1 Cursul de teoria probabilităţilor an II

1.1 Lecţia 1

Mă bazez pe faptul că s-a făcut deja un curs de Teoria măsurii. Deci se ştie ce e (Ω, K, μ)

Aşadar (Ω, K, P) este un spaţiu cu o măsură specială, care are proprietatea că $P(\Omega) = 1$

Consecinţă: $P(A^c) = 1 - P(A)$.

Notez consecvent complementara cu un ^{c} :$A^c = \Omega \backslash A$

Caracteristici ale unui spaţiu probabilizat - sau, poftim, cîmp de probabilitate, accept şi denumirea asta:

1. Ω este mulţimea evenimentelor elementare. Ele reprezintă mulţimea rezultatelor posibile ale unui experiment mental.

Exemple.

- arunc un zar de n ori: $\Omega = \mathbb{Z}_6^n$; dacă arunc n zaruri e la fel

- arunc o monedă de n ori: $\Omega = \mathbb{Z}_2^n$;dacă arunc n monede e la fel

- aştept pînă iese un 6 la aruncarea cu un zar: $\Omega = \mathbb{Z}_6^n$;

- se ia un număr întreg "la întîmplare": $\Omega = \mathbb{Z}$; dacă se iau n numere întregi atunci $\Omega = \mathbb{Z}^n$

- se ia un punct oarecare în plan/spaţiu: $\Omega = \mathbb{R}^2/\mathbb{R}^3$

- se ia un arc de cerc la întîmplare: Ω este cercul trigonometric

2. K este $\sigma-$algebra evenimentelor. Un eveniment este o colecţie de evenimente elementare. Dacă Ω este cel mult numărabil, este bine: atunci se ia $K = P(\Omega)$. Adică $\sigma-$algebra totală. Dacă nu, e mai complicat. Multe probbilităţi au proprietatea că P($\{\varpi\}$) = 0 pentru orice $\varpi \in \Omega$. Astea nu se pot construi pe $P(\Omega)$ fiindcă este o teoremă a lui Ulam care spune că dacă o măsură pe $P(\Omega)$ neglijează toate punctele şi Ω este nenumărabilă, atunci măsura e identic nulă. Exemplu tipic $\Omega = \mathbb{R}, K = B(\mathbb{R}) = \sigma(Top(\mathbb{R}))$, sau $\Omega = \mathbb{R}^n, K = B(\mathbb{R}^n) = \sigma(Top(\mathbb{R}^n)) = (B(\mathbb{R})^{\otimes n}$ dar asta fără demonstraţie)

3. Cea mai simplă probabilitate cu putinţă este şi cea mai simplă măsură cu putinţă: este δ_a definită de $\delta_a(A) = 1_A(a)$

Fapt 31. Dacă $(P_n)_n$ sunt probabilităţi pe spaţiul măsurabil (Ω, K) şi dacă $p_n \geq 0$ au proprietatea că $\Sigma p_n = 1$,atunci $P := \Sigma_n p_n P_n$ este de asemenea probabilitate

Fapt 32. Deci dacă $(p_n)_n$ sunt ponderi de sumă 1, atunci $P = \Sigma_n p_n \delta_{x_n}$ este o probabilitate. Ea se numeşte **probabilitate discretă.**

Fapt 33. Dacă Ω este cel mult numărabilă, atunci toate probabilităţile pe $(\Omega, P(\Omega))$ sunt discrete.

Fapt 34. Dacă Ω este finită, atunci probabilitatea $P = \frac{\Sigma_{\varpi \in \Omega} \delta_{\varpi}}{|\Omega|}$ se numeşte probabilitatea clasică sau repartiţia uniformă şi se notează cu $U(\Omega)$. Aşadar $U(\Omega)(A) = \frac{|A|}{|\Omega|} = \frac{\text{nr cazuri favorabile}}{\text{nr cazuri posibile}}$

Fapt 35. Dacă P, Q sunt două probabilităţi pe (Ω, K) atunci mulţimea $E = E(P, Q) := \{A \in K : P(A) = Q(A)\}$ satisface axiomele

(i) $\varnothing \in E$: (ii) $A \in E \implies A^c \in E$: (iii) $A_n \in E \forall n$ disjuncte $\implies \cup_n A_n \in E$

Seamănă cu axiomele $\sigma-$algebrei. O asemenea familie de mulţimi se numeşte U-sistem. Cum intersecţia de U-sisteme este de asemenea un U-sistem, are sens să definim $U(M)$ ca fiind intersecţia tuturor U-sistemelor care conţin familia de mulţimi $M \subset P(\Omega)$. Faptul de bază este

$P_{|M} = Q_{|M} \implies P_{|U(M)} = Q_{|U(M)}$

TEOREMĂ. Dacă M este stabil la intersecţii finite, atunci $U(M) = \sigma(M)$

Fapt 36. JARGON. Dacă se zice: se ia un punct **la întîmplare** dintr-o mulţime finită se subînţelege că probabilitatea la care ne referim este cea uniformă. Dacă mulţimea este un compact din $\Omega \subset \mathbb{R}^n$ de măsură Lebesgue nenulă şi finită se subînţelege că probabilitatea ca punctul să fie într-o mulţime boreliană A este egală cu $\mathrm{P}(A) = \frac{\lambda^n(A\cap\Omega)}{\lambda^n(\Omega)}$; aici λ^n este măsura Lebesgue în R^n. Dacă $n = 1$, ea se numeşte lungime; dacă n=2 se numeşte arie iar dacă n=3 se numeşte volum.

Pariu cinstit: $\mathrm{P}(A) = \frac{1}{2}$; favorabil dacă $\mathrm{P}(A) \geq \frac{1}{2}$; nefavorabil dacă $\mathrm{P}(A) \leq \frac{1}{2}$

Exerciţii.

1. Se aruncă un zar de 3 ori. Care e probabilitatea ca numerele să fie toate diferite? Să vină în ordine? să nu existe două numere vecine?

2. Un segment de lungime 1 se împarte la întîmplare în trei segmente. Care e probabilitatea ca ele să vină în ordine crescătoare ? descrescătoare? să se poată forma cu ele un triunghi?

3. Se ia la întîmplare un număr de 3 cifre. Care e probabilitatea să fie par? divizibil cu 3? cu 4? dar ca să aibă cifrele puse în ordine crescătoare?

4. Se ia un număr natural la întîmplare. Care e probabilitatea să fie par (NONSENS!!)

1.2 Lecţia 2. Variabile aleatoare

Fie (Ω, K) şi (E,E) două spaţii măsurabile. O funcţie $X : \Omega \to E$ care este $K - E$ măsurabilă se numeşte variabilă aleatoare cu valori în E. Dacă scriem doar "variabilă aleatoare" fără alte epitete se subînţelege că (E,E) $= (\mathbb{R}, B(\mathbb{R}))$

Dacă (E,E) $= (\mathbb{R}^n, B(\mathbb{R}^n))$ spunem că X este un vector aleator n-dimensional. Dacă n=2 spunem că X este un punct aleator.

Proprietăţi.

1. Dacă (F,F) este un spaţiu măsurabil, $X : \Omega \to E$ este o variabilă aleatoare şi f :E→F este măsurabilă, atunci $f(X)$ este o variabilă aleatoare cu valori în F

2. Dacă X, Y sunt variabile aleatoare şi $f : R^2 \to R$ este măsurabilă (de exemplu dacă e continuă) atunci $f(X, Y)$ este de asemenea variabilă aleatoare. Deci $X+Y, X-Y, XY, X/Y, \max(X,Y), \min(X,Y), 1_B(X)$ etc sunt toate variabile aleatoare

3. Fie $L(\Omega, K)$ mulţimea variabilelor aleatoare de pe spaţiul măsurabil (Ω, K). Atunci

(i) $L(\Omega, K)$ este o algebră de funcţii reale care este şi latice

(ii) Dacă $(X_n)_n$ este un şir de variabile aleatoare şi $\liminf X_n \in R$ (respectiv $\limsup X_n \in R$) atunci $\liminf X_n$ este o variabilă aleatoare (respectiv $\limsup X_n$ este variabilă aleatoare).

4. X este un vector aleator n-dimensional dacă şi numai dacă toate componentele sale X_k sunt variabile aleatoare

Repartiţia unei variabile aleatoare. Deşi mă enervează ff tare, în ultima vreme am ajuns să accept şi denumirea "distribuţia" în loc de "repartiţia". Totuşi, distribuţie înseamnă altceva. Francezii fac diferenţa (repartiţie = loi, ruşii la fel, repartiţie=raspredelenie, nemţii la fel (verteilung) , polonezii la fel (rozklad) iar românii de la ASE şi politehnică s-au imbecilizat şi zic distribuţie.

Fie (Ω, K) şi (E,E) două spaţii măsurabile şi $X : \Omega \to$E o variabilă aleatoare. Probabilitatea $P \circ X^{-1} : E \to [0,1]$ definită prin $P \circ X^{-1}(B) = P\left(X^{-1}(B)\right)$ se numeşte repartiţia lui X.

1. Scopul suprem al teoriei probabilităţilor este să se găsească repartiţia unei variabile aleatoare

2. Dacă discutăm de variabile aleatoare fără alte epitete, repartiţiile lor sunt probabilităţi pe dreaptă. Mulţimea lor o notăm cu $\Pr(\mathbb{R}, B(\mathbb{R}))$. În acest caz repartiţia $P \circ X^{-1}$ o notăm cu F_X (mai aproape de notaţiile internaţionale). Deci $F_X(B) = P(X \in B)$. În loc să scriem $F_X((-\infty, x])$ scriem $F_X(x)$ şi spunem că avem de a face cu **funcţia de repartiţie** a lui X. Cu puţină atenţie **nu există pericol de confuzie**: Dacă $F \in \Pr(\mathbb{R}, B(\mathbb{R}))$, atunci $F(x) := F((-\infty, x])$

3. Orice funcţie $F : \mathbb{R} \to \mathbb{R}$ crescătoare, continuă la dreapta cu proprietatea că $F(-\infty) = 0, F(\infty) = 1$ este o funcţie de repartiţie. Adică există o probabilitate $\mu \in \Pr(\mathbb{R}, B(\mathbb{R}))$ astfel ca $F(x) = \mu((-\infty, x])$ pt orice $x \in \mathbb{R}$ (demonstraţia este cam lunguţă - eu le-am zis că iese din teorema lui Caratheodory).

4. Reguli de calcul. $F((a,b]) = F(b) - F(a)\,; F(\{a\}) = F(a) - F(a-0)\,;$ F$((a,b)) = F(b-0) - F(a)\,; F((x,\infty)) = 1 - F(x) = \overline{F(x)} = \underline{F(x)}$ etc. Important pentru cultura generală: $\overline{F(x)}$ se numeşte **coada** sau **funcţia de supravieţuire** şi se notează de actuari cu ${}_xp_0$.

5. Două repartiţii F şi G coincid dacă şi numai dacă $F(x) = G(x)$ pt orice x. Demonstraţia e uşoară: $B(\mathbb{R}) = U(M)$ unde $M = \{(-\infty, x] : x \in \mathbb{R}\}$ este stabilă la intersecţii finite.

Deci $F(x) = G(x)$ pt orice $x \implies F_{|M} = G_{|M} \implies F_{|U(M)} = G_{|U(M)} \implies F_{|\sigma(M)} = G_{|\sigma(M)} \implies F = G$

5. Dacă F_X este repartiţie discretă, spunem că X este discretă; dacă funcţie de repartiţie F_X este continuă spunem că X este continuă. Orice repartiţie $F \in \Pr(\mathbb{R}, B(\mathbb{R}))$ se scrie unic sub forma $F = pF_c + qF_d$ unde $p, q \geq 0, p + q = 1, F_{\acute{c}}$ este continuă şi F_d este discretă. În vorbe: orice repartiţie este o mixtură dintre una discretă şi una continuă.

6. Dacă se poate scrie $F(x) = \int p 1_{(-\infty, x)} d\lambda$ spunem că F este **absolut continuă**. O condiţie suficientă este ca F să fie continuă şi derivabilă pe porţiuni. Atunci $p = F'$. Un mod de a scrie acest lucru acceptat internaţional este: $d\acute{F}_X(x) = p(x)\,dx$.

7. Repartiţiile didactice sunt sau discrete sau absolut continui. Cele clasice poartă cîte un nume: uniformă, binomială, Poisson, etc

8. Ce e de făcut cînd se dă comanda "Să se calculeze repartiţia lui X"?

81. În general încercăm să calculăm P(X $\leq x$)

82. Dacă X este discretă încercăm să calculăm P$(X = x)$

83. Dacă X este absolut continuă, încercăm să calculăm densitatea lui X.

9. **Cuantilele**. Se notează cu $F^{-1}(p)$. A nu se confunda cu inversa: de regulă o funcţie de repartiţie $\mathbb{R} \to [0,1]$ NU este bijectiva.. Cuantila superioară $F^{-1+}(p) = \sup\{x : F(x) \leq p\}$; Cuantila inferioara $F^{-1-}(p) = \inf\{x : F(x) \geq p\}$

General: $x \in F^{-1}(p) \iff F(x-0) \leq p \leq F(x) \iff x \in [F^{-1-}(p), F^{-1+}(p)]$. Cu excepţia unei mulţimi cel mult numărabile de valori $p \in (0,1)$ cuantila superioară coincide cu cea inferioară

10. **Simularea**. Dacă U este o variabilă aleatoare repartizată Uniform$(0,1)$ atunci $X = F^{-1}(U)$ este o variabilă aleatoare cu repartiţia F. Adică $U\tilde{\ }U(0,1) \implies F^{-1}(U)\tilde{\ }F$

11. **Predicţie**. Ne asumăm un risc ε şi vrem să prezicem un interval $I = (a,b]$ unde va fi variabila aleatoare care sa aibă probabilitatea mai mare sau egală cu 1 - ε.Acesta este un **interval de predicţie de risc** $\leq \varepsilon$. Intervalul standard de risc ε este $I = (F^{-1}\left(\frac{\varepsilon}{2}\right), F^{-1}\left(1-\frac{\varepsilon}{2}\right)]$. Cu cît e mai scurt, cu atît e mai bun. De exemplu, dacă $X\tilde{\ }Exp(1)$,atunci $F^{-1}(p) = -\ln(1-p) = \ln\frac{1}{1-p}$. Deci un interval de predicţie de risc 1% este $I = (\ln\frac{1}{1-\frac{1}{200}}, \ln\frac{1}{\frac{1}{200}}] = (.00512, 5.\,298\,3]$

iar unul de risc 0.01% ar fi $(.00005, 9.\,903\,5]$

Exercitii 1. Schema bilei revenite (Bin(n,p)) ; schema bilei nerevenite (Hyper (a,n,k)) ; Poisson$(\lambda) = \lim_n$Bin$\left(n, \frac{\lambda}{n}\right)$

2. $U\tilde{\ }$U$\left(0, \frac{1}{\lambda}\right) \implies -\ln U\tilde{\ }Exp(\lambda)$;

3. Reguli de lucru cu densităţi. Dacă $p = p_X$ este densitatea lui X atunci

$p_{aX+b}(x) = \frac{1}{|a|}p\left(\frac{x-b}{a}\right); p_{X^2}(x) = \frac{1_{[0,\infty)}(x)}{2}\left(p\left(\sqrt{x}\right) + p\left(-\sqrt{x}\right)\right); p_{e^X}(x) = \frac{1_{[0,\infty)}(x)}{x}p(\ln x); p_{\ln X}(x) = e^x p(e^x)$ etc

4. Ceva drăguţ.Fie $Y = \widehat{X}\,(\mathrm{mod}\,k)$. Ce repartiţie are Y? Zicem $Y\tilde{\ }\begin{pmatrix} 0 & 1 & 2 & 3 \\ P_0 & P_1 & P_2 & P_3 \end{pmatrix}$

Daca $X\tilde{\ }Negbin(1,p)$ sau Geometric(p) este foarte uşor. Cele k valori sunt descrescătoare la negbin, la geometric e un pic diferit. $\mathrm{P}_0 = p\left(q^3 + q^7 + ...\right) = \frac{pq^3}{1-q^4}.P_1 = \frac{p}{1-q^4}, P_2 = \frac{pq}{1-q^4}, P_3 = \frac{pq^2}{1-q^4}$. Cel mai probabil este 1. Dacă $X\tilde{\ }$Bin(n,p) e mai complicat: se face cu radacinile de ordin k ale unităţii $\varepsilon_j = \cos\frac{2j\pi}{k} + i\sin\frac{2j\pi}{k}, j = 0,1,...,k-1$ şi se speculează faptul că $\varepsilon_\alpha\varepsilon_\beta = \varepsilon_{\alpha\oplus\beta}$ unde $\alpha \oplus \beta$ este adunarea din $\mathbb{Z}_k$

$(q+p)^n = q^n + C_n^1 q^{n-1}p + C_n^2 q^{n-2}p^2 + ... + C_n^{k-1}q^{n-k+1}p^{k-1} + ...$

$(q+p\varepsilon_1)^n = q^n + \varepsilon_1 C_n^1 q^{n-1}p + \varepsilon_2 C_n^2 q^{n-2}p^2 + ... + \varepsilon_{k-1}C_n^{k-1}q^{n-k+1}p^{k-1} + ...$

$(q+p\varepsilon_2)^n = q^n + \varepsilon_2 C_n^1 q^{n-1}p + \varepsilon_4 C_n^2 q^{n-2}p^2 + ... + \varepsilon_{k-2}C_n^{k-1}q^{n-k+1}p^{k-1} + ...$

....

$(q+p\varepsilon_{k-1})^n = q^n+\varepsilon_{k-2}C_n^1 q^{n-1}p+\varepsilon_{k-3}C_n^2 q^{n-2}p^2+...+\varepsilon_1 C_n^{k-1}q^{n-k+1}p^{k-1}+$ ....

Adunăm

$k\mathrm{P}(Y=0) = (q+p)^n + (q+p\varepsilon_1)^n + ... + (q+p\varepsilon_{k-1})^n$

$k\mathrm{P}(Y=1) = (q+p)^n + \varepsilon_{k-1}(q+p\varepsilon_1)^n + ... + \varepsilon_1(q+p\varepsilon_{k-1})^n$

etc

Iese ceva interesant doar în cazurile k=2,3,4. Cazul k=2 este banal

Cazul k=3: $\varepsilon_1 = \cos\frac{2\pi}{3}+i\sin\frac{2\pi}{3} = \frac{-1+i\sqrt{3}}{2}, \varepsilon_2 = \cos\frac{4\pi}{3}+i\sin\frac{4\pi}{3} = \frac{-1-i\sqrt{3}}{2}$

$3\mathrm{P}(Y=0) = (q+p)^n + (q+p\varepsilon_1)^n + (q+p\varepsilon_2)^n = 1+2\operatorname{Re}(q+p\varepsilon_1)^n$

$3\mathrm{P}(Y=1) = (q+p)^n + \varepsilon_2(q+p\varepsilon_1)^n + \varepsilon_1(q+p\varepsilon_2)^n = 1+2\operatorname{Re}\varepsilon_2(q+p\varepsilon_1)^n$

$3\mathrm{P}(Y=2) = (q+p)^n + \varepsilon_1(q+p\varepsilon_1)^n + \varepsilon_2(q+p\varepsilon_2)^n = 1+2\operatorname{Re}\varepsilon_1(q+p\varepsilon_1)^n$

Scriem $q+p\varepsilon_1 = q+p\frac{-1+i\sqrt{3}}{2} = q-\frac{p}{2}+i\frac{p\sqrt{3}}{2} = re^{i\phi}$ cu $r=\sqrt{p^2-pq+q^2} = \sqrt{1-3pq}$

De unde $(q+p\varepsilon_1)^n = (1-3pq)^{\frac{n}{2}}e^{in\phi}, \varepsilon_2(q+p\varepsilon_1)^n = (1-3pq)^{\frac{n}{2}}e^{i\left(n\phi+\frac{4\pi}{3}\right)}, \varepsilon_1(q+p\varepsilon_1)^n = (1-3pq)^{\frac{n}{2}}e^{i\left(n\phi+\frac{2\pi}{3}\right)}$ şi rezultă

$3\mathrm{P}(Y=0) = 1+2(1-3pq)^{\frac{n}{2}}\cos n\phi$

$3\mathrm{P}(Y=1) = 1+2(1-3pq)^{\frac{n}{2}}\cos\left(n\phi+\frac{4\pi}{3}\right)$

$3\mathrm{P}(Y=2) = 1+2(1-3pq)^{\frac{n}{2}}\cos\left(n\phi+\frac{2\pi}{3}\right)$

Ce e interesant este că niciodată cele trei probabilităţi nu pot coincide.

Cazul k=4. $\varepsilon_1 = i, \varepsilon_2=-1, \varepsilon_3=-i$. Fie $P_j = P(Y=j)$.

$4P_0 = (q+p)^n+(q+p\varepsilon_1)^n+(q+p\varepsilon_2)^n+(q+p\varepsilon_3)^n \quad = 1+(q+pi)^n+(q-p)^n+(q-pi)^n$

$4P_1 = (q+p)^n+\varepsilon_3(q+p\varepsilon_1)^n+\varepsilon_2(q+p\varepsilon_2)^n+\varepsilon_1(q+p\varepsilon_3)^n \quad = 1-i(q+pi)^n-(q-p)^n+i(q-pi)^n$

$4P_2 = (q+p)^n+\varepsilon_2(q+p\varepsilon_1)^n+(q+p\varepsilon_2)^n+\varepsilon_2(q+p\varepsilon_3)^n \quad = 1-(q+pi)^n+(q-p)^n-(q-pi)^n$

$4P_3 = (q+p)^n+\varepsilon_1(q+p\varepsilon_1)^n+\varepsilon_2(q+p\varepsilon_2)^n+\varepsilon_3(q+p\varepsilon_3)^n \quad = 1+i(q+pi)^n-(q-p)^n-i(q-pi)^n$

sau, daca scriem $q+pi = e^{i\phi}$

$4P_0 = 1+(q-p)^n+2\operatorname{Re}(q+pi)^n \quad = 1+(q-p)^n+2(1-3pq)^{\frac{n}{2}}\cos n\phi$

$4P_1 = 1-(q-p)^n+2\operatorname{Re}i(q-pi)^n \quad = 1-(q-p)^n-2(1-3pq)^{\frac{n}{2}}\cos\left(n\phi+\frac{\pi}{2}\right)$

$4P_2 = 1+(q-p)^n-2\operatorname{Re}(q-pi)^n \quad = 1+(q-p)^n-2(1-3pq)^{\frac{n}{2}}\cos(n\phi)$

$4P_3 = 1-(q-p)^n+2\operatorname{Re}i(q+pi)^n \quad = 1-(q-p)^n+2(1-3pq)^{\frac{n}{2}}\cos\left(n\phi+\frac{\pi}{2}\right)$

Rezultatul final este

$Y\tilde{}\left(\begin{array}{cccc} 0 & 1 & 2 & 3 \\ \frac{1+(q-p)^n+2(1-3pq)^{\frac{n}{2}}\cos n\phi}{4} & \frac{1-(q-p)^n+2(1-3pq)^{\frac{n}{2}}\sin n\phi}{4} & \frac{1+(q-p)^n-2(1-3pq)^{\frac{n}{2}}\cos n\phi}{4} & \frac{1-(q-p)^n-2(1-3pq)^{\frac{n}{2}}\,\mathrm{s}}{4} \end{array}\right.$

Ar putea fi toate egale?

daca $\cos n\phi = 0, \sin n\phi = 1$ rezultă $Y\tilde{}\begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1+(q-p)^n}{4} & \frac{1-(q-p)^n+2(1-3pq)^{\frac{n}{2}}}{4} & \frac{1+(q-p)^n}{4} & \frac{1-(q-p)^n-2(1-3pq)^{\frac{n}{2}}}{4} \end{pmatrix}$

Nu pot fi egale niciodata

Dar daca $X\tilde{}$Poisson(λ) şi $Y=\widehat{X}\,(\operatorname{mod} k)$?

Reţeta este aceeaşi

$ke^{\lambda}P_0 = e^{\lambda}+e^{\lambda\varepsilon_1}+...+e^{\lambda\varepsilon_{k-1}}, ke^{\lambda}P_1 = e^{\lambda}+\varepsilon_{k-1}e^{\lambda\varepsilon_1}+...+\varepsilon_1 e^{\lambda\varepsilon_{k-1}}, ke^{\lambda}P_2 = e^{\lambda}+\varepsilon_{k-2}e^{\lambda\varepsilon_1}+...+\varepsilon_2 e^{\lambda\varepsilon_{k-1}}$,etc

Dacă $k = 3$ avem

$3P_0 = 1 + e^{\lambda(\varepsilon_1-1)} + e^{\lambda(\varepsilon_2-1)} = 1 + 2\operatorname{Re} e^{\lambda(\varepsilon_1-1)}$

$3P_1 = 1 + \varepsilon_2 e^{\lambda(\varepsilon_1-1)} + \varepsilon_1 e^{\lambda(\varepsilon_2-1)} = 1 + 2\operatorname{Re}\varepsilon_2 e^{\lambda(\varepsilon_1-1)}$,

$3P_2 = 1 + \varepsilon_1 e^{\lambda(\varepsilon_1-1)} + \varepsilon_2 e^{\lambda(\varepsilon_2-1)} = 1 + 2\operatorname{Re}\varepsilon_1 e^{\lambda(\varepsilon_1-1)}$

Dar $e^{\lambda(\varepsilon_1-1)} = e^{\lambda\left(\frac{-1+i\sqrt{3}}{2}-1\right)} = e^{-\frac{3\lambda}{2}}e^{\frac{i\sqrt{3}\lambda}{2}}, \varepsilon_1 e^{\lambda(\varepsilon_1-1)} = e^{-\frac{3\lambda}{2}}e^{\frac{i\sqrt{3}\lambda}{2}+\frac{2\pi}{3}}, \varepsilon_2 e^{\lambda(\varepsilon_1-1)} = e^{-\frac{3\lambda}{2}}e^{\frac{i\sqrt{3}\lambda}{2}+\frac{4\pi}{3}}$

Deci

$P_0 = \frac{1+2e^{-\frac{3\lambda}{2}}\cos\frac{\sqrt{3}\lambda}{2}}{3}$

$P_1 = \frac{1+2e^{-\frac{3\lambda}{2}}\cos\left(\frac{\sqrt{3}\lambda}{2}+\frac{2\pi}{3}\right)}{3}$

$P_2 = \frac{1+2e^{-\frac{3\lambda}{2}}\cos\left(\frac{\sqrt{3}\lambda}{2}+\frac{4\pi}{3}\right)}{3}$

Şi nici acum nu pot fi niciodată egale. Abaterea de la $\frac{1}{3}$ este intotdeauna mai mică decît $2e^{-\frac{3\lambda}{2}}$

De exemplu, dacă

$\lambda = \frac{2\pi}{3\sqrt{3}} \implies Y\tilde{}\begin{pmatrix} 0 & 1 & 2 \\ 0.38768 & 0.22464 & 0.38768 \end{pmatrix}; \lambda = 1 \implies Y\tilde{}\begin{pmatrix} 0 & 1 & 2 \\ 0.4297 & 0.18701 & 0.38327 \end{pmatrix}$

$\lambda = 2 \implies Y\tilde{}\begin{pmatrix} 0 & 1 & 2 \\ 0.328 & 0.30763 & 0.36437 \end{pmatrix}; \lambda = 3 \implies Y\tilde{}\begin{pmatrix} 0 & 1 & 2 \\ 0.32699 & 0.33319 & 0.33982 \end{pmatrix}$

$\lambda = \frac{8\pi}{3\sqrt{3}} \implies Y\tilde{}\begin{pmatrix} 0 & 1 & 2 \\ 0.3331 & 0.3338 & 0.3331 \end{pmatrix}$

5. Simulari. Sa se simuleze o variabila aleatoare repartizată Pareto: $F(x) = \min\left(\left(\frac{a}{x}\right)^n, 1\right)$.Etc.

$F(x) = 1 - \left(\frac{a}{x}\right)^n$ dacă $x \geq a$. Cuantilele sunt soluţiile ecuaţiilor $1 - \left(\frac{a}{x}\right)^n = p \iff \frac{a}{x} = (1-p)^{\frac{1}{n}} \iff x = a(1-p)^{-\frac{1}{n}}$

Deci $X = a(1-U)^{-\frac{1}{n}}$. Cum U şi 1 - U au aceeaşi repartiţie $\mathrm{U}(0,1)$,putem pune $X = aU^{-\frac{1}{n}}$

6. Simulaţi $X\tilde{}Bin(1,p)$. De exemplu $X = \min\left(\left[\frac{U}{1-p}\right], 1\right)$

7. Simulaţi $X\tilde{}Hyper(a,n,k)$. Aici e mai greu, trebuie simulată o urnă. Instrucţiunea sample(n,k) din R face aceasta: simulează extragerea a k bile din n bile fără revenire. Un scurt script care simulează repartiţia în cauză este următorul

```
a=10;n=15;t=a+n;k=19
omega=sample(t,k);x=length(which(omega<=a));x
```

Se simulează extragerea a 15 bile din 25 de bile, din care 10 sunt albe şi 15 sunt negre.

După 100 de simulări am obţinut abelul de frecvenţe table(x)=$\begin{pmatrix} 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 8 & 16 & 25 & 29 & 17 & 4 \end{pmatrix}$

pe care îl putem compara cu adevărata repartiţie $X\tilde{}\frac{1}{100}\left(\begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \\ \frac{10}{326876} & \frac{375}{81719} & \frac{23625}{163438} & \frac{136500}{81719} & \frac{716625}{81719} & \frac{171990}{7429} \end{matrix}\right.$

$X\tilde{}\frac{1}{100}\left(\begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \\ 0.00001 & 0.00459 & 0.14455 & 1.6704 & 8.7694 & 23.151 & 32.154 & 23.624 & 8.8589 & 1.5312 & 0.0 \end{matrix}\right.$

1.3 Lecţia 3. Integrala

Dacă (Ω, K, P) este un spaţiu probabilizat şi X este o variabilă aleatoare, $\int XdP$ se notează cu $\mathrm{E}X$.

Proprietăţi specifice

1. $\int 1_A dP = P(A)$

2. $\int 1 dP = 1$

3. Inegalitatea lui Jensen. Dacă $f : I \to \mathbb{R}$ este convexă, $X : \Omega \to I \subset \mathbb{R}$ este o variabilă aleatoare, atunci $\mathrm{E}f(X) \geq f(\mathrm{E}X)$

4. Inegalitatea normelor. Dacă $1 \leq p \leq q \leq \infty$ atunci $\|X\|_p \leq \|X\|_q$. Ca atare $\mathrm{L}^p(\Omega, K, P) \subset \mathrm{L}^q(\Omega, K, P) \subset \mathrm{L}^\infty(\Omega, K, P)$

5. Dacă Supp(P) nu este finit, atunci toate spaţiile L^p sunt diferite iar incluziunea $\mathrm{L}^\infty(\Omega, K, P) \subset \bigcap\limits_{p<\infty} \mathrm{L}^p(\Omega, K, P)$ este stricta

6. Cum calculăm $\mathrm{E}X$.

Formula de transport

$$\mathrm{E}g(X) = \int g dF_X = \begin{cases} \sum\limits_{x \in I} g(x)\, p(x) & \text{dacă} \quad X \text{ este discretă, } \mathrm{F}_X = \sum\limits_{x \in I} p(x)\, \delta_x \\ \int_{-\infty}^{\infty} g(x)\, p(x)\, dx & \text{dacă} \quad X \text{ este absolut continuă, } \mathrm{dF}_X(x) = p(x)\, dx \end{cases}$$

7. Comutarea dintre medie şi limită

7.1. Teorema lui Beppo-Levi:

Dacă $X_n \nearrow X, X_n \geq Z, EZ \neq -\infty \Longrightarrow EX_n \nearrow EX$.

Dacă $X_n \searrow X, EX_n \geq Z, EZ \neq \infty \Longrightarrow EX_n \searrow EX$.

7.2. Teorema Lui Lagrange: $X_n \to X, |X_n| \geq Z, E|Z| \neq \infty \Longrightarrow EX_n \to EX$.

8. **Dar de ce să calculăm media**?

A. Din necesităţi de predicţie. Nu întotdeauna ştim să calculăm funcţia de repartiţie ca să dăm un interval de predicţie la un risc dat, ε. Sunt mai multe cazurile în care putem calcula $\mathrm{E}X$

B. Fiindcă trebuie să putem aproxima o variabilă aleatoare cu un număr (de exemplu o primă de asigurare sau un preţ). Se pune problema să găsim **cea mai bună constantă cu care să putem aproxima o variabilă aleatoare**. Dacă punem problema aşa, ne trebuie un criteriu de optim: după ce cunosc că o constantă a este mai bună ca altă constantă, b? Un asemenea criteriu este dat de regulă de o distanţă: calculăm distanţa de la constanta a la variabila aleatoare X şi vrem să fie cît mai mică. Pe diverse submulţimi de variabile aleatoare există mai multe asemenea distanţe deja studiate. Cele mai la îndemînă sunt distanţele din $\mathrm{L}^p(\Omega, K, P) : \mathrm{d}_p(a, X) = \|X - a\|_p$. Aşadar, dacă ştim că variabila $X \in L^p(\Omega, K, P)$ putem căuta acea constantă a care minimizează funcţia $d_p(x) = \|X - a\|_p$.

Formal, cea mai bună constantă din punctul de vedere L^p este $a = \arg\min d_p$

Dacă $p = 1$,obţinem $a = Median(X)$

Dacă $p = 2$ obţinem $a = EX$

Dacă $p = \infty$ obţinem $a = Midrange(X) = \frac{1}{2}(m + M)$ unde

m=essinf$X := \sup\{x : F_X(x) = 0\}$, M = esssup $X := \inf\{x : F_X(x) = 1\}$

Cît de bună este aproximarea?

8. **Inegalitatea lui Cebîşev.** Cantitatea $d_2^2(EX) = E(X-EX)^2 = EX^2 - E^2X$ se numeşte varianţa lui X şi se notează cu Var(X) sau $\sigma^2(X)$. Atunci $\sigma(X)$ se numeşte abaterea medie pătratică a lui X.

În plus, are loc inegalitatea $\mathrm{P}(|X-EX| \geq k\sigma(X)) \leq \frac{1}{k^2}$

Aceasta este inegalitatea lui Cebîşev.

Scrisă pozitiv, ea devine

$\mathrm{P}(EX - k\sigma(X) \leq X \leq EX + k\sigma(X)) \geq 1 - \frac{1}{k^2}$

formă care este simplă şi uşor de înţeles. Deşi inegalitatea în sine este foarte slabă.

Exerciţii 1. Dintr-o urnă cu bilele numerotate $0, \frac{1}{n}, \frac{2}{n}, ..., \frac{n-1}{n}$se scot două bile (cu sau fără înlocuire). Fie X maximul lor şi Y minimul lor. Să se calculeze EX,EY,Var$X$,Var Y. Treceţi la limită

2. Se iau două numere la întîmplare între 0 şi 1. Fie X maximul lor şi Y minimul lor. Aceeaşi problemă ca şi mai sus

3. Diverse probleme luate absolut oricum: se ia o funcţie F care sa fie pe undeva crescătoare, se potriveşte din x şi y ca să fie funcţie de repartitie, i se calculează apoi densitatea (dacă are - 99% că vi se va da o funcţie continuă) şi i se calculează media şi varianţa

1.4 Lecţia 4. Procedee de calcul al mediei

1. Funcţia generatoare de momente (mgf)

Momentul de ordin n este prin definiţie numărul μ_n =EX^n.Funcţia generatoare de momente este $\mathrm{m}_X(t)$ =Ee^{tX}.Notăm $D_X = \{t \in R : m_X(t) \neq \infty\}$. Dacă $D_X = R$ variabila e cu coadă foarte scurtă, dacă D_X =$\{0\}$ este cu coadă lungă. Pentru variabile aleatoare pozitive D_X conţine întotdeauna intervalul $(-\infty, 0]$. Dacă există $h > 0$ astfel ca $\mathrm{m}_X(t) < \infty$ pentru orice $x < h$ atunci spunem că X are coadă medie. În general X are coadă medie dacă există $h > 0$ astfel ca E$e^{t|X|} < \infty$ pentru orice $x < h$

1.1 Dacă X are coadă medie, atunci are toate momentele absolute finite: $X \in \bigcap\limits_{1 \leq p < \infty} L^p$

Evident, dacă există k ca $\mathrm{E}|X|^k = \infty$ şi $t > 0$, atunci $\mathrm{m}_X(t) = E\sum\limits_{n=0}^{\infty} t^n \frac{|X|^n}{n!} = \sum\limits_{n=0}^{\infty} t^n \frac{E|X|^n}{n!}$ (Beppo - Levi) $\geq t^k \frac{E|X|^k}{k!} = \infty$.

1.2. Dacă X are coadă medie atunci putem scrie $\acute{\mathrm{m}}_X(t) = \sum\limits_{n=0}^{\infty} t^n \frac{EX^n}{n!}$ (putem comuta limita cu suma datorită teoremei lui Lebesgue, seria fiind dominată de $e^{|tX|}$. Dacă putem dezvolta în serie pe m_X obţinem momentele lui X.

1.3. Exemple.

1.3.1. X˜Exp(λ) $\Longrightarrow$ $m_X(t) = \frac{\lambda}{\lambda - t}, 1D_X = (-\infty, \lambda)$; are coadă medie. Dezvoltăm în serie

$\mathrm{m}_X(t) = 1 + \frac{t}{\lambda} + \frac{t^2}{\lambda^2} + ...$ şi identificăm seria cu $\mathrm{m}_X(t) = 1 + t\mu_1 + \frac{t^2\mu_2}{2!} + ...$; rezultă $\mu_n = \frac{n!}{\lambda^n}$

1.3.2.. $X\tilde{}\mathrm{N}(0,1) \iff m_X(t) = e^{\frac{t^2}{2}}$ (e uşor: $m_X(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tx-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2-2tx+t^2}{2}+\frac{t^2}{2}}dx = e^{\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2}}dx = e^{\frac{t^2}{2}} \cdot 1$)

Deci $m_X(t) = e^{\frac{t^2}{2}} = 1 + \frac{t^2}{2} + \frac{t^4}{2^2 2!} + ... + \frac{t^{2n}}{n!2^n} + ... = 1 + t\mu_1 + \frac{t^2\mu_2}{2!} + \frac{t^3}{3!}\mu_3 + ... + \frac{t^{2n}}{(2n)!}\mu_{2n} + ...$

Deci $\mu_{2n+1} = 0$ pentru orice număr natural n iar $\frac{\mu_{2n}}{(2n)!} = \frac{1}{n!2^n} \iff \mu_{2n} = \frac{(2n)!}{n!2^n} = 1 \cdot 3 \cdot 5 \cdot ... \cdot (2n-1) = (2n-1)!!$

1.3.3. Repartiţia normală. Dacă $Z\tilde{}\mathrm{N}(0,1)$ atunci variabila $X = \mu + \sigma Z$ are densitatea $p_X(x) = \frac{1}{\sqrt{2\pi}}e^{+\frac{(x-\mu)^2}{2\sigma^2}}$, funcţia generatoare de momente $\mathrm{m}_X(t) = e^{t\mu+\frac{t^2\sigma^2}{2}}$, funcţia caracteristică $\phi_X(t) = e^{it\mu-\frac{t^2\sigma^2}{2}}$. Aceasta este repartiţia normală de medie μ şi varianţă σ^2. Ea se notează cu $\mathrm{N}(\mu, \sigma^2)$.

2. Funcţia generatoare de cumulanţi (cgf)

Are formula $\mathrm{k}_X = \ln m_X$

E interesantă fiindcă $\mathrm{k}'_X(0) = \mu_1$,$\mathrm{k}''_X(0) = \mu_2 - \mu_1^2 = \mathrm{Var}(X)$,$\mathrm{k}'''_X(0) = \mathrm{E}(X-\mu_1)^3$

3. Funcţia caracteristică

Corectează marele defect al mgf-ului: acela că D_X nu e una şi aceeaşi mulţime pentru orice X.

$\phi_X(t) = \mathrm{E}e^{itX} = \mathrm{E}\cos(tX) + i\mathrm{E}\sin(tX)$

În plus există şi

Teorema de unicitate: $\phi_X = \phi_Y \iff F_X = F_Y$

Pe vremuri o demonstram şi pe asta.

4. Funcţia generatoare

Se foloseşte pentru variabile aritmetice, adică $X : \Omega \to \mathbb{N}$

$g_X : [0,1] \to \mathbb{R}_+$, $g_X(x) = Ex^X = \sum_{n=0}^{\infty} P(X=n)\, x^n.$

Proprietatea ei cea mai importantă este că prid derivare se obţin momentele factoriale

$g_X^{(n)}(1-0) = \mathrm{E}X(X-1)...(X-n+1)$

Orice funcţie analitică $g : [0,1] \to \mathbb{R}_+$ cu proprietatea că $g(1) = 1$ şi care are toate derivatele pozitive este funcţia generatoare a unei probabiliăţi pe $(\mathbb{N}, P(\mathbb{N}))$.

Consecinţe

$g_X(0) = P(X=0)$; $EX = g'_X(1-0)$;$\mathrm{Var}(X) = g''_X(1-0) + g'_X(1) - g'_X(1)^2$

Exemple

4.1. Repartiţia Hipergeometrică $\mathrm{Hyper}(a,n,k) = \sum_{j=0}^{k} \frac{C_a^j C_n^{k-j}}{C_{a+n}^k}\delta_j$.Funcţia sa generatoare $g_{a,n,k}$ nu are o formă analitică cunoscută: $g_{a,n,k}(x) = \sum_{j=0}^{k} \frac{C_a^j C_n^{k-j}}{C_{a+n}^k}x^j$

Dar se poate folosi cu succes datorită recurenţei $g'_{a,n,k}(x) = k\frac{a}{a+n}g_{a-1,n,k-1}(x)$ şi a faptului că orice funcţie generatoare are proprietatea că $g(1) = 1 : g'(1) =$

$EX = k\frac{a}{a+n}, g''(1) =EX(X-1) = k(k-1)\frac{a}{a+n}\frac{a-1}{a+n-1}$. Astfel Var$(X)$ = $k(k-1)\frac{a}{a+n}\frac{a-1}{a+n-1} + k\frac{a}{a+n} - \left(k\frac{a}{a+n}\right)^2 = k\frac{a}{a+n}\left(\frac{(k-1)(a-1)}{a+n-1} + 1 - k\frac{a}{a+n}\right)$ sau ,cu notaţiile $p = \frac{a}{a+n}, q = \frac{n}{a+n}$ găsim

E$X = kp$, Var$(X) = kpq\frac{a+n-k}{a+n-1} = kpq\left(1 - \frac{k-1}{a+n-1}\right)$

4.2. Repartiţia Binomială Bin(k,p) are funcţia generatoare foarte simplă: $g(x) = (q+px)^k$ unde $q = 1-p$. Deci

E$X = kp$, Var$(X) = kpq$

Dacă raportul $\frac{k}{a+n}$ este mic, se poate aproxima Hipergeometrica cu Binomiala.

4.3. Repartiţia Poisson (λ) : $\mathrm{g}_X(x) = e^{\lambda(x-1)}$. deci

EX = Var$(X) = \lambda$. Este legea evenimentelor rare: dacă $kp = \lambda$ şi p este foarte mic, se poate aproxima Binomiala cu Poisson. Momentele factoriale formează o progresie geometrică.

4.4. Repartiţiile geometrică şi Negbin: $g_{Geoma(p)}(x) = \frac{px}{1-qx}$ şi $g_{Negbin(1,p)}(x) = \frac{p}{1-qx}$

5. Formula de integrare prin părţi. Lucrul cu cozile.

5.1. Fie $X \geq 0$ o variabilă aleatoare şi $g : \mathbb{R} \to \mathbb{R}$ o funcţie continuă derivabilă pe porţiuni. Presupunem că $g(X) \in L^1$.

Atunci este valabilă formula

E$g(X) = g(0) + \int_0^\infty g'(y)\underline{F}(y)\,dy$ unde $\underline{F}(y) = P(X > y)$ este coada (dreaptă) a lui Y.

Exemple: E$X = \int_0^\infty \underline{F}(y)\,dy, E(X-a)_+ = \int_a^\infty \underline{F}(y)\,dy, E\min(X,a) = \int_0^a \underline{F}(y)\,dy$ (aici se subînţelege că $a \geq 0$!)

Demonstraţie. Plecăm de la formula Leibniz Newton. Aici λ este măsura Lebesgue

$g(X) - g(0) = \int_0^X g'(y)\,dy = \int g'(y)1_{[0,X]}(y)\,d\lambda(y) = \int g'(y)1_{[0,X)}(y)\,d\lambda(y)$ (măsura Lebesgue neglijează punctele!!)

$= \int g'(y)1_{[0,\infty)}(y)1_{(y,\infty)}(X)\,d\lambda(y)$

Deci , aplicînd teorema lui Fubini (am presupus în ipoteză că $g(X) \in L^1(P)$ $\Longleftrightarrow g'(y)1_{(y,\infty)}(X) \in L^1(P \otimes \lambda)$)

E$g(X) - g(0) = \int \left(\int g'(y)1_{[0,\infty)}(y)1_{(y,\infty)}(X)\,d\lambda(y)\right)dP = \int \left(\int g'(y)1_{[0,\infty)}(y)1_{(y,\infty)}(X)\,dP\right)d\lambda(y)$

$= \int g'(y)1_{[0,\infty)}(y)\left(\int 1_{(y,\infty)}(X)\,dP\right)d\lambda(y) = \int_0^\infty g'(y)\left(\int 1_{(y,\infty)}(X)\,dP\right)d\lambda(y) = \int_0^\infty g'(y)P(X > y)\,dy$.

5.2. Fie $X \geq 0$ o variabilă aleatoare şi $g : \mathbb{R} \to \mathbb{R}$ o funcţie continuă derivabilă pe porţiuni. Presupunem că $g(X) \in L^1$.

Atunci este valabilă formula

E$g(X) = \int_0^\infty (g'(y)\underline{F}(y) - g'(-y)F(-y))\,dy$ unde $\underline{F}(y) = P(X > y), F(y) = P(X \leq y)$

Demonstraţie. Scriem $X = X_+ - X_-$. Atunci $g(X) = g(X_+) + g(-X_-)$ şi aplicăm formula de dinainte.

Exemplu: E$X = \int_0^\infty (\underline{F}(y) - F(-y))\,dy$. Interpretare geometrică.

Excerciţii. Calculaţi diverse medii şi varianţe.

Arătaţi că dacă X este o variabilă aleatoare aritmetică, $\mathrm{E}X = \sum_{n=0}^{\infty} P(X > n)$, $\mathrm{E}X^2 = \sum_{n=0}^{\infty} (2n+1) P(X > n)$ etc

1.5 Lecţia 5. Vectori aleatori

Dacă $X : \Omega \to \mathbb{R}^n$ este un vector aleator, repartiţia sa F_X este o probabilitate pe $(\mathbb{R}^n, B(\mathbb{R}^n))$

Cum o descriem?

1. Funcţia de repartiţie. Se defineşte la fel ca în cazul unidimensional, cu aceeaşi convenţie

$F_X(x) = P(X \le x) := P(X_1 \le x_1, X_2 \le x_2, ..., X_n \le x_n)$; $F_X(B) = P(X \in B)$

Nu e uşor de lucrat cu ea. Oricum, rezultatul unidimensional $F_X = F_Y \iff F_X(x) = F_Y(x)$ pt orice $x \in \mathbb{R}^n$ se păstrează.

Motivul? Mulţimea M a intervalelor $(-\infty, x]$ este stabilă la intersecţii finite (evident $(-\infty, x] \cap (-\infty, y] = (-\infty, \min(x, y)]$!!) şi generează mulţimile din $B(\mathbb{R}^n) = B(\mathbb{R})^{\otimes n}$

Cazul $n = 2$. X=(X_1, X_2),x=(x_1, x_2), $F_{\mathrm{X}}(\mathrm{x}) = P(X_1 \le x_1, X_2 \le x_2) = F_{\mathrm{X}}((-\infty, x_1] \times (-\infty, x_2])$ deci

$F_{\mathrm{X}}((a_1, b_1] \times (a_2, b_2]) = F_{\mathrm{X}}(a_1, a_2) + F_{\mathrm{X}}(b_1, b_2) - F_{\mathrm{X}}(a_1, b_2) - F_{\mathrm{X}}(a_2, b_1)$

Proprietăţi. $F : \mathbb{R}^2 \to [0,1]$ este o funcţie de repartiţie bidimensională dacă şi numai dacă

α. F$(-\infty, y)$ =F$(x, -\infty) = 0$ pentru orice $x, y \in \mathbb{R}$

$\beta . x \rightarrowtail F(x, \infty), y \rightarrowtail F(\infty, y)$ sunt funcţii de repartiţie unidimensionale

$\gamma . F(a_1, a_2) + F(b_1, b_2) - F(a_1, b_2) - F(a_2, b_1) \ge 0$ pentru orice $a_1 \le b_1, a_2 \le b_2$.

Exemplu.

1\. 1. Dacă $F_X = \delta_{(a,b)}$ (repartiţia Dirac în punctul a, b) atunci $F_X(x, y) = 1_{[a,\infty) \times [b,\infty)}(x, y) = 1_{[a,\infty)}(x) 1_{[b,\infty)}(y)$. Consecinţă: dacă $\mathrm{F}_X = \sum_j p_j \delta_{(a_j, b_j)}$ atunci $F_X(x, y) = \sum_j p_j 1_{[a_j,\infty)}(x) 1_{[b_j,\infty)}(y)$

1.2. Se iau două numere la întîmplare , U şi V intervalul $[0,1]$. Fie $X = \min(U, V)$, $Y = \max(U, V)$.

Atunci P$(X \le x, Y \le y)$ =P$(\min(U, V) \le x, \max(U, V) \le y)$ =P$(((U \le x) \cup (V \le x)) \cap (U \le y, V \le y))$

=P$((U \le \min(x, y), V \le y) \cup (U \le y, V \le \min(x, y)))$

=P$(U \le \min(x, y), V \le y)$ +P$(U \le y, V \le \min(x, y))$ −P$(U \le \min(x, y), V \le \min(x, y)) = 2y \min(x, y) - \min(x, y)^2$

Deci $F(x, y) = \begin{cases} 2xy - y^2 & \text{dacă} \quad 0 \le x \le y \le 1 \\ y^2 & \text{dacă} \quad 0 \le y \le x \le 1 \end{cases}$.

Dacă $x \ge 1$ sau $y \ge 1$ avem $F(x, y) = 2 \min(x, y) - \min(x, y)^2$

O funcţie de repartiţie $G : [0,1]^2 \to [0,1]$ cu proprietatea că $G(x, 1) = G(1, x) = x$ se numeşte **copulă**.

De exemplu $G(x, y) = xy$ sau $G(x, y) = \min(x, y)$ sunt copule.

1.3 Exerciţiu. Arătaţi că $F(x,y) = \max(x,y)$ NU este o funcţie de repartiţie. Într-adevăr, dacă $a_1 = a_2, b_1 = b_2$ atunci

$\max(a_1, a_2) + \max(b_1, b_2) - \max(a_1, b_2) - \max(a_2, b_1) = a + b - b - b = a - b < 0.$

2. Repartiţiile didactice sunt cele discrete sau absolut continue.

Repartiţia este discretă dacă, la fel ca în cazul unidimensional, se poate scrie sub forma $F_X = \sum\limits_{x \in A} p(x)\delta_x$ unde $A \subset \mathbb{R}^n$ este o mulţime cel mult numărabilă. Un criteriu suficient ca acest lucru să se întîmple ete ca $X(\Omega)$ să fie o mulţime cel mult numărabilă.

Repartiţia este absolut continuă dacă există o funcţie nenegativă şi măsurabilă $p : \mathbb{R}^n \to \mathbb{R}_+$ cu proprietatea că $F_X(B) := P(X \in B) = \int p 1_B d\lambda^n$. Orice funcţie $p \geq 0$ cu proprietatea că $\int p d\lambda^n = 1$ este o densitate.

2.1. O urnă are a_1 bile de culoarea 1,a_2 bile de culoarea 2,...,a_n bile de culoarea n. Se extrag k bile. X_1 sunt de culoarea 1,X_2 sunt de culoarea 2, ..., X_n sunt de culoarea n.

Care este repartiţia vectorului $X = (X_j)_{1 \leq j \leq n}$?

Dacă bila nu se pune la loc, evident $\mathrm{P}(X = x) = \frac{C_{a_1}^{x_1} C a_2^{x_2} ... C_{a_n}^{x_n}}{C_{a_1 + a_2 + ... + a_n}^{k}}$ dacă $x_1 + ... + x_n = k, x_j \geq 0 \quad \forall 1 \leq j \leq n$ sau $\mathrm{P}(X = x) = 0$ altfel.

Dacă bila se pune la loc, atunci $\mathrm{P}(X = x) = \frac{k!}{x_1! ... x_n!} p_1^{x_1} p_2^{x_2} ... p_n^{x_n}$ în aceleaşi condiţii. Aici $p_j = \frac{a_j}{a_1 + a_2 + ... + a_n}$.

2.2. Se aruncă un punct X la întîmplare într-o mulţime $B \subset \mathbb{R}^n$ măsurabilă. Presupunem că $\lambda^n(B) \in (0, \infty)$. Atunci

$F_X(A) = P(X \in A) = \frac{\lambda^n(A \cap B)}{\lambda^n(B)}$ este absolut continuu şi are densitatea $p = \frac{1_B}{\lambda_{(B)}}$.

3. Majoritatea vectorilor aleatori au repartiţii care nu sunt nici discrete, nici absolut continue.

De exemplu, dacă luăm o variabilă aleatoare ξ absolut continuă şi formăm vectorul $X = (\xi, \xi)$, acesta va avea o repartiţie care nu este discretă (căci $\mathrm{P}(X = x) = 0$ pentru orice x) dar nici absolut continuă, deoarece este concentrată pe diagonala $\Delta = \{(x,x) : x \in \mathbb{R}\}$. Dacă $F(t) = P(\xi \leq t)$ este funcţia de repartiţie a lui ξ, atunci $F_X(x,y) = P(\xi \leq \min(x,y)) = F(\min(x,y))$

4. **Repartiţiile marginale**. Fie X un vector aleator n-dimensional. Repartiţia componentei X_j se numeşte marginala j a lui F_X. Dacă am şti funcţia de repartiţie F_X, atunci este evident că

$F_{X_j}(x_j) = P(X_j \leq x_j) = P(X_1 \leq \infty, ..., X_{j-1} \leq \infty, X_j \leq x_j, X_{j+1} \leq \infty, ..., X_n \leq \infty) = F_X(\infty, ..., \infty, x_j, \infty, ..., \infty)$

Dacă X este absolut continuu, atunci X_j sunt de asemenea absolut continui. Dacă $F_X = p \cdot \lambda^n$, atunci $\mathrm{p}_{X_j}(x_j) = \int ... \int p(x_1, ..., x_{j-1}, x_j, x_{j+1}, ..., x_n) dx_1 ... dx_{j-1} dx_{j+1} ... dx_n$

De exemplu

- în cazul $n = 2 : p_{X_1}(x_1) = \int p(x_1, x_2) dx_2, p_{X2}(x_2) = \int p(x_1, x_2) dx_1$

- în cazul $n = 3 : p_{X_1}(x_1) = \int\int p(x_1, x_2, x_3) dx_2 dx_3, p_{X_2}(x_2) = \int\int p(x_1, x_2, x_3) dx_1 dx_3, p_{X_3}(x_3) = \int\int p(x_1, x_2, x_3) dx_1 dx_2$

Reciproc nu este adevărat, vezi punctul 3 de mai sus. Deşi ξ este absolut continuă, vectorul (ξ,ξ) nu este.

4.1. O urnă are a$_1$ bile de culoarea 1,a$_2$ bile de culoarea 2,...,a$_n$ bile de culoarea n. Se extrag k bile. X_1 sunt de culoarea 1,X_2 sunt de culoarea 2, ..., X_n sunt de culoarea n.

Care este repartiţiile marginale ale vectorului $X=(X_j)_{1\le j\le n}$?

R. Dacă nu se pune bila la loc sunt Hypergeometric$(a_j,N-a_j,k)$ cu $N=a_1+a_2+...+a_n$

Dacă se pune bila la loc sunt Binomial (k,p_j)

4.2. X este repartizat uniform pe discul unitate $D=\{(x,y):x^2+y^2\le 1\}$. Care sunt densităţile marginale?

R. $p_{X_1}(x)=\int 1_D(x,y)\,dy=\frac{1}{\pi}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}dy=\frac{2\sqrt{1-y^2}}{\pi}1_{(-1,1)}(y)\,,p_{X_2}=p_{X_1}$

5. **Aproximarea cu constante**. Lucrurile se complică deoarece funcţia $f(X,t)=d(t,X)$ căreia trebuie să îi găsim punctul de minim argmin$f$ se poate defini în multe feluri. Dacă însă X este din L^2 (adică toate componentele sale sunt din L^2) şi luăm distanţa euclidiană + norma din L^2 atunci $f(X,t)=\sqrt{E\sum_{k=1}^{n}(\ X_j-t_j)^2}$ şi lucurile se petrec la fel ca în cazul unidimensional: Argmin$f$ =Argminf^2 =Argmin$\left(\sum_{k=1}^{n}(\ X_j-t_j)^2\right)=(EX_1,...,EX_n)$

Acest vector se notează cu EX.Atunci $f(X,\mathrm{E}X)=\sqrt{\sum_{k=1}^{n}\mathrm{Var}(X_j)}:=\sigma^2$

Analogul inegalităţii lui Cebîşev: P$(\|X-EX\|\ge k\sigma)\le\frac{1}{k^2}$

(Acelaşi argument: σ^2 =E$\|X-EX\|^2\ge$ E$\left(\|X-EX\|^2 1_{(\|X-EX\|\ge k\sigma)}\right)\ge k^2\sigma_2P(\|X-EX\|\ge k\sigma))$

Exemplu: O urnă are a bile de albe, n bile negre, r bile roşii. Se extrag k bile. X sunt albe, Y sunt negre. Faceţi o predicţie asupra vectorului $Z=(X,Y)$

R1. Bila nu se pune înapoi. E$Z=\left(kp_a,k\frac{n}{a+n+r}\right)$, Var$(X)=kp_aq_a\frac{a+n+r-k}{a+n+r-1}$,Var$(Y)=kp_nq_n\frac{a+n+r-k}{a+n+r-1}$

Am notat cu $p_a=\frac{a}{a+n+r},q_a=1-p_a=\frac{n+r}{a+n+r},p_n=\frac{n}{a+n+r},q_n=1-p_n=\frac{a+r}{a+n+r}$

Atunci P$\left(\|Z-(kp_a,kp_n)\|\ge 5\sqrt{kp_aq_a\frac{a+n+r-k}{a+n+r-1}+kp_nq_n\frac{a+n+r-k}{a+n+r-1}}\right)\le\frac{1}{25}=$ 4%

Concret: $a=n=r=1200=k\implies p_a=p_n=\frac{1}{3}$,E$Z=(400,400)\,,\sigma=\frac{2400\sqrt{2}}{3\sqrt{3599}}=18.859\cdot$

Deci P$(\|Z-(400,400)\|\ge 94.295)\le 4\%$. Putem paria, cu riscul de a pierde pariul mai mic decît 4% că Z se află în interiorul cercului de centru $(400,400)$ şi rază 95.

R2. Bila se pune înapoi. Acum E$Z=\left(kp_a,k\frac{n}{a+n+r}\right)$ dar Var$(X)=kp_aq_a$,Var$(Y)=kp_nq_n$

În aceleaşi ipoteze de dinainte $\sigma = \sqrt{\frac{4}{9}1200} = 23.094$ deci $\mathrm{P}(\|Z-(400,400)\| \geq 115.47) \leq 4\%$

6. **Portofolii**. `Covarianţă`. Vectorul $X = (X_j)_{1\leq j\leq n}$ poate fi interpretat ca fiind o piaţă financiară, în care X_j să reprezinte profitul obţinut în urma unui leu investit. Dacă în această piaţă se investesc s_j lei în poziţia j şi $S = \sum_{j=1}^{n} s_j$, atunci vectorul $s = (s_1, ..., s_n)$ se numeşte *portofoliu de volum S*. Variabila aleatoare $\xi = \sum_{j=1}^{n} s_j X_j$ se numeşte *valoarea* portofoliului, media sa $\mathrm{E}\xi$ este *randamentul ortofoliului* iar varianţa sa $\mathrm{Var}\xi$ este numit de unii *riscul* portofoliului.

Randamentul este uşor de calculat: $\mathrm{E}\xi = \sum_{j=1}^{n} s_j \mathrm{E}X_j$. Varianţa lui introduce în calcul aşa numita *matrice de covarianţă*.

Definiţie. Matricea $(c_{i,j})_{1\leq i,j\leq n} = C = \mathrm{Cov}(X) = (EX_iX_j - EX_iEX_j)_{1\leq i,j\leq n}$ se numeşte *matricea de covarianţă* a vectorului X.

Fapte evidente. $\mathrm{Var}\xi = \sum_{i,j=1}^{n} s_i s_j c = s'Cs$. Matricea C este semipozitiv definită iar funcţia $V(s) = s'Cs$ este convexă, deoarece matricea sa hessiană $\left(\frac{\partial^2}{\partial s_i \partial s_j}\right)_{i,j}$ coincide cu C.

Demonstraţie: Evident

7. **Formula de transport**. Formal, este aceeaşi din cazul unidimensional. Dacă $f : \mathbb{R}^n \to \mathbb{R}$ este o funcţie măsurabilă şi $f(X) \in \mathrm{L}^1$, atunci $\mathrm{E}f(X) = \int f dF_X$

În cele două cazuri didactice - discret sau absolut continuu - e cam la fel. Doar dacă X este absolut continuu apare variaţi $\mathrm{E}f(X) = \int fpd\lambda^n$.

Exemple. $\mathrm{E}X_j = \int_{\mathbb{R}^n} x_j p(x_1, ..., x_n)\, dx_1...dx_n$,$\mathrm{E}X_iX_j = \int_{\mathbb{R}^n} x_i x_j p(x_1, ..., x_n)\, dx_1...dx_n$ etc

8. Funcţia generatoare de momente: $\mathrm{m}_X(t) = Ee^{\sum_{j=1}^{n} t_j X_j}$. Are sens să se lucreze cu ea dacă $Ee^{\delta \sum_{j=1}^{n} |X_j| \neq \infty}$ *pentru* un $\delta > 0$.

Atunci $\mathrm{E}X = \nabla_{m_X}(0)$,$\mathrm{E}X_iX_j = \frac{\partial^2 \mathrm{m}_X}{\partial t_i \partial t_j}(0) = H_{m_X}(0)$ iar $\mathrm{Cov}(X) = H_{m_X}(0) - \nabla_{m_X}(0)\nabla_{m_X}(0)'$. H se numeşte matricea hessiană - din anul I de facultate.

Funcţia caracteristică: $\phi_X(t) = Ee^{i\sum_{j=1}^{n} t_j X_j}$

Funcţia generatoare (are sens doar pentru vectori stocastici aritmetici, adică cu valori în $\mathbb{N}^n$.

$g_X(x) = \sum_j x_1^{j_1} x_2^{j_2} ... x_n^{j_n} P(X = (j_1, j_2, ..., j_n))$

Atunci $\mathrm{E}X = \nabla g(1)$, iar dacă $i \neq j$,atunci $\mathrm{E}X_iX_j = \frac{\partial^2 g}{\partial t_i \partial t_j}(1)$.Dacă $i = j$ avem $\frac{\partial^2 g}{\partial t_i^2}(1) = \mathrm{E}X(X-1) = \mathrm{E}X^2 - \mathrm{E}X$ sau, în scriere matriceală $\mathrm{H}_g(1) = (\mathrm{E}X_iX_j - \delta_{i,j}\mathrm{E}X_i)_{i,j}$.

Obţinem formula $\mathrm{Cov}(X) = H_g(1) - \nabla_g(1)\nabla_g(1)' + (\delta_{i,j}EX_i)_{i,j}$

Exerciţii. Să se calculeze $\mathrm{E}X$,$\mathrm{Cov}(X)$ în cît mai multe cazuri.

9. Portofoliu optim - varianta Markowitz. Utilitatea portofoliului este U(s) =E$s'X - \frac{r}{2}$Var$(s'x) = s'\mu - \frac{r}{2}s'Cx$ unde $\mu =$ EX, C=Cov(X) , $r \geq 0$ este *coeficientul de aversiune la risc.*

Exemplu de calcul. La exemplul 2.1, varianta cu bila pusă la loc, se zice că X are repartiţia multinomială Mult(k,$p_1, p_2, ..., p_n$).Atunci $g_X(x) = \sum_{j_1,...,j_n \geq 0, \sum_{\alpha=1}^{n} j_\alpha = k} \frac{k!}{j_1!...j_n!} p_1^{j_1} p_2^{j_2} ... p_n^{j_n} =$

$(p_1x_1 + p_2x_2 + ... + p_nx_n)^k$ (de aceea i se spune repartiţia multinomială!) deci

$$\nabla g(x) = k(p_1x_1 + p_2x_2 + ... + p_nx_n)^{k-1} \begin{pmatrix} p_1 \\ p_2 \\ ... \\ p_k \end{pmatrix} \Longrightarrow \mathrm{E}X = k \begin{pmatrix} p_1 \\ p_2 \\ ... \\ p_k \end{pmatrix}$$

$$Hg(x) = k(k-1)(p_1x_1 + p_2x_2 + ... + p_nx_n)^{k-2} \begin{pmatrix} p_1^2 & p_1p_2 & ... & p_1p_n \\ p_2p_1 & p_2^2 & ... & p_2p_n \\ ... & ... & ... & ... \\ p_np_1 & p_np_2 & ... & p_n^2 \end{pmatrix} \Longrightarrow$$

$$Hg(1) = k(k-1) \begin{pmatrix} p_1^2 & p_1p_2 & ... & p_1p_n \\ p_2p_1 & p_2^2 & ... & p_2p_n \\ ... & ... & ... & ... \\ p_np_1 & p_np_2 & ... & p_n^2 \end{pmatrix}$$

Deci $\mathrm{Cov}(X) = Hg(1) + (np_i\delta_{i,j})_{i,j} - \mathrm{E}X(EX)' = k \begin{pmatrix} p_1q_1 & -p_1p_2 & ... & -p_1p_n \\ -p_2p_1 & kp_2q_2 & ... & -p_2p_n \\ ... & ... & ... & ... \\ -p_np_1 & -p_np_2 & ... & kp_nq_n \end{pmatrix}$

Deci în cazul unui portofoliu $\xi = \sum s_jX_j$

Var$\xi = k\left(\sum_{j=1}^{n} s_j^2 p_j q_j - \sum_{i \neq j=1}^{n} s_i s_j p_i p_j\right)$

Exerciţii.

1. Se dă densitatea $p(x,y) = (ax+by)1_{(0,1)}(x)1_{(0,1)}(y)$. Găsiţi condiţia ca p să fie densitate, media şi covarianţa. Acelaşi lucru pentru cît mai multe exemple găsite de dv.

2. Acelaşi lucru pentru $p(x,y) = c1_C(x,y)$ unde C este o mulţime de arie finită din plan. De exemplu $C =$co$(0,1,i)$ sau $C =$co$(0,1,i,2+2i)$

1.6 Lecţia 6. Probabilitate condiţionată.

1. Definiţie. Fie $(\Omega, \boldsymbol{K}, P)$ un spaţiu probabilizat şi $A \in \boldsymbol{K}$ un eveniment de probabilitate nenulă, P$(A) \neq 0$.

Atunci definim P$(B|A) = \frac{P(A \cap B)}{P(A)}$ şi citim "probabilitatea lui B condiţionată de A"

2. Daca notăm P$_A(B)$ =P$(B|A)$ obţinem un nou spaţiu probabilizat $(\Omega, \mathcal{L}, P_A)$. Integrala faţă de P$_A$ se calculează după formula $\int X dP_A = \frac{\int X1_A dP}{P(A)}$ şi se notează E$(X|A)$)demonstraţia imediată cu metoda celor 4 paşi)

O variabilă aleatoare$Y$ cu proprietatea că P(Y $\in B$) $= P(X \in B|A)$ se notează (abuziv!) $Y = (X|A)$

Exemplu. **Timpul remanent de viaţă.** Dacă T este o variabilă aletoare pozitivă, o putem interpreta ca fiind timpul de viaţă al unui organism sau aparat. Atunci variabila aleatoare $T_x = (T - x|T > x)$ se numeste timpul remanent de viaţă la vîrsta x. Aşadar P$(T_x > t) = \frac{P(T>t+x)}{P(T>x)}$ şi se obişnuieşte a se nota cu ${}_tp_x$.

3. Formula lui Bayes. Fie $(H_n)_n$ o partiţie cel mult numărabilă a lui Ω (adică $m \neq n \implies H_m H_n = \varnothing$ şi $\bigcup\limits_n H_n = \Omega$) formată evenimente neneglijabile, adică P$(H_n) \neq 0$. Presupunem cunoscute probabilităţile P$(A|H_n)$. Atunci

P$(H_i|A) = \frac{P(A|H_i)P(H_i)}{\sum\limits_n P(A|H_n)P(H_n)}$

Evenimentele H_n se numesc *ipoteze*, probabilităţile P$(A|H_n)$ formează *modelul*, probabilităţile $(\mathrm{P}(H_n))_n$ se numesc probabilităţile *apriori* iar probabilităţile $(P(H_n|A))_n$ se numesc probabilităţile *aposteriori.*

Exemplu. Populaţia României se compune din 50% bisericoşi, 40% neutri şi 10% atei. Dintre bisericoşi 50% respectă postul mare, dintre neutri îl respectă 20% iar dintre atei 1%. Cineva este văzut că mănîncă un cîrnat chiar în vinerea mare. Ce este el? bisericos? neutru? ateu?

Soluţie. Cele trei ipoteze sunt H_1 =bisericos, H_2 =neutru, H_3 = ateu. Probabilităţile apriori sunt P$(H_1) = .5$,P$(H_2) = .4$,P$(H_3) = .1$

A = individul care mănîncă un cîrnat în vinerea mare

Modelul este P$(A|H_1) = .5$,P$(A|H_2) = .8$,P$(A|H_3) = .99$

Atunci

P$(H_1|A) = \frac{P(A|H_1)P(H_1)}{\sum\limits_n P(A|H_n)P(H_n)} = \frac{.5 \cdot .5}{.5 \cdot .5 + .8 \cdot .4 + .99 \cdot .1} = 0.373\,69$,P$(H_2|A) = \frac{.8 \cdot .4}{.5 \cdot .5 + .8 \cdot .4 + .99 \cdot .1} = 0.478\,33$,P$(H_3|A) = \frac{.99 \cdot .1}{.5 \cdot .5 + .8 \cdot .4 + .99 \cdot .1} = 0.147\,98$

Cel mai probabil este că păcătosul este neutru, apoi că este un bisericos împins de Satana iar cel mai puţin probabil e să fie ateu.

4. **Probabilitatea condiţionată de o partiţie cel mult numărabilă.** Fie $\boldsymbol{P} = (A_n)_n$ o partiţie cel mult numărabilă a lui Ω. Atunci $P(B|\boldsymbol{P})$ este prin definiţie variabila aleatoare $Y = \sum\limits_n P(B|A_n) 1_{A_n}$ cu convenţia de bun simţ că dacă P$(A_n) = 0$, atunci $P(B|A_n) 1_{A_n} = 0$.

Observaţie importantă: E$P(B|\boldsymbol{P}) = P(B)$: într-adevăr, E$Y = \sum\limits_n P(B|A_n)E1_{A_n} = \sum\limits_n P(B \cap A_n) = P(B)$

Observaţie importantă: dacă $\mathcal{F}$ este $\sigma-$algebra generată de $\boldsymbol{P}$, atunci Y este $\mathcal{F}-$măsurabilă.

Mai mult, dacă X este o variabilă aleatoare, atunci notăm, abuziv, $Y = (X|\boldsymbol{P}) = \sum\limits_n (Y|A_n) 1_{A_n}$ o variabilă aleatoare cu proprietatea că P$(Y \in B|\boldsymbol{P}) = P(X \in B|\boldsymbol{P}) = \sum\limits_n P(X \in B|A_n) 1_{A_n}$.

Observaţie importantă: E$(X|\boldsymbol{P}) = \sum\limits_nE(Y|A_n) P(A_n)$ =EY şi, în plus, Y este $\mathcal{F}-$măsurabilă.

5. Repartiţia **condiţionată**. Fie $Z = (X, Y)$ un vector aleator în care $Y\tilde{}\begin{pmatrix} y_1 & y_2 & ... \\ p_1 & p_2 & ... \end{pmatrix}$ este discretă. Atunci notăm

$(X|Y)$ variabila aleatoare $\sum\limits_n (X|Y = y_n)\, 1_{(Y=y_n)}$. Dacă $\mathcal{F}$ este $\sigma-$algebra generată de Y, atunci $(X|Y)$ este de asemenea $\mathcal{F}-$măsurabilă.

6. Media conditionata. **Abordarea riguroasa**. Fie $\mathcal{F}$o $\sigma-$algebră inclusă în $\boldsymbol{K}$ şi fie $X \in L^1(\Omega, \boldsymbol{K}, P)$.Atunci E$(X|\mathcal{F})$ este *prin definiţie* o variabila aletoare, Y, cu proprietatea că Y este $\mathcal{F}-$măsurabilă şi E$(X1_A)$ =E$(Y1_A)$ pentru orice $A \in \mathcal{F}$. Probabilitatea condiţionată a mulţimii $B \in \boldsymbol{K}$**,** P$(B|\mathcal{F})$ este *prin definiţie* variabila aletoare, $Y = E(1_B|\mathcal{F})$, adică o varibilă aleatoare $\mathcal{F}-$măsurabilă cu proprietatea că E$(Y1_A) = P(A \cap B)$ pentru orice $A \in \mathcal{F}$. Dacă $\mathcal{F}$ este $\sigma-$algebra generată de o partiţie cel mult numărabilă, ***P*,** atunci verificaţi (este imediat!) că E$(X|\mathcal{F}) = \sum\limits_{H \in \boldsymbol{P}}$ E$(X|H)\, 1_H$ iar P$(B|\mathcal{F}) = \sum\limits_{H \in \boldsymbol{P}}$ P$(B|H)\, 1_H$

Exemplu. $\Omega = \mathbb{Z}_n^2, \boldsymbol{P} = \{\{j\} \times \mathbb{Z}_n : j \in \mathbb{Z}_n\}, B = \{(i, j) \in \Omega : i \leq j\}$. Atunci P$(B|\sigma(\boldsymbol{P})) = \sum\limits_{j=0}^{n-1} \frac{n-j}{n} 1_{\{j\} \times \mathbb{Z}_n}$

Dacă $\mathcal{F} = \sigma(Y)$ unde Y este o variabilă aleatoare, atunci orice variabilă aleatoare Z care este $\mathcal{F}-$măsurabilă este de forma $X = h(Y)$. $\Big($Verificaţi pentru X =1_A, apoi pentru $X = \sum\limits_n a$

Aşadar E$(X|\mathcal{F})$ trebuie să fie o funcţie de forma $h(Y)$ cu proprietatea că E$(Xg(Y))$ =E$(h(Y)\, g(Y))$pentru orice g măsurabilă şi mărginită (verificaţi! în definiţia iniţială $g = 1_A, A \in \sigma(Y) \iff A =_{Y^{-1}(B)}$cu B o mulţime boreliană de pe dreapta reală, deci $g = 1_B; etc$). De aceea se obişnuieşte notaţia E$(X|Y)$ în loc de E$(X|\sigma(Y))$

Definiţie. Variabila aleatoare E$(X|Y)$ se numeşte regresia lui Y asupra lui X.

7. Proprietăţi ale mediei condiţionate

7.0 Liniaritate, monotonie

7.1. Dacă $\mathcal{F}_1 \subset \mathcal{F}_2$ atunci E(E$(X|\mathcal{F}_2)\,|\mathcal{F}_1)$ =E$(X|\mathcal{F}_1)$ - proprietatea de proiecţie. Caz particular: E(E$(X|\mathcal{F}_2)$) =EX

7.2 Dacă $Y \in L^\infty(\mathcal{F})$ atunci E$(XY|\mathcal{F}) = Y$E$(X|\mathcal{F})$ $-$ variabilele $\mathcal{F}-$ măsurabile se comportă precum constantele

7.3. Jensen. E$(f(X)\,|\mathcal{F}) \geq f(E(X|\mathcal{F}))$ dacă f este convexă şi $f(X) \in L^1$

7.4. Beppo - Levi. Dacă $X_n \nearrow X$ şi $X_1 \in L^1$ atunci E$(X_n|\mathcal{F}) \nearrow$E$(X|\mathcal{F})$ cu excepţia posibilă a unei mulţimi neglijabile

7.5 Proprietatea de optim:dacă $X \in L^2$, atunci E$(X - E(X|\mathcal{F}))^2 \leq$ E$(X - Y)^2$ pentru orice $Y \in L^2(\mathcal{F})$

Caz particular: E$(X - E(X|Y))^2 \leq$ E$(X - g(Y))^2$ pentru orice g funcţie măsurabilă cu proprietatea că $g(Y) \in L^2$.

De aceea, $E(X|Y)$ se mai numeşte *estimatorul Bayesian exact* dat de Y asupra lui X.

7. Formule pentru repartiţia condiţionată.

Uneori - în cazurile didactice - repartiţia condiţionată - deci şi media/varianţa condiţionată etc se pot calcula.

7.1. Cazul discret. Vectorul este $Z = (X, Y)$ .Fie $p_{X,Y}(i, j) = P(X = i, Y = j)$.

Atunci $p_{X|Y=j}(i) = \frac{P(X=i,Y=j)}{P(Y=j)} = \frac{p_{X,Y}(i,j)}{\Sigma_\alpha p_{X,Y}(\alpha,j)}$, $p_{Y|X=i}(j) = \frac{P(X=i,Y=j)}{P(X=i)} = \frac{p_{X,Y}(i,j)}{\Sigma_\beta p_{X,Y}(i,\beta)}$

7.2. Cazul absolut continuu. Densitatea lui Z este $p_{X,Y}$.Atunci:

densitatea lui X este $p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x,t)\,dt$

densitatea lui Y este $p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(s,y)\,ds$

densitatea lui $(X|Y = y)$ este $p_{X|Y=y}(x) = \frac{p_{X,Y}(x,y)}{\int_{-\infty}^{\infty} p_{X,Y}(x,t)dt}$

densitatea lui $(Y|X = x)$ este $p_{Y|X=x}(y) = \frac{p_{X,Y}(x,y)}{\int_{-\infty}^{\infty} p_{X,Y}(s,y)ds}$

Justificare. Stricto senso, nu se poate scrie $\mathrm{P}(Y \in B|X = x)$ deoarece $\mathrm{P}(X = x) = 0$.Dar are sens să gîndim $\mathrm{P}(Y \leq y|X = x)$ ca fiind $\lim_{\varepsilon\to 0}\frac{P(Y\leq y|x-\varepsilon\leq X\leq x+\varepsilon)}{P(x-\varepsilon\leq X\leq x+\varepsilon)}$ $=\lim_{\varepsilon\to 0}\frac{\int_{x-\varepsilon}^{x+\varepsilon}\int_{-\infty}^{y} p_{X,Y}(u,v)dudv}{\int_{x-\varepsilon}^{x+\varepsilon}\int_{-\infty}^{\infty} p_{X,Y}(u,v)dudv}$ =lin

Ştim că dacă funcţia f este continuă în x, atunci $\lim_{\varepsilon\to 0}\frac{\int_{x-\varepsilon}^{x+\varepsilon} f(u)du}{2\varepsilon} = f(x)$ .Presupunem că acesta este cazul şi, atunic avem mai departe $\mathrm{P}(Y \leq y|X = x) = \frac{\int_{-\infty}^{y} f(x,v)dv}{\int_{-\infty}^{\infty} f(x,v)dv}$.

Dacă o derivăm în y,găsim densitatea $p_{Y|X=x}(y) = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,v)dv}$

În aceste două situaţii este valabilă $\Re$

$\mathrm{E}(g(X)|Y = y) = \sum g(x)\, p_{X|Y=y}(x) = \frac{\sum g(x) p_{X|Y}(x,y)}{\Sigma_\alpha p_{X,Y}(\alpha,y)}$ (cazul discret)

$\mathrm{E}(g(X)|Y = y) = \int g(x)\, p_{X|Y=y}(x)\,dx = \int\limits_{-\infty}^{\infty} \frac{g(x)f(x,y)}{\int_{-\infty}^{\infty} f(x,v)dv}dx$ (cazul absolut continuu)

1. (a) i. A. Exerciţii

1. Se da $p(x, y) = (ax + by)\, 1_{(0,1)}(x, y)$. Condiţia ca p să fie densitate lui $Z = (X, Y)$, repartiţiile marginale, media, covarianţa, portofoliul optim = Argmin$U$ pentru $U(s, t) = E\xi - rVar(\xi)$ cu $\xi = sX + tY$

2. Aceeaşi oproblemă pentru $p(x, y, z) = (ax + by + cz)\, 1_{(0,1)}(x, y, z)$

3. Sau pentru $\mathrm{P}\circ Z^{-1}$ =Uniform$\left(\left\{(i, j) \in \mathbb{Z}_4^2 : i \leq j\right\}\right)$, sau Uniform$\left(\left\{(i, j) \in \mathbb{Z}_4^2 : i + j \leq 4\right\}\right)$

4. Sau pentru X˜Polinomial$(4, p, q, r)$

5. Sau cît mai multe exerciţii în două sau trei dimensiuni inventate de dv

$\mathcal{F}$

1.7 Lecţia 7. Independenţa

1.7.1 A. Două obiecte

Definiţii. Fie (Ω, K, P) un spaţiu probabilizat şi $A, B \in K$.

1. Spunem că A este independent de B (şi scriem $A \amalg B$) dacă şi numai dacă $P(A \cap B) = P(A)\, P(B)$

Observaţie. Dacă $\mathrm{P}(A) \neq 0$, atunci faptul că $A \amalg B$ este echivalent cu $\mathrm{P}(B|A) = P(B)$. Intuitiv, apariţia lui A nu afectează probabilitatea de apariţie a evenimentului B. Sau, obţinerea unei informaţii asupra lui A nu are nici o valoare: nu schimbă datele problemei.

Observaţie. Evenimentul sigur $A = \Omega$ sau cel imposibil $A = \varnothing$ sunt independente de orice alt eveniment.

2. Fie $\boldsymbol{M}_1$ şi $\boldsymbol{M}_2$ două familii de evenimente incluse în K. Spunem că $\boldsymbol{M}_1$ şi $\boldsymbol{M}_2$sunt independente (şi notăm $\boldsymbol{M}_1 \amalg \boldsymbol{M}_2$) dacă $A \amalg B$ pentru orice $A \in \boldsymbol{M}_1$ şi $B \in \boldsymbol{M}_2$.

3. **Propoziţia 1**. Dacă $\boldsymbol{M}_1 \amalg \boldsymbol{M}_2$, atunci U-sistemele generate vor fi din nou independente: $U(\boldsymbol{M}_1) \amalg U(\boldsymbol{M}_2)$. In particular, dacă $\boldsymbol{M}_1$ şi $\boldsymbol{M}_2$ sunt stabile la intersecţii finite, atunci $\sigma-$algebrele generate vor fi de asemenea independente: $\sigma(\boldsymbol{M}_1) \amalg \sigma(\boldsymbol{M}_2)$.

4. Fie $X : \Omega \to E, Y : \Omega \to F$ două variabile aleatoare, unde $(E, \boldsymbol{E})$ şi $(F, \boldsymbol{F})$ sunt două spaţii măsurabile oarecare. Atunci spunem că X este independentă de Y (şi scriem $X \amalg Y$) dacă şi numai dacă $\sigma-$algebrele generate sunt independente, adică dacă $X^{-1}(\boldsymbol{E}) \amalg X^{-1}(\boldsymbol{F})$.

Formulări echivalente

a. $X \amalg Y$

b. $\sigma(X) \amalg \sigma(Y)$

c. $\mathrm{P}(X \in A, Y \in B) = P(X \in A) P(Y \in B)$ pentru orice $A \in \boldsymbol{E}, B \in \boldsymbol{F}$.

d. $\mathrm{P} \circ (X, Y)^{-1} = \mathrm{P} \circ (X)^{-1} \otimes \mathrm{P} \circ (Y)^{-1}$

5. **Cum verificăm dacă** $X \amalg Y$?

a. Algoritm general : verificăm egalitatea $\mathrm{P}(X \in A, Y \in B) = P(X \in A) P(Y \in B)$ pentru orice $A \in \boldsymbol{E}, B \in \boldsymbol{F}$.

b. În cazul variabilelor aleatoare reale:de verificat că $\mathrm{P}(X \leq x, Y \leq y) = P(X \leq x) P(Y \leq y)$ pentru orice $x, y \in \mathbb{R}$.

Adică dacă $F_{X,Y} = F_X \otimes F_Y$.

c. Dacă variabilele sunt discrete: de verificat că $\mathrm{P}(X = x, Y = y) = P(X = x) P(Y = y)$ pentru orice $x \in E, y \in F$

d. În cazul real: dacă X este discretă, de verificat că $\mathrm{P}(X = x, Y \leq y) = P(X = x) P(Y \leq y)$ pentru orice $x, y \in \mathbb{R}$

e. În cazul real: dacă (X, Y) are densitatea p,de verificat că se poate scrie $p(x, y) = p_1(x) p_2(y)$

Adică dacă $p_{X,Y} = p_X \otimes p_Y$.Dacă notăm $\mathrm{supp}(p) = Cl((x, y) : p(x, y) > 0)$, o condiţie necesară ca acest lucru să se întîmple este ca $\mathrm{supp}(p)$ să fie de forma $A \times B$.

6. **Proprietăţi**.

a. Dacă X, Y sunt independente, atunci $f(X)$ şi $g(Y)$ sunt de asemenea independente pentru orice f, g măsurabile.

Precis: dacă $X : \Omega \to E, Y : : \Omega \to F$ sunt independente şi $f : E \to E', g : F \to F'$ sunt măsurabile, atunci $f(X) \amalg g(Y)$

Consecinţe: Dacă $E = \mathbb{R}^n, E = \mathbb{R}^m$ atunci $1_A(X) \amalg 1_B(Y)$ pentru orice $A \in \boldsymbol{E}, B \in \boldsymbol{F}$,$a'X \amalg b'Y$ pentru orice $a \in \mathbb{R}^n, b \in \mathbb{R}^m$ etc

b. Dacă $X \in L^1$ şi $Y \in L^1$ sunt independente, atunci $XY \in L^1$; (deoarece $\mathrm{E}|XY|$ $=\mathrm{E}|X|\mathrm{E}|Y|$,vezi formula de transport) şi sunt şi necorelate. Independenţa implică necorelare. Reciproc, dacă sunt necorelate, nu rezultă că sunt independente - de exemplu dacă (X, Y) este repartizat uniform pe discul unitate . Totuşi, dacă X şi Y au doar două valori, necorelarea implică independenţă.

c. Dacă X, Y sunt independente, $g : \mathbb{R}^2 \to \mathbb{R}$ este măsurabilă, $g(X, Y) \in L^1$, atunci $\mathrm{E}g(X, Y|X = x) = \int g(x, y)\, dF_Y(y)$

Ca atare avem de exemplu, $\mathrm{P}(X \geq Y) = \mathrm{E}(\mathrm{P}(X \geq Y|X)) = \mathrm{E}F_Y(X) = \int F_Y(x)\, dF_X(x)$ etc.

Dacă, de exemplu, X este repartizată $\mathrm{Exp}(\lambda)$,$\mathrm{P}(X \geq Y) = \mathrm{EP}(X \geq Y|Y) = \mathrm{Ee}^{-\lambda Y_+} = \mathrm{m}_{Y_+}(\lambda)$, *unde* m_{Y_+} este mgf-ul lui Y_+.

d. Dacă X este independentă de Y, atunci $\mathrm{E}(X|Y) = \mathrm{E}X$. Sau, reformulat: Dacă X este independentă de $\sigma-$algebra $\boldsymbol{F}$,atunci $\mathrm{E}(X|\boldsymbol{F}) = \mathrm{E}X$ Într-adevăr, măsurabilitatea este evidentă, la fel ca şi egalitatea $\mathrm{E}(X 1_B) = \mathrm{E}((\mathrm{E}X)\, 1_B) = P(B)\mathrm{E}X$. Reciproca este falsă: dacă $\mathrm{E}(X|Y) = \mathrm{E}X$ nu rezultă că X II Y. De exemplu dacă (X, Y) ˜ $\begin{pmatrix} (0,0) & (1,0) & (2,0) & (1,1) & (2,0) & (2,1) & (2,2) \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{pmatrix}$, atunci $(X|Y = 0)$ ˜ $(X|Y = 0)$ ˜Uniform$(\mathbb{Z}_3)$, $(X|Y = 1) = 1$ deci $\mathrm{E}(X|Y) = \mathrm{E}X = 1$ deşi evident că X˜ $\begin{pmatrix} 0 & 1 & 2 \\ \frac{2}{9} & \frac{5}{9} & \frac{2}{9} \end{pmatrix}$ nu este independentă de Y˜Uniform$(\mathbb{Z}_3)$.

1.7.2 B. O mulţime oarecare obiecte

Definiţii

1. **Evenimente.** Evenimentele $(A_\alpha)_{\alpha \in I}$, unde I este o mulţime oarecare de indici sunt independente dacă $\mathrm{P}\left(\bigcap_{\beta \in J} A_\beta\right) = \prod_{\beta \in J} \mathrm{P}(A_\beta)$ pentru orice mulţime de indici $J \subset I$ care este finită.

2. **Familii de mulţimi**. Fie $(\boldsymbol{M}_\alpha)_{\alpha \in I}$ o familii de mulţimi din $\boldsymbol{K}$. Ele sunt independente dacă $\mathrm{P}\left(\bigcap_{\beta \in J} A_\beta\right) = \prod_{\beta \in J} \mathrm{P}(A_\beta)$ pentru orice orice mulţime finită de indici $J \subset I$ şi $A_\beta \in \boldsymbol{M}_\alpha$.

Propoziţia 2. Dacă mulţimile $(\boldsymbol{M}_\alpha)_{\alpha \in I}$ sunt independente, atunci şi U-sistemele generate vor fi independente. Dacă familiile $\boldsymbol{M}_\alpha$ sunt şi stabile la intersecţii finite, atunci $(\sigma(\boldsymbol{M}_\alpha))_{\alpha \in I}$ vor fi independente.

3. **Variabile aleatoare**. Variabilele aleatoare $(X_\alpha)_{\alpha \in I}$ sunt independente dacă $\sigma-$algebrele generate $(\sigma(X_\alpha))_{\alpha \in I}$ sunt independente.

4. **Proprietatea de asociativitate**. Dacă $(\boldsymbol{F}_\alpha)_{\alpha \in I}$ sunt $\sigma-$ algebre independente şi $\Pi = (\Delta_\beta)_{\beta \in J}$ este o partiţie a mulţimii de indici I , atunci $\sigma-$ algebrele $\left(\sigma\left(\bigcup_{\alpha \in \Delta_\beta} \boldsymbol{F}_\alpha\right)\right)_{\beta \in J}$ vor fi de asemenea independente. Sau, în termeni de variabile aleatoare: Dacă $(X_\alpha)_{\alpha \in I}$ sunt variabile aleatoare independente, $\Pi = (\Delta_\beta)_{\beta \in J}$ este o partiţie a mulţimii de indici I , $g_\beta : \mathbb{R}^{\Delta_\beta} \to \mathbb{R}$ sunt măsurabile, atunci şi variabilele aleatoare $\left(g_\beta\left((X_\alpha)_{\alpha \in \Delta_\beta}\right)\right)_\beta$ vor fi de asemenea

independente

1.7.3 C. Şiruri de obiecte

Cum verificăm dacă un şir de variabile aleatoare sunt independente?

Suficient să verificăm dacă $\mathrm{P}(X_1 \leq x_1, ..., X_n \leq x_n)$ =$\mathrm{P}(X_1 \leq x_1)\mathrm{P}(X_2 \leq x_2)$ ...$\mathrm{P}(X_n \leq x_n)$ pentru orice n. Dacă unele din ele sunt discrete, e suficient să punem "=" în loc de "$\leq$"

Probabilitatea produs. Teorema lui Kolmogorov.

Fie $(\Omega_n, K_n, P_n)_n$ un şir de spaţii probabilizate, $\Omega = \prod\limits_{n\geq 1} \Omega_n, K = \bigotimes\limits_{n\geq 1} \boldsymbol{K}_n$. Pe acest spaţiu măsurabil există o unică probabilitate cu u proprietatea că $\mathrm{P}(A_1 \times A_2 \times .....) = \prod\limits_{n\geq 1} \mathrm{P}_n (A_n)$. Ea se numeşte peobabilitatea produs şi se notează cu $\mathrm{P} = \bigotimes\limits_{n\geq 1} P_n$.

Fapt 1: variabilele aleatoare $(X_n)_{n\geq 1}$ sunt independente dacă $(X_1, ..., X_n)$ sunt independente pentru orice $n \geq 2$

Dacă ele au aceeaşi repartiţie, se numesc independente şi identic repartizate şi scriem "fie $(X_n)_{n\geq 1}$" iid.

Fapt 2: Fie $(E_n, \boldsymbol{E}_n)_n$ spaţii măsurabile şi $F_n : E_n \to [0,1]$ un şir de repartiţii. Atunci există un spaţiu probabilizat (Ω, K, P) şi un şir de variabile aleatoare $X_n : \Omega \to E_n$ care sunt independente şi au exact repartiţiile F_n.

Se ia $\Omega = \prod\limits_{n\geq 1} E_n, K = \bigotimes\limits_{n\geq 1} \boldsymbol{E}_n, P = \bigotimes\limits_{n\geq 1} P_n, \boldsymbol{X}_n ((x_k)_k) = x_n$.

Fapt 3. Dacă $(E_n, \boldsymbol{E}_n) = (\mathbb{R}, B(\mathbb{R}))$ atunci se poate lua mai concret $\Omega = [0,1]^\infty, K = B([0,1])^{\not{\otimes}\infty}, P = \lambda^\infty$,(unde λ este măsura Lebesgue pe $[0,1]$) şi $\boldsymbol{X}_n ((x_k)_k) = F_n^{-1}(x_n)$, unde F_n^{-1} sunt cuantilele lui funcţiei de repartiţie X_n.

Aşadar enunţul "Fie $(X_n)_n$ un şir de variabile aletoare iid cu repartiţia F are întotdeaun sens"

Exemplu. Fie $\Omega = (0,1), \boldsymbol{K} = B((0,1)), p \geq 2$ un număr natural. Orice număr $\varpi \in \Omega$ se poate scrie în baza p sub forma

$\varpi = \sum\limits_{n=1}^{\infty} X_n(\varpi) p^{-n}$. Atunci variabilele aleatoare $(X_n)_n$ sunt iid repartizate Uniform($\mathbb{Z}_p$) . În plus, variabilele aleatoare $\left\{X_1, X_n, ..., X_n, X_{n+1} = \left(\sum\limits_{k=1}^{n} X_k\right) (\mathrm{mod}\, p)\right\}$ sunt n+1 variabile aleatoare au proprietatea că NU sunt independente, deşi oricare n din ele sunt independente.

Fapt 4. Dacă $(X_n)_n$ sunt variabile aleatoare independente din L^2, *necorelate*, atunci $\mathrm{Var}\left(\sum\limits_{k=1}^{n} X_k\right) = \sum\limits_{k=1}^{n} \mathrm{Var}(X_n)$. Deci cu atît mai mult egalitatea este valabilă dacă ele sunt independente. Ca atare $\mathrm{Var}\left(\frac{1}{n}\sum\limits_{k=1}^{n} X_k\right) = \frac{1}{n}\sum\limits_{k=1}^{n} \mathrm{Var}(X_n)$. În cazul particular cînd $(X_n)_n$ sunt i.i.d., atunci $\mathrm{Var}\left(\frac{1}{n}\sum\limits_{k=1}^{n} X_k\right) = \frac{Var(X_1)}{n} \underset{n\to\infty}{\to} 0$

Exerciţii.

1. Dacă $X \amalg Y$ atunci $\mathrm{P}(X > x, Y > y) = \mathrm{P}(X > x)\mathrm{P}(Y > y)$. Interpretare daca X, Y sunt timpi de viata.

2. Dacă $\mathrm{P}(X > x, Y > y) \geq \mathrm{P}(X > x)\mathrm{P}(Y > y)$ pentru orice x, y variabilele X şi Y se numesc pozitiv dependente, iar dacă este valabilă inegalitatea inversă, se numesc negativ dependente. Verificaţi că dacă $X \amalg Y$ atunci $\min(X, Y)$ şi $\max(X, Y)$ sunt pozitiv dependente.

3. Dacă (X, Y) are densitatea $p(x, y) = (ax + by) 1_{(0,1)^2(x,y)}$ cu $a, b \geq 0, a + b = 2$ atunci X şi Y sunt negativ dependente

4. Verificaţi dacă $[X]$ şi $\{X\}$ sunt independente dacă X este repartizată exponenţial sau Uniform(a, b)

5. Verificaţi că dacă $X = (X_j)_{1 \leq j \leq n}$ este un vector cu toate componentele independente, atunci $\mathrm{m}_X(t) = \prod_{j=1}^{n} \mathrm{m}_{X_j}(t_j)$ şi $\phi_X(t) = \prod_{j=1}^{n} \phi_{X_j}(t_j)$; dacă are valori din $\mathbb{N}^n$, atunci $\mathrm{g}_X(x) = \prod_{j=1}^{n} \mathrm{g}_{X_j}(x_j)$

6. Verificaţi că dacă $X = (X_j)_{1 \leq j \leq n}$ este un vector cu toate componentele independente şi $S = \sum_{j=1}^{n} X_j$, atunci $\mathrm{m}_S(t) = \prod_{j=1}^{n} \mathrm{m}_{X_j}(t)$ şi $\phi_X(t) = \prod_{j=1}^{n} \phi_{X_j}(t)$; dacă are valori din $\mathbb{N}^n$, atunci $\mathrm{g}_X(x) = \prod_{j=1}^{n} \mathrm{g}_{X_j}(x)$. Ce rezultă în cazurile particulare cînd $X_j \tilde{} N(\mu_j, \sigma_j^2)$, $X_j \tilde{}$Binomial(n_j, p), $X_j \tilde{}$Poisson(λ_j), $X_j \tilde{}$Exponential(λ)?

7. Fie X, Y variabile aleatoare mărginite. Presupunem că $\mathrm{E}X^mY^n = \mathrm{E}X^m\mathrm{E}Y^n$ pentru orice $m, n \geq 1$. Atunci $X \amalg Y$

Hint. $\phi_{X,Y}(s, t) = \mathrm{E}e^{i(sX+tY)} = \sum_n \frac{i^n}{n!}\mathrm{E}(sX + tY)^n = \sum_n \frac{i^n}{n!} \sum_{k=0}^{n} C_n^k s^{n-k} t^k \mathrm{E}X^{n-k}Y^k = \sum_n \frac{i^n}{n!} \sum_{k=0}^{n} C_n^k s^{n-k} t^k \mathrm{E}X^{n-k}\mathrm{E}Y^k$

$\phi_X(s)\phi_Y(t) = \mathrm{E}e^{isX}\mathrm{E}e^{isX} = \sum_m \frac{i^m}{m!} s^m t^k \mathrm{E}X^m \sum_k \frac{i^k}{k!}\mathrm{E}Y^k = \sum_{m,k} \frac{i^{m+k}}{(m+k)!} C_{m+k}^k s^m t^k \mathrm{E}X^m\mathrm{E}Y^k$

Cele două repartiţii au aceeaşi funcţie caracteristică, deci coincid (Teorema de unicitate)

De ce am pus condiţia să fie mărginite? Ca să putem comuta media cu suma.

8. Vectorul $\mathrm{Z} = (X, Y)$ este repartizat $\begin{pmatrix} (-1,0) & (0,1) & (1,0) & (0,-1) \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$.

Gasiţi repartiţiile marginale, media,covarianţa, funcţia de repartiţie, funcţia de supravieţuire $\underline{F}(x, y) = P(X > x, Y > y)$, $(Y|X)$, $(X|Y)$, $\mathrm{E}(X^mY^n) - \mathrm{E}(X^m)\mathrm{E}(Y^n)$ *etc*

1.8 Lecţia 8. Convergenţa variabilelor aleatoare. Legea Numerelor Mari

1. În orice spaţiu cu măsură $(E, \boldsymbol{E}, \mu)$ putem construi spaţiile Banach $\mathrm{L}^p(E, \boldsymbol{E}, \mu) = \left\{X : E \to \mathbb{R} : \int |X|^p \, d\mu \neq \infty\right\}, 1 \leq p \leq \infty$.

Spaţiile probabilizate nu fac excepţie: Fie $(X_n)_n$, X variabile aleatoare din $\mathrm{L}^p(\Omega, \boldsymbol{K}, \mathrm{P})$. Spunem că X_n converge la X în L^p (şi scriem $X_n \xrightarrow{L^p} X$) dacă $\|X_n - X\|_p \to 0$

Ce este special la spaţiile probabilizate este inegalitatea normelor. Ea arată că

Propoziţie. Dacă $1 \leq p \leq q$ şi $(X_n)_n, X$ sunt variabile aleatoare din L^q,şi $X_n \xrightarrow{L^q} X$ atunci $X_n \xrightarrow{L^p} X$

2. De asemenea, în orice spaţiu cu măsură $(E, \boldsymbol{E}, \mu)$ se introduce convergenţa în măsură:

Definiţie. Spunem că X_n converge la X în măsură dacă $\mu(\{|X_n - X| > \varepsilon\}) \longrightarrow 0$ pentru orice $\varepsilon > 0$ şi scriem $X_n \xrightarrow{\mu} X$. În general nu este nici o legătură între aceste tipuri de convergenţă: de exemplu, pe $(\mathbb{R}, \boldsymbol{B}(\mathbb{R}), \lambda)$ şirul de funcţii măsurabile $f_n(x) = e^{-nx}1_{(0,\infty)}(x)$ converge la 0 în orice $L^p(\mathbb{R}, \boldsymbol{B}(\mathbb{R}), \lambda)$ cu $1 \leq p < \infty$ şi totuşi pentru orice $\varepsilon > 0$ este evident că $\lambda(\{x : |f_n(x)| > \varepsilon\}) = \infty$.

În spaţiile probabilizate există o implicaţie

Propoziţie. **Fie** $(X_n)_n, X$ variabile aleatoare din $L^p(\Omega, \boldsymbol{K}, \mathrm{P}), 1 \leq p \leq \infty$. Atunci $X_n \xrightarrow{L^p} X \implies X_n \xrightarrow{P} X$

Într-adevăr, $\mathrm{P}(|X_n - X| > \varepsilon) \leq \frac{\|X_n - X\|_p^p}{\varepsilon^p} \longrightarrow 0$ dacă $n \to \infty$.

3. La fel, în orice spaţiu cu măsură $(E, \boldsymbol{E}, \mu)$ se introduce şi convergenţa aproape peste tot, notată a.p.t:

spunem că X_n converge aproape peste tot dacă $\mu(\{\liminf X_n \neq \limsup X_n\}) = 0$.Aceeaşi definiţie este valabilă şi în cazul spaţiilor probabilizate, numai că aici notăm a.s. (= aproape sigur = almost surely)în loc de a.p.t.. Aşadar

$X_n \xrightarrow{a.s.} X \overset{Def}{\iff} \mu(\{\liminf X_n = \limsup X_n = X\}) = 1$

În general, pe un spaţiu cu măsură oarecare, convergenţa a.p.t. NU garantează convergenţa în măsură (vezi exemplul de la 2). Totuşi, în cazul spaţiilor probabilizate avem

Propoziţie. Dacă $X_n \xrightarrow{a.s.} X$ atunci $X_n \xrightarrow{P} X$

Demonstraţie. Evident, înlocuind X_n cu $X_n - X$ putem presupune că $X = 0$. Fie $A = \left\{X_n \xrightarrow{a.s.} 0\right\}$

Aşadar ştim că $\mathrm{P}(A^c) = 0$ $\mathrm{P}(A^c) = P(\exists \varepsilon > 0, \forall n > 0, \exists n_1 > n \text{ ca } |X_n| > \varepsilon) = P\left(\exists m > 0, \forall n > 0, \exists k > 0 \text{ ca } |X_{n+k}| > \frac{1}{m}\right) = P\left(\bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \left\{|X_{n+k}| > \frac{1}{m}\right\}\right)$

Cum şirul de mulţimi $\left(\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \left\{|X_{n+k}| > \frac{1}{m}\right\}\right)_m$ este crescător, aplicînd proprietatea de continuitate a probabilităţii găsim că $\mathrm{P}(A^c) = \lim_m P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \left\{|X_{n+k}| > \frac{1}{m}\right\}\right) = 0$, aşadar $P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \left\{|X_{n+k}| > \frac{1}{m}\right\}\right) = 0$ pentru orice $m \geq 1$. Dar şirul de mulţimi $\left(\bigcup_{k=1}^{\infty} \left\{|X_{n+k}| > \frac{1}{m}\right\}\right)_n$ este descrescător, aşadar $P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \left\{|X_{n+k}| > \frac{1}{m}\right\}\right) = \lim_n P\left(\bigcup_{k=1}^{\infty} \left\{|X_{n+k}| > \frac{1}{m}\right\}\right) = 0$;deci cu atît mai mult $\lim_n P\left(\left\{|X_n| > \frac{1}{m}\right\}\right) = 0$, adică $X_n \xrightarrow{P} 0$.

Reciproc nu este adevărat, după cum ne putem convinge cu exemplul următor: orice număr natural n se poate scrie sub forma $n = 2^k + j$ unde $0 \leq j \leq 2^k - 1$. Luăm $\Omega = (0,1)$, P =măsura Lebesgue,$A_n = \left(\frac{j}{2^k}, \frac{j+1}{2^k}\right)$, $X_n = 1_{A_n}$. De exemplu primele mulţimi A_n sunt $(0,1), \left(0,\frac{1}{2}\right), \left(\frac{1}{2},1\right), \left(0,\frac{1}{4}\right), \left(\frac{1}{4},\frac{2}{4}\right), \left(\frac{2}{4},\frac{3}{4}\right), \left(\frac{3}{4},1\right), ...$

Atunci $\limsup X_n = 1\,(a.s.)$, $\liminf X_n = 0\,(a.s.)$ şi totuşi $X_n \xrightarrow{P} 0$

4. **Legea slabă a numerelor mari**. Dacă $(X_n)_n \in L^2$ sunt necorelate atunci $\frac{X_1+...+X_n}{n} \xrightarrow{P} \mu$, unde μ =EX_1.

Evident: după ce centrăm variabilele X_n înlocuindu-le cu $Y_n = X_n - \mu$: $\left\|\frac{X_1+...+X_n}{n} - \mu\right\|_2^2$ =Var$\left(\frac{Y_1+...+Y_n}{n}\right) = \frac{\sigma^2}{n} \to 0$ unde am notat cu σ^2 =Var(X_1).

De ce se numeşte legea slabă? Pentru că asigură DOAR convergenţa în probabilitate, care este mai slabă decît cea aproape sigură.

5. **Legea tare a numerelor mari**.Dacă $(X_n)_n \in L^1$ sunt i.i.d. atunci $\frac{X_1+...+X_n}{n} \xrightarrow{a.s.} \mu$, unde μ =EX_1. Mai mult, convergenţa are loc şi în L^1.

Sunt mai multe demonstraţii clasice. Cea mai puternică se bazează pe

TEOREMA ERGODICA BIRKHOFF. Fie $(E, \boldsymbol{E}, \Pi)$ un spaţiu probabilizat şi $t : E \to E$ o funcţie măsurabilă cu proprietăţile:

a. $\Pi\left(t^{-1}(A)\right) = \Pi(A)$ pentru orice $A \in \boldsymbol{E}$.

b. Dacă $t^{-1}(A) = A$, atunci $\Pi(A) \in \{0,1\}$

Fie, de asemenea, $f \in L^1(E, \boldsymbol{E}, \Pi)$.

Atunci $\frac{f+f(t)+f\left(t^2\right)+f\left(t^3\right)+...f\left(t^{n-1}\right)}{n}$ converge $\Pi-$aproape sigur şi în L^1 la $\int f d\Pi$.Am notat $t^2 = t \circ t, f^3 = t \circ t \circ t, ...$

Observaţie. O funcţie t cu proprietăţile a. şi b. se numeşte **ergodică**.

La noi vom lua E=$\mathbb{R}^\infty$, $\boldsymbol{E}$=$B(\mathbb{R})^\infty$, $\Pi = F^\infty$ unde $F = P \circ X_n^{-1}$ este repartiţia comună a variabilelor i.i.d $(X_n)_n$, $t : E \to E$ va fi shiftul $t(x_1, x_2, x_3, ..) = (x_2, x_3, x_4, ...)$ care evident are proprietatea a.

Acceptăm că are proprietatea b. Apoi considerăm vectorul aleator $X : \Omega \to E, X = (X_n)_n$ şi observăm că P$\circ X^{-1} = \Pi$. Aşadar pentru orice funcţie măsurabilă $f \in L^1(E, \boldsymbol{E}, \Pi)$ avem $\frac{f+f(t)+f\left(t^2\right)+f\left(t^3\right)+...f\left(t^{n-1}\right)}{n} \xrightarrow{a.s.} \int f d\Pi = \int f d\left(P \circ X^{-1}\right)$ =E$f(X)$. Deci $\frac{f+f(t)+f\left(t^2\right)+f\left(t^3\right)+...f\left(t^{n-1}\right)}{n}(X) \xrightarrow{a.s.}Ef(X)$.

Legea tare este un caz foarte particular în care se ia $f\left((x_n)_n\right) = x_1$:atunci $\frac{f+f(t)+f\left(t^2\right)+f\left(t^3\right)+...f\left(t^{n-1}\right)}{n}(X) = \frac{X_1+...+X_n}{n}$.

Demonstrăm acum că shiftul t are proprietatea b.

LEMA. 5.1.

1. Fie $(E, \boldsymbol{E}, \Pi)$ un spaţiu probabilizat, $t : E \to E$ o funcţie măsurabilă cu proprietatea că P$\left(t^{-1}(A)\right) = P(A)$ pentru orice $A \in \boldsymbol{M}$ şi $\boldsymbol{M} \subset \boldsymbol{E}$ o familie de submulţimi. Presupunem că P$(A \cap t^{-n}(B)) \to P(A)P(B)$ pentru orice $A, B \in \boldsymbol{M}$. Atunci P$(A \cap t^{-n}(B)) \to P(A)P(B)$ pentru orice $A, B \in u(\boldsymbol{M})$ unde $u(\boldsymbol{M})$ este u-sistemul generat.

2. Dacă $\boldsymbol{M}$ este stabil la intersecţii finite, atunci afirmaţia este valabilă pentru pentru orice $A, B \in \sigma(\boldsymbol{M})$

3. Dacă în plus $\sigma(\boldsymbol{M}) = \boldsymbol{E}$, atunci egalitatea $t^{-1}(A) = A$ implică P$(A) \in \{0,1\}$, adică t este aplicaţie ergodică.

Demonstraţie.

1. Fie $A \in \boldsymbol{M}$ fixat şi $\boldsymbol{C}_A = \{B \in \boldsymbol{E} : P(A \cap t^{-n}(B)) \to P(A)P(B)\}$. Atunci $\boldsymbol{C}_A$ este un u-sistem care conţine pe $\boldsymbol{M}$. (singurul lucru mai delicat este să arătăm că dacă $(B_k)_k$ este un şir de elemente disjuncte din $\boldsymbol{C}_A$, atunci şi reuniunea $\bigcup\limits_k B_k$ este tot în $\boldsymbol{C}_A$. Adică dacă $P(A \cap t^{-n}(B_k)) \to P(A)P(B_k)$, atunci şi $\sum\limits_k P(A \cap t^{-n}(B_k)) \to \sum\limits_k P(A)P(B_k)$, lucru care rezultă din teorema Lebesgue cu dominarea $P(A \cap t^{-n}(B_k)) \le P(t^{-n}(B_k)) = P(B_k)$

În concluzie dacă $A \in \boldsymbol{M}$ şi $B \in \sigma(\boldsymbol{M})$, atunci $P(A \cap t^{-n}(B)) \to P(A)P(B)$.

Fie $\boldsymbol{C} = \{A \in \boldsymbol{E} : P(A \cap t^{-n}(B)) \to P(A)P(B) \text{ pentru orice } B \in u(\boldsymbol{M})\}$. Atunci $\boldsymbol{C}$ conţine pe $\boldsymbol{M}$ şi este de asemenea un u-sistem. (folosim acum dominarea $P(A_k \cap t^{-n}(B)) \le P(A_k)$!)

2. Evident

3. Dacă $t^{-1}(B) = B$, atunci $t^{-n}(B) = B$ deci convergenţa $P(A \cap t^{-n}(B)) \to P(A)P(B)$ devine P$(A \cap B) = P(A)P(B)$. Dacă luăm $A = B$ (e voie, căci $\sigma(\boldsymbol{M}) = \boldsymbol{E}$!) atunci avem P$(B) = P^2(B)$.

Observaţie. O funcţie t cu proprietatea din Lemă se numeşte **funcţie mixing**.

LEMA 5.2. Shiftul canonic $t : \mathbb{R}^\infty \to \mathbb{R}^\infty$ este mixing faţă de probabilitatea $\Pi = F^\infty$.

Demonstraţie.

$\sigma-$algebra produs este generată de mulţimea $\boldsymbol{M} = \{A_1 \times A_2 \times ... \times A_k \times \mathbb{R} \times \mathbb{R} \times .... : k \ge 1, A_j \in B(\mathbb{R})\}$. Dacă $A, B \in \boldsymbol{M}$ şi $B = B_1 \times B_2 \times ... \times B_k \times \mathbb{R} \times \mathbb{R}\times$, atunci $t^{-n}(B) = \mathbb{R} \times \mathbb{R} \times ... \times \mathbb{R} \times B_1 \times B_2 \times ... \times B_k \times \mathbb{R} \times \mathbb{R} \times ....$, (pe primele n poziţii este $\mathbb{R}$); deci pentru n destul de mare $A \cap t^{-n}(B) = A_1 \times A_2 \times ... \times A_k \times \mathbb{R} \times \mathbb{R} \times ... \times \mathbb{R} \times B_1 \times B_2 \times ... \times B_k \times \mathbb{R} \times ....$,aşadar $P(A \cap t^{-n}(B))$ este chiar egal cu $P(A)P(B)$

Cu aceasta legea numerelor mari ca aplicaţie a teoremei lui Birkhof este complet demonstrată.

1.8.1 6. Consecinţe ale Legii tari a numerelor mari.

6.1. **Metoda Monte Carlo**

Dacă vrem să calculăm o integrală de forma $\int f 1_C d\lambda^n$ unde $C \subset \mathbb{R}^n$ este o mulţime cu proprietatea că $\lambda^n(C) \in (0,\infty)$ putem observa că ea se scrie sub forma $\int f 1_C d\lambda^n = \lambda^n(C) Ef(X)$ cu X un vector repartizat uniform pe C.Atunci avem $\int f 1_C d\lambda^n = \lambda^n(C)\left(a.s. - \lim_N \frac{f(X_1)+...f(X_N)}{N}\right)$ unde $(X_n)_n$ este un şir de v.a. i.i.d cu repartiţia Uniform(C)

Problema dificilă este cum se poate simula un asemenea şir. Uneori este simplu.

6.2. **Teorema fundamentală a statisticii.**

Fie $(X_n)_n$ un şir de variabile aleatoare i.i.d. Variabila aleatoare $F_n(x) = \frac{1_{(X_1 \le x)} + 1_{(X_2 \le x)} + ... + 1_{(X_n \le x)}}{n}$ se numeşte funcţia de repartiţie empirică după n observaţii. Legea tare a numerelor mari ne spune că $F_n(x) \xrightarrow{a.s.} F(x)$ unde $F(x) = E1_{(X_1 \le x)} = P(X_1 \le x)$

Interpretare: *funcţiile de repartiţie empirice converg a.s. la funcţia de repartiţie teoretică.*

Teorema lui Glivenco: $\sup_x |F_n(x) - F(x)| \xrightarrow{a.s.} 0$

Interpretare: *funcţiile de repartiţie empirice converg UNIFORM a.s. la funcţia de repartiţie adevărată.*

Exemplu. Se aruncă două puncte la întîmplare într-un cerc de rază 1. Să se calculeze distanţa medie dintre ele.

Soluţie analitică. Fie $Z = (X, Y)$ şi $Z' = (X', Y')$ doi vectori iid repartizaţi Uniform(C) cu $C = \left\{(x, y) : x^2 + y^2 \le 1\right\}$ şi fie $D = \|Z - Z'\|$

Atunci avem, din formula de transport $\mathrm{E}\|Z - Z'\| = \mathrm{E}\sqrt{(X - X')^2 + (Y - Y')^2} =$
$\frac{1}{\pi^2} \int \sqrt{(x - x')^2 + (y - y')^2} 1_C(x, y) 1_C(x', y') \, dx dx' dy dy'$

$= \frac{1}{\pi^2} \int_{-1}^{1} \int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{(x - x')^2 + (y - y')^2} dy dy' dx dx'$

Este o integrală care nu se poate calcula

Simulăm un vector repartizat Uniform(C) prin metoda următoare, numită a **respingerii**:

- luăm X, Y independente repartizate Uniform$(-1, 1)$.
- verificăm dacă $X^2 + Y^2 \le 1$.
  Daca DA, reţinem $Z = (X, Y)$
  Daca NU, repetăm.

Algoritmul dă rezultate bune dacă putem încadra mulţimea C într-un dreptunghi de forma $I \times J$ cu $I, J \subset \mathbb{R}$ intervale mărginite în aşa fel încît raportul $\frac{\lambda^2(C)}{\lambda(I)\lambda(J)}$ să nu fie prea mic.

Formal, se ia U~Uniform$\left((0,1)^2\right)$ (e simplu) si se pune $Z = \left(U | U^{-1}(C)\right)$

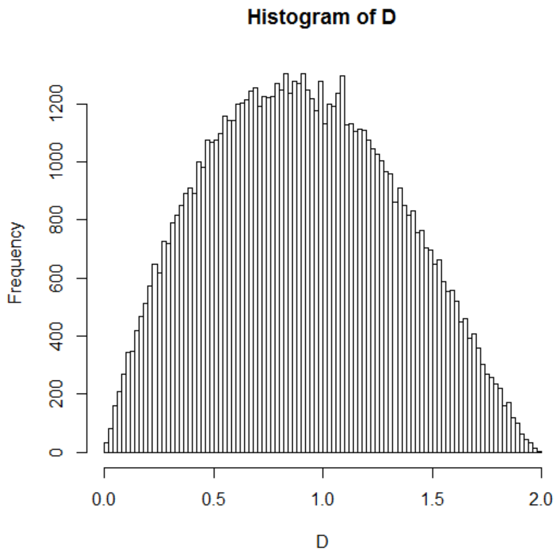
Dupa cum am vazut in lectia 6.2 Z este un vector aleator cu proprietatea ca $\mathrm{P}(Z \in B) = P(U \in B | U \in C) = \frac{\lambda^2(U \cap C)}{\lambda^2(C)}$

Un script R care simuleaza N asemenea vectori aleatori este

```
n=10000
u=2*runif(n)-1;v=2*runif(n)-1;w=which(u^2+v^2<1);x=u[w];y=v[w]
up=2*runif(n)-1;vp=2*runif(n)-1;wp=which(up^2+vp^2<1);xp=up[wp];yp=vp[wp]
N=min(length(x),length(xp))
X=x[1:N];Y=y[1:N];Xp=xp[1:N];Yp=yp[1:N]
D=sqrt((X-Xp)^2+(Y-Yp)^2);summary(D)
```

Explicaţie. n este numărul de vectori (U_j, V_j) ~Uniform$\left((0,1)^2\right)$ simulaţi; în urma testului $U_j^2 + V_j^2 \le 1$ din ei se reţin vectorii (X_j, Y_j); se repetă cu vectorii $\left(U_j', V_j'\right)$ din care se reţin vectorii $\left(X_j', Y_j'\right)$.

N este numărul de vectori aleatori care au trecut testul de apartenenţă în C. Cu ajutorul lor se calculează distanţele $D_j = \sqrt{\left(X_j - X_j'\right)^2 + \left(Y_j - Y_j'\right)^2}$ şi li se face apoi sumarul.



Rezultate: Min = $\min D_j$ este minimul empiric: cu legea tare a numerelor mari tinde aproape sigur la $ess \inf D$ =0

1st Qu $= F_n^{-1}\left(\frac{1}{4}\right)$, unde F_n este repartiţia empirică a lui D după N probe;

tinde la $F^{-1}\left(\frac{1}{4}\right)$ unde F este repartiţia lui D

Median = mediana empirică a lui D, tinde la mediana teoretică

Mean = media empirică a lui D,tinde la ED

3rd Qu. $= F_n^{-1}\left(\frac{3}{4}\right) \xrightarrow{a.s.} F^{-1}\left(\frac{3}{4}\right)$

Max $= \max(D_j) \xrightarrow{a.s.} ess \sup D = 2$

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.004957	0.575500	0.889100	0.906700	1.226000	1.985000
0.01565	0.58110	0.89570	0.91090	1.22700	1.98500
0.01315	0.57970	0.89000	0.90500	1.21900	1.99500
0.002263	0.573700	0.888600	0.904800	1.221000	1.992000
0.001696	0.575100	0.892900	0.906600	1.224000	1.993000

Dar daca am arunca punctele în pătratul unitate? Formal, am avea de calculat E$D = \int_0^1\int_0^1\int_0^1\int_0^1 \sqrt{(x-x')^2+(y-y')^2}dydy'dxdx'$ care pare mai simplă; nu cred că se poate face.

Simularea este mult mai simplă

x=runif(n);y=runif(n);xp=runif(n);yp=runif(n);d=sqrt((x-xp)^2+(y-yp)^2;summary(d)

Vedeţi ce iese.

1.9 Lecţia 9. Convoluţia, convergenţa repartiţiilor şi Teorema Limită Centrală

1.9.1 Convoluţia

Fie X, Y două variabile aleatoare. Dacă se ştie repartiţia lor (notată cu F şi G), atunci nu se poate spune nimic despre repartiţia sumei $X+Y$, decît dacă ştim repartiţia comună a vectorului (X, Y) .În cazul cel mai simplu, dacă presupunem că X şi Y sunt independente, atunci repartiţia vectorului (X, Y) este $F \otimes G$. În acest caz repartiţia sumei $S = X + Y$ se notează cu $F * G$ şi se numeşte **convoluţia** dintre F şi G.

Formula de transport ne arată că dacă $f : \mathbb{R} \to \mathbb{R}$ este măsurabilă mărginită, atunci

$\mathrm{E} f(X+Y) = \int f(x+y)\, dF \otimes G(x,y) = \int\int f(x+y)\, dF(x)\, dG(y) = \int\int f(x+y)\, dG(y)\, dF(x)$

Aşadar am descoperit formula

$$\int f dF * G = \int\int f(x+y)\, dF(x)\, dG(y) = \int\int f(x+y)\, dG(y)\, dF(x)$$

care poate fi luată ca o definiţie a convoluţiei.

Proprietăţi.

1. $(F * G)(B) = \int\int F(B-y)\, dG(y) = \int\int G(B-x)\, dF(x)$; deci funcţie de repartiţie se calculează după formula

$(F * G)(t) = \int\int F(t-y)\, dG(y) = \int\int G(t-x)\, dF(x)$

2. $F * \delta_0 = F, (F * \delta_a)(x) = F(x-a)$

3. $F * (aG + bH) = aF * G + bF * H$ (convoluţia este distributivă faţă de adunare)

4. Spaţiul măsurilor cu semn mărginite de pe dreapta reală echipat cu adunarea, înmulţirea cu scalari şi convoluţia $(M_b(\mathbb{R}, B(\mathbb{R})), +, *)$ este o algebră comutativă, cu elementul neutru δ_0.În cadrul acestei algebre familia Prob$(\mathbb{R}, B(\mathbb{R}))$ a probabilităţilor este o mulţime convexă

5. $\delta_a * \delta_b = \delta_{a+b}, (p\delta_a + q\delta_b)^{*n} = \sum\limits_{k=0}^{n} C_n^k p^{n-k} q^k \delta_{(n-k)a+kb}$

6. **Proprietatea de netezire**. Dacă G este absolut continuă, atunci şi $F * G$ este absolut continuă indiferent de F. Dacă G are densitatea g, atunci $F * G$ are densitatea $p_{F*G}(t) = \int g(t-x)\, dF(x)$

7. **Convoluţia densităţilor**. Dacă F are densitatea p şi G are densitatea q atunci $F*G$ are densitatea $p_{F*G}(t) = \int q(t-x)\, p(x)\, dx = \int p(t-y)\, q(y)\, dy$

8. Dacă $\mathrm{m}_F(t) = \int e^{tx} dF(x)$, $\mathrm{m}_G(t) = \int e^{tx} dG(x)$ sunt mgf-urile repartiţiilor Fsi G atunci $\mathrm{m}_{F*G} = \mathrm{m}_F \mathrm{m}_G$. Aceeaşi afirmaţie este valabilă pentru funcţiile caracteristice ϕ_F, ϕ_G sau pentru funcţiile generatoare (dacă F şi G au suport pe mulţimea numerelor n g_F, g_G

9. Bin(m,p) $*$Bin(n,p) =Bin$(m+n,p)$,Poisson(a) $*$Poisson(b) =Poisson$(a+b)$,

$\mathrm{N}\left(\mu_1, \sigma_1^2\right) * \mathrm{N}\left(\mu_2, \sigma_2^2\right) = \mathrm{N}\left(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}^2\right)$, $\mathrm{Gamma}(m, a) * \mathrm{Gamma}(n, a) = \mathrm{Gamma}(m + n, a$

(prin definiţie, dacă m este număr natural $\mathrm{Gamma}(m, a) = \mathrm{Exp}(a)^{*n}$.

10. $\mathrm{Uniform}(a, b) * \mathrm{Uniform}(c, d) = \mathrm{Trapez}(a + c, a + d, b + c, b + d)$

$$F_{\mathrm{Uniform}(0,1)^{*n}}(x) = \frac{x^n - C_n^1 (x-1)_+^n + C_n^2 (x-2)_+^n - C_n^3 (x-3)_+^n + \ldots}{n!}$$

Demonstraţiile punctelor 1-8 sunt banale şi lăsate ca exerciţii.

9. La binomială şi Poisson se face funcţia generatoare, la Gamma nu avem nimic de demonstrat iar la normală folosim funcţia caracteristică sau mgf-ul: $\mathrm{m}_{N(\mu,\sigma^2)}(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$

10. Observăm că $\mathrm{Uniform}(a, b) = \delta_a * \mathrm{Uniform}(0, b - a)$, $\mathrm{Uniform}(c, d) = \delta_c * \mathrm{Uniform}(0, d - c)$ deci

$\mathrm{Uniform}(a, b) * \mathrm{Uniform}(c, d) = \delta_{a+c} * \mathrm{Uniform}(0, \alpha) * \mathrm{Uniform}(0, \beta)$ cu $\alpha = b - a, \beta = d - c$.

Fie X, Y independente, X~Uniform$(0, \alpha)$, Y~Uniform$(0, \beta)$. Atunci $\mathrm{P}(X + Y \leq t) = \frac{\lambda^2((0,\alpha)\times(0,\beta)\cap D_t)}{\alpha\beta}$ unde $D_t = \{(x, y) : x + y \leq t\}$. Calculînd aria aceasta (un triunghi triunghic isoscel din care se mai scad eventual alte două triunghiuri dreptunghice isoscele) rezultă $\mathrm{F}_{X+Y}(x) = \frac{1}{2\alpha\beta}\left(x^2 - (x - \alpha)_+^2 - (x - \beta)_+^2 + (x - \alpha - \beta)_+^2\right)$

cu densitatea

$p_{X+Y}(x) = \frac{1}{\alpha\beta}\left(x - (x - \alpha)_+ - (x - \beta)_+ + (x - \alpha - \beta)_+\right)$

Graficul ei are forma unui trapez isoscel :de exemplu dacă $\alpha \leq \beta$ avem

$p_{X+Y}(x) = \frac{1}{\alpha\beta}\left(x 1_{(\alpha,\beta)} + \alpha 1_{(\alpha,\beta)} + (\alpha + \beta - x) 1_{(\beta,\alpha+\beta)}\right)$

Analog,

$F_{U(0,\alpha)*U(0,\beta)*U(0,\gamma)}(x) = \frac{1}{6\alpha\beta\gamma}\left(x^3 - (x - \alpha)_+^3 - (x - \beta)_+^3 - (x - \gamma)_+^3 + (x - \alpha - \beta)_+^3 + (x - \alpha - \beta)_+^3 + (\alpha\right.$

a carei densitate deja are formă de clopot.

Pentru a calcula convoluţia $\mathrm{Uniform}(0, 1)^{*n}$, calculăm $\lambda^n\left([0, 1]^n \cap D_t\right)$ unde $D_t = \left\{(x_k)_{1\leq k\leq n} \in \mathbb{R}_+^n : x_1 + x_2 + ... + x_n \leq t\right\}$

Observăm că $[0, 1]^n \cap D_t = D_t - \bigcup_{k=1}^{n} \Delta_{t,j}$ unde $\Delta_{t,j} = \left\{(x_k)_{1\leq k\leq n} \in \mathbb{R}_+^n : x_1 + x_2 + ... + x_n \leq t, x_j \geq 1\right\}$

Dacă scriem $x_j = 1 + y_j$, atunci vedem că putem scrie $\Delta_{t,j} = e_j + \{(x_1, ..., x_{j-1}, y_j, x_{j+1}, ..., x_n) \in \mathbb{R}_+^n : x_1 + x_2$ $e_j = (0, ..., 0, 1, 0, ...0)$ este versorul canonic de pe poziţia j. Altfel scris, $\Delta_{t,j} = e_j + D_{t-1}$

Folosind principiul includerii şi excluderii avem şi observînd că $\Delta_{t,j} \cap \Delta_{t,i} = e_i + e_j + D_{t-2}, \Delta_{t,j} \cap \Delta_{t,i} \cap \Delta_{t,k} = e_i + e_j + e_k + D_{t-3}$ etc găsim că

$$\lambda^n\left([0, 1]^n \cap D_t\right) = \lambda^n\left(D_t - \bigcup_{k=1}^{n} \Delta_{t,j}\right) = \lambda^n(D_t) - \lambda^n\left(\bigcup_{k=1}^{n} \Delta_{t,j}\right)$$

$$= \lambda^n(D_t) - \left(\sum_i \lambda^n(\Delta_{t,i}) - \sum_{i,j:i\neq j} \lambda^n(\Delta_{t,i} \cap \Delta_{t,j}) + \sum_{i,j,k::i\neq j\neq k\neq i} \lambda^n(\Delta_{t,i} \cap \Delta_{t,j} \cap \Delta_{t,k}) - ...\right)$$

găsim formula (10) , deoarece se verifică imediat că $\lambda^n(D_t) = \frac{t^n}{n!}$

1.9.2 Convergenţa repartiţiilor

Ce se întîmplă dacă facem mai multe convoluţii de tipul $F, F^{*2}, F^{*3}, ...$? Legea numerelor mari spune că în cazul i.i.d. media aritmetică $\overline{X}_n = \frac{X_1+X_2+...+X_n}{n}$ converge aproape sigur la media teoretică $\mu = E\mathrm{X}_1 = \int xdF(x)$, unde F este repartiţia lui X_1. Cum se reflectă acest lucru în repartiţia lui $\overline{X}_n$?

Oricum, dacă $X_n \xrightarrow{a.s.} X$ şi $h : \mathbb{R} \to \mathbb{R}$ este o funcţie continuă, atunci $h(X_n) \xrightarrow{a.s.} h(X)$.

Cum am putea traduce acest lucru în termeni de repartiţii?

Dacă am putea comuta media cu limita, ar trebui să rezulte că $\mathrm{E}h(X_n) \to \mathrm{Eh}(X)$.

Un caz în care acest lucru este absolut sigur este dacă h este continuă **şi mărginită**.

Să notăm cu F_n repartiţiile lui X_n şi cu F repartiţia lui X. Deci dacă h este continuă şi mărginită avem
$X_n \xrightarrow{a.s.} X \Longrightarrow h(X_n) \xrightarrow{a.s.} h(X) \Longrightarrow \mathrm{E}h(X_n) \to \mathrm{Eh}(X) \iff \int hdF_n \to \int hdF$.

Aceasta este definiţia la care s-au oprit matematicienii.

Definiţie. Fie $(F_n)_n$ un şir de repartiţii. Spunem că F_n converge la F (**slab**) dacă

$$\int hdF_n \to \int hdF \text{ pentru orice } h : \mathbb{R} \to \mathbb{R} \text{ continuă mărginită}$$

Notaţia acceptată este "$F_n \Longrightarrow F$"

Vom da fără demonstraţie cele două teoreme mari privind convergenţa slabă

Teorema 1.(Portmanteau**).**$F_n \Longrightarrow F$ dacă şi numai dacă $F_n(x) \to F(x)$ în orice punct de continuitate al lui F, adică în toţi acei x cu proprietatea că $F(x-0) = F(x)$

Comentariu. Trebuie pusă condiţia neplăcută de continuitate deoarece este evident că $\delta_{\frac{1}{n}} \Longrightarrow \delta_0$,dar nu este adevărat că funcţiile de repartiţie converg peste tot: funcţia de repartiţie a lui $\delta_{\frac{1}{n}}$ este $1_{\left[\frac{1}{n},\infty\right)}$ şi ea converge la $1_{(0,\infty)}$ care, nefiind continuă la dreapta NU este o funcţie de repartiţie. De altfel se poate arăta că $F_n \Longrightarrow F$ dacă şi numai dacă $F_n(x) \to F(x)$ pe o mulţime $\Gamma \subset \mathbb{R}$ densă.

Teorema 2.(de convergenţă a funcţiilor caracteristice**): .**$F_n \Longrightarrow F$ dacă şi numai dacă $\phi_{F_n} \to \phi_F$ adică $\int e^{itx}dF_n(x) \to \int e^{itx}dF(x)$ pentru orice $t \in \mathbb{R}$

Acum dăm un criteriu suficient pentru a asigura convergenţa slabă, folositor în cazurile didactice

Propoziţia 3. (criteriul convergenţei densităţilor)

Să presupunem că $F_{\acute{n}}$ au densităţile p_n faţă de o măsură $\sigma-$finită, μ.Să mai presupunem că p_n converge la o funcţie p despre care **ştim deja** că este densitate. Atunci $F_n \Longrightarrow p \cdot \mu$

Demonstraţie. $|F_n(x) - F(x)| = \left|\int_{-\infty}^{x} (p_n(z) - p(z))\, d\mu(z)\right| \leq \int |p_n - p|\, d\mu$

Scriem $|p_n - p| = |p - p_n| = 2(p - p_n)_+ - (p - p_n)$ şi avem în continuare

$\int |p_n - p|\, d\mu = \int \left(2\,(p - p_n)_+ - (p - p_n)\right) d\mu = 2\int (p - p_n)_+\, d\mu - \int (p - p_n)\, d\mu = 2\int (p - p_n)_+\, d\mu - 1 + 1 = 2\int (p - p_n)_+\, d\mu$

Dar $(p - p_n)_+$ converge la 0 şi este dominat de p. Din teorema lui Lebesgue $\int (p - p_n)_+\, d\mu$ va converge de asemenea la 0.

Cazuri particulare: măsura Lebesgue sau măsura cardinal pe o mulţime numărabilă.

Corolar. Teorema de aproximare Poisson. Dacă $F_n = Bin\,(n, p_n)$ şi $np_n \to \lambda$, atunci $F_n \implies Poisson\,(\lambda)$

Într-adevăr, $C_n^k p^k\,(1-p)^{n-k}$ converge la $\frac{\lambda^k}{k!}e^{-\lambda}$.

Corolar al teoremei lui Glivenko. Repartiţiile discrete sunt dense în familia tuturor repartiţiilor de pe dreaptă.

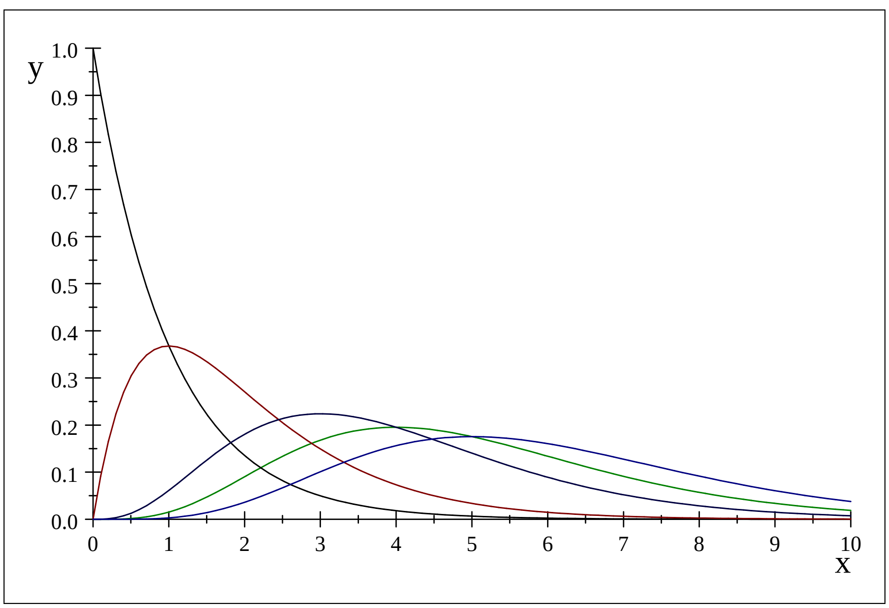
Într-adevăr, repartiţiile empirice sunt discrete şi converg la repartiţia teoretică.

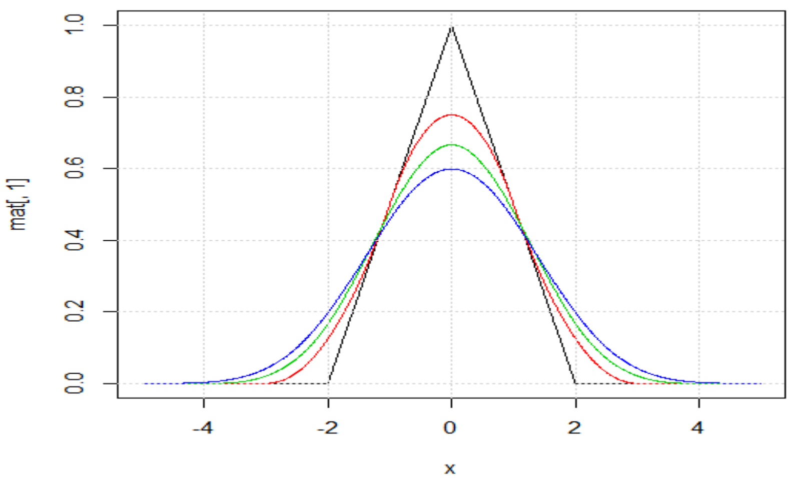
Notaţie. Să presupunem că variabilele X_n au repartiţiile F_n şi $F_n \implies F$. Atunci se mai foloseşte notaţia $X_n \xrightarrow{D} F$, care se citeşte "X_n converge în repartiţie la F". Sau, dacă X este o variabilă aleatoare cu repartiţia F, atunci se mai poate citi "X_n converge în repartiţie la X". Evident că dacă X_n converge aproape sigur la X, atunci X_n converge în repartiţie la X.

1.9.3 Teorema limită centrală

Convoluţiile tind să aibă o formă de clopot.

Iată, de exemplu densităţile lui F^{*n} dacă $F = Exp\,(1)$ pentru $n = 1, 2, 3, 4, 5$ (negru = densitatea lui F,maro = densitatea lui $F * F$, indigo = densitatea lui $F * F * F$,verde =densitatea lui F^{*4},albastru = densitatea lui F^{*5})





sau densitatile repartiţiilor $\mathrm{U}(-1,1)^{*n}$ pentru $n=2,3,4,5$ (Negru pentru $n=2$, roşu pentru $n=3$, verde pentru $n=4$, albastru pentru $n=5$)

Explicatia se află în

TEOREMA LIMITĂ CENTRALĂ

Fie $(X_n)_n$ un şir de variabile aleatoare i.i.d din L^2 cu media μ şi varianţa σ^2.

Atunci $\frac{X_1+X_2+...+X_n-n\mu}{\sigma\sqrt{n}} \xrightarrow{D} \mathrm{N}(0,1)$

Sau, mai explicit

$\mathrm{P}\left(\frac{X_1+X_2+...+X_n-n\mu}{\sigma\sqrt{n}} \le x\right) \longrightarrow \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}dy$ pentru orice x număr real.

Demonstraţie. Vom folosi teorema de convergenţă a funcţiilor caracteristice: va trebui să să arătăm că $\phi_{s_n}(t) \longrightarrow \phi_{N(0,1)}(t)$ pentru orice t.Am notat cu s_n raportul $\frac{X_1+X_2+...+X_n-n\mu}{\sigma\sqrt{n}} = \frac{Y_1+Y_2+...+Y_n}{\sigma\sqrt{n}}$ unde $Y_n = X_n-\mu$ sunt variabilele centrate; evident $Y_n \in L^2$.

Fie $\phi(t) = \mathrm{E}e^{itY_n}, \phi:\mathbb{R}\to\mathbb{C}$. Funcţia ϕ este derivabilă de două ori, $\phi(0) = 1, \phi'(0) = 0, \phi''(0) = i^2E(X_n-\mu)^2 = -\sigma^2$.

Este important de reţinut că putem scrie $\phi(t) = \phi(0)+t\phi'(0)+\frac{t^2}{2}\phi''(0)+o(t)t^2$ unde $o:\mathbb{R}\to\mathbb{C}$ este o funcţie continuă şi $o(0)=0$.Într-devăr, $o(t) = \frac{\phi(t)-\phi(0)-t\phi'(0)-\frac{t^2}{2}\phi''(0)}{t^2}$ este evident continuă; din regula L'Hospital $\lim\limits_{t\to 0} o(t) = \frac{1}{2}\lim\limits_{t\to 0}\left(\frac{\phi'(t)-\phi'(0)}{t} - \phi''(0)\right) = 0$ - căci am presupus că $\phi'(t)$ este derivabilă în 0.

Atunci $\phi_{s_n}(t) = \mathrm{E}e^{it\frac{Y_1+Y_2+...+Y_n}{\sigma\sqrt{n}}} = \left(Ee^{i\frac{t}{\sigma\sqrt{n}}Y_1}\right)^n$ (am presupus variabilele aleatoare i.i.d!)

$= \phi^n\left(\frac{t}{\sigma\sqrt{n}}\right) = \left(\phi(0) + \frac{t}{\sigma\sqrt{n}}\phi'(0) + \frac{t^2}{2n\sigma^2}\phi''(0) + o\left(\frac{t}{\sigma\sqrt{n}}\right)\frac{t^2}{n\sigma}\right)^n =$

$\left(1 - \frac{t^2}{2n} + o\left(\frac{t}{\sigma\sqrt{n}}\right)\frac{t^2}{n\sigma}\right)^n = (1-\alpha_n)^n = \left((1-\alpha_n)^{\frac{1}{\alpha}\cancel{n}}\right)^{n\alpha_n}$ (cu $\alpha_n = \frac{t^2}{2n} - o\left(\frac{t}{\sigma\sqrt{n}}\right)\frac{t^2}{n\sigma}$

Deci $\lim\phi^n\left(\frac{t}{\sigma\sqrt{n}}\right) = \lim e^{-n\alpha_n} = \lim e^{-n\left(\frac{t^2}{2n} - o\left(\frac{t}{\sigma\sqrt{n}}\right)\frac{t^2}{n\sigma}\right)} = e^{-\frac{t^2}{2} - \lim o\left(\frac{t}{\sigma\sqrt{n}}\right)t^2} = e^{-\frac{t^2}{2}} = \phi_{N(0,1)}(t)\,.QED.$

Convergenţa funcţiilor de repartiţie nu implică şi convergenţa densităţilor. Reciproca este evident adevărată, după cum am văzut mai sus.

Totuşi, în anumite ipoteze se poate demonstra - mult mai greu şi cu un aparat matematic foarte sofisticat - şi

TEOREMA LIMITĂ LOCALĂ.

Presupunem că variabilele aleatoare de mai sus, X_n, sunt şi absolut continue. Fie p_n densitatea sumelor centrate şi normate $s_n = \frac{Y_1+Y_2+...+Y_n}{\sigma\sqrt{n}}$. Să presupunem că există un n_0 cu proprietatea că p_{n_0} este mărginită. Atunci $p_n(x) \longrightarrow \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2}}$: densităţile tind la densitatea normală. Aceasta ar fi explicaţia fenomenului arătat în cele două grafice de mai sus.

Aplicaţie în aproximarea unor repartiţii.

Binomial(n,p) $\cong$N$\left(np, \sqrt{npq}\right)$ dacă np şi npq sunt "mari"

Gamma(n,λ) $\cong$N$\left(\frac{n}{\lambda}, \left(\frac{\sqrt{n}}{\lambda}\right)^2\right)$ dacă $\frac{n}{\lambda}$ şi $\frac{\sqrt{n}}{\lambda}$ sunt "mari"

Poisson(λ) $\cong$N$\left(n\lambda, \sqrt{n\lambda}^2\right)$ dacă $n\lambda$ este "mare"

etc

Aplicaţie în calculul primelor de asigurare.

Se dă un portofoliu omogen de n asiguraţi de tip F.

Traducere matematică: Fie $(X_k)_{1\le k\le n}$ variabile aleatoare i.i.d. nenegative din L^2. Ele e numesc daune.

Asiguratorul îşi asumă riscul de default ε. Ce primă de asigurare să ceară?

Traducere matematică. Care să fie cel mai mic număr Π cu proprietatea că $P(X_1 + X_2 + ... + X_n \ge n\Pi) \le \varepsilon$?

Sriem $P(X_1 + X_2 + ... + X_n \ge n\Pi) = P\left(\frac{X_1+X_2+...+X_n-n\mu}{\sigma\sqrt{n}} \ge \frac{n\Pi-n\mu}{\sigma\sqrt{n}}\right) = P\left(s_n \ge \frac{\sqrt{n}(\Pi-\mu)}{\sigma}\right)$

Dacă n este destul de mare (ce înseamnă acest lucru este o altă discuţie, de loc simplă!), atunci putem aproxima ultima probabilitate cu $N(0,1)\left(\left(\frac{\sqrt{n}(\Pi-\mu)}{\sigma}, \infty\right)\right) = \underline{\Phi}\left(\frac{\sqrt{n}(\Pi-\mu)}{\sigma}\right)$, unde prin tradiţie se notează $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}dy$, adică funcţia de repartiţie a normalei standard. Cum repartiţia normală este simetrică faţă de origine, $\underline{\Phi}(a) = \Phi(-a)$ pentru orice $a \in \mathbb{R}$. De aceea putem scrie $P(X_1 + X_2 + ... + X_n \ge n\Pi) \cong \Phi\left(-\frac{\sqrt{n}(\Pi-\mu)}{\sigma}\right)$ şi punem condiţia ca $\Phi\left(-\frac{\sqrt{n}(\Pi-\mu)}{\sigma}\right) = \varepsilon \iff -\frac{\sqrt{n}(\Pi-\mu)}{\sigma} = \Phi^{-1}(\varepsilon) \iff \frac{\sqrt{n}(\Pi-\mu)}{\sigma} = -\Phi^{-1}(\varepsilon) = \Phi^{-1}(1-\varepsilon)$ de unde găsim $\frac{\Pi-\mu}{\sigma} = \frac{\Phi^{-1}(1-\varepsilon)}{\sqrt{n}}$ adică

Principiul de calcul medie-abatere:

$\Pi = \mu + \frac{\sigma}{\sqrt{n}} \Phi^{-1}(1-\varepsilon)$

Pe această bază putem, de exemplu aprecia că, pentru un portofoliu omogen de 100 asiguraţi de tip Binomial$\left(100, \frac{1}{4}\right)$ prima va fi de tipul $\Pi = 25 + \frac{\sqrt{100 \cdot \frac{1}{4} \cdot \frac{3}{4}}}{\sqrt{100}} \Phi^{-1}(1-\varepsilon) = 25 + 0.4\,331 \Phi^{-1}(1-\varepsilon)$

Concret, pentru un risc de 1% am avea $\Phi^{-1}(1-\varepsilon) = \Phi^{-1}(0.99) = 2.326348$ deci $\Pi = 25 + 2.\,326\,3 = 27.\,326$

Putem compara cu probabilitatea exactă: $X_1 + ... + X_{100}$ are repartiţia Binomial$\left(100, \frac{1}{4}\right)^{*100}$ =Binomial$\left(10000, \frac{1}{4}\right)$;

$n\Pi = 2732.6$

Probabilitatea de default este (în limbajul R) $\psi = 1$ - pbinom(2732.6,10000,1/4)= 5.236055e-08