guess of the value of the next random number after observing the first n random numbers. As you learn the pattern and your guesses become more accurate, not only will the distribution of X_n converge to the distribution of X, but the outcomes of X_n will converge to the outcomes of X.

$$X_n \xrightarrow{p} X, \;\; X_n \xrightarrow{P} X, \;\; \underset{n\to\infty}{\operatorname{plim}} X_n = X. \qquad (2)$$

For random elements $\{X_n\}$ on a separable metric space (S, d), convergence in probability is defined similarly by[6]

$$\forall \varepsilon > 0, \Pr\big(d(X_n, X) \geq \varepsilon\big) \to 0.$$

Properties

- Convergence in probability implies convergence in distribution.[proof]
- In the opposite direction, convergence in distribution implies convergence in probability when the limiting random variable X is a constant.[proof]
- Convergence in probability does not imply almost sure convergence.[proof]
- The continuous mapping theorem states that for every continuous function $g(\cdot)$, if $X_n \xrightarrow{p} X$, then also $g(X_n) \xrightarrow{p} g(X)$.
- Convergence in probability defines a topology on the space of random variables over a fixed probability space. This topology is metrizable by the *Ky Fan metric*:[7]

  $$d(X,Y) = \inf\big\{\varepsilon > 0 : \; \Pr\big(|X-Y| > \varepsilon\big) \leq \varepsilon\big\}$$

  or alternately by this metric

  $$d(X,Y) = \mathbb{E}\left[\min(|X-Y|, 1)\right].$$

Almost sure convergence

This is the type of stochastic convergence that is most similar to pointwise convergence known from elementary real analysis.

Definition

To say that the sequence X_n converges **almost surely** or **almost everywhere** or **with probability 1** or **strongly** towards X means that

$$\Pr\Big(\lim_{n\to\infty} X_n = X\Big) = 1.$$

This means that the values of X_n approach the value of X, in the sense (see almost surely) that events for which X_n does not converge to X have probability 0. Using the probability space $(\Omega, \mathcal{F}, \Pr)$ and the concept of the random variable as a function from Ω to **R**, this is equivalent to the statement

$$\Pr\Big(\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\Big) = 1.$$

Using the notion of the limit superior of a sequence of sets, almost sure convergence can also be defined as follows:

$$\Pr\Big(\limsup_{n\to\infty} \big\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon\big\}\Big) = 0 \quad \text{for all} \quad \varepsilon > 0.$$

Almost sure convergence is often denoted by adding the letters *a.s.* over an arrow indicating convergence:

$$X_n \xrightarrow{\text{a.s.}} X. \qquad (3)$$

For generic random elements $\{X_n\}$ on a metric space (S, d), convergence almost surely is defined similarly:

Examples of almost sure convergence

Example 1

Consider an animal of some short-lived species. We record the amount of food that this animal consumes per day. This sequence of numbers will be unpredictable, but we may be *quite certain* that one day the number will become zero, and will stay zero forever after.

Example 2

Consider a man who tosses seven coins every morning. Each afternoon, he donates one pound to a charity for each head that appeared. The first time the result is all tails, however, he will stop permanently.

Let X_1, X_2, … be the daily amounts the charity received from him.

We may be *almost sure* that one day this amount will be zero, and stay zero forever after that.

$$\Pr\Big(\omega \in \Omega : d\big(X_n(\omega), X(\omega)\big) \underset{n\to\infty}{\longrightarrow} 0\Big) = 1$$

However, when we consider *any finite number* of days, there is a nonzero probability the terminating condition will not occur.

Properties

- Almost sure convergence implies convergence in probability (by Fatou's lemma), and hence implies convergence in distribution. It is the notion of convergence used in the strong law of large numbers.
- The concept of almost sure convergence does not come from a topology on the space of random variables. This means there is no topology on the space of random variables such that the almost surely convergent sequences are exactly the converging sequences with respect to that topology. In particular, there is no metric of almost sure convergence.

Sure convergence or pointwise convergence

To say that the sequence of random variables (X_n) defined over the same probability space (i.e., a random process) converges **surely** or **everywhere** or **pointwise** towards X means

$$\lim_{n\to\infty} X_n(\omega) = X(\omega), \ \forall \omega \in \Omega.$$

where Ω is the sample space of the underlying probability space over which the random variables are defined.

This is the notion of pointwise convergence of a sequence of functions extended to a sequence of random variables. (Note that random variables themselves are functions).

$$\left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\} = \Omega.$$

Sure convergence of a random variable implies all the other kinds of convergence stated above, but there is no payoff in probability theory by using sure convergence compared to using almost sure convergence. The difference between the two only exists on sets with probability zero. This is why the concept of sure convergence of random variables is very rarely used.

Convergence in mean

Given a real number $r \geq 1$, we say that the sequence X_n converges **in the *r*-th mean** (or **in the L^r-norm**) towards the random variable X, if the r-th absolute moments $\mathrm{E}(|X_n|^r)$ and $\mathrm{E}(|X|^r)$ of X_n and X exist, and

$$\lim_{n\to\infty} \mathrm{E}(|X_n - X|^r) = 0,$$

where the operator E denotes the expected value. Convergence in r-th mean tells us that the expectation of the r-th power of the difference between X_n and X converges to zero.

This type of convergence is often denoted by adding the letter L^r over an arrow indicating convergence:

$$X_n \xrightarrow{L^r} X. \qquad (4)$$

The most important cases of convergence in r-th mean are:

- When X_n converges in r-th mean to X for $r = 1$, we say that X_n converges **in mean** to X.
- When X_n converges in r-th mean to X for $r = 2$, we say that X_n converges **in mean square** (or **in quadratic mean**) to X.

Convergence in the r-th mean, for $r \geq 1$, implies convergence in probability (by Markov's inequality). Furthermore, if $r > s \geq 1$, convergence in r-th mean implies convergence in s-th mean. Hence, convergence in mean square implies convergence in mean.

It is also worth noticing that if

$$X_n \xrightarrow{L^r} X, \qquad (4)$$

then

$$\lim_{n\to\infty} E[|X_n|^r] = E[|X|^r]$$

Properties

Provided the probability space is complete:

- If $X_n \xrightarrow{p} X$ and $X_n \xrightarrow{p} Y$, then $X = Y$ almost surely.
- If $X_n \xrightarrow{\text{a.s.}} X$ and $X_n \xrightarrow{\text{a.s.}} Y$, then $X = Y$ almost surely.
- If $X_n \xrightarrow{L^r} X$ and $X_n \xrightarrow{L^r} Y$, then $X = Y$ almost surely.
- If $X_n \xrightarrow{p} X$ and $Y_n \xrightarrow{p} Y$, then $aX_n + bY_n \xrightarrow{p} aX + bY$ (for any real numbers a and b) and $X_nY_n \xrightarrow{p} XY$.
- If $X_n \xrightarrow{\text{a.s.}} X$ and $Y_n \xrightarrow{\text{a.s.}} Y$, then $aX_n + bY_n \xrightarrow{\text{a.s.}} aX + bY$ (for any real numbers a and b) and $X_nY_n \xrightarrow{\text{a.s.}} XY$.

- If $X_n \xrightarrow{L^r} X$ and $Y_n \xrightarrow{L^r} Y$, then $aX_n + bY_n \xrightarrow{L^r} aX + bY$ (for any real numbers a and b).
- None of the above statements are true for convergence in distribution.

The chain of implications between the various notions of convergence are noted in their respective sections. They are, using the arrow notation:

$$\begin{matrix} \xrightarrow{L^s} & \underset{s>r\geq 1}{\Rightarrow} & \xrightarrow{L^r} & & \\ & & \Downarrow & & \\ \xrightarrow{\text{a.s.}} & \Rightarrow & \xrightarrow{p} & \Rightarrow & \xrightarrow{d} \end{matrix}$$

These properties, together with a number of other special cases, are summarized in the following list:

- Almost sure convergence implies convergence in probability:[8][proof]

  $$X_n \xrightarrow{\text{a.s.}} X \quad \Rightarrow \quad X_n \xrightarrow{p} X$$

- Convergence in probability implies there exists a sub-sequence (k_n) which almost surely converges:[9]

  $$X_n \xrightarrow{p} X \quad \Rightarrow \quad X_{k_n} \xrightarrow{\text{a.s.}} X$$

- Convergence in probability implies convergence in distribution:[8][proof]

  $$X_n \xrightarrow{p} X \quad \Rightarrow \quad X_n \xrightarrow{d} X$$

- Convergence in *r*-th order mean implies convergence in probability:

  $$X_n \xrightarrow{L^r} X \quad \Rightarrow \quad X_n \xrightarrow{p} X$$

- Convergence in *r*-th order mean implies convergence in lower order mean, assuming that both orders are greater than or equal to one:

  $$X_n \xrightarrow{L^r} X \quad \Rightarrow \quad X_n \xrightarrow{L^s} X, \qquad \text{provided } r \geq s \geq 1.$$

- If X_n converges in distribution to a constant *c*, then X_n converges in probability to *c*:[8][proof]

  $$X_n \xrightarrow{d} c \quad \Rightarrow \quad X_n \xrightarrow{p} c, \qquad \text{provided } c \text{ is a constant.}$$

- If X_n converges in distribution to *X* and the difference between X_n and Y_n converges in probability to zero, then Y_n also converges in distribution to *X*:[8][proof]

  $$X_n \xrightarrow{d} X,\ |X_n - Y_n| \xrightarrow{p} 0 \quad \Rightarrow \quad Y_n \xrightarrow{d} X$$

- If X_n converges in distribution to *X* and Y_n converges in distribution to a constant *c*, then the joint vector (X_n, Y_n) converges in distribution to (X, c):[8][proof]

  $$X_n \xrightarrow{d} X,\ Y_n \xrightarrow{d} c \quad \Rightarrow \quad (X_n, Y_n) \xrightarrow{d} (X, c) \qquad \text{provided } c \text{ is a constant.}$$

  Note that the condition that Y_n converges to a constant is important, if it were to converge to a random variable *Y* then we wouldn't be able to conclude that (X_n, Y_n) converges to (X, Y).

- If X_n converges in probability to *X* and Y_n converges in probability to *Y*, then the joint vector (X_n, Y_n) converges in probability to (X, Y):[8][proof]

  $$X_n \xrightarrow{p} X,\ Y_n \xrightarrow{p} Y \quad \Rightarrow \quad (X_n, Y_n) \xrightarrow{p} (X, Y)$$

- If X_n converges in probability to *X*, and if $\mathbf{P}(|X_n| \leq b) = 1$ for all *n* and some *b*, then X_n converges in *r*th mean to *X* for all $r \geq 1$. In other words, if X_n converges in probability to *X* and all random variables X_n are almost surely bounded above and below, then X_n converges to *X* also in any *r*th mean.
- **Almost sure representation**. Usually, convergence in distribution does not imply convergence almost surely. However, for a given sequence $\{X_n\}$ which converges in distribution to X_0 it is always possible to find a new probability space (Ω, *F*, P) and random variables $\{Y_n, n = 0, 1, ...\}$ defined on it such that Y_n is equal in distribution to X_n for each $n \geq 0$, and Y_n converges to Y_0 almost surely.[10][11]
- If for all *ε* > 0,

  $$\sum_n \mathbb{P}\left(|X_n - X| > \varepsilon\right) < \infty,$$

  then we say that X_n *converges almost completely*, or *almost in probability* towards *X*. When X_n converges almost completely towards *X* then it also converges almost surely to *X*. In other words, if X_n converges in probability to *X* sufficiently quickly (i.e. the above sequence of tail probabilities is summable for all $\varepsilon > 0$), then X_n also converges almost surely to *X*. This is a direct implication from the Borel–Cantelli lemma.

- If S_n is a sum of *n* real independent random variables:

$$S_n = X_1 + \cdots + X_n$$

then S_n converges almost surely if and only if S_n converges in probability.

- The dominated convergence theorem gives sufficient conditions for almost sure convergence to imply L^1-convergence:

$$\left.\begin{matrix} X_n \xrightarrow{\text{a.s.}} X \\ |X_n| < Y \\ \mathrm{E}(Y) < \infty \end{matrix}\right\} \quad \Rightarrow \quad X_n \xrightarrow{L^1} X \qquad (5)$$

- A necessary and sufficient condition for L^1 convergence is $X_n \xrightarrow{P} X$ and the sequence (X_n) is uniformly integrable.

See also

- Proofs of convergence of random variables
- Convergence of measures
- Convergence in measure
- Continuous stochastic process: the question of continuity of a stochastic process is essentially a question of convergence, and many of the same concepts and relationships used above apply to the continuity question.
- Asymptotic distribution
- Big O in probability notation
- Skorokhod's representation theorem
- The Tweedie convergence theorem
- Slutsky's theorem
- Continuous mapping theorem

Notes

1. Bickel et al. 1998, A.8, page 475
2. van der Vaart & Wellner 1996, p. 4
3. Romano & Siegel 1985, Example 5.26
4. Durrett, Rick (2010). *Probability: Theory and Examples*. p. 84.
5. van der Vaart 1998, Lemma 2.2
6. Dudley 2002, Chapter 9.2, page 287
7. Dudley 2002, p. 289
8. van der Vaart 1998, Theorem 2.7
9. Gut, Allan (2005). *Probability: A graduate course*. Theorem 3.4: Springer. ISBN 978-0-387-22833-4.
10. van der Vaart 1998, Th.2.19
11. Fristedt & Gray 1997, Theorem 14.5

References

- Bickel, Peter J.; Klaassen, Chris A.J.; Ritov, Ya'acov; Wellner, Jon A. (1998). *Efficient and adaptive estimation for semiparametric models*. New York: Springer-Verlag. ISBN 978-0-387-98473-5.
- Billingsley, Patrick (1986). *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics (2nd ed.). Wiley.
- Billingsley, Patrick (1999). *Convergence of probability measures* (https://archive.org/details/convergenceofpro0000bill) (2nd ed.). John Wiley & Sons. pp. 1–28 (https://archive.org/details/convergenceofpro0000bill/page/1). ISBN 978-0-471-19745-4.
- Dudley, R.M. (2002). *Real analysis and probability*. Cambridge, UK: Cambridge University Press. ISBN 978-0-521-80972-6.
- Fristedt, Bert; Gray, Lawrence (1997). *A Modern Approach to Probability Theory*. New York: Springer Science+Business Media. doi:10.1007/978-1-4899-2837-5 (https://doi.org/10.1007%2F978-1-4899-2837-5). ISBN 978-1-4899-2837-5.
- Grimmett, G.R.; Stirzaker, D.R. (1992). *Probability and random processes* (2nd ed.). Clarendon Press, Oxford. pp. 271–285. ISBN 978-0-19-853665-9.
- Jacobsen, M. (1992). *Videregående Sandsynlighedsregning (Advanced Probability Theory)* (3rd ed.). HCØ-tryk, Copenhagen. pp. 18–20. ISBN 978-87-91180-71-2.
- Ledoux, Michel; Talagrand, Michel (1991). *Probability in Banach spaces*. Berlin: Springer-Verlag. pp. xii+480. ISBN 978-3-540-52013-9. MR 1102015 (https://www.ams.org/mathscinet-getitem?mr=1102015).
- Romano, Joseph P.; Siegel, Andrew F. (1985). *Counterexamples in probability and statistics*. Great Britain: Chapman & Hall. ISBN 978-0-412-98901-8.
- van der Vaart, Aad W.; Wellner, Jon A. (1996). *Weak convergence and empirical processes*. New York: Springer-Verlag. ISBN 978-0-387-94640-5.
- van der Vaart, Aad W. (1998). *Asymptotic statistics*. New York: Cambridge University Press. ISBN 978-0-521-49603-2.
- Williams, D. (1991). *Probability with Martingales*. Cambridge University Press. ISBN 978-0-521-40605-5.
- Wong, E.; Hájek, B. (1985). *Stochastic Processes in Engineering Systems*. New York: Springer–Verlag.

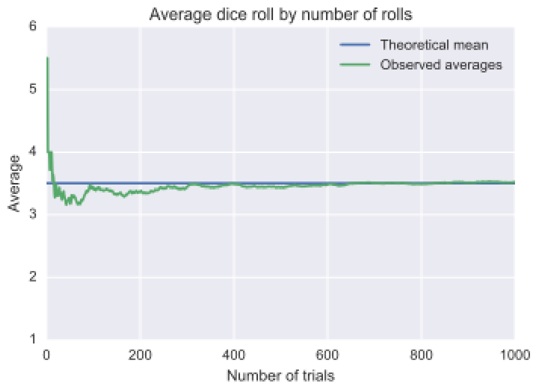
- https://www.ma.utexas.edu/users/gordanz/notes/weak.pdf

WIKIPEDIA

Law of large numbers

In probability theory, the **law of large numbers** (**LLN**) is a theorem that describes the result of performing the same experiment a large number of times. According to the law, the average of the results obtained from a large number of trials should be close to the expected value and will tend to become closer to the expected value as more trials are performed.[1]

The LLN is important because it guarantees stable long-term results for the averages of some random events.[1][2] For example, while a casino may lose money in a single spin of the roulette wheel, its earnings will tend towards a predictable percentage over a large number of spins. Any winning streak by a player will eventually be overcome by the parameters of the game. It is important to remember that the law only applies (as the name indicates) when a *large number* of observations is considered. There is no principle that a small number of observations will coincide with the expected value or that a streak of one value will immediately be "balanced" by the others (see the gambler's fallacy).



An illustration of the law of large numbers using a particular run of rolls of a single die. As the number of rolls in this run increases, the average of the values of all the results approaches 3.5. Although each run would show a distinctive shape over a small number of throws (at the left), over a large number of rolls (to the right) the shapes would be extremely similar.

Contents

Examples

For example, a single roll of a fair, six-sided die produces one of the numbers 1, 2, 3, 4, 5, or 6, each with equal probability. Therefore, the expected value of the average of the rolls is:

$$\frac{1+2+3+4+5+6}{6} = 3.5$$

According to the law of large numbers, if a large number of six-sided dice are rolled, the average of their values (sometimes called the sample mean) is likely to be close to 3.5, with the precision increasing as more dice are rolled.

It follows from the law of large numbers that the empirical probability of success in a series of Bernoulli trials will converge to the theoretical probability. For a Bernoulli random variable, the expected value is the theoretical probability of success, and the average of n such variables (assuming they are independent and identically distributed (i.i.d.)) is precisely the relative frequency.

For example, a fair coin toss is a Bernoulli trial. When a fair coin is flipped once, the theoretical probability that the outcome will be heads is equal to $\frac{1}{2}$. Therefore, according to the law of large numbers, the proportion of heads in a "large" number of coin flips "should be" roughly $\frac{1}{2}$. In particular, the proportion of heads after n flips will almost surely converge to $\frac{1}{2}$ as n approaches infinity.

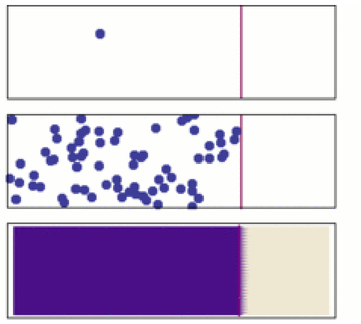
Although the proportion of heads (and tails) approaches 1/2, almost surely the absolute difference in the number of heads and tails will become large as the number of flips becomes large. That is, the probability that the absolute difference is a small number approaches zero as the number of flips becomes large. Also, almost surely the ratio of the absolute difference to the number of flips will approach zero. Intuitively, the expected difference grows, but at a slower rate than the number of flips.

Another good example of the LLN is the Monte Carlo method. These methods are a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results. The larger the number of repetitions, the better the approximation tends to be. The reason that this method is important is mainly that, sometimes, it is difficult or impossible to use other approaches.[3]

Limitation

The average of the results obtained from a large number of trials may fail to converge in some cases. For instance, the average of n results taken from the Cauchy distribution or some Pareto distributions (α<1) will not converge as n becomes larger; the reason is heavy tails. The Cauchy distribution and the Pareto distribution represent two cases: the Cauchy distribution does not have an expectation,[4] whereas the expectation of the Pareto distribution (α<1) is infinite.[5] Another example is where the random numbers equal the tangent of an angle uniformly distributed between −90° and +90°. The median is zero, but the expected value does not exist, and indeed the average of n such variables have the same distribution as one such variable. It does not converge in probability toward zero (or any other value) as n goes to infinity.

History



Diffusion is an example of the law of large numbers. Initially, there are solute molecules on the left side of a barrier (magenta line) and none on the right. The barrier is removed, and the solute diffuses to fill the whole container.
Top: With a single molecule, the motion appears to be quite random.
Middle: With more molecules, there is clearly a trend where the solute fills the container more and more uniformly, but there are also random fluctuations.
Bottom: With an enormous number of solute molecules (too many to see), the randomness is essentially gone: The solute appears to move smoothly and systematically from high-concentration areas to low-concentration areas. In realistic situations, chemists can describe diffusion as a deterministic macroscopic phenomenon (see Fick's laws), despite its underlying random nature.

The Italian mathematician Gerolamo Cardano (1501–1576) stated without proof that the accuracies of empirical statistics tend to improve with the number of trials.[6] This was then formalized as a law of large numbers. A special form of the LLN (for a binary random variable) was first proved by Jacob Bernoulli.[7] It took him over 20 years to develop a sufficiently rigorous mathematical proof which was published in his *Ars Conjectandi* (The Art of Conjecturing) in 1713. He named this his "Golden Theorem" but it became generally known as "**Bernoulli's Theorem**". This should not be confused with Bernoulli's principle, named after Jacob Bernoulli's nephew Daniel Bernoulli. In 1837, S.D. Poisson further described it under the name "*la loi des grands nombres*" ("the law of large numbers").[8][9] Thereafter, it was known under both names, but the "law of large numbers" is most frequently used.

After Bernoulli and Poisson published their efforts, other mathematicians also contributed to refinement of the law, including Chebyshev,[10] Markov, Borel, Cantelli and Kolmogorov and Khinchin. Markov showed that the law can apply to a random variable that does not have a finite variance under some other weaker assumption, and Khinchin showed in 1929 that if the series consists of independent identically distributed random variables, it suffices that the expected value exists for the weak law of large numbers to be true.[11][12] These further studies have given rise to two prominent forms of the LLN. One is called the "weak" law and the other the "strong" law, in reference to two different modes of convergence of the cumulative sample means to the expected value; in particular, as explained below, the strong form implies the weak.[11]

Forms

There are two different versions of the **law of large numbers** that are described below. They are called the ***strong law** of large numbers* and the ***weak law** of large numbers*.[13][1] Stated for the case where X_1, X_2, ... is an infinite sequence of independent and identically distributed (i.i.d.) Lebesgue integrable random variables with expected value $E(X_1) = E(X_2) = ... = \mu$, both versions of the law state that – with virtual certainty – the sample average

$$\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$$

converges to the expected value:

$$\overline{X}_n \to \mu \quad \text{as } n \to \infty. \qquad \textbf{(law. 1)}$$

(Lebesgue integrability of X_j means that the expected value $E(X_j)$ exists according to Lebesgue integration and is finite. It does *not* mean that the associated probability measure is absolutely continuous with respect to Lebesgue measure.)

Based on the assumption of finite variance $\operatorname{Var}(X_i) = \sigma^2$ (for all i) and no correlation between random variables, the variance of the average of n random variables

$$\operatorname{Var}(\overline{X}_n) = \operatorname{Var}(\tfrac{1}{n}(X_1 + \cdots + X_n)) = \frac{1}{n^2}\operatorname{Var}(X_1 + \cdots + X_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

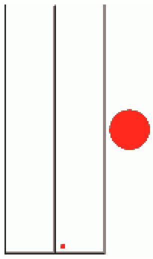
Sometimes an assumption of finite variance $\operatorname{Var}(X_1) = \operatorname{Var}(X_2) = \ldots = \sigma^2 < \infty$ is **not necessary**. Large or infinite variance will make the convergence slower, but the LLN holds anyway. This assumption is often used because it makes the proofs easier and shorter.

Mutual independence of the random variables can be replaced by pairwise independence in both versions of the law.[14]

The difference between the strong and the weak version is concerned with the mode of convergence being asserted. For interpretation of these modes, see Convergence of random variables.

Weak law

The **weak law of large numbers** (also called Khinchin's law) states that the sample average converges in probability towards the expected value[15]



Simulation illustrating the law of large numbers. Each frame, a coin that is red on one side and blue on the other is flipped, and a dot is added in the corresponding column. A pie chart shows the proportion of red and blue so far. Notice that while the proportion varies significantly at first, it approaches 50% as the number of trials increases.

$$\overline{X}_n \xrightarrow{P} \mu \qquad \text{when } n \to \infty. \tag{law. 2}$$

That is, for any positive number ε,

$$\lim_{n\to\infty} \Pr\Big(\,|\overline{X}_n - \mu| > \varepsilon\,\Big) = 0.$$

Interpreting this result, the weak law states that for any nonzero margin specified, no matter how small, with a sufficiently large sample there will be a very high probability that the average of the observations will be close to the expected value; that is, within the margin.

As mentioned earlier, the weak law applies in the case of i.i.d. random variables, but it also applies in some other cases. For example, the variance may be different for each random variable in the series, keeping the expected value constant. If the variances are bounded, then the law applies, as shown by Chebyshev as early as 1867. (If the expected values change during the series, then we can simply apply the law to the average deviation from the respective expected values. The law then states that this converges in probability to zero.) In fact, Chebyshev's proof works so long as the variance of the average of the first n values goes to zero as n goes to infinity.[12] As an example, assume that each random variable in the series follows a Gaussian distribution with mean zero, but with variance equal to $2n/\log(n+1)$, which is not bounded. At each stage, the average will be normally distributed (as the average of a set of normally distributed variables). The variance of the sum is equal to the sum of the variances, which is asymptotic to $n^2/\log n$. The variance of the average is therefore asymptotic to $1/\log n$ and goes to zero.

There are also examples of the weak law applying even though the expected value does not exist.

Strong law

The **strong law of large numbers** (also called Kolmogorov's law) states that the sample average converges almost surely to the expected value[16]

$$\overline{X}_n \xrightarrow{\text{a.s.}} \mu \qquad \text{when } n \to \infty. \tag{law. 3}$$

That is,

$$\Pr\left(\lim_{n\to\infty} \overline{X}_n = \mu\right) = 1.$$

What this means is that the probability that, as the number of trials n goes to infinity, the average of the observations converges to the expected value, is equal to one.

The proof is more complex than that of the weak law.[17] This law justifies the intuitive interpretation of the expected value (for Lebesgue integration only) of a random variable when sampled repeatedly as the "long-term average".

Almost sure convergence is also called strong convergence of random variables. This version is called the strong law because random variables which converge strongly (almost surely) are guaranteed to converge weakly (in probability). However the weak law is known to hold in certain conditions where the strong law does not hold and then the convergence is only weak (in probability). See #Differences between the weak law and the strong law.

The strong law of large numbers can itself be seen as a special case of the pointwise ergodic theorem.

The strong law applies to independent identically distributed random variables having an expected value (like the weak law). This was proved by Kolmogorov in 1930. It can also apply in other cases. Kolmogorov also showed, in 1933, that if the variables are independent and identically distributed, then for the average to converge almost surely on *something* (this can be considered another statement of the strong law), it is necessary that they have an expected value (and then of course the average will converge almost surely on that).[18]

If the summands are independent but not identically distributed, then

$$\overline{X}_n - \operatorname{E}\left[\overline{X}_n\right] \xrightarrow{\text{a.s.}} 0,$$

provided that each X_k has a finite second moment and

$$\sum_{k=1}^{\infty} \frac{1}{k^2} \operatorname{Var}[X_k] < \infty.$$

This statement is known as *Kolmogorov's strong law*, see e.g. Sen & Singer (1993, Theorem 2.3.10).

An example of a series where the weak law applies but not the strong law is when X_k is plus or minus $\sqrt{k/\log\log\log k}$ (starting at sufficiently large k so that the denominator is positive) with probability 1/2 for each.[18] The variance of X_k is then $k/\log\log\log k$. Kolmogorov's strong law does not apply because the partial sum in his criterion up to $k=n$ is asymptotic to $\log n/\log\log\log n$ and this is unbounded.

If we replace the random variables with Gaussian variables having the same variances, namely $\sqrt{k/\log\log\log k}$, then the average at any point will also be normally distributed. The width of the distribution of the average will tend toward zero (standard deviation asymptotic to $1/\sqrt{2\log\log\log n}$), but for a given ε, there is probability which does not go to zero with n, while the average sometime after the nth trial will come back up to ε. Since the width of the distribution of the average is not zero, it must have a positive lower bound p(ε), which means there is a probability of at least p(ε) that the average will attain ε after n trials. It will happen with probability p(ε)/2 before some m which depends on n. But even after m, there is still a probability of at least p(ε) that it will happen. (This seems to indicate that p(ε)=1 and the average will attain ε an infinite number of times.)

Differences between the weak law and the strong law

The *weak law* states that for a specified large n, the average $\overline{X}_n$ is likely to be near μ. Thus, it leaves open the possibility that $|\overline{X}_n - \mu| > \varepsilon$ happens an infinite number of times, although at infrequent intervals. (Not necessarily $|\overline{X}_n - \mu| \neq 0$ for all n).

The *strong law* shows that this almost surely will not occur. In particular, it implies that with probability 1, we have that for any $\varepsilon > 0$ the inequality $|\overline{X}_n - \mu| < \varepsilon$ holds for all large enough n.[19]

The strong law does not hold in the following cases, but the weak law does.[20][21][22]

1. Let X be an exponentially distributed random variable with parameter 1. The random variable $\sin(X)e^X X^{-1}$ has no expected value according to Lebesgue integration, but using conditional convergence and interpreting the integral as a Dirichlet integral, which is an improper Riemann integral, we can say:

$$E\left(\frac{\sin(X)e^X}{X}\right) = \int_0^\infty \frac{\sin(x)e^x}{x}e^{-x}dx = \frac{\pi}{2}$$

2. Let x be geometric distribution with probability 0.5. The random variable $2^X(-1)^X X^{-1}$ does not have an expected value in the conventional sense because the infinite series is not absolutely convergent, but using conditional convergence, we can say:

$$E\left(\frac{2^X(-1)^X}{X}\right) = \sum_1^\infty \frac{2^x(-1)^x}{x}2^{-x} = -\ln(2)$$

3. If the cumulative distribution function of a random variable is

$$1 - F(x) = \frac{e}{2x\ln(x)}, x \geq e$$

$$F(x) = \frac{e}{-2x\ln(-x)}, x \leq -e$$

then it has no expected value, but the weak law is true.[23][24]

Uniform law of large numbers

Suppose $f(x,\theta)$ is some function defined for $\theta \in \Theta$, and continuous in θ. Then for any fixed θ, the sequence $\{f(X_1,\theta), f(X_2,\theta), \ldots\}$ will be a sequence of independent and identically distributed random variables, such that the sample mean of this sequence converges in probability to $E[f(X,\theta)]$. This is the *pointwise* (in θ) convergence.

The **uniform law of large numbers** states the conditions under which the convergence happens *uniformly* in θ. If[25][26]

1. Θ is compact,
2. $f(x,\theta)$ is continuous at each $\theta \in \Theta$ for almost all xs, and measurable function of x at each θ.
3. there exists a dominating function $d(x)$ such that $E[d(X)] < \infty$, and

$$\|f(x,\theta)\| \leq d(x) \quad \text{for all } \theta \in \Theta.$$

Then $E[f(X,\theta)]$ is continuous in θ, and

$$\sup_{\theta\in\Theta}\left\|\frac{1}{n}\sum_{i=1}^n f(X_i,\theta) - \mathrm{E}[f(X,\theta)]\right\| \xrightarrow{\text{a.s.}} 0.$$

This result is useful to derive consistency of a large class of estimators (see Extremum estimator).

Borel's law of large numbers

Borel's law of large numbers, named after Émile Borel, states that if an experiment is repeated a large number of times, independently under identical conditions, then the proportion of times that any specified event occurs approximately equals the probability of the event's occurrence on any particular trial; the larger the number of repetitions, the better the approximation tends to be. More precisely, if E denotes the event in question, p its probability of occurrence, and $N_n(E)$ the number of times E occurs in the first n trials, then with probability one,[27]

$$\frac{N_n(E)}{n} \to p \text{ as } n \to \infty.$$

This theorem makes rigorous the intuitive notion of probability as the long-run relative frequency of an event's occurrence. It is a special case of any of several more general laws of large numbers in probability theory.

Chebyshev's inequality. Let X be a random variable with finite expected value μ and finite non-zero variance σ^2. Then for any real number $k > 0$,

$$\Pr(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Proof of the weak law

Given $X_1, X_2, \ldots$ an infinite sequence of i.i.d. random variables with finite expected value $E(X_1) = E(X_2) = \ldots = \mu < \infty$, we are interested in the convergence of the sample average

$$\overline{X}_n = \tfrac{1}{n}(X_1 + \cdots + X_n).$$

The weak law of large numbers states:

Theorem: $\overline{X}_n \xrightarrow{P} \mu \qquad \text{when } n \to \infty.$ **(law. 2)**

Proof using Chebyshev's inequality assuming finite variance

This proof uses the assumption of finite variance $\operatorname{Var}(X_i) = \sigma^2$ (for all i). The independence of the random variables implies no correlation between them, and we have that

$$\operatorname{Var}(\overline{X}_n) = \operatorname{Var}(\tfrac{1}{n}(X_1 + \cdots + X_n)) = \frac{1}{n^2}\operatorname{Var}(X_1 + \cdots + X_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

The common mean μ of the sequence is the mean of the sample average:

$$E(\overline{X}_n) = \mu.$$

Using Chebyshev's inequality on $\overline{X}_n$ results in

$$\operatorname{P}(\left|\overline{X}_n - \mu\right| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

This may be used to obtain the following:

$$\operatorname{P}(\left|\overline{X}_n - \mu\right| < \varepsilon) = 1 - \operatorname{P}(\left|\overline{X}_n - \mu\right| \geq \varepsilon) \geq 1 - \frac{\sigma^2}{n\varepsilon^2}.$$

As n approaches infinity, the expression approaches 1. And by definition of convergence in probability, we have obtained

$$\overline{X}_n \xrightarrow{P} \mu \qquad \text{when } n \to \infty. \qquad \textbf{(law. 2)}$$

Proof using convergence of characteristic functions

By Taylor's theorem for complex functions, the characteristic function of any random variable, X, with finite mean μ, can be written as

$$\varphi_X(t) = 1 + it\mu + o(t), \quad t \to 0.$$

All X_1, X_2, ... have the same characteristic function, so we will simply denote this φ_X.

Among the basic properties of characteristic functions there are

$$\varphi_{\frac{1}{n}X}(t) = \varphi_X(\tfrac{t}{n}) \quad \text{and} \quad \varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) \quad \text{if } X \text{ and } Y \text{ are independent.}$$

These rules can be used to calculate the characteristic function of $\overline{X}_n$ in terms of φ_X:

$$\varphi_{\overline{X}_n}(t) = \left[\varphi_X\left(\frac{t}{n}\right)\right]^n = \left[1 + i\mu\frac{t}{n} + o\left(\frac{t}{n}\right)\right]^n \to e^{it\mu}, \quad \text{as} \quad n \to \infty.$$

The limit $e^{it\mu}$ is the characteristic function of the constant random variable μ, and hence by the Lévy continuity theorem, $\overline{X}_n$ converges in distribution to μ:

$$\overline{X}_n \xrightarrow{\mathcal{D}} \mu \qquad \text{for} \qquad n \to \infty.$$

μ is a constant, which implies that convergence in distribution to μ and convergence in probability to μ are equivalent (see Convergence of random variables.) Therefore,

$$\overline{X}_n \xrightarrow{P} \mu \qquad \text{when } n \to \infty. \qquad \textbf{(law. 2)}$$

This shows that the sample mean converges in probability to the derivative of the characteristic function at the origin, as long as the latter exists.

Consequences

The law of large numbers provides an expectation of an unknown distribution from a realization of the sequence, but also any feature of the probability distribution.[1] By applying Borel's law of large numbers, one could easily obtain the probability mass function. For each event in the objective probability mass function, one could approximate the probability of the event's occurrence with the proportion of times that any specified event occurs. The larger the number of repetitions, the better the approximation. As for the continuous case: $C = (a - h, a + h]$, for small positive h. Thus, for large n:

$$\frac{N_n(C)}{n} \approx p = P(X \in C) = \int_{a-h}^{a+h} f(x)dx \approx 2hf(a)$$

With this method, one can cover the whole x-axis with a grid (with grid size 2h) and obtain a bar graph which is called a histogram.

See also

- Asymptotic equipartition property
- Central limit theorem
- Infinite monkey theorem
- Law of averages
- Law of the iterated logarithm
- Law of truly large numbers
- Lindy effect
- Regression toward the mean
- Sortition

Notes

1. Dekking, Michel (2005). *A Modern Introduction to Probability and Statistics* (https://archive.org/details/modernintroducti00fmde). Springer. pp. 181 (https://archive.org/details/modernintroducti00fmde/page/n191)–190. ISBN 9781852338961.
2. Yao, Kai; Gao, Jinwu (2016). "Law of Large Numbers for Uncertain Random Variables". *IEEE Transactions on Fuzzy Systems*. **24** (3): 615–621. doi:10.1109/TFUZZ.2015.2466080 (https://doi.org/10.1109%2FTFUZZ.2015.2466080). ISSN 1063-6706 (https://www.worldcat.org/issn/1063-6706). S2CID 2238905 (https://api.semanticscholar.org/CorpusID:2238905).
3. Kroese, Dirk P.; Brereton, Tim; Taimre, Thomas; Botev, Zdravko I. (2014). "Why the Monte Carlo method is so important today". *Wiley Interdisciplinary Reviews: Computational Statistics*. **6** (6): 386–392. doi:10.1002/wics.1314 (https://doi.org/10.1002%2Fwics.1314).
4. Dekking, Michel (2005). *A Modern Introduction to Probability and Statistics* (https://archive.org/details/modernintroducti00fmde). Springer. pp. 92 (https://archive.org/details/modernintroducti00fmde/page/n102). ISBN 9781852338961.
5. Dekking, Michel (2005). *A Modern Introduction to Probability and Statistics* (https://archive.org/details/modernintroducti00fmde). Springer. pp. 63 (https://archive.org/details/modernintroducti00fmde/page/n74). ISBN 9781852338961.
6. Mlodinow, L. *The Drunkard's Walk.* New York: Random House, 2008. p. 50.
7. Jakob Bernoulli, *Ars Conjectandi: Usum & Applicationem Praecedentis Doctrinae in Civilibus, Moralibus & Oeconomicis*, 1713, Chapter 4, (Translated into English by Oscar Sheynin)
8. Poisson names the "law of large numbers" (*la loi des grands nombres*) in: S.D. Poisson, *Probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilitiés* (Paris, France: Bachelier, 1837), p. 7 (https://archive.org/details/recherchessurla02poisgoog/page/n30). He attempts a two-part proof of the law on pp. 139–143 and pp. 277 ff.
9. Hacking, Ian. (1983) "19th-century Cracks in the Concept of Determinism", *Journal of the History of Ideas*, 44 (3), 455-475 JSTOR 2709176 (https://www.jstor.org/stable/2709176)
10. Tchebichef, P. (1846). "Démonstration élémentaire d'une proposition générale de la théorie des probabilités" (https://zenodo.org/record/1448850). *Journal für die reine und angewandte Mathematik*. **1846** (33): 259–267. doi:10.1515/crll.1846.33.259 (https://doi.org/10.1515%2Fcrll.1846.33.259). S2CID 120850863 (https://api.semanticscholar.org/CorpusID:120850863).
11. Seneta 2013.
12. Yuri Prohorov. "Law of large numbers" (https://www.encyclopediaofmath.org/index.php/Law_of_large_numbers). *Encyclopedia of Mathematics*.
13. Bhattacharya, Rabi; Lin, Lizhen; Patrangenaru, Victor (2016). *A Course in Mathematical Statistics and Large Sample Theory*. Springer Texts in Statistics. New York, NY: Springer New York. doi:10.1007/978-1-4939-4032-5 (https://doi.org/10.1007%2F978-1-4939-4032-5). ISBN 978-1-4939-4030-1.
14. Etemadi, N.Z. (1981). "An elementary proof of the strong law of large numbers". *Wahrscheinlichkeitstheorie Verw Gebiete*. **55** (1): 119–122. doi:10.1007/BF01013465 (https://doi.org/10.1007%2FBF01013465). S2CID 122166046 (https://api.semanticscholar.org/CorpusID:122166046).
15. Loève 1977, Chapter 1.4, p. 14
16. Loève 1977, Chapter 17.3, p. 251
17. "The strong law of large numbers – What's new" (http://terrytao.wordpress.com/2008/06/18/the-strong-law-of-large-numbers/). Terrytao.wordpress.com. Retrieved 2012-06-09.
18. Yuri Prokhorov. "Strong law of large numbers" (https://www.encyclopediaofmath.org/index.php/Strong_law_of_large_numbers). *Encyclopedia of Mathematics*.
19. Ross (2009)

20. Lehmann, Erich L; Romano, Joseph P (2006-03-30). *Weak law converges to constant* (https://books.google.com/books?id=K6t5qn-SEp8C&q=%22even%20if%20the%20mean%20does%20not%20exist%22&pg=PA432). ISBN 9780387276052.
21. "A NOTE ON THE WEAK LAW OF LARGE NUMBERS FOR EXCHANGEABLE RANDOM VARIABLES" (https://web.archive.org/web/20160701234328/http://www.mathnet.or.kr/mathnet/kms_tex/31810.pdf) (PDF). Dguvl Hun Hong and Sung Ho Lee. Archived from the original (http://www.mathnet.or.kr/mathnet/kms_tex/31810.pdf) (PDF) on 2016-07-01. Retrieved 2014-06-28.
22. "weak law of large numbers: proof using characteristic functions vs proof using truncation VARIABLES" (https://math.stackexchange.com/q/266870).
23. Mukherjee, Sayan. "Law of large numbers" (https://web.archive.org/web/20130309032810/http://www.isds.duke.edu/courses/Fall09/sta205/lec/lln.pdf) (PDF). Archived from the original (http://www.isds.duke.edu/courses/Fall09/sta205/lec/lln.pdf) (PDF) on 2013-03-09. Retrieved 2014-06-28.
24. J. Geyer, Charles. "Law of large numbers" (http://www.stat.umn.edu/geyer/8112/notes/weaklaw.pdf) (PDF).
25. Newey & McFadden 1994, Lemma 2.4
26. Jennrich, Robert I. (1969). "Asymptotic Properties of Non-Linear Least Squares Estimators" (https://doi.org/10.1214%2Faoms%2F1177697731). *The Annals of Mathematical Statistics*. **40** (2): 633–643. doi:10.1214/aoms/1177697731 (https://doi.org/10.1214%2Faoms%2F1177697731).
27. An Analytic Technique to Prove Borel's Strong Law of Large Numbers Wen, L. Am Math Month 1991 (https://www.jstor.org/discover/10.2307/2323947?uid=3738032&uid=2&uid=4&sid=21103621939777)

References

- Grimmett, G. R.; Stirzaker, D. R. (1992). *Probability and Random Processes, 2nd Edition*. Clarendon Press, Oxford. ISBN 0-19-853665-8.
- Richard Durrett (1995). *Probability: Theory and Examples, 2nd Edition*. Duxbury Press.
- Martin Jacobsen (1992). *Videregående Sandsynlighedsregning (Advanced Probability Theory) 3rd Edition*. HCØ-tryk, Copenhagen. ISBN 87-91180-71-6.
- Loève, Michel (1977). *Probability theory 1* (4th ed.). Springer Verlag.
- Newey, Whitney K.; McFadden, Daniel (1994). *Large sample estimation and hypothesis testing*. Handbook of econometrics, vol. IV, Ch. 36. Elsevier Science. pp. 2111–2245.
- Ross, Sheldon (2009). *A first course in probability* (8th ed.). Prentice Hall press. ISBN 978-0-13-603313-4.
- Sen, P. K; Singer, J. M. (1993). *Large sample methods in statistics*. Chapman & Hall, Inc.
- Seneta, Eugene (2013), "A Tricentenary history of the Law of Large Numbers", *Bernoulli*, **19** (4): 1088–1121, arXiv:1309.6488 (https://arxiv.org/abs/1309.6488), doi:10.3150/12-BEJSP12 (https://doi.org/10.3150%2F12-BEJSP12), S2CID 88520834 (https://api.semanticscholar.org/CorpusID:88520834)

External links

- "Law of large numbers" (https://www.encyclopediaofmath.org/index.php?title=Law_of_large_numbers), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]
- Weisstein, Eric W. "Weak Law of Large Numbers" (https://mathworld.wolfram.com/WeakLawofLargeNumbers.html). *MathWorld*.
- Weisstein, Eric W. "Strong Law of Large Numbers" (https://mathworld.wolfram.com/StrongLawofLargeNumbers.html). *MathWorld*.
- Animations for the Law of Large Numbers (https://web.archive.org/web/20081110071309/http://animation.yihui.name/prob:law_of_large_numbers) by Yihui Xie using the R package animation (https://cran.r-project.org/package=animation)
- Apple CEO Tim Cook said something that would make statisticians cringe (http://www.businessinsider.com/law-of-large-numbers-tim-cook-2015-2). "We don't believe in such laws as laws of large numbers. This is sort of, uh, old dogma, I think, that was cooked up by somebody [..]" said Tim Cook and while: "However, the law of large numbers has nothing to do with large companies, large revenues, or large growth rates. The law of large numbers is a fundamental concept in probability theory and statistics, tying together theoretical probabilities that we can calculate to the actual outcomes of experiments that we empirically perform. *explained Business Insider*

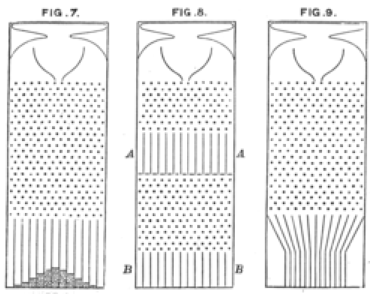
WIKIPEDIA

De Moivre–Laplace theorem

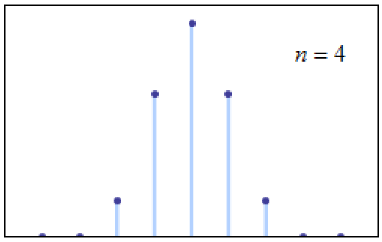
In probability theory, the **de Moivre–Laplace theorem**, which is a special case of the central limit theorem, states that the normal distribution may be used as an approximation to the binomial distribution under certain conditions. In particular, the theorem shows that the probability mass function of the random number of "successes" observed in a series of n independent Bernoulli trials, each having probability p of success (a binomial distribution with n trials), converges to the probability density function of the normal distribution with mean np and standard deviation $\sqrt{np(1-p)}$, as n grows large, assuming p is not 0 or 1.

The theorem appeared in the second edition of *The Doctrine of Chances* by Abraham de Moivre, published in 1738. Although de Moivre did not use the term "Bernoulli trials", he wrote about the probability distribution of the number of times "heads" appears when a coin is tossed 3600 times.[1]

This is one derivation of the particular Gaussian function used in the normal distribution.



Within a system whose bins are filled according to the binomial distribution (such as Galton's "bean machine", shown here), given a sufficient number of trials (here the rows of pins, each of which causes a dropped "bean" to fall toward the left or right), a shape representing the probability distribution of k successes in n trials (see bottom of Fig. 7) matches approximately the Gaussian distribution with mean np and variance $np(1-p)$, assuming the trials are independent and successes occur with probability p.



Consider tossing a set of n coins a very large number of times and counting the number of "heads" that result each time. The possible number of heads on each toss, k, runs from 0 to n along the horizontal axis, while the vertical axis represents the relative frequency of occurrence of the outcome k heads. The height of each dot is thus the probability of observing k heads when tossing n coins (a binomial distribution based on n trials). According to the de Moivre–Laplace theorem, as n grows large, the shape of the discrete distribution converges to the continuous Gaussian curve of the normal distribution.

Contents

Theorem

As n grows large, for k in the neighborhood of np we can approximate[2][3]

$$\binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}}\, e^{-\frac{(k-np)^2}{2npq}}, \qquad p+q=1,\ p,q>0$$

in the sense that the ratio of the left-hand side to the right-hand side converges to 1 as $n \to \infty$.

Proof

The theorem can be more rigorously stated as follows: $(X-np)/\sqrt{npq}$, with X a binomially distributed random variable, approaches the standard normal as $n \to \infty$, with the ratio of the probability mass of X to the limiting normal density being 1. This can be shown for an arbitrary nonzero and finite point c. On the unscaled curve for X, this would be a point k given by

$$k = np + c\sqrt{npq}$$

For example, with c at 3, k stays 3 standard deviations from the mean in the unscaled curve.

The normal distribution with mean μ and standard deviation σ is defined by the differential equation (DE)

$$f'(x) = -\frac{x-\mu}{\sigma^2} f(x) \text{ with initial condition set by the probability axiom } \int_{-\infty}^{\infty} f(x)\,dx = 1.$$

The binomial distribution limit approaches the normal if the binomial satisfies this DE. As the binomial is discrete the equation starts as a difference equation whose limit morphs to a DE. Difference equations use the discrete derivative (https://calculus.subwiki.org/wiki/discrete%20derivative), $p(k+1)-p(k)$, the change for step size 1. As $n \to \infty$, the discrete derivative becomes the continuous derivative. Hence the proof need show only that, for the unscaled binomial distribution,

$$\frac{f'(x)}{f(x)} \cdot \left(-\frac{\sigma^2}{x-\mu}\right) \to 1 \text{ as } n \to \infty.$$

The required result can be shown directly:

$$\begin{aligned}\frac{f'(x)}{f(x)}\frac{npq}{np-k} &= \frac{p(n,k+1)-p(n,k)}{p(n,k)}\frac{\sqrt{npq}}{-c}\\ &= \frac{np-k-q}{kq+q}\frac{\sqrt{npq}}{-c}\\ &= \frac{-c\sqrt{npq}-q}{npq+cq\sqrt{npq}+q}\frac{\sqrt{npq}}{-c}\\ &\to 1\end{aligned}$$

The last holds because the term $-cnpq$ dominates both the denominator and the numerator as $n \to \infty$.

As k takes just integral values, the constant c is subject to a rounding error. However, the maximum of this error, $0.5/\sqrt{npq}$, is a vanishing value.[4]

Alternate Proof

The proof consists of transforming the left-hand side (in the statement of the theorem) to the right-hand side by three approximations.

First, according to Stirling's formula, the factorial of a large number n can be replaced with the approximation

$$n! \simeq n^n e^{-n}\sqrt{2\pi n} \qquad \text{as } n \to \infty.$$

Thus

$$\begin{aligned}\binom{n}{k}p^k q^{n-k} &= \frac{n!}{k!(n-k)!}p^k q^{n-k}\\ &\simeq \frac{n^n e^{-n}\sqrt{2\pi n}}{k^k e^{-k}\sqrt{2\pi k}(n-k)^{n-k}e^{-(n-k)}\sqrt{2\pi(n-k)}}p^k q^{n-k}\\ &= \sqrt{\frac{n}{2\pi k(n-k)}}\frac{n^n}{k^k(n-k)^{n-k}}p^k q^{n-k}\\ &= \sqrt{\frac{n}{2\pi k(n-k)}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}\end{aligned}$$

Next, the approximation $\frac{k}{n} \to p$ is used to match the root above to the desired root on the right-hand side.

$$\begin{aligned}\binom{n}{k}p^k q^{n-k} &\simeq \sqrt{\frac{1}{2\pi n\frac{k}{n}\left(1-\frac{k}{n}\right)}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}\\ &\simeq \frac{1}{\sqrt{2\pi npq}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k} \qquad p+q=1\end{aligned}$$

Finally, the expression is rewritten as an exponential and the Taylor Series approximation for ln(1+x) is used:

$$\ln(1+x) \simeq x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

Then

$$\begin{aligned}
\binom{n}{k}p^k q^{n-k} &\simeq \frac{1}{\sqrt{2\pi npq}}\exp\left\{\ln\left(\left(\frac{np}{k}\right)^k\right)+\ln\left(\left(\frac{nq}{n-k}\right)^{n-k}\right)\right\}\\
&= \frac{1}{\sqrt{2\pi npq}}\exp\left\{-k\ln\left(\frac{k}{np}\right)+(k-n)\ln\left(\frac{n-k}{nq}\right)\right\}\\
&= \frac{1}{\sqrt{2\pi npq}}\exp\left\{-k\ln\left(\frac{np+x\sqrt{npq}}{np}\right)+(k-n)\ln\left(\frac{n-np-x\sqrt{npq}}{nq}\right)\right\}\\
&= \frac{1}{\sqrt{2\pi npq}}\exp\left\{-k\ln\left(1+x\sqrt{\frac{q}{np}}\right)+(k-n)\ln\left(1-x\sqrt{\frac{p}{nq}}\right)\right\} \qquad p+q=1\\
&= \frac{1}{\sqrt{2\pi npq}}\exp\left\{-k\left(x\sqrt{\frac{q}{np}}-\frac{x^2q}{2np}+\cdots\right)+(k-n)\left(-x\sqrt{\frac{p}{nq}}-\frac{x^2p}{2nq}-\cdots\right)\right\}\\
&= \frac{1}{\sqrt{2\pi npq}}\exp\left\{\left(-np-x\sqrt{npq}\right)\left(x\sqrt{\frac{q}{np}}-\frac{x^2q}{2np}+\cdots\right)+\left(np+x\sqrt{npq}-n\right)\left(-x\sqrt{\frac{p}{nq}}-\frac{x^2p}{2nq}-\cdots\right)\right\}\\
&= \frac{1}{\sqrt{2\pi npq}}\exp\left\{\left(-np-x\sqrt{npq}\right)\left(x\sqrt{\frac{q}{np}}-\frac{x^2q}{2np}+\cdots\right)-\left(nq-x\sqrt{npq}\right)\left(-x\sqrt{\frac{p}{nq}}-\frac{x^2p}{2nq}-\cdots\right)\right\}\\
&= \frac{1}{\sqrt{2\pi npq}}\exp\left\{\left(-x\sqrt{npq}+\frac{1}{2}x^2q-x^2q+\cdots\right)+\left(x\sqrt{npq}+\frac{1}{2}x^2p-x^2p-\cdots\right)\right\}\\
&= \frac{1}{\sqrt{2\pi npq}}\exp\left\{-\frac{1}{2}x^2q-\frac{1}{2}x^2p-\cdots\right\}\\
&= \frac{1}{\sqrt{2\pi npq}}\exp\left\{-\frac{1}{2}x^2(p+q)-\cdots\right\}\\
&\simeq \frac{1}{\sqrt{2\pi npq}}\exp\left\{-\frac{1}{2}x^2\right\}\\
&= \frac{1}{\sqrt{2\pi npq}}e^{\frac{-(k-np)^2}{2npq}}
\end{aligned}$$

Each "$\simeq$" in the above argument is a statement that two quantities are asymptotically equivalent as n increases, in the same sense as in the original statement of the theorem—i.e., that the ratio of each pair of quantities approaches 1 as $n \to \infty$.

Trivia

- *The Wall* is an example of a television game show that uses the De Moivre–Laplace theorem.[5]

See also

- Poisson distribution is an alternative approximation of the binomial distribution for large values of *n*.

Notes

1. Walker, Helen M (1985). "De Moivre on the law of normal probability" (http://www.york.ac.uk/depts/maths/histstat/demoivre.pdf) (PDF). In Smith, David Eugene (ed.). *A source book in mathematics* (https://archive.org/details/sourcebookinmath0000smit/page/78). Dover. p. 78 (https://archive.org/details/sourcebookinmath0000smit/page/78). ISBN 0-486-64690-4. "But altho' the taking an infinite number of Experiments be not practicable, yet the preceding Conclusions may very well be applied to finite numbers, provided they be great, for Instance, if 3600 Experiments be taken, make *n* = 3600, hence ½*n* will be = 1800, and ½√*n* 30, then the Probability of the Event's neither appearing oftner than 1830 times, nor more rarely than 1770, will be 0.682688."
2. Papoulis, Athanasios; Pillai, S. Unnikrishna (2002). *Probability, Random Variables, and Stochastic Processes* (4th ed.). Boston: McGraw-Hill. ISBN 0-07-122661-3.
3. Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. Volume 1. Wiley. Section VII.3. ISBN 0-471-25708-7.
4. Thamattoor, Ajoy (2018). "Normal limit of the binomial via the discrete derivative". *The College Mathematics Journal*. **49** (3): 216–217. doi:10.1080/07468342.2018.1440872 (https://doi.org/10.1080%2F07468342.2018.1440872). S2CID 125977913 (https://api.semanticscholar.org/CorpusID:125977913).
5. Roeder, Oliver (November 17, 2017). "What if God Were a Giant Game of Plinko?" (https://fivethirtyeight.com/features/what-if-god-were-a-giant-game-of-plinko/). *FiveThirtyEight*. Retrieved November 24, 2017.

LANTURI MARKOV
Lect. Dr. Costache Luminita

1 Lanţ Markov omogen. Criteriu de ergodicitate. Repartiţie staţionară

Definiţia 1. Fie $(X_n)_{n\geq 0}$ un şir de v. a. discrete, luând valori într-o mulţime finită sau numărabilă S, numită **spaţiul stărilor**.

a) Şirul $(X_n)_{n\geq 0}$ formează **un lanţ Markov** dacă pentru orice $n \in \mathbb{N}^*$ şi $i_0, \ldots, i_n \in S$, avem

$$P(X_n = i_n / X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) = P(X_n = i_n / X_{n-1} = i_{n-1})$$

b) Un lanţ Markov se numeşte **omogen** dacă probabilităţile

$$P(X_n = j / X_{n-1} = i) = p_{ij}$$

sunt independente de n.

In cazul când lanţul Markov are un număr finit de stări $1, 2, \ldots, N$, el se numeşte **finit**.

Matricea $P = \begin{pmatrix} p_{11} & \ldots & p_{1N} \\ \vdots & & \\ p_{N1} & \ldots & p_{NN} \end{pmatrix}$ se numeşte **matricea probabilităţilor de trecere (tranziţie)** a lanţului Markov considerat.

Proprietăţi

1. Linia i a matricei P, formată din elementele $p_{i1}p_{i2}\ldots p_{iN}$ conţine probabilităţile de trecere din starea i în stările $1, 2, \ldots, N$.

2. P este o **matrice sochastică** , adică este formată din elemente pozitive, iar suma elementelor fiecărei linii este 1.

3. Notând $p_{ij}(n) = P(X_n = j / X_0 = i)$, probabilitatea de a trece după n paşi din starea i în starea j, se obţine **matricea de trecere după** n **paşi** $P(n) = (p_{ij}(n))_{i,j}$ care verifică relaţia $P(n) = P^n$.

In particular, $P(n) = P(n-1)P$ şi $P(n) = PP(n-1)$, de unde rezultă ecuaţiile directe ale lui Kolmogorov $p_{ij}(n) = \sum\limits_k p_{ik}(n-1)p_{kj}, i, j \in S$, precum şi ecuaţiile inverse $p_{ij}(n) = \sum\limits_k p_{ik}p_{kj}(n-1), i, j \in S$.

4. Repartiţia v. a. $X_n \sim \begin{pmatrix} 1 & 2 & \ldots N \\ p_1(n) & p_2(n) & \ldots p_N(n) \end{pmatrix}$ este definită de **vectorul de probabilitate** $p(n) = (p_1(n), p_2(n), \ldots, p_N(n))$, unde
$p(n) = p(0)P^n$

Definiţia 2. Un lanţ Markov cu matricea de trecere P este **ergodic** dacă $\lim\limits_{n\to\infty} P^n = \Pi$, unde Π este o matrice stochastică , având toate liniile egale cu un anumit vector de probabilitate $\sigma = (\sigma_1\sigma_2\ldots\sigma_n)$ numit **repartiţia staţionară** a procesului.

Criteriu de ergodicitate Dacă $\exists n > 0$ astfel încât matricea P^n să aibă toate elementele strict pozitive, atunci lanţul este ergodic.

Găsirea repartiţiei staţionare Fie P matricea de trecere a unui lanţ Markov ergodic. Atunci distribuţia limită este unicul vector de probabilitate σ satisfăcând ecuaţia vectorială $\sigma P = \sigma$.

Observaţia 1. Dacă σ este repartiţia staţionară a unui lanţ Markov ergodic, atunci şirul distribuţiilor $p(n)$ la momentul n satisface relaţia

$$\lim_{n\to\infty} p(n) = \sigma$$

Lanțuri Markov

Lucruri de bază de știut (prin exerciții)

① În sesiune, dacă un student învață azi, atunci el va învăța și mâine cu probabilitatea 0,7, iar dacă el nu învață azi, atunci nu va învăța nici mâine cu probabilitatea 0,4.

a) Determinați probabilitatea ca dacă el învață azi, poimâine să nu învețe.

b) Determinați repartiția staționară.

Soluție: a) Lanțul Markov are 2 stări: I (învață) și N (nu învață).

Atunci matricea probabilităților de trecere, notată cu P este:

$$P = \begin{array}{c|cc} & I & N \\ \hline I & 0,7 & 0,3 \\ N & 0,6 & 0,4 \end{array}$$

$I \xrightarrow{0,7} I$ (adică elem. de pe poz. (1,1) este 0,7)

$N \xrightarrow{0,4} N$ (elem. de pe poz. (2,2) este 0,4)

completate astfel încât pe fiecare linie suma să fie 1

(P fiind matrice stochastică)

De azi până poimâine sunt 2 zile, deci trec 2 pași.

(la fel dacă am zice de joi până sâmbătă sau de luni până miercuri etc.)

Pentru că trebuie să determinăm probabilitatea ca ~~dacă~~ azi învață, atunci să nu învețe poimâine, aceasta înseamnă că avem de aflat elementul $p_{12}(2)$, adică elementul din matricea P^2 aflat pe poziția (1,2) (trecerea din starea I în starea N după 2 pași)

$$P^2 = \begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix}\begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix} = \begin{pmatrix} & 0,33 \\ & \end{pmatrix} = p_{12}(2)$$

(Elementul căutat se obține înmulțind linia 1 cu coloana 2)

Nici nu trebuie calculată toată matricea P^2.

b) Pentru a determina repartiția staționară trebuie verificat că lanțul este ergodic, adică $\exists n>0$ a.î. P^n să aibă toate elementele strict pozitive. În cazul nostru, $n=1$, $P=\begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix}$ are toate elem. strict pozitive, deci lanțul este ergodic.

Atunci repartiția staționară se determină din ecuația $\sigma P=\sigma$, unde σ este vectorul $\sigma=(\sigma_1 \ \ \sigma_2)$ cu $\sigma_1+\sigma_2=1$ (cf. Def. 10.2).

Așadar, $(\sigma_1 \ \ \sigma_2)\begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix}=(\sigma_1 \ \ \sigma_2) \Rightarrow$

$$\Rightarrow \begin{pmatrix} 0,7\sigma_1+0,6\sigma_2 \\ 0,3\sigma_1+0,4\sigma_2 \end{pmatrix}^t=(\sigma_1 \ \ \sigma_2) \Rightarrow \begin{cases} 0,7\sigma_1+0,6\sigma_2=\sigma_1 \\ 0,3\sigma_1+0,4\sigma_2=\sigma_2 \\ \sigma_1+\sigma_2=1 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} 0,6\sigma_2=0,3\sigma_1 \\ \sigma_1+\sigma_2=1 \end{cases} \Rightarrow \begin{cases} \sigma_1=2\sigma_2 \\ \sigma_1+\sigma_2=1 \end{cases} \Rightarrow 3\sigma_2=1 \Rightarrow \sigma_2=\frac{1}{3}$$

$$\sigma_1=\frac{2}{3}.$$

$\Rightarrow \sigma=\left(\frac{2}{3} \ \ \frac{1}{3}\right)$ repartiția staționară

9. Un lanţ Markov cu 2 stări poate fi folosit să modeleze o varietate largă de sisteme ce alterneazaă între stările ON şi OFF. După fiecare unitate de timp în starea OFF, sistemul trece pe ON cu probabilitatea p. După fiecare unitate de timp în starea ON, sistemul trece în OFF cu probabilitatea q. Folosind 0 şi 1 să notăm stările OFF şi ON, care este matricea probabilităţilor de trecere? Stiind că sistemul este în starea OFF la momentul 0, care este probabilitatea ca sistemul să fie în starea OFF la momentul $n = 33$?

R: $P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$, $p_{00}(33) = \frac{q+[1-(p+q)]^{33}p}{p+q}$

(L. Costache - Lecţii de matematici speciale, p. 147)

② Rezolvarea ex. 9 / pg. 146

Lanțul Markov cu 2 stări OFF (0) și ON (1) are matricea probabilităților de trecere

$$P = \begin{array}{c|cc} & 0 & 1 \\ 0 & \multicolumn{2}{c}{\multirow{2}{*}{}} \\ \end{array} \quad P = \begin{array}{cc} & \begin{array}{cc} 0 & \;\;1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \end{array}$$

OFF $\xrightarrow{p}$ ON $\quad (0 \xrightarrow{p} 1)$

ON $\xrightarrow{q}$ OFF $\quad (1 \xrightarrow{q} 0)$

Pentru că n este mare ($n=33$), nu vom ridica matricea P la puterea 33, ci vom calcula (cu noțiuni de la Algebră, semestrul I) matricea P^n.

Calculăm mai întâi valorile proprii din $\det(P-\lambda I_2)=0 \Rightarrow$

$$\Rightarrow \begin{vmatrix} 1-p-\lambda & p \\ q & 1-q-\lambda \end{vmatrix} = 0 \Rightarrow \lambda_1 = 1, \lambda_2 = 1-p-q$$

Determinăm vectorii proprii corespunzători din

$$Px = \lambda_1 x \Rightarrow \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \begin{cases} (1-p)x_1 + px_2 = x_1 \\ qx_1 + (1-q)x_2 = x_2 \end{cases} \Rightarrow x_1 = x_2 \Rightarrow V_{\lambda_1} = sp\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$$

$$Py = \lambda_2 y \Rightarrow \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (1-p-q)\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \begin{cases} (1-p)y_1 + py_2 = (1-p-q)y_1 \\ qy_1 + (1-q)y_2 = (1-p-q)y_2 \end{cases} \Rightarrow V_{\lambda_2} = sp\left\{\begin{pmatrix} \frac{-p}{p+q} \\ \frac{q}{p+q} \end{pmatrix}\right\}$$

Scriem matricea vectorilor proprii: $T = \begin{pmatrix} 1 & \frac{-p}{p+q} \\ 1 & \frac{q}{p+q} \end{pmatrix}$

Se calculează T^{-1} și obținem $T^{-1} = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ -1 & 1 \end{pmatrix}$

Atunci $P = TDT^{-1}$, unde $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

$$P^n = TD^nT^{-1} = \begin{pmatrix} 1 & -\frac{p}{p+q} \\ 1 & \frac{q}{p+q} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1-p-q)^n \end{pmatrix} \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ -1 & 1 \end{pmatrix} =$$

stările

$$= \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 \qquad\qquad\qquad\qquad & \qquad\qquad\qquad\qquad 1 \end{array} \\ \begin{pmatrix} \frac{q}{p+q} + \frac{p}{p+q}(1-p-q)^n & \frac{p}{p+q}\left[1-(1-p-q)^n\right] \\ \frac{q}{p+q}\left[1-(1-p-q)^n\right] & \frac{p}{p+q} + \frac{q}{p+q}(1-p-q)^n \end{pmatrix} \end{array}$$

Probabilitatea cerută este $p_{00}(33)$, adică elementul de pe poziția $(1,1)$ din matricea P^{33}, deci $\boxed{p_{00}(33) = \frac{q}{p+q} + \frac{p}{p+q}(1-p-q)^{33}}$

prima linie, prima coloană

Observație: 1) Dacă s-ar cere și în acest caz repartiția staționară, după ce se arată în prealabil că lanțul este ergodic (eventual după ce s-au specificat în enunț anumite condiții asupra lui p și q), atunci, deoarece deja s-a aflat P^n, e mai ușor să determinăm π din Def. 10.2, adică $\lim\limits_{n\to\infty} P^n = \Pi$, unde Π e o matrice stochastică cu toate liniile egale cu π

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix} \quad \left(\text{pt. că } \lim_{n\to\infty} (\underbrace{1-p-q}_{<1})^n = 0\right)$$

Deci $\sigma = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix}$.

Evident σ putea fi aflat și ca la ex. precedent, din ecuația $\sigma P = \sigma$ (cu $\sigma = (\sigma_1 \ \sigma_2)$, $\sigma_1 + \sigma_2 = 1$)

2) Dacă s-ar da $p(0) = (p_1, p_2)$ și s-ar cere $p(56)$, ar înseamna că avem $X_0 \sim \begin{pmatrix} 0 & 1 \\ p_1 & p_2 \end{pmatrix}$ ← stările OFF și ON.

$p_1 \ p_2$: $p(0)$

Din proprietatea 4, $p(56) = p(0)P^{56} = (p_1, p_2) \cdot P^{56} =$

$$= \left(p_1\left(\frac{q}{p+q} + \frac{p}{p+q}(1-p-q)^{56}\right) + p_2\left(\frac{q}{p+q}\left(1-(1-p-q)^{56}\right)\right),\ p_1 \cdot \frac{p}{p+q}\left[1-(1-p-q)^{56}\right] + p_2\left(\frac{p}{p+q} + \frac{q}{p+q}(1-p-q)^{56}\right)\right).$$

Lanţuri Markov

R Vodican

Definiţie Fie $N \in \mathbb{N}^*$, fixat şi $(X_n)_{n\in\mathbb{N}}$ un şir de variabile aleatoare care iau valori în mulţimea $\{1,2,\dots,N\} \overset{not}{=} S$, pe care o vom numi spaţiul stărilor. Spunem că şirul $(X_n)_{n\in\mathbb{N}}$ este un lanţ Markov omogen finit, dacă

$$(*) \quad P(X_{n+1}=i_{n+1} \mid X_n=i_n, X_{n-1}=i_{n-1}, \dots, X_0=i_0) = $$
$$= P(X_{n+1}=i_{n+1} \mid X_n=i_n) \overset{not}{=} p_{i_n i_{n+1}}$$

pentru orice n pentru care probabilitatea condiţionată are sens şi pentru $\forall \{i_0, i_1, \dots, i_n\} \subseteq S$.

p_{ij} se numeşte probabilitatea de trecere, într-un singur pas, din starea i în starea j, iar matricea $P=(p_{ij})_{i,j\in S}$ se numeşte matricea probabilităţilor de trecere într-un singur pas.

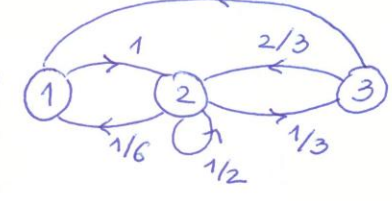
Fie $p_i \overset{not}{=} P(X_0=i)$, $i\in S$. Atunci, vectorul $\Pi_0 \overset{not}{=} (p_i)_{i\in S}$ se numeşte repartiţia iniţială (Π_0 conţine probabilităţile ca lanţul să pornească dintr-o anumită stare).

X_n reprezintă starea lanţului Markov la momentul n, iar relaţia $(*)$ arată că „viitorul" (starea lanţului la momentul $n+1$) depinde doar de „prezent" (starea lanţului la momentul n)

Unui lanţ Markov finit şi omogen i se poate asocia un graf orientat $G=(S,\Gamma)$, unde S este spaţiul stărilor lanţului, iar $\Gamma \subseteq S\times S$ este o relaţie definită prin: $(i,j)\in\Gamma \Leftrightarrow p_{ij}>0$. Arcele orientate vor desemna trecerea lanţului din starea i în starea j cu probabilitatea $p_{ij}\,(>0)$ pentru $i,j\in S$.

Exemplu Să se deseneze graful asociat lanţului Markov a cărui matrice de trecere este

$$P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix} \Rightarrow$$



<u>Propoziția 1</u> Un lanț Markov este complet determinat de repartiția inițială și de probabilitățile de trecere.
(Propoziția afirmă că, dacă sunt cunoscute repartiția inițială și probabilitățile de trecere într-un pas, atunci se poate calcula $P(X_0=i_0, X_1=i_1, \ldots, X_n=i_n)$)

<u>Propoziția 2</u> Dacă $\Pi_0=(p_1, p_2, \ldots, p_N)$ este repartiția inițială a unui lanț Markov și $P=(p_{ij})_{i,j\in S}$ ($S \stackrel{not}{=} \{1, \ldots, N\}$) matricea probabilităților de trecere într-un pas, atunci:

1) $p_i \geq 0$, $\forall i=\overline{1,N}$ și $\sum_{i=1}^{N} p_i = 1$;

2) $p_{ij} \geq 0$, $\forall i, j \in \{1, \ldots, N\}$ și $\sum_{j=1}^{N} p_{ij} = 1$, $\forall i=\overline{1,N}$

Justificările sunt simple exerciții de calcul:

1) $\sum_{i=1}^{N} p_i = \sum_{i=1}^{N} P(X_0=i) = P\left(\bigcup_{i=1}^{N}(X_0=i)\right) = P(\text{ev. sigur}) = 1$

2) Fie $i\in\{1,\ldots,N\}$. Atunci: $\sum_{j=1}^{N} p_{ij} = \sum_{j=1}^{N} P(X_n=j \mid X_{n-1}=i) =$

$$= \sum_{j=1}^{N} \frac{P(X_n=j, X_{n-1}=i)}{P(X_{n-1}=i)} = \frac{P\left(\left\{\bigcup_{j=1}^{n}(X_n=j)\right\}\cap(X_{n-1}=i)\right)}{P(X_{n-1}=i)} =$$

$$\underset{\uparrow}{=} \frac{P(X_{n-1}=i)}{P(X_{n-1}=i)} = 1.$$

$\left\{\bigcup_{j=1}^{n}(X_n=j)\right\} = \text{ev. sigur}$

<u>Definiție</u> O matrice $P=(p_{ij})_{1\leq i,j\leq N}$ cu elemente numere reale <u>astfel</u> încât $p_{ij}\geq 0$ $\forall i,j\in\{1,\ldots,N\}$ și $\sum_{j=1}^{N} p_{ij} = 1$, $\forall i=\overline{1,N}$ se numește <u>matrice stochastică</u>

Așadar, din <u>Prop 2</u> $\Rightarrow$ matricea probabilităților de trecere într-un pas asociată unui lanț Markov este o matrice stochastică.

<u>Observație</u> Produsul a două matrice stochastice

este o matrice stochastică: Fie $A=(a_{ij})_{1\le i,j\le N}$ și $B=(b_{ij})_{1\le i,j\le N}$ două matrice stochastice și $AB \overset{not}{=} C = (c_{ij})_{1\le i,j\le N}$. Atunci, pentru $i,j\in\{1,2,\dots,N\}$:

$$c_{ij}=\sum_{k=1}^{N} a_{ik}b_{kj} \Rightarrow c_{ij}\ge 0 \text{ și } \sum_{j=1}^{N} c_{ij}=\sum_{j=1}^{N}\left(\sum_{k=1}^{N} a_{ik} b_{kj}\right)=\sum_{k=1}^{N} a_{ik}\underbrace{\left(\sum_{j=1}^{N} b_{kj}\right)}_{\substack{=1\\ \uparrow \\ B \text{ matr. stochastică}}} \underset{\substack{\uparrow\\ A \text{ matr. stochastică}}}{=} 1$$

<u>Propoziția 3</u> Fie S o mulțime cel mult numărabilă de indici, un șir $(p_i)_{i\in S}$ și o matrice $P=(p_{ij})_{i,j\in S}$ a. î. $p_i\ge 0$, $\forall i\in S$, $\sum_{i\in S} p_i=1$ și $p_{ij}\ge 0$, $\forall i,j\in S$, $\sum_{j\in S} p_{ij}=1$, $\forall i\in S$. Atunci există un câmp de probabilitate $(\Omega,\mathcal{K},P)$ și un lanț Markov $(X_n)_{n\in\mathbb{N}}$ pe acest câmp pentru care S este spațiul stărilor, $\pi_0=(p_i)_{i\in S}$ este repartiția inițială a lanțului, iar P matricea probabilităților de trecere într-un pas.

<u>Probabilități de trecere după n pași</u>

Fie $(X_n)_{n\in\mathbb{N}}$ un lanț Markov omogen finit cu spațiul stărilor $S=\{1,2,\dots,N\}$ și $P=(p_{ij})_{1\le i,j\le N}$ matricea probabilităților de trecere într-un pas.

<u>Definiție</u> Pentru $n\in\mathbb{N}^*$: $p_{ij}^{(n)} \overset{not}{=} P(X_{n+k}=j \mid X_k=i)$ se numește <u>probabilitatea de trecere din starea</u> i în starea j <u>în n pași</u>

<u>Propoziția 4</u> Probabilitățile de trecere în n pași verifică relația

$$p_{ij}^{(n+1)}=\sum_{k=1}^{N} p_{ik}^{(n)} p_{kj}\,,\ i,j\in S$$

numită ecuația <u>CHAPMAN - KOLMOGOROV</u>

Din Prop 4 $\Rightarrow$ matricea probabilităților de trecere în n pași $P^{(n)} = (p_{ij}^{(n)})_{i,j\in S}$ verifică relațiile $P^{(n)} = P^n$ și $P^{(n+1)} = P^{(n)} \cdot P$ ∎

Propoziția 5 Dacă $(X_n)_{n\in\mathbb{N}}$ este un lanț Markov pentru care $\exists\, k>0$ a.î. P^k să aibă toate elementele strict pozitive, atunci există $\lim\limits_{n\to\infty} P^n \overset{not}{=} Q$, unde Q este o matrice stochastică, cu toate liniile identice. Fie π vectorul reprezentând una din aceste linii. Atunci π se numește repartiția staționară a lanțului Markov.

Proprietate Atunci când există, repartiția staționară este unica repartiție σ pentru care $\sigma P = \sigma$.

Exerc 1) Se consideră două urne A și B. Urna A are două bile albe, iar urna B are 3 bile negre. Din fiecare urnă se extrage câte o bilă, schimbându-se bilele între ele: cea extrasă din urna A se pune în urna B și invers.

Se spune că sistemul de urne se află în starea „i" dacă în urna A sunt „i" bile negre. Să se justifice că procesul descris este un lanț Markov și să se scrie matricea probabilităților de trecere.

Stările sistemului:

	urna A	urna B
starea 0	○ ○	● ● ●
starea 1	○ ●	○ ● ●
starea 2	● ●	○ ○ ●

Probabilitățile de trecere:

$p_{00} = 0$ (e imposibil ca schimbând două bile între ele sistemul să rămână în starea 0)

$p_{01} = 1$ (din starea 0, schimbând două bile între ele, sistemul nu poate trece decât în starea 1)

$p_{02} = 0$ (e imposibil ca printr-o singură schimbare de bile să se treacă din starea 0 în starea 2)

Să notăm:
„$a \in A_i$": se extrage o bilă albă din urna A aflată în starea i"
„$n \in A_i$": se extrage o bilă neagră din urna A aflată în starea i"
Asemănător, vom folosi notațiile „$a \in B_i$" și „$n \in B_i$"
Atunci: $p_{10} = P(a \in B_1, n \in A_1) \underset{\uparrow}{=} P(n \in A_1) \cdot P(a \in B_1) =$
$(a \in B_1), (n \in A_1)$ evenim. indep.
$= \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

$p_{11} = P((a \in A_1, a \in B_1) \cup (n \in A_1, n \in B_1)) = P(a \in A_1, a \in B_1) +$
$+ P(n \in A_1, n \in B_1) = P(a \in A_1) \cdot P(a \in B_1) + P(n \in A_1) \cdot P(n \in B_1) =$
$= \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{2}$

$p_{12} = P(a \in A_1, n \in B_1) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$

$p_{20} = 0$ (sistemul de urne nu poate trece din starea 2 în starea 0)

$p_{21} = P(n \in A_2, a \in B_2) = \frac{2}{2} \cdot \frac{2}{3} = \frac{2}{3}$

$p_{22} = P(n \in A_2, n \in B_2) = \frac{2}{2} \cdot \frac{1}{3} = \frac{1}{3}$.

Matricea probabilităților de trecere într-un pas este

$$P = \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1/6 & 1/2 & 1/3 \\ 0 & 2/3 & 1/3 \end{pmatrix}$$

Faptul că starea viitoare a sistemului de urne este influențată doar de starea actuală, nu și de cele anterioare ne permite să afirmăm că șirul de v.a. $(X_n)_{n \in \mathbb{N}}$, unde X_n reprezintă starea sistemului de urne la momentul n, este un lanț Markov.

O altă justificare a acestei afirmații este faptul că matricea P este o matrice stochastică (v. Prop 3)

Ex2 O stație de deservire poate servi clienți la momentele 0, 1, 2, ... Să notăm cu X_n variabila aleatoare ale cărei valori reprezintă numărul clienților sosiți în stație în intervalul de timp $(n, n+1)$. Presupunem că X_n, $n \in \mathbb{N}$ sunt variabile aleatoare, identic repartizate cu $P(X_n = k) \overset{not}{=} p_k$, $\forall n, k \in \mathbb{N}$. și că în stație există loc de așteptare pentru m clienți, inclusiv clientul servit. Persoanele care ajung în stație și găsesc m clienți, pleacă. Așadar, X_n are matricea de repartiție:

$$X_n : \begin{pmatrix} 0 & 1 & 2 & \dots & k & \dots \\ p_0 & p_1 & p_2 & \dots & p_k & \dots \end{pmatrix} \text{ cu } \sum_{k=0}^{\infty} p_k = 1$$

Fie Y_n numărul clienților prezenți în stație la momentul n (inclusiv clientul servit). Deoarece Y_{n+1} depinde doar de Y_n și X_n, Y_n sunt v. a. independente $\Rightarrow (Y_n)_{n \in \mathbb{N}}$ este un lanț Markov. Să se determine matricea probabilităților de trecere într-un pas corespunzătoare acestui lanț Markov.

Să observăm că Y_{n+1} reprezintă suma între numărul clienților prezenți în stație la momentul n, mai puțin cel servit, și numărul clienților care sosesc în intervalul $(n, n+1)$, dacă această sumă este $\leq m$ și $Y_{n+1} = m$, în caz contrar. Mai exact:

$$Y_{n+1} = \begin{cases} Y_n - 1 + X_n, & \text{dacă } 0 \leq X_n + Y_n - 1 \leq m \text{ și } 1 \leq Y_n \leq m \\ X_n, & \text{dacă } Y_n = 0 \text{ și } 0 \leq X_n \leq m \\ m, & \text{altfel} \end{cases}$$

(Y_n ia valori în mulțimea $\{0, 1, \dots, m\}$)

Fie $P = (p_{ij})_{0 \leq i,j \leq m}$ matricea probabilităților de trecere într-un pas asociată lanțului Markov $(Y_n)_{n \in \mathbb{N}}$.

Pentru $j \in \{0, 1, \dots, m\}$: $p_{0j} = P(Y_{n+1} = j \mid Y_n = 0) =$

$$= \begin{cases} P(X_n = j) = p_j & \text{pentru } 0 \leq j \leq m-1 \\ P(X_n \geq m) = \sum^{\infty} p_i, & \text{pentru } j = m \end{cases}$$

Pentru $i \in \{1, -, m\}$ și $j \in \{0, 1, -, m\}$

$$p_{ij} = P(Y_{n+1} = j \mid Y_n = i) =$$

$$= \begin{cases} P(X_n = j-(i-1)) = p_{j-i+1}, \text{ pentru } i-1 \le j \le m-1 \\ P(X_n \ge m+1-i) = \sum_{j=m}^{\infty} p_{j-i+1}, \text{ pentru } j = m \\ 0, \text{ altfel} \end{cases}$$

Dacă notăm $q_m = \sum_{k=m}^{\infty} p_k$, obținem:

$$P = \begin{pmatrix} p_0 & p_1 & p_2 \cdots & p_{m-1} & q_m \\ p_0 & p_1 & p_2 \cdots & p_{m-1} & q_m \\ 0 & p_0 & p_1 \cdots & p_{m-2} & q_{m-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & p_1 & q_2 \\ 0 & 0 & 0 \cdots & p_0 & q_1 \end{pmatrix}$$

Ex 3 Un lanț Markov are două stări: 1 și 2. Matricea probabilităților de trecere într-un pas corespunzătoare este $P = \begin{pmatrix} p & q \\ q & p \end{pmatrix} \left(\stackrel{not}{=} \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \right)$

cu $p+q=1$, $0<p,q<1$. Să se determine matricea probabilităților de trecere în trei pași și să se calculeze $P(X(0)=1 \mid X(2)=1)$, știind că $P(X(0)=1)=\alpha$ și $P(X(0)=2)=\beta$.

Matricea probabilităților de trecere în trei pași este P^3. Vom calcula doar P^2 de ale cărei elemente vom avea nevoie pentru a obține $P(X(0)=1 \mid X(2)=1)$.

$$P^2 = \begin{pmatrix} p^2+q^2 & 2pq \\ 2pq & p^2+q^2 \end{pmatrix} \left(\stackrel{not}{=} \begin{pmatrix} p_{11}^{(2)} & p_{12}^{(2)} \\ p_{21}^{(2)} & p_{22}^{(2)} \end{pmatrix} \right)$$

Folosind definiția probabilității condiționate $\Rightarrow$

$$\Rightarrow P(X(0)=1 \mid X(2)=1) = \frac{P(X(0)=1,\ X(2)=1)}{P(X(2)=1)} = \frac{\overbrace{P(X(2)=1 \mid X(0)=1)}^{p_{11}^{(2)}}}{P(X(2)=1)} \cdot$$

$$\cdot \frac{\underbrace{P(X(0)=1)}_{\alpha}}{} = \frac{(p^2+q^2)\alpha}{P(X(2)=1)}$$

Pentru a calcula $P(X(2)=1)$ vom folosi formula probabilității totale și matricea probabilităților de trecere în doi pași P^2.

$P(X(2)=1) = P(X(2)=1 \mid X(0)=1) \cdot P(X(0)=1) + P(X(2)=1 \mid X(0)=2) \cdot P(X(0)=2) = p_{11}^{(2)} \cdot \alpha + p_{21}^{(2)} \cdot \beta = (p^2+q^2) \cdot \alpha + 2pq \cdot \beta$. Așadar,

$$P(X(0)=1 \mid X(2)=1) = \frac{(p^2+q^2)\alpha}{(p^2+q^2)\alpha + 2pq \cdot \beta}$$

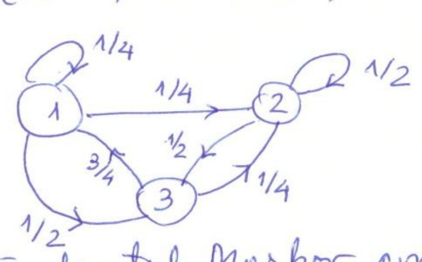
<u>Observație</u> Cum $X(0)$ se poate afla în starea 1 sau în starea 2 $\Rightarrow \alpha+\beta=1$

<u>Ex 4</u> Un lanț Markov are matricea de trecere

$$P = \begin{pmatrix} 1/4 & 1/4 & 1/2 \\ 0 & 1/2 & 1/2 \\ 3/4 & 1/4 & 0 \end{pmatrix}$$ și repartiția inițială $\pi_0 : (1/3 \ 1/3 \ 1/3)$

Se cere: a) Să se reprezinte graful asociat lanțului; b) să se determine distribuția sistemului după un pas; c) să se calculeze $P(X_1=2, X_2=3, X_3=1 \mid X_0=1)$ și d) să se calculeze $P(X_0=3, X_1=2, X_2=1)$

a) Graful asociat:



$(X_n)_{n\in N}$

b) Să presupunem că lanțul Markov considerat $(X_n)_{n\in N}$ are stările 1, 2 și 3. Din distribuția inițială $\Rightarrow P(X_0=1) = P(X_0=2) = P(X_0=3) = 1/3$. Să notăm $P = (p_{ij})_{1 \le i,j \le 3}$.

Atunci: $p_1 \overset{not}{=} P(X_1=1) \underset{\substack{\uparrow \\ \text{form. probab. totale}}}{=} \sum_{i=1}^{3} P(X_1=1 \mid X_0=i)\cdot P(X_0=i) =$

$= \sum_{i=1}^{3} p_{i1}\cdot P(X_0=i) =$ „primul element al liniei $\pi_0 \cdot P$".

La fel se arată că, pentru $j\in\{2,3\}$: $p_j \overset{not}{=} P(X_1=j) =$

$=$ „elementul j al liniei $\pi_0 P$"

Așadar, distribuția sistemului după un pas este:

$\pi_0 P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$

c) $P(X_1=2, X_2=3, X_3=1 \mid X_0=1) = \dfrac{P(X_0=1, X_1=2, X_2=3, X_3=1)}{P(X_0=1)} =$

$(X_n)_n$ lanț M. $\longrightarrow P(X_3=1 \mid X_2=3) = p_{31}$

$= \dfrac{\overbrace{P(X_3=1 \mid X_2=3, X_1=2, X_0=1)}^{=P(X_3=1\mid X_2=3)=p_{31}} \cdot P(X_2=3, X_1=2, X_0=1)}{P(X_0=1)} =$

$= \dfrac{p_{31}\cdot P(X_2=3 \mid X_1=2, X_0=1)\cdot P(X_1=2, X_0=1)}{P(X_0=1)} = \dfrac{p_{31}\cdot \overbrace{P(X_2=3\mid X_1=2)}^{p_{23}}}{\cancel{P(X_0=1)}}\cdot P(X_1=2 \mid X_0=1)\cdot \cancel{P(X_0=1)} = p_{31}\cdot p_{23}\cdot p_{12} = \frac{3}{4}\cdot\frac{1}{2}\cdot\frac{1}{4} = \frac{3}{32}.$

d) $P(X_0=3, X_1=2, X_2=1) = P(X_2=1 \mid X_1=2, X_0=3)\cdot P(X_1=2, X_0=3)$

$= \underbrace{P(X_2=1 \mid X_1=2)}_{=p_{21}=0}\cdot P(X_1=2, X_0=3) = 0$

Ex5 Un student merge la facultate cu tramvaiul sau cu autobuzul. Dacă într-o zi merge cu tramvaiul, a doua zi merge cu autobuzul. Dacă într-o zi merge cu autobuzul, a doua zi merge cu tramvaiul sau cu autobuzul (pentru a alege, aruncă o monedă și dacă obține stema, a doua zi va merge cu autobuzul, iar dacă va obține banul, va merge cu tramvaiul). Se cere: a) dacă în prima zi a mers cu autobuzul, care este probabilitatea ca în cea de-a 100-a zi să meargă cu tramvaiul? b) pe termen lung, să se stabilească ce mijloc de transport este mai probabil să aleagă studentul?

a) Fie X_n variabila aleatoare care ia valoarea 0 sau 1 după cum în ziua n, studentul circulă cu tramvaiul sau cu autobuzul. Conform enunțului, valorile lui X_{n+1} depind doar de cele ale lui X_n, așadar $(X_n)_{n\in\mathbb{N}^*}$ este un lanț Markov cu două stări:

starea 1: circulă cu tramvaiul

starea 2: circulă cu autobuzul

Conform enunțului, matricea probabilităților de trecere este

$$P=\begin{pmatrix}0 & 1\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix}$$

Dacă în prima zi, studentul circulă cu autobuzul, repartiția inițială va fi $(0,1)$, iar repartiția din cea de-a 100-a zi va fi $(0,1)\cdot P^{99}$.

Să calculăm P^n $(n\geq 1)$. Vom arăta că matricea P diagonalizează și apoi, vom calcula P^n folosind acest fapt.

$\det(P-\lambda I_2)=\frac{1}{2}(2\lambda^2-\lambda-1)=(\lambda-1)\left(\lambda+\frac{1}{2}\right)\Rightarrow P$ are doar valori proprii simple $\Rightarrow P$ diagonalizează $\Rightarrow$

$\Rightarrow \exists T\in M_2(\mathbb{R})$ nesingulară a.î. $T^{-1}PT=\begin{pmatrix}1 & 0\\ 0 & -1/2\end{pmatrix}\overset{not}{=}D$

(Reamintim că putem alege T a.î. să aibă pe coloane vectorii proprii ai lui P corespunzători valorilor proprii 1, respectiv $-1/2$).

$$\underset{\substack{\parallel\\ 1}}{P-\lambda_1 I_2}=P-I_2=\begin{pmatrix}-1 & 1\\ \frac{1}{2} & -\frac{1}{2}\end{pmatrix}\Rightarrow (P-I_2)\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}\Leftrightarrow x=y$$

$\Rightarrow \begin{pmatrix}x\\ y\end{pmatrix}=x\begin{pmatrix}1\\ 1\end{pmatrix}\Rightarrow \left\{v_1=\begin{pmatrix}1\\ 1\end{pmatrix}\right\}$ o bază pentru $\operatorname{Ker}(P-I_2)$

$$P-\lambda_2 I_2=P+\frac{1}{2}I_2=\begin{pmatrix}1/2 & 1\\ 1/2 & 1\end{pmatrix}\Rightarrow \left(P+\frac{1}{2}I_2\right)\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}\Leftrightarrow$$

$\Leftrightarrow x+2y=0\Rightarrow \begin{pmatrix}x\\ y\end{pmatrix}=y\begin{pmatrix}-2\\ 1\end{pmatrix}\Rightarrow \left\{v_2=\begin{pmatrix}-2\\ 1\end{pmatrix}\right\}$ o bază pentru $\operatorname{Ker}\left(P+\frac{1}{2}I_2\right)$.

Așadar, pentru $T = (v_1 | v_2) = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$ avem

$$P = TDT^{-1} \underset{\text{inducție}}{\Rightarrow} P^n = TD^nT^{-1} \underset{\uparrow}{=} \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (-\frac{1}{2})^n \end{pmatrix} \cdot$$

$$T^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$$

$$\cdot \frac{1}{3}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 + 2\cdot(-\frac{1}{2})^n & 2 - 2\cdot(-\frac{1}{2})^n \\ 1 - (-\frac{1}{2})^n & 2 + (-\frac{1}{2})^n \end{pmatrix}$$

$\Rightarrow$ repartiția din cea de-a 100 a zi este:

$$(0, 1)\, P^{99} = \frac{1}{3}\left(1 - \left(-\frac{1}{2}\right)^{99},\ 2 + \left(-\frac{1}{2}\right)^{99}\right) =$$

$= \left(\frac{1}{3} + \frac{1}{3\cdot 2^{99}},\ \frac{2}{3} - \frac{1}{3\cdot 2^{99}}\right) \Rightarrow$ probabilitatea ca în ziua 100 să circule cu tramvaiul este $\frac{1}{3} + \frac{1}{3\cdot 2^{99}}$

b) Dacă (a, b) este o repartiție inițială, repartiția peste o perioadă lungă de timp va fi $\lim_{n\to\infty} (a, b) P^n =$

$= (a, b)\cdot \frac{1}{3}\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = (a+b)\left(\frac{1}{3}, \frac{2}{3}\right) \Rightarrow$ pe o perioadă mai lungă de timp, este mai probabil ca studentul să circule cu autobuzul.

<u>Obs</u> Repartiția $\left(\frac{1}{3}\ \ \frac{2}{3}\right)$ este oricare dintre liniile matricei $\lim_{n\to\infty} P^n = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$. Așa cum am mai menționat, această repartiție se numește <u>repartiția staționară</u> a lanțului Markov. Ea nu există pentru orice lanț Markov (O condiție suficientă este precizată la pag 4 a acestui seminar).

Când nu avem nevoie de matricea P^n, putem determina repartiția staționară (dacă există) din relația $\sigma P = \sigma$, așa cum vom proceda în problema următoare.

Ex 6 Departamentul de servicii al unei universități poate cumpăra calculatoare de la două firme: A și B. Notăm cu 1 starea de funcționare a unui calculator și cu 2 starea de nefuncționare. Matricele stochastice de funcționare zilnică a unui calculator în cele două firme sunt:

$$P = \begin{pmatrix} 0,95 & 0,05 \\ 0,9 & 0,1 \end{pmatrix} \text{ pentru firma A și}$$

$$Q = \begin{pmatrix} 0,98 & 0,02 \\ 0,85 & 0,15 \end{pmatrix} \text{ pentru firma B.}$$

De la care firmă este mai avantajos să se cumpere calculatoarele având în vedere că prețul este același?

Având în vedere că elementele matricelor P și Q sunt strict pozitive, conf. prop. 5, pentru ambele lanțuri Markov există distribuții staționare. Fie π, respectiv σ aceste distribuții.

Dacă $\pi = (a, b)$, din $\pi P = \pi \Rightarrow \begin{cases} 0,05a - 0,9b = 0 \\ a + b = 1 \end{cases} \Leftrightarrow$

$$\Leftrightarrow a = \frac{18}{19},\ b = \frac{1}{19} \Rightarrow \pi = \left(\frac{18}{19}, \frac{1}{19}\right)$$

Dacă $\sigma = (c, d)$, din $\sigma Q = \sigma \Rightarrow \begin{cases} c \cdot 0,98 + d \cdot 0,85 = c \\ c + d = 1 \end{cases} \Leftrightarrow$

$$\Leftrightarrow c = \frac{85}{87},\ d = \frac{2}{87} \Rightarrow \sigma = \left(\frac{85}{87}, \frac{2}{87}\right)$$

Cum $\frac{85}{87} > \frac{18}{19} \Rightarrow$ un calculator de la firma B are șanse, pe termen lung, să funcționeze mai mult decât un calculator de la firma A.

LANTURI MARKOV. PROBLEME
Lect. Dr. Costache Luminita

1. Să se calculeze matricea probabilităţilor de trecere după 2 paşi, respectiv 3 paşi pentru lanţul Markov a cărui matrice de trecere este $P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$

2. Fie un lanţ Markov dat de matricea de trecere $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$. Să se calculeze probabilităţile staţionare stabilind mai întâi că lanţul este ergodic.

3. Vremea în ţara vrăjitorului din Oz este un lanţ Markov omogen cu 3 stări :s_1 : "ploaie", s_2 : "vreme bună ", s_3 : "zăpadă : şi matricea probabilităţilor de trecere $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$. Să se calculeze :

   a) probabilitatea ca după 3 zile de la o zi cu vreme bună să avem o zi cu zăpadă ;

   b) repartiţia staţionară .

4. Calculaţi repartiţia staţionară a unui lanţ Markov cu matricea de trecere $P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$.

5. Un cumpărător preferă două tipuri de mezeluri: salam de Sibiu şi şuncă de Praga. Dacă azi cumpără salam de Sibiu există o şansă de 0,75 să cumpere şi mâine, iar dacă azi cumpără şuncă de Praga există o şansă de 0,6 să cumpere şi mâine. Determinaţi probabilitatea ca dacă vineri cumpără salam, duminică să cumpere şuncă .

6. Presupunem că o persoană face cumpărături de la Carrefour şi Kaufland. Dacă azi cumpără de la Carrefour, mâine va cumpăra de la Kauffland cu probabilitatea 0,45, iar dacă azi cumpără de la Kaufland, mâine va cumpăra de la Carrefour cu probabilitatea 0,6. Dacă marţi cumpără de la Carrefour, care e probabilitatea ca joi să cumpere tot de acolo?

7. Un şofer îşi alimentează maşina de la OMV sau Lukoil. Dacă azi pune benzină de la Lukoil, mâine va pune de la OMV cu probabilitatea 0,6, iar dacă azi pune benzină de la OMV, mâine va pune de la Lukoil cu probabilitatea 0,3.Dacă azi e miercuri şi a alimentat de la Lukoil, care

este probabilitatea ca pentru excursia de sâmbătă să alimenteze de la OMV?

8. In Evul Mediu, universităţile Harvard, Dartmouth şi Yale admiteau numai studenţi bărbaţi. Presupunem că în acea perioadă $80^0/_0$ dintre fiii celor de la Harvard mergeau tot la Harvard şi restul la Yale, $40^0/_0$ dintre fiii celor de la Yale mergeau la Yale şi restul se împărţeau în mod egal între Harvard şi Dartmouth şi $70^0/_0$ dintre fiii celor de la Dartmouth mergeau tot la Dartmouth, $20^0/_0$ la Harvard şi $10^0/_0$ la Yale. Scrieţi matricea probabilităţilor de trecere şi calculaţi probabilitatea ca dacă bunicul a fost la Yale, nepotul să urmeze cursurile Universităţii Harvard.

9. Intr-un anumit oraş există 3 profesii P_1, P_2, P_3. Copiii unui tată cu profesia P_1, P_2, P_3 păstrează aceeaşi profesie cu probabilitatea $\frac{3}{5}, \frac{2}{3}$, respectiv $\frac{1}{4}$. iar dacă nu o păstrează , atunci aleg una dintre celelalte două profesii cu probabilităţi egale. Să se calculeze : a) repartiţia celor 3 profesii în generaţia următoare , dacă în generaţia actuală $20^0/_0$ din adulţi au profesia P_1, $30^0/_0$ din adulţi au profesia P_2 şi $50^0/_0$ din adulţi au profesia P_3; b) repartiţia staţionară .

10. Un meteorolog a dezvoltat următorul model pentru prezicerea vremii. Dacă plouă azi, proabilitatea este 0,6 că va ploua şi mâine şi 0,4 că nu va mai ploua mâine. Dacă nu plouă azi, probabilitatea este 0,2 că va ploua şi mâine şi 0,8 că nu va ploua mâine.

    a) Găsiţi matricea de trecere P pentru acest lanţ Markov şi matricea de trecere după 2 paşi.

    b) Dacă plouă azi, care este probabilitatea că va ploua şi poimâine?

    c) Dacă nu plouă azi, care este probabilitatea ca să plouă şi poimâine?

11. Fie un lanţ Markov a cărui matrice de trecere este $P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}$.

    a) Să se calculeze matricea probabilităţilor de trecere după 2, respectiv 3 paşi.

    b) Lanţul este ergodic?

    c) Dacă este ergodic, atunci să se determine probabilităţile limită .

12. Presupunem că înainte de a fi făcută o evidenţă a legăturii dintre fumat şi bolile respiratorii, $40^0/_0$ din adulţii de sex masculin erau fumători şi $60^0/_0$ erau nefumători. La un an dupaă această evidenţă a fost făcută public, $30^0/_0$ dintre fumători s-au oprit din fumat, în timp ce $10^0/_0$ din nefumători au început să fumeze.

a) Scrieţi matricea de trecere a lanţului Markov cu 2 stări;

b) Reprezentaţi distribuţia iniţială a fumătorilor şi nefumătorilor ca un vector probabilitate;

c) Găsiţi vectorul probabilitate ce descrie distribuţia fumătorilor şi nefumătorilor după un an;

d) Găsiţi vectorul probabilitate ce descrie distribuţia fumătorilor şi nefumătorilor după 2 ani.

13. In modelul stochastic de învăţare bazat pe teoria selectării stimulilor propus de W.K. Estes în 1950, se consideră un lanţ Markov cu 2 stări. Astfel starea 1 semnifică faptul că subiectul a învăţat, de exemplu, să primească o alună sau să evite un şoc electric. Starea 2 semnifică faptul că subiectul nu a învăţat încă . Se presupune că de îndată ce subiectul a învăţat el nu mai uită , iar dacă nu a învăţat încă , el va reuşi cu probabilitatea α să înveţe după fiecare încercare. Să se determine matricea de trecere şi să se calculeze $p_{21}(n)$.

14. Matricea probabilităţilor de trecere ale unui lanţ Markov omogen este $P = \begin{pmatrix} 0,1 & 0,5 & 0,4 \\ 0,6 & 0,2 & 0,2 \\ 0,3 & 0,4 & 0,3 \end{pmatrix}$, iar $X_0 \sim \begin{pmatrix} 1 & 2 & 3 \\ 0,7 & 0,2 & 0,1 \end{pmatrix}$. Calculaţi repartiţia variabilei X_1 şi probabilitatea ca la momentele $n = 0, 1, 2$ lanţul să se găsească în stările 1,2,2 respectiv.

15. Un sistem de telecomunicaţii transmite cifrele 0 şi 1. Fiecare cifră trece prin mai multe stadii de prelucrare, în fiecare stadiu existând probabilitatea p ca să fie transmisă corect şi probabilitatea $q = 1 - p$ ca ea să fie transmisă greşit. Fie X_k cifra care intră în stadiul k de prelucrare.

a) Scrieţi matricea P a lanţului Markov omogen cu stările $\{0, 1\}$ astfel obţinut şi calculaţi $P^n, n \in \mathbb{N}^*$, precum şi $P(X_2 = 1/X_0 = 1)$, $P(X_7 = 0/X_3 = 1)$

b) Determinaţi repartiţia staţionară

c) Dacă $X_0 \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$, calculaţi repartiţia v. a. X_n şi $P(X_0 = 0/X_n = 1)$

16. Un lanţ Markov cu 2 stări poate fi folosit să modeleze o varietate largă de sisteme ce alterneazaă între stările ON şi OFF. După fiecare unitate de timp în starea OFF, sistemul trece pe ON cu probabilitatea p. După fiecare unitate de timp în starea ON, sistemul trece în OFF cu probabilitatea q. Folosind 0 şi 1 să notăm stările OFF şi ON, care este matricea probabilităţilor de trecere? Stiind că sistemul este în

starea OFF la momentul 0, care este probabilitatea ca sistemul să fie în starea OFF la momentul $n = 33$?

17. Fie lanţul Markov cu 2 stări E_1 şi E_2 cu probabilităţile de trecere $p_{11} = p_{22} = p, p_{12} = p_{21} = q(0 < p < 1, p + q = 1)$ şi repartiţia iniţială $P(X_0 = 1) = \alpha, P(X_0 = 2) = \beta(\alpha + \beta = 1)$. Se cer:

    a) să se determine probabilităţile de trecere după n paşi $(p_{jk}(n))$;

    b) să se determine probabilităţile limită ;

    c) să se determine probabilitatea $P(X_0 = 1/X_n = 1)$.

SELECTII. ESTIMATORI
Lect. Dr. Costache Luminita

Statistica matematică se ocupă cu gruparea, analiza şi interpretarea datelor referitoare la anumite fenomene şi unele previziuni privind producerea lor în viitor. Spre deosebire de teoria probabilităţilor al cărei demers este deductiv, în care se obţin concluzii dintr-o repartiţie, statistica matematică urmează calea inductivă care face trecerea în sens invers, de la un rezultat la o repartiţie. Statistică matematică este solicitată de fizică , biologie, medicină , arheologie. Inceputurile statisticii matematice sunt prefigurate în special în lucrările lui Laplace. Construirea ei, spre sfârşitul secolului XIX, a fost pregătită de teoria probabilităţilor, statistica demografică , teoria erorilor, unele contribuţii ale psihologiei şi sociologiei. Crearea ei este legată de şcoala anlo-saxonă de statistică matematică , de numele lui Galton (1829-1911) şi Pearson (1857-1936).

1 Teoria selecţiei

Orice mulţime de obiecte care au cel puţin o proprietate comună şi este supusă unei prelucrări statistice se numeşte **populaţie statistică** . Numărul elementelor mulţimii, care se numeşte **volumul populaţiei**, poate fi finit sau infinit. Proprietăţile (trăsăturile) comune ale elementelor unei populaţii statistice se numesc **caracteristici**. Dacă o anumită caracteristică a elementelor se schimbă cantitativ se numeşte **variabilă** . Caracteristicile variabile pot fi discrete sau continue.

Fie X o v. a. reprezentând o anumită caracteristică numerică (durata de viaţă , venitul, număr de defecţiuni etc.) a unei populaţii statistice. Rezultatele numerice $x_1, x_2, \ldots, x_n$ obţinute prin n măsurători (interogări, observări etc.) formează o **selecţie empirică de volum** n a v. a. X. Elementele mulţimii $\{x_1, x_2, \ldots, x_n\}$ numite **variabile de selecţie** au o dublă interpretare:

1) considerate după efectuarea selecţiei, variabilele $x_1, x_2, \ldots, x_n$ reprezintă valori concrete pe care le folosim ca informaţii asupra v. a. X şi le notăm $X_1 = x_1, \ldots, X_n = x_n$;

2) considerate înainte de efectuarea selecţiei, variabilele $X_1, \ldots, X_n$ pot fi privite ca v. a. independente şi identic repartizate din punct de vedere probabilistic cu v. a. X

Se consideră că fiecare v. a. $X_j, j = \overline{1,n}$ se realizează cu aceeaşi probabilitate egală cu $\frac{1}{n}$. Pe baza acestui fapt se construieşte v. a. de selecţie (variabila empirică).

Variabila aleatoare de selecţie are următoarea repartiţie
$X^* \sim \begin{pmatrix} x_1 & \ldots & x_n \\ \frac{1}{n} & \ldots & \frac{1}{n} \end{pmatrix}$, în care $P(X^* = x_k) = \frac{1}{n}, \forall k = \overline{1,n}$

Presupunând că după efectuarea a n observaţii am obţinut : de n_1 ori a apărut $X_1, \ldots$, de n_k ori a apărut X_k, unde $n_1 + \ldots + n_k = n$, atunci $X^* \sim \begin{pmatrix} x_1 & \ldots & x_k \\ \frac{n_1}{n} & \ldots & \frac{n_k}{n} \end{pmatrix}$

Funcţia empirică de repartiţie se va nota cu F^* sau F_n^*.

Media de selecţie este $\overline{x} = \frac{x_1+\ldots+x_n}{n}$

Dispersia de selecţie este $S^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2$

Teorema 1. *Dacă* $E(X) = m$ *şi* $D^2(X) = \sigma^2$, *atunci* $E(\overline{X}) = m$ *şi* $D^2(\overline{X}) = \frac{\sigma^2}{n}$.

2 Selecţii dintr-o populaţie normală

2.1 Repartiţia mediei de selecţie dintr-o populaţie normală

Teorema 2. *Dacă* $\{X_1, X_2, \ldots, X_n\}$ *este o selecţie de volum* n *dintr-o populaţie normală* $N(m, \sigma)$, *atunci media de selecţie* $\overline{X} = \frac{X_1+X_2+\ldots+X_n}{n}$ *urmează o repartiţie normală* $N(m, \frac{\sigma}{\sqrt{n}})$.

Corolarul 1. *In condiţiile anterioare* $\frac{\overline{X}-m}{\frac{\sigma}{\sqrt{n}}}$ *are o repartiţie normală standard* $N(0,1)$.

Teorema 3. *Dacă* $\{X_1, X_2, \ldots, X_n\}$ *reprezintă o selecţie de volum* n *dintr-o populaţie normală* $N(0,1)$, *atunci v. a.* $Y = \sum_{k=1}^{n} X_k^2$ *urmează o repartiţie* χ^2 *cu* n *grade de libertate.*

2.2 Repartiţia dispersiei de selecţie pentru selecţii dintr-o populaţie normală

Fiind dată o populaţie statistică şi $X_1, X_2, \ldots, X_n$ o selecţie de volum n se pot defini:

a) media de selecţie $\overline{X} = \frac{X_1+X_2+\ldots+X_n}{n}$

b) dispersia de selecţie cu ajutorul căreia putem evalua dispersia populaţiei:

1) dacă media m a populaţiei generale este cunoscută atunci dispersia de selecţie este dată de $s_0^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - m)^2$

2) dacă media m a populaţiei generale nu este cunoscută , dispersia de selecţie este $S^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2$

3) în cazul selecţiilor de volum mic ($n < 30$) este indicat să fie folosită dispersia modificată de selecţie $s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \overline{x})^2$, $s^2 = \frac{n}{n-1}S^2$

Teorema 4. *Dacă* $\{X_1, X_2, \ldots, X_n\}$ *este o selecţie dintr-o populaţie normală* $N(m, \sigma)$*, atunci v. a.* $\frac{n \cdot s_0^2}{\sigma^2}$ *are o repartiţie* χ^2 *cu n grade de libertate.*

Teorema 5. *Dacă* $\{X_1, X_2, \ldots, X_n\}$ *este o selecţie dintr-o populaţie normală* $N(m, \sigma)$*, atunci*

(1) statisticile $\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$ *şi* $S^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{X})^2$ *sunt independente;*

(2) statistica $V = \frac{n \cdot S^2}{\sigma^2}$ *are o repartiţie* χ^2 *cu* $n-1$ *grade de libertate.*

Teorema 6. *Dacă* $\{X_1, X_2, \ldots, X_n\}$ *este o selecţie dintr-o populaţie normală* $N(m, \sigma)$*, atunci v. a.* $t = \frac{\overline{X} - m}{\frac{s}{\sqrt{n}}}$ *are o repartiţie Student cu* $n-1$ *grade de libertate*

3 Elemente de teoria estimaţiei

3.1 Estimator nedeplasat, consistent, eficient

Operaţia prin care determinăm valorile parametrilor repartiţiilor se numeşte **estimare**.

Considerăm o v. a. X dintr-o populaţie Γ având repartiţia $f(x, \theta)$. Fie $\{X_1, X_2, \ldots, X_n\}$ o selecţie de volum n din Γ.

Definiţia 1. Un **estimator** $\theta^*(X_1, X_2, \ldots, X_n)$ pentru θ este o statistică (adică o variabilă aleatoare de variabile aleatoare) care ia valori numerice (adică valoarea lui $\theta^*(X_1, X_2, \ldots, X_n)$ îl aproximează pe θ).

Definiţia 2. $\theta^*(X_1, X_2, \ldots, X_n)$ se numeşte **estimator consistent** pentru θ dacă $\lim\limits_{n\to\infty} P(|\theta^*(X_1, X_2, \ldots, X_n) - \theta| < \varepsilon) = 1, \forall \varepsilon > 0$, adică $\theta^*(X_1, X_2, \ldots, X_n)$ converge în probabilitate către θ.

Definiţia 3. $\theta^*(X_1, X_2, \ldots, X_n)$ se numeşte **estimator nedeplasat** pentru θ dacă $E(\theta^*(X_1, X_2, \ldots, X_n)) = \theta$.

Definiţia 4. Dacă

$$\lim_{n\to\infty} E(\theta^*(X_1, X_2, \ldots, X_n)) = \theta$$

$$\lim_{n\to\infty} D^2(\theta^*(X_1, X_2, \ldots, X_n)) = 0$$

spunem că $\theta^*(X_1, X_2, \ldots, X_n)$ este un **estimator corect** al parametrului θ.

Definiţia 5. Dacă

$$E(\theta^*(X_1, X_2, \ldots, X_n)) = \theta$$

$$\lim_{n\to\infty} D^2(\theta^*(X_1, X_2, \ldots, X_n)) = 0$$

spunem că $\theta^*(X_1, X_2, \ldots, X_n)$ este un **estimator absolut corect** al parametrului θ.

Teorema 7. *Dacă* $\theta^*(X_1, X_2, \ldots, X_n)$ *este un estimator pentru* θ *astfel încât* $E(\theta^*(X_1, X_2, \ldots, X_n)) = \theta$ *şi* $\lim\limits_{n\to\infty} D^2(\theta^*(X_1, X_2, \ldots, X_n)) = 0$, *atunci* $\theta^*(X_1, X_2, \ldots, X_n)$ *este un estimator consistent pentru* θ.

Teorema 8. ***(Rao-Cramer)*** *Dacă* $\theta^*(X_1, X_2, \ldots, X_n)$ *este un estimator nedeplasat pentru* θ, *atunci* $D^2(\theta^*) \geq \frac{1}{nI(\theta)}$, *unde* $I(\theta) = -E\left(\frac{\partial^2 \ln f(x,\theta)}{\partial \theta^2}\right) =$
$= E\left[\left(\frac{\partial \ln f(x,\theta)}{\partial \theta}\right)^2\right]$

Definiţia 6. Dacă $\theta^*(X_1, X_2, \ldots, X_n)$ este un estimator nedeplasat astfel încât $D^2(\theta^*) = \frac{1}{nI(\theta)}$, atunci el se numeşte **eficient**.

Corolarul 2. *Orice estimator eficient este consistent.*

3.2 Metoda verosimilităţii maxime

Metoda verosimilităţii maxime este una dintre cele mai vechi metode de estimare. Ea a fost folosită de Gauss când a dezvoltat metoda celor mai mici pătrate şi reintrodusă în 1912 de Fischer.

Se consideră o selecţie $\{x_1, \ldots, x_n\}$ de volum n dintr-o populaţie în care caracteristica sub cercetare X este o v.a. având densitatea de probabilitate $f(x, \theta)$ presupusă cunoscută , iar parametrul θ este necunoscut şi poate lua valori într-o mulţime $\Theta \subset \mathbb{R}^k$.

Definiţia 7. Vom numi **funcţie de verosimilitate** corespunzătoare valorilor $x_1, \ldots, x_n$, o funcţie $L(x_1, \ldots, x_n; \theta)$, considerată ca funcţie de θ, definită prin $L(x_1, \ldots, x_n; \theta) = \prod\limits_{i=1}^{n} f(x_i; \theta)$, unde $f(x_i; \theta)$ este fie densitatea de probabilitate a v. a. X, fie repartiţia sa , adică $f(x; \theta) = P(X = x)$, dacă X este discretă .

Definiţia 8. **Estimatorul de verosimilitate maximă** pentru θ este acea valoare $\theta^* = \theta^*(x_1, \ldots, x_n)$ cu proprietatea că $L(x_1, \ldots, x_n; \theta^*) =$
$= \max\limits_{\theta \in \Theta} L(x_1, \ldots, x_n; \theta)$.

Intrucât funcţiile L şi $\ln L$ au aceleaşi puncte de maxim rezultă că , dacă $\theta = (\theta_1, \ldots, \theta_k)$, atunci $\theta^*(x_1, \ldots, x_n) = (\theta_1(x_1, \ldots, x_n), \ldots, \theta_k(x_1, \ldots, x_n))$ trebuie să verifice sistemul de ecuaţii

$$\frac{\partial}{\partial \theta_j} \ln L(x_1, \ldots, x_n; \theta_1, \ldots, \theta_k) = 0, j = \overline{1, k}$$

Observaţia 1. Orice estimator eficient al parametrului θ este un estimator de verosimilitate maximă .

INTERVALE DE INCREDERE. IPOTEZE PARAMETRICE
Lect. Dr. Costache Luminita

1 Intervale de încredere

Fie dată densitatea de probabilitate $f(x,\theta)$. In urma efectuării unei selecţii de volum n, $\{x_1,\ldots,x_n\}$, putem determina două statistici $\underline{\theta}(x_1,\ldots,x_n)$ şi $\overline{\theta}(x_1,\ldots,x_n)$ astfel încât $(\underline{\theta},\overline{\theta})$ să conţină valoarea adevărată a parametrului θ cu probabilitatea dată $p=1-\alpha$, adică $P(\underline{\theta}<\theta<\overline{\theta})=p$, unde p nu depinde de θ.

Intervalul $(\underline{\theta},\overline{\theta})$ care acoperă pe θ cu probabilitatea dată $1-\alpha$ se numeşte **interval de încredere**. Cu cât intervalul de încredere este mai mic şi probabilitatea $1-\alpha$ este mai apropiată de 1, cu atât avem o indicaţie mai precisă despre θ. Intervalul încredere $(\underline{\theta},\overline{\theta})$ este **bilateral**. Intervalul de încredere poate fi şi **unilateral superior** sau **inferior**: $P(\theta<\overline{\theta})=1-\alpha$ sau $P(\underline{\theta}<\theta)=1-\alpha$.

Fie $F(x)$ o funcţie de repartiţie şi $\alpha\in(0,1)$. Se numeşte **α-cuantilă** a repartiţiei F un număr $c\in\mathbb{R}$ pentru care $F(c)=\alpha$.

In cele ce urmează α-cuantilele repartiţiilor $N(0,1),\chi^2(n),T(n)$ vor fi notate respectiv $z_\alpha,\chi^2_\alpha(n),t_\alpha(n)$.

1.1 Intervale de încredere pentru media şi dispersia repartiţiei normale

1. σ cunoscut, m necunoscut

$$m\in\left(\overline{x}-\frac{\sigma}{\sqrt{n}}z_{1-\frac{\alpha}{2}},\overline{x}+\frac{\sigma}{\sqrt{n}}z_{1-\frac{\alpha}{2}}\right)$$

2. m cunoscut, σ necunoscut

$$\sigma^2\in\left(\frac{n}{\chi^2_{1-\frac{\alpha}{2}}(n)}\cdot s_0^2,\frac{n}{\chi^2_{\frac{\alpha}{2}}(n)}\cdot s_0^2\right)$$

3. m şi σ necunoscuţi

$$\sigma^2\in\left(\frac{n-1}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\cdot s^2,\frac{n-1}{\chi^2_{\frac{\alpha}{2}}(n-1)}\cdot s^2\right)$$

$$m\in\left(\overline{x}-\sqrt{\frac{s^2}{n}}t_{1-\frac{\alpha}{2}}(n-1),\overline{x}+\sqrt{\frac{s^2}{n}}t_{1-\frac{\alpha}{2}}(n-1)\right)$$

2 Verificarea ipotezelor statistice parametrice

In statistica matematică prin ipoteză înţelegem o presupunere ce se poate face asupra unuia sau mai multor parametri ce caracterizează anumite repartiţii sau asupra tipului legilor de repartiţie ce caracterizează anumite populaţii.

Ipoteza statistică poate fi parametrică atunci când se referă la parametrul unei funcţii de repartiţie cunoscute şi neparametrică atunci când se referă la forma funcţiei de repartiţie necunoscută .

Considerăm o v.a. a cărei densitate de probabilitate $f(x, \theta)$ depinde de un singur parametru θ. Ne propunem să verificăm ipoteza că parametrul θ are valoarea θ_0. Notăm această ipoteză $H_0 : \theta = \theta_0$ şi o vom numi **ipoteza nulă** faţă de ipoteza $H_1 : \theta = \theta_1$ cu care este confruntată şi este numită **ipoteză alternativă** .

Verificarea ipotezei nule H_0 se face pe baza unei selecţii $\{x_1, \dots, x_n\}$ cu ajutorul valorii observate a unei statistici. Metoda aleasă pentru verificarea unei ipoteze statistice se numeşte **test** sau **criteriu statistic**.

Fie θ un parametru asociat v. a. X şi ipoteza nulă $H_0 : \theta = \theta_0$ care trebuie testată (verificată) astfel încât probabilitatea **unei erori de prima speţă** (respingerea lui H_0 atunci când ea este adevărată) să fie egală cu α. Numărul $\alpha \in (0, 1)$ se numeşte **nivelul de semnificaţie** al testului. Etapele aplicării testului sunt următoarele:

a) Alegerea **ipotezei alternative** H_1 care poate fi **unilaterală** $H_1 : \theta < \theta_0, \theta > \theta_0$ sau **bilaterală** $H_1 : \theta \neq \theta_0$;

b) Alegerea unei statistici $T = T(x_1, \dots, x_n)$ astfel încât, în ipoteza H_0, repartiţia lui T să fie cunoscută ;

c) In funcţie de ipoteza alternativă H_1 şi de nivelul de semnificaţie α, fixarea unei regiuni critice de forma

$$R_{cr} = \{T/T < c_\alpha\}, R_{cr} = \{T/T > c_{1-\alpha}\}$$

în cazul unui test unilateral, sau

$$R_{cr} = \{T/T < c_{\frac{\alpha}{2}}\} \cup R_{cr} = \{T/T > c_{1-\frac{\alpha}{2}}\}$$

în cazul unui test bilateral. Cu $c_\alpha, c_{1-\alpha}, c_{\frac{\alpha}{2}}, c_{1-\frac{\alpha}{2}}$ s-au notat cuantilele repartiţiei lui T în ipoteza H_0; deci, în această ipoteză , probabilitatea unei erori de prima speţă este $P(T \in R_{cr}) = \alpha$.

d) Se calculează valoarea $T(x_1, \dots, x_n)$ luată de statistica T pe elementele unei anumite selecţii empirice $x_1, \dots, x_n$

e) Se respinge ipoteza H_0 dacă şi numai dacă $T(x_1, \dots, x_n) \in R_{cr}$

2.1 Verificarea ipotezei asupra mediei m a unei populaţii normale cu σ^2 cunoscut

1.1. Testul bilateral
$H_0 : m = m_0$
$H_1 : m = m_1 \neq m_0$
Regiunea critică este $R_{cr} : |\frac{\overline{x}-m_0}{\frac{\sigma}{\sqrt{n}}}| > z_{1-\frac{\alpha}{2}}$
1.2. Testul unilateral stânga
$H_0 : m = m_0$
$H_1 : m = m_1 < m_0$
Regiunea critică este $R_{cr} : \frac{\overline{x}-m_0}{\frac{\sigma}{\sqrt{n}}} < z_\alpha$
1.3. Testul unilateral dreapta
$H_0 : m = m_0$
$H_1 : m = m_1 > m_0$
Regiunea critică este $R_{cr} : \frac{\overline{x}-m_0}{\frac{\sigma}{\sqrt{n}}} > z_{1-\alpha}$

2.2 Verificarea ipotezei asupra mediei m a unei populaţii normale cu σ^2 necunoscut

2.1. Testul bilateral
$H_0 : m = m_0$
$H_1 : m = m_1 \neq m_0$
Regiunea critică este $R_{cr} : |\frac{\overline{x}-m_0}{\frac{s}{\sqrt{n}}}| > t_{1-\frac{\alpha}{2}}(n-1)$
2.2. Testul unilateral stânga
$H_0 : m = m_0$
$H_1 : m = m_1 < m_0$
Regiunea critică este $R_{cr} : \frac{\overline{x}-m_0}{\frac{s}{\sqrt{n}}} < t_\alpha(n-1)$
2.3. Testul unilateral dreapta
$H_0 : m = m_0$
$H_1 : m = m_1 > m_0$
Regiunea critică este $R_{cr} : \frac{\overline{x}-m_0}{\frac{s}{\sqrt{n}}} > t_{1-\alpha}(n-1)$

2.3 Verificarea ipotezei asupra dispersiei unei populaţii normale

Fie ipoteza $H_0 : \sigma^2 = \sigma_0^2$ şi alternativa ei $H_1 : \sigma^2 = \sigma_1^2$
3.1. Testul unilateral stânga
$H_0 : \sigma^2 = \sigma_0^2$
$H_1 : \sigma^2 = \sigma_1^2 < \sigma_0^2$
Regiunea critică este $R_{cr} : \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_\alpha(n-1)$
3.2. Testul unilateral dreapta
$H_0 : \sigma^2 = \sigma_0^2$
$H_1 : \sigma^2 = \sigma_1^2 > \sigma_0^2$

Regiunea critică este $R_{cr}: \frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{1-\alpha}(n-1)$

3.3. Testul bilateral

$H_0: \sigma^2 = \sigma_0^2$

$H_1: \sigma^2 = \sigma_1^2 \neq \sigma_0^2$

Regiunea critică este $R_{cr}: \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{\frac{\alpha}{2}}(n-1) \cup \frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{1-\frac{\alpha}{2}}(n-1)$

PROBLEME ESTIMATORI
Lect. Dr. Costache Luminita

1. Repartiţia valorilor defecţiunilor unor aparate de măsură şi control a fost analizată pe baza a 100 de observaţii care au furnizat următoarele valori cu privire la numărul de porniri (puneri în funcţiune) după care acestea se defectează :

x_i	1	3	6	8	10
n_i	15	20	30	10	25

Stabiliţi:

a) media şi dispersia de selecţie;

b) funcţia de repartiţie a selecţiei şi valorile ei în punctele $x = 2$ şi $x = 7$.

2. Pentru a cerceta prezenţa studenţilor la un anumit curs s-a ales un eşantion de n studenţi şi s-a înregistrat numărul absenţelor acestora la patru cursuri consecutive.

Nr. studenţi n_i	50	20	15	8	7
Nr. absenţe x_i	0	1	2	3	4

a) Să se scrie repartiţia empirică şi funcţia de repartiţie empirică $F_n^*(x)$;

b) Să se calculeze media şi dispersia de selecţie;

c) Să se calculeze $F_{100}^*(3)$

3. Se controlează greutatea unor pachete şi, pentru aceasta, se extrage o selecţie de volum n, care dă următoarele valori:

Pachetul	1	2	3	4	5	6	7	8	9	10
Greutatea	21,5	21,6	21,75	22	22,45	22,6	23,2	23,4	23,5	23.65

Să se calculeze $F_{10}^*(20), F_{10}^*(22), F_{10}^*(23)$. Să se scrie repartiţia empirică şi să se calculeze media de selecţie.

4. Repartiţia valorilor rezistenţei la rupere a unor fire de bumbac (în kg) a fost analizată pe baza a 100 de observaţii. Valorile observate împreună cu frecvenţele lor absolute sunt date în tabelul următor:

x_i (valorile rezistenţei la rupere în kg)	0,5	0,65	0,75	0,8	0,9	1	1,12	1,35
n_i (frecvenţe absolute)	2	4	5	26	30	25	6	2

Se cer:

a) Valoarea medie şi dispersia de selecţie a rezistenţei la rupere;

b) $F^*(0,78), F^*(0,9)$, unde $F^*(x)$ este funcţia empirică de repartiţie

5. Dintr-o populaţie normală cu media $m = 12$, probabilitatea ca media de selecţie corespunzătoare unei selecţii de volum $n = 16$ să nu depăşească 14 (în modul) este de 0,24. Care este probabilitatea ca

o singură observaţie din această selecţie să aibă o valoare mai mare decât 14?

6. Dintr-o populaţie normală se extrag toate selecţiile posibile de volum $n = 25$. Dacă $5^0/_0$ dintre ele au medii care diferă de media populaţiei cu cel puţin 5 unităţi în valoare absolută , aflaţi abaterea medie pătratică a populaţiei.

7. Arătaţi că media de selecţie este un estimator eficient pentru parametrul λ al repartiţiei Poisson.

8. Arătaţi că media de selecţie este un estimator eficient pentru media m a repartiţiei normale.

9. Să se estimeze prin metoda verosimilităţii maxime parametrul θ al repartiţiilor:

   a) $f(x,\theta) = (2-\theta)x^{(\ln 2)^{-1}\ln\frac{\theta-1}{2-\theta}}, 1 < \theta < 2$

   b) $f(x,\theta) = \frac{\theta}{1-\theta}x^{\frac{2\theta-1}{1-\theta}}, \frac{1}{2} < \theta < 1$

   c) $f(x,\theta) = \frac{1}{\sqrt{2\pi\theta}}\mathrm{e}^{-\frac{1}{2\theta}(x-\theta)^2}, 0 < x < 1, \theta > 0$

   d) $f(x,\theta) = \theta(1-\theta)^x, x = 0, 1, 2\ldots, 0 < \theta < 1$

   e) $f(x,\theta) = (1+\theta)x^\theta, 0 < x < 1, \theta > 0$

   f) $f(x,\theta) = \frac{\theta x_0^\theta}{x^{\theta+1}}, x > x_0(x_0$ constantă dată $), \theta > 0$

   g) $f(x,\theta) = \frac{x^{\alpha-1}\mathrm{e}^{-\frac{x}{\theta}}}{\Gamma(\alpha)\theta^\alpha}, x > 0, \theta > 0, \alpha$ dat

   h) $f(x,\theta) = \theta\mathrm{e}^{-\theta x}, x \geq 0, \theta > 0$

   i) $f(x,\theta) = \frac{x}{\theta}\cdot\mathrm{e}^{-\frac{x^2}{2\theta}}, x > 0, \theta > 0.$

10. Fie X o v. a. cu densitatea $f(x,\theta) = \frac{A}{x^{\frac{\theta+1}{\theta}}}, x \geq 1, \theta > 0$ şi să considerăm o selecţie aleatoare $X_1, \ldots, X_n$ din populaţia de caracteristică X.

    a) Să se determine A în funcţie de θ;

    b) Să se afle estimatorul de verosimilitate al parametrului θ şi să se studieze proprietăţile acestuia.

INTERVALE DE INCREDERE. IPOTEZE PARAMETRICE. PROBLEME

Lect. Dr. Costache Luminita

1. Rectorul Universităţii Politehnice Bucureşti vrea să ştie care este media vârstei studenţilor. Din anii trecuţi se cunoaşte că abaterea standard este de 2 ani. Un sondaj asupra a 50 de studenţi arată că media este de 23,2 ani. Cu un nivel de semnificaţie de 0,05, să se determine un interval de încredere pentru medie. Se presupune că populaţia are caracteristica normală .

2. Intr-o fabrică de ţesut s-a constatat că rezistenţa la rupere a unui anumit fir de bumbac are o repartiţie normală cu media necunoscută m şi abaterea medie pătratică $\sigma = 36$g (calitatea standard). Pentru a cerceta calitatea unui lot de fire de bumbac în ceea ce priveşte rezistenţa la rupere s-a făcut o selecţie de volum $n = 9$ fire, obţinându-se media de selecţie $\overline{X} = 195$ g. Să estimeze rezistenţa la rupere m a lotului de fire controlat, printr-un interval de încredere $1 - \alpha = 0,95$.

3. Fie X v. a. normală care reprezintă greutatea unor ouă . O selecţie de volum $n = 200$ a dat rezultatele următoare :

Greutatea (g)	37,5	42,5	47,5	52,5	57,5	62,5	67,5	72,5
Nr. observaţiilor	8	14	26	44	68	20	14	6

Stiind că dispersia lui X este $\sigma^2 = 64$, aflaţi un interval de încredere $95^0/_0$ şi $90^0/_0$ pentru media m.

4. Fie X o v. a. care reprezintă numărul de zile după care un anumit produs alimentar este absorbit pe piaţă . In urma unei selecţii de volum $n = 8$ în rândul centrelor de desfacere, s-au obţinut următoarele rezultate cu privire la numărul de zile după care se epuizează produsul respectiv :

Centru de desfacere i	1	2	3	4	5	6	7	8
Nr. de zile x_i	40	54	40	50	38	40	50	60

Presupunând că X are o repartiţie normală să se afle:

a) un interval de încredere $95^0/_0$ pentru viteza media cu care este absorbit produsul pe piaţă ;

b) un interval de încredere $95^0/_0$ pentru abaterea medie pătratică a numărului de zile după care se epuizează produsul respectiv.

5. Durabilitatea unor motoare de automobile poate fi considerată o v. a. normală cu media $m = 200000$ km şi dispersia 50000^2 km. Se face o schimbare în procesul de producţie prin introducerea unei metode noi de fabricaţie. O selecţie de volum $n = 100$ de motoare a dat

$\overline{X} = 220000$ km. Considerând $\alpha = 0,05$ se poate afirma că noua metodă duce la creşterea durabilităţii motoarelor?

6. Douăzeci de determinări a procentului de NaCl într-o anumită soluţie au condus $\overline{X} = 0,7^0/_0$ şi $s = 0,03^0/_0$. Stiind că procentul de NaCl într-o anumită soluţie este o v. a. normală , să se verifie la pragul de semnificaţie $\alpha = 0,05$ ipoteza $H_0 : m = 0,8^0/_0$ faţă de alternativa $H_1 : m < 0,8^0/_0$.

7. Precizia unui cântar electronic se verifică cu ajutorul dispersiei măsurătorilor efectuate asupra unui etalon. Dispersia măsurătorilor nu trebuie să depăşească valoarea nominală $\sigma_0^2 = 0,04$. S-au efectuat $n = 11$ cântăriri ale unui etalon şi s-au obţinut rezultatele :

i	1	2	3	4	5	6	7	8	9	10	11
x_i	100,6	99,6	100	100,1	100,3	100	99,9	100,2	100,4	100,6	100,5

   Să se verifice, la un prag de semnificaţie $\alpha = 0,05$, dacă cântarul asigură precizia standard stabilită , presupunând că datele de selecţie sunt observaţii asupra unei v. a. normale.

   Pentru analiza preciziei unor măsurători s-au făcut $n = 16$ măsurıtori şi s-a stabilit $s^2 = 0,56$. Verificaţi ipoteza $H_0 : \sigma^2 = 0,41$ faţă de alternativa $H_1 : \sigma^2 \neq 0,41$ la pragul de semnificaţie $\alpha = 0,1$, ştiind că populaţia în studiu este normală .

8. Pentru a estima precizia unui termometru se realizează 15 măsurători independente asupra temperaturii unui lichid menţinut constant la $20°C$. Presupunem că rezulttele măsurătorilor sunt realizări ale variabilelor aleatoare normale $X_k, k = \overline{1,15}$ de medie $m = 20$ şi σ necunoscută . Construiţi un interval de încredere $1 - \alpha = 0,99$ pentru σ^2, ştiind că $\frac{1}{15}\sum\limits_{k=1}^{15}(x_k - 20)^2 = 18$.

9. O fabrică de acumulatori afirmă că durata de funcţionare a acumulatorilor este 300 de zile. Un laborator cercetează 4 acumulatori şi obţine rezultatele 298,290,306,302. Aceste rezultate indică faptul că acest tip de acumulatori are o durată de funcţionare mai mică decât afirmă fabrica ($\alpha = 0,02$,$t_{0,02}(3) = -4,541$)?

10. S-a stabilit experimental că nivelul colesterolului în organismul unui adult este o v. a. normală . O selecţie aleatoare de volum $n = 41$ adulţi a dat un nivel mediu observat al colesterolului $\overline{X} = 213$ cu $s^2 = 48,4$. Să se testeze ipoteza $H_0 : m = m_0 = 200$ cu alternativa $H_1 : m = m_1 > 200$ la un prag de semnificaţie $\alpha = 0,05$.

11. Pentru stabilirea rezistenţei la rupere a unor cabluri s-au efectuat $n =$ $= 36$ măsurători şi s-a stabilit media de selecţie $\overline{X} = 500$ kg. Stiind că

rezistenţa la rupere este o v. a. normală cu $\sigma^2 = 100$ kg, să se verifice ipoteza $H_0 : m = 496$ kg faţă de alternativele : a) $H_1 : m \neq 496$ kg; b) $H_1 : m < 496$ kg la pragul de semnificaţie $\alpha = 0,01$.

12. S-a stabilit că greutatea unor ouă pentru a fi importate trebuie să fie de $m_0 = 50$ g. O cercetare selectivă asupra unui volum $n = 150$ ouă dintr-un lot importat a determinat o greutate medie observată de $\overline{X} = 43$ g. Se cere să se verifice la un prag de semnificaţie $\alpha = 0,01$, ipoteza $H_0 : m = m_0 = 50$ g faţă de ipoteza alternativă $H_1 : m < 50$ g, dacă greutatea ouălelor este o v. a. normală $N(m, 16)$.

13. Durata de funcţionare a unui tip oarecare de bec electric de 100 waţi poate fi considerată ca o v. a. X repartizată normal cu media $m = = 1500$ şi $\sigma^2 = 200^2$. O selecţie de volum $n = 25$ de astfel de becuri dă o durată medie de funcţionare de 1380 ore. La pragul $\alpha = 0,01$, să se verifice ipoteza $H_0 : m = m_0 = 1500$ faţă de $H_1 : m = m_1 < 1500$.

14. O firmă producătoare de becuri afirmă că durata medie de viaţă a becurilor produse este de 170 ore. Un reprezentant al Oficiului pentru Protecţia Consumatorilor cercetează un eşantion aleator de $n = 100$ de becuri, obţinând o durată medie observată de viaţă de 158 ore şi o abatere standard $s = 30$ ore. Determinaţi un interval de încredere $1 - \alpha = 0,99$ pentru durata medie de viaţă m, în ipoteza că durata de viaţă a becurilor este o v. a. normală . Poate fi acuzată firma producătoare de publicitate mincinoasă ?

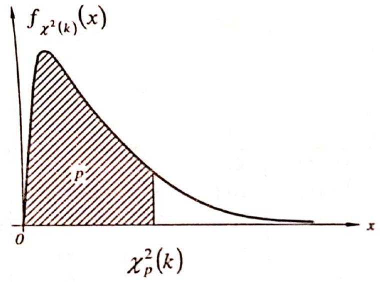
15. Consumul nominal de benzină al unui anumit motor de maşină este de 10 l la 100 km. Se aleg la întâmplare 25 de motoare fabricate după o tehnologie modernizată , obţinându-se media $\overline{x} = 9,3$ l şi varianţa $s^2 = 4l^2$. Presupunând că selecţia provine dintr-o populaţie normală , folosiţi un test unilateral cu nivelul de semnificaţie $\alpha = 0,05$, pentru a testa ipoteza că noua tehnologie nu a influenţat consumul de benzină

16. Dintr-o populaţie normală o slecţie de volum $n = 30$ a dat rezultatele:

x_i	0	1	2	3	4	5	6
n_i	2	3	6	7	7	3	2

Să se verifice ipoteza $H_0 : \sigma^2 = \sigma_0^2 = 1$ faţă de alternativa $H_1 : \sigma^2 = = \sigma_1^2 > 1$ la pragul de semnificaţie $\alpha = 0,05$.

Tabelul A5. Cuantilele repartiției χ^2 hi-pătrat



$$f_{\chi^2(k)}(x) = \frac{1}{2^{k/2}\Gamma(k/2)} \cdot x^{\frac{k-2}{2}} \cdot e^{-\frac{x}{2}}, \quad x > 0; \quad k > 0.$$

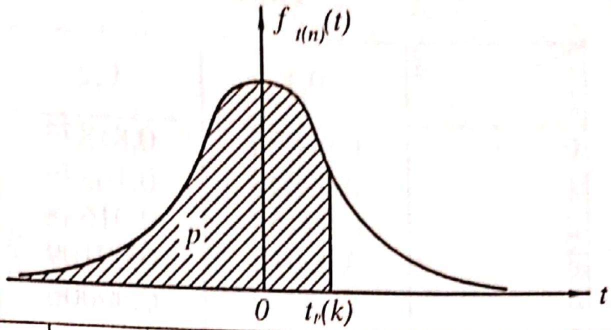
k \ p	0,005	0,010	0,025	0,05	0,10	0,20	0,30	0,70	0,80	0,90	0,95	0,975	0,990	0,995	0,999
1	$0{,}0^4393$	$0{,}0^3157$	$0{,}0^3982$	$0{,}0^2393$	0,0158	0,0642	0,1480	1,0700	1,6400	2,7100	3,8400	5,0200	6,6300	7,8800	10,800
2	0,0100	0,0201	0,0506	0,1030	0,2110	0,4460	0,7130	2,4100	3,2200	4,6100	5,9900	7,3800	9,2100	10,600	13,8000
3	0,0717	0,1150	0,2160	0,3520	0,5840	1,0050	1,4240	3,6600	4,6400	6,2500	7,8100	9,3500	11,3000	12,8000	16,3000
4	0,2070	0,2970	0,4840	0,7110	1,0600	1,6500	2,1900	4,8800	5,9900	7,7800	9,4900	11,1000	13,3000	14,9000	18,5000
5	0,4120	0,5540	0,8310	1,1500	1,6100	2,3400	3,0000	6,0600	7,2900	9,2400	11,1000	12,8000	15,1000	16,7000	20,5000
6	0,6760	0,8720	1,2400	1,6400	2,2000	3,0700	3,8300	7,2300	8,5600	10,6000	12,6000	14,4000	16,8000	18,5000	22,5000
7	0,9890	1,2400	1,6900	2,1700	2,8300	3,8200	4,6700	8,3800	9,8000	12,0000	14,1000	16,0000	18,5000	20,3000	24,3000
8	1,3400	1,6500	2,1800	2,7300	3,4900	4,5900	5,5300	9,5200	11,0000	13,4000	15,5000	17,5000	20,1000	22,0000	26,1000
9	1,7300	2,0900	2,7000	3,3300	4,1700	5,3800	6,3900	10,7000	12,2000	14,7000	16,9000	19,0000	21,7000	23,6000	27,9000
10	2,1600	2,5600	3,2500	3,9400	4,8700	6,1800	7,2700	11,8000	13,4000	16,0000	18,3000	20,5000	23,2000	25,2000	29,6000
11	2,6000	3,0500	3,8200	4,5700	5,5800	6,9900	8,1500	12,9000	14,6000	17,3000	19,7000	21,9000	24,7000	26,8000	31,3000
12	3,0700	3,5700	4,4000	5,2300	6,3000	7,8100	9,0300	14,0000	15,8000	18,5000	21,0000	23,3000	26,2000	28,3000	32,9000
13	3,5700	4,1100	5,0100	5,8900	7,0400	8,6300	9,9300	15,1000	17,0000	19,8000	22,4000	24,7000	27,7000	29,8000	34,5000
14	4,0700	4,6600	5,6300	6,5700	7,7900	9,4700	10,8000	16,2000	18,2000	21,1000	23,7000	26,1000	29,1000	31,3000	36,1000
15	4,6000	5,2300	6,2600	7,2600	8,5500	10,3000	11,7000	17,3000	19,3000	22,3000	25,0000	27,5000	30,6000	32,8000	37,7000
16	5,1400	5,8100	6,9100	7,9600	9,3100	11,2000	12,6000	18,4000	20,5000	23,5000	26,3000	28,8000	32,0000	34,3000	39,3000

Tabelul A5 (continuare)

k \ p	0,005	0,010	0,025	0,05	0,10	0,20	0,30	0,70	0,80	0,90	0,95	0,975	0,990	0,995	0,999
17	5,70	6,41	7,56	8,67	10,1	12,0	13,5	19,5	21,6	24,8	27,6	30,2	33,4	35,7	40,8
18	6,26	7,01	8,23	9,39	10,9	12,9	14,4	20,6	22,8	26,0	28,9	31,5	34,8	37,2	42,3
19	6,84	7,63	8,91	10,10	11,7	13,7	15,4	21,7	23,9	27,2	30,1	32,9	36,2	38,6	43,8
20	7,43	8,26	9,59	10,90	12,4	14,6	16,3	22,8	25,0	28,4	31,4	34,2	37,6	40,0	45,3
21	8,03	8,90	10,30	11,60	13,2	15,4	17,2	23,9	26,9	29,6	32,7	35,5	38,9	41,4	46,8
22	8,64	9,54	11,00	12,30	14,0	16,3	18,1	24,9	27,3	30,8	33,9	36,8	40,3	42,8	48,3
23	9,26	10,20	11,70	13,10	14,8	17,2	19,0	26,0	28,4	32,0	35,2	38,1	41,6	44,2	49,7
24	9,89	10,90	12,40	13,80	15,7	18,1	19,9	27,1	29,6	33,2	36,4	39,4	43,0	45,6	51,2
25	10,50	11,50	13,10	14,60	16,5	18,9	20,9	28,2	30,7	34,4	37,7	40,6	44,3	46,9	52,6
26	11,20	12,20	13,80	15,40	17,3	19,8	21,8	29,2	31,8	35,6	38,9	41,9	45,6	48,3	54,1
27	11,80	12,90	14,60	16,20	18,1	20,7	22,7	30,3	32,9	36,7	40,1	43,2	47,0	49,6	55,5
28	12,50	13,60	15,30	16,90	18,9	21,6	23,6	31,4	34,0	37,9	41,3	44,5	48,3	51,0	56,9
29	13,10	14,30	16,00	17,70	19,8	22,5	24,6	32,5	35,1	39,1	42,6	45,7	49,6	52,3	58,3
30	13,80	15,00	16,80	18,50	20,6	23,4	25,5	33,5	36,3	40,3	43,8	47,0	50,9	53,7	59,7
35	17,20	18,50	20,60	22,50	24,8	27,3	30,2	38,9	41,8	46,1	49,8	53,2	57,3	60,3	66,6
40	20,70	22,20	24,40	26,50	29,1	32,3	34,9	44,2	47,3	51,8	55,8	59,3	63,7	66,8	73,1
45	24,30	25,90	28,40	30,60	33,4	36,9	39,9	49.5	52,7	57,5	61,7	65,4	70,0	73,2	80,1
50	28,00	29,70	32,40	34,80	37,7	41,4	44,3	54,7	58,2	63,2	67,5	71,4	76,2	79,5	86,7
75	47,20	49,50	52,90	56,10	59,8	64,5	68,1	80,9	85,1	91,1	96,2	100,3	106,4	110,3	118,6
100	67,30	70,10	74,20	77,90	82,4	87,9	92,1	106,9	111,7	118,5	124,3	129,6	135,6	140,2	143,4

Tabelul A3. Cuantilele repartiţiei t – Student

$$F\left(t_p\right)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\int_{-\infty}^{t_p}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}dx$$



n	$t_{0,60}$	$t_{0,70}$	$t_{0,80}$	$t_{0,90}$	$t_{0,95}$	$t_{0,975}$	$t_{0,99}$	$t_{0,995}$
1	0,325	0,727	1,376	3,078	6,314	12,706	31,821	63,657
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925
3	0,277	0,584	0978	1,638	2,353	3,182	4,541	5,841
4	0,271	0,569	0,911	1,533	2,132	2,776	3,747	4,604
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	4,032
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	2,977
15	0,258	0,536	0,866	1,341	1,753	2,131	2,602	2,947
16	0,258	0,535	0,865	1,337	1,746	2,120	2,583	2,921
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898
18	0,257	0,534	0,862	1,330	1,734	2,101	2,552	2,878
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845
21	0,257	0,532	0,859	1,323	1,721	2,080	2,518	2,831
22	0,256	0,532	0,858	1,321	1,717	2,074	2,508	2,819
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807
24	0,256	0,531	0,857	1,318	1,711	2,064	2,492	2,797
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787
26	0,256	0,531	0,856	1,315	1,706	2,056	2,479	2,779
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763
29	0,256	0,530	0,854	1,311	1,699	2,045	2,462	2,756
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750
40	0,255	0,529	0,851	1,303	1,681	2,021	2,423	2,704
60	0,254	0,527	0,848	1,296	1,671	2,000	2,390	2,660
120	0,254	0,526	0,845	1,289	1,658	1,980	2,358	2,617
∞	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576
	$-t_{0,40}$	$-t_{0,30}$	$-t_{0,20}$	$-t_{0,10}$	$-t_{0,05}$	$-t_{0,025}$	$-t_{0,01}$	$-t_{0,005}$