probabilități

PROBABILITĂȚI

$P(\emptyset)=0 \qquad P(\Omega)=1 \qquad P(A\cap B)=P(A)P(B)$ $\perp\!\!\!\perp$ $\Rightarrow$ independență

$P\left(\bigcup_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} P(A_i)$, $A_i\cap A_j=\emptyset$ $i\neq j$

$P(A^c)=1-P(A) \qquad P(A)+P(A^c)=1$

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)-\sum_{i<j} P(A_i\cap A_j)+\sum_{i<j<k} P(A_i\cap A_j\cap A_k)+\dots+(-1)^{n-1}P(A_1\cap A_2\cap\dots\cap A_n)$$ $\leftarrow$ $\boxed{\text{Formula lui Poincare}}$

$$P(A)=\frac{\text{nr. caz. fav}}{\text{nr. caz. nefav}}$$ $\leftarrow$ $\boxed{\text{modelul Laplace}}$

ALGEBRĂ COMBINAȚIONALĂ

$|A\cup B|=|A|+|B|$ $\leftarrow$ $\boxed{\text{Formula sumei}}$

$$|A_1\cup A_2\cup\dots\cup A_n|=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\le i_1<\dots<i_k\le k}|A_{i_1}\cap\dots\cap A_{i_k}|$$

$\uparrow$ $\boxed{\text{Principiul includerii - excluderii}}$

$|A\times B|=|A|\cdot|B|$ $\leftarrow$ $\boxed{\text{Formula produsului}}$

$|A^n|=|A|^n$

$\hookrightarrow |A_1\times A_2\times\dots\times A_n|=|A_1|\cdot|A_2|\cdot\ldots\cdot|A_n|$

$|A^c|=|\Omega|-|A|$

PROBABILITĂȚI CONDIȚIONATE

$$P(A|B) = \frac{P(A\cap B)}{P(B)}, \quad P(B) > 0$$

$$P(A\cap B) = P(B)\cdot P(A|B)$$
$$= P(A)\cdot P(B|A)$$

$$P(A_1\cap A_2\cap\ldots\cap A_n) = P(A_1)\,P(A_2|A_1)\,P(A_3|A_1\cap A_2)\cdot\ldots\cdot P(A_n|A_1\cap A_2\cap\ldots\cap A_{n-1})$$

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i), \quad \Omega = B_1\cup B_2\cup\ldots\cup B_n$$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

Formula lui Bayes

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}$$

$$P(A|\underbrace{B,C}_{B\cap C}) = \frac{P(B|A,C)P(A|C)}{P(B|C)}$$

$P(A\cap B|C) = P(A/C)P(B|C) \Rightarrow$ independență cond.

VARIABILE ALEATOARE DISCRETE

$$P(X\in A) = P(\{\omega / X(\omega)\in A\}) = P(X^{-1}(A)) = (P\circ X^{-1})(A) = \sum_{x\in A\cap X(\Omega)} (X=x)$$

Funcția de repartiție

$F:\mathbb{R}\to[0,1] \quad F(x) = P(X\le x),\ \forall x\in\mathbb{R}$

$\Rightarrow \nearrow$ + cont. la dr

$\lim\limits_{x\to-\infty} F(x) = 0$ și $\lim\limits_{x\to+\infty} F(x) = 1$

$$P(X>x) = 1-P(X\le x) = 1-F(x)$$

$$P(x<X\le y) = P(X\le y) - P(X\le x) = F(y)-F(x)$$

$$P(X=x) = P(X\le x) - P(X<x) = F(x) - \lim_{y\nearrow x} F(y)$$

Funcția de masă

$p_X(x) = P(X=x),\ \forall x\in\mathbb{R} \qquad p_X(x)\ge 0 \qquad \sum\limits_{x\in X(\Omega)} p_X(x) = 1$

$\Rightarrow P(X\in A) = \sum\limits_{x\in A\cap X(\Omega)} p_X(x)$

$$F(x) = \sum_{\substack{y\le x\\ y\in X(\Omega)}} p_X(y),\quad A = (-\infty, x]$$

Variabile aleatoare Bernoulli

$$X(\omega) = \begin{cases} 1, & \omega\in A\\ 0, & \omega\notin A\end{cases} \qquad F(x) = \begin{cases} 0, & x<0\\ 1-p, & 0\le x<1\\ 1, & x\ge 1\end{cases}$$

$X\sim B(p)$

$$p_X(x) = p^x(1-p)^{1-x}$$

Variabile aleatoare binomiale

$X \sim B(n, p)$ $\quad p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$

(extragere cu întoarcere)

Variabile aleatoare hipergeometrice

$X \sim HG(n, N, M)$ $\quad p_X(k) = \dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$

(extragere fără întoarcere)

Repartiție discretă uniformă

$X \sim U(\Delta)$ $\quad p_X(x) = \dfrac{1}{|\Delta|}$

$$P(x \in A) = \frac{|A|}{|\Delta|}$$

(val la a j-a extragere)

Variabile aleatoare repartizate geometric

$X \sim Geom(p)$ $\quad p_X(k) = (1-p)^{k-1} p$

$$P(X = n+k \mid X \geq n) = P(X = k)$$

$P(X \geq s+t \mid X \geq s) = P(X \geq t)$ ← $\boxed{\text{Lipsa de memorie}}$

Variabile aleatoare repartizate negativ binomial

$X \sim NB(r, p)$ $\quad p_X(k) = \binom{k-1}{r-1} (1-p)^{k-r} p^r$

(o secv. de lung k are r val de 1)

Scanned with CamScanner

Variabile aleatoare Poisson

$X \sim P(\lambda)$ $\qquad p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$

(nr. de cazuri de boală rară dintr-o regiune)

$P(X=x, Y=y) = P(X=x)P(Y=y) \Rightarrow$ independență

Media

$$E[X] = \sum_x x F(x)$$

$\sum_x |x| f(x)$ diverg $\Rightarrow$ media nu este definită

$E[aX+bY] = aE[X] + bE[Y]$ $\qquad E[a] = a$

$X \perp\!\!\!\perp Y \Rightarrow E[XY] = E[X]E[Y]$

$Y = g(X) \Rightarrow E[Y] = E[g(X)] = \sum_x g(x) p_X(x)$

Varianța

$Var(X) = E[(X-E[X])^2] = E[X^2] - (E[X])^2 = \sigma^2$

$Var(X+a) = Var(X)$

$Var(aX) = a^2 Var(X)$ $\qquad Var(aX+b) = a^2 Var(X)$

$X \perp\!\!\!\perp Y \Rightarrow Var(X+Y) = Var(X) + Var(Y)$

Momente de ordin k

$E[X^k] = \sum_x x^k F(x)$ $\qquad E[(X-a)^k]$ centrat în a

$E[(X-E[X])^k]$ centrat

Abaterea standard

$$SD(X) = \sqrt{Var(X)} = \sqrt{\cdot}$$

Covarianța

$$Cov(X,Y) = E[(X-E[X])(Y-E[Y])]$$
$$= E[XY] - E[X]E[Y] = Cov(Y,X)$$

$$Cov(X,X) = Var(X)$$

$X \perp\!\!\!\perp Y \Rightarrow Cov(X,Y) = 0 \Rightarrow$ necorelate

$$Cov(X,a) = 0$$

$$Cov(aX+b, Y) = a\,Cov(X,Y)$$

$$Cov(X+Y, Z) = Cov(X,Z) + Cov(Y,Z)$$

$$\Rightarrow Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$$

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$$

Corelația

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

coeficient de corelație

$$|\rho(X,Y)| \le 1$$

VARIABILE ALEATOARE CONTINUE

$$P(X\in A) = \int_A f(x)\,dx\ ,\quad \forall A \subseteq \mathbb{R}$$

$$P(a \le X \le b) = \int_a^b f(x)\,dx \qquad P(x\in\mathbb{R}) = \int_{-\infty}^{+\infty} f(x)\,dx = 1$$

Densitatea de repartiție

$f:\mathbb{R}\to\mathbb{R}$! nu este o probabilitate

$\Rightarrow f(x) \ge 0$ și $\int_{-\infty}^{+\infty} f(x)\,dx = 1,\ \forall x\in\mathbb{R}$

$$P(X=a) = \int_a^a f(x)\,dx = 0$$

$$P(a\le X\le b) = P(a<X\le b) = P(a\le X<b) = P(a<X<b)$$

Funcția de repartiție

$F:\mathbb{R}\to[0,1] \qquad F(x) = P(X\le x) = \int_{-\infty}^{x} f(t)\,dt,\ \forall x\in\mathbb{R}$

f continuă $\Rightarrow F'(x) = f(x)$

$$P(X>a) = 1 - P(X\le a) = 1 - F(a) = \int_a^{+\infty} f(x)\,dx$$

Media

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\,dx$$

$\int_{-\infty}^{+\infty} |x| f(x)\,dx > +\infty \Rightarrow$ media nu există

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x)\,dx$$

Varianța

$$Var(X) = E[X^2] - E[X]^2 = \int_{-\infty}^{+\infty} x^2 f(x)\,dx - \left(\int_{-\infty}^{+\infty} x f(x)\,dx\right)^2$$

Momente de ordin k

$$E[X^k] = \int_{-\infty}^{+\infty} x^k f(x)\,dx$$

Variabile aleatoare repartizate uniform

$U \sim \mathcal{U}((a,b))$

$V = a + (b-a)U$

$\Rightarrow V \sim \mathcal{U}((a,b))$

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{altfel} \end{cases}$$

$$E[U] = \frac{a^2+ab+b^2}{3}$$

$$Var(U) = \frac{(b-a)^2}{12}$$

$$F(x) = \begin{cases} 0, & x \le a \\ \frac{x-a}{b-a}, & a < x < b \\ 0, & x \geq b \end{cases}$$

X v.a. cu fct. de repartiție F cont $+ \nearrow$:

- $U \sim \mathcal{U}((0,1)) \Rightarrow F^{-1}(U)$ are aceeași repartiție ca X
- $F(X) \sim \mathcal{U}((0,1))$

$\boxed{\text{T. fundamentală a simulării}}$

Variabile repartizate logistic

$$F(x) = \frac{e^x}{1+e^x}, \quad x \in \mathbb{R} \qquad U \sim \mathcal{U}((0,1))$$

$$\Rightarrow F^{-1}(U) = \ln\left(\frac{U}{1-U}\right) \sim \text{logistic}$$

Variabile repartizate exponențial

$X \sim Exp(\lambda)$

$f(x) = \lambda e^{-\lambda x}, \ x > 0$

$E[X] = \frac{1}{\lambda}$

$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$

$Var(X) = \frac{1}{\lambda^2}$

lipsa de memorie

$P(X \geq a) = e^{-\lambda a}$

Variabile repartizate normal

$X \sim N(\mu, \sigma^2)$

$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ x \in \mathbb{R}$

$E[X] = \mu$

$Var(X) = \sigma^2$

$P(|X-\mu| \leq \sigma) \simeq 0.68$

$P(|X-\mu| \leq 2\sigma) \simeq 0.95$

$P(|X-\mu| \leq 3\sigma) \simeq 0.997$

Variabile repartizate normal standard

$X \sim N(\underbrace{0}_{E[X]}, \underbrace{1}_{Var(X)})$

$f(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$F(x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$

$= 1 - \Phi(x)$

$Y \sim N(\mu, \sigma^2) = \mu + \sigma Z, \ Z \sim N(0,1)$

$Y \sim N(\mu, \sigma^2) \Rightarrow \frac{Y-\mu}{\sigma} \sim N(0,1) \leftarrow$ normalizare

REPARTIȚII COMUNE

V. a. discrete

$X, Y : \Omega \to \mathbb{R}$

$p_{X,Y}(x,y) = P(X=x, Y=y)$

$F_{X,Y}(x,y) = P(X \le x, Y \le y)$

$P((X,Y) \in A) = \sum_{(x,y)\in A} P(X=x, Y=y) = \sum_{(x,y)\in A} p_{X,Y}(x,y)$

V.a. continue

$X, Y : \Omega \to \mathbb{R}$

$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)\,du\,dv$

$f_{X,Y}(x,y) = \frac{d^2}{dx\,dy} F_{X,Y}(x,y)$

$P((X,Y) \in A) = \iint_{(x,y)\in A} f_{X,Y}(x,y)\,dx\,dy$

$A = [a,b] \times [c,d] \Rightarrow P((X,Y) \in A) = P(a \le X \le b, c \le Y \le d) =$

$= \int_{a}^{b}\int_{c}^{d} f_{X,Y}(x,y)\,dx\,dy$

REPARTIȚII MARGINALE

V.a. discrete

$p_X(x) = P(X=x) = \sum_{y} p_{X,Y}(x,y)$

$p_Y(y) = P(Y=y) = \sum_{x} p_{X,Y}(x,y)$

V.a. continue

$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dy$

$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx$

REPARTIȚII CONDIȚIONATE

V.a. discrete

$$p_{X/A}(x) = P(X=x/A) = \frac{P(\{X=x\}\cap A)}{P(A)}$$

$$\sum_x p_{X/A}(x) = 1$$

$$p_{X/Y}(x|y) = P(X=x/Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)} =$$

$$= \frac{p_X(x)\, p_{Y/X}(y/x)}{p_Y(y)} = \frac{p_X(x)\, p_{Y/X}(y/x)}{\sum_{x'} p_X(x')\, p_{Y/X}(y/x')}$$

Formula lui Bayes

$$p_{X/Y}(x|y) = p_Y(y)\cdot p_{X/Y}(x|y)$$

$$= p_X(x)\cdot p_{Y/X}(y/x)$$

$$p_X(x) = \sum_y p_Y(y)\, p_{X/Y}(x|y) \leftarrow \boxed{\text{Formula prob. totale}}$$

V.a. continue

$$P(x\in B/A) = \int_B f_{X/A}(x)\,dx$$

$$P(x\in B | x\in A) = \frac{P(x\in B, x\in A)}{P(x\in A)} = \frac{P(X\in A\cap B)}{P(X\in A)} = \frac{\int_{A\cap B} f_X(x)\,dx}{P(X\in A)}$$

$$f_{X/X\in A}(x) = \begin{cases} \frac{f_X(x)}{P(X\in A)}, & x\in A \\ 0, & \text{altfel} \end{cases} = \frac{f_X(x)}{P(X\in A)}\cdot \mathbb{1}_A(x)$$

$$f_X(x) = \sum_{i=1}^{n} f_{X/A_i}(x)\, P(A_i) \leftarrow \boxed{\text{Formula prob. totale}}$$

$$f_{X/Y}(x/y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{Y/X}(y/x)\, f_X(x)}{\int_{-\infty}^{+\infty} f_{Y/X}(y/x')\, f_X(x')\,dx'}$$

$\uparrow$ $\boxed{\text{Formula lui Bayes}}$

$$f_{X/Y}(x/y) = f_{X/Y}(x/y)\, f_Y(y) = f_{Y/X}(y/x)\, f_X(x)$$

$$P(X\in A / Y=y) = \int_A f_{X/Y}(x,y)\,dx$$

$f_{X,Y}(x,y) = f_X(x)\, f_Y(y) \Rightarrow$ independență $\underline{u}$

$\Leftrightarrow f_{X/Y}(x/y) = f_X(x)$

$f_{Y/X}(y/x) = f_Y(y)$

$\Rightarrow P(X\in A, Y\in B) = P(X\in A)\,P(Y\in B)$

$F_{X,Y}(x,y) = F_X(x)\, F_Y(y)$

V.a. hibride

X discret
Y discret $\Rightarrow$ $P(Y=y/X=x) = \dfrac{P(X=x/Y=y)\,P(Y=y)}{P(X=x)}$

$$P(X=x) = \sum_y P(X=x/Y=y)\,P(Y=y)$$

X discret
Y cont $\Rightarrow$ $f_{Y/X}(y/x) = \dfrac{P(X=x/Y=y)\, f_Y(y)}{P(X=x)}$

$$P(X=x) = \int P(X=x/Y=y)\, f_Y(y)\,dy$$

X cont
Y discret $\Rightarrow$ $P(Y=y/X=x) = \dfrac{f_{X/Y}(x/y)\,P(Y=y)}{f_X(x)}$

$$f_X(x) = \sum_y f_{X/Y}(x/y)\,P(Y=y)$$

$$P(A/X=x) = \frac{f_{X/A}(x)P(A)}{P(A)f_{X/A}(x)+P(A^c)f_{X/A^c}(x)}$$

$\left|\begin{array}{l} X \text{ cont} \\ Y \text{ cont} \end{array}\right. \Rightarrow f_{Y/X} = \dfrac{f_{X/Y}(x/y)\,f_Y(y)}{f_X(x)}$

$$f_X(x) = \int f_{X/Y}(x/y) f_Y(y)\, dy$$

MEDIA CONDIȚIONATĂ

V.a. discrete

$$E[X/A] = \sum_x x P(X=x/A)$$

$$E[X/Y=y] = \sum_x x P(X=x/Y=y) = \sum_x x\, p_{X/Y}(x/y)$$

V.a. continue

$$E[X/A] = \int_{-\infty}^{+\infty} x f_{X/A}(x)\, dx$$

$$E[X/Y=y] = \int_{-\infty}^{+\infty} x f_{X/Y}(x/y)\, dx$$

INEGALITĂȚI

Cauchy-Schwarz

$$E[XY] \leq \sqrt{E[X^2]E[Y^2]}$$

Jensen

g. convexă $\Rightarrow E[g(x)] \geq g(E[x])$

g. concavă $\Rightarrow E[g(x)] \leq g(E[x])$

g. afină $\Rightarrow E[g(x)] = g(E[x])$

Markov

$$P(X \geq a) \leq \frac{E[X]}{a}, \quad \forall X \geq 0, a > 0$$

Chebyshev

$$P(|X-\mu| \geq a) \leq \frac{\sigma^2}{a^2}, \quad \forall a > 0$$

$$\mu = E[X] \quad \sigma^2 = Var(X)$$

Chernoff

$$P(X \geq a) \leq \frac{E[e^{tX}]}{e^{ta}}, \quad \forall a > 0, t > 0$$

LEGEA NR. MARI

$X_1, X_2, \dots, X_n$ v.a. $\perp\!\!\!\perp$ și identic distribuite (iid)

$E[X_1] = \mu \quad Var(X_1) = \sigma^2$

$\overline{X_n} = \frac{X_1 + \dots + X_n}{n} \leftarrow$ media eșantionului

$E[\overline{X_n}] = \mu \quad Var(\overline{X_n}) = \frac{\sigma^2}{n}$

Legea tare a nr. mari

$$P(\overline{X_n} \to \mu) = 1, \text{ i.e. } \overline{X_n}(\omega) \xrightarrow[n \to +\infty]{} \mu, \ \forall \omega \in \Omega_0, P(\Omega_0) = 1$$

Legea slabă a nr. mari

$$P(|\overline{X_n} - \mu| > \varepsilon) \xrightarrow[n \to +\infty]{} 0, \quad \forall \varepsilon > 0$$

$(X_n)_n, \ X_n \xrightarrow{P} X \Leftrightarrow P(|X_n - X| > \varepsilon) \xrightarrow[n \to +\infty]{} 0, \ \forall \varepsilon > 0$

convergență în probabilitate

Teorema limită centrală

$$Z_n = \frac{(X_1+\dots+X_n) - E[X_1+\dots+X_n]}{\sqrt{Var(X_1+\dots+X_n)}}$$

$$S_n = X_1+\dots+X_n \Rightarrow Z_n = \frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

$\hookrightarrow$ var normalizată $\sim N(0,1)$

$E[Z_n]=0 \quad Var(Z_n)=1$

$$\lim_{n\to+\infty} P(Z_n \le z) = \Phi(z), \ \forall z \in \mathbb{R}$$

$$P(S_n \le c) \simeq P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le \frac{c-n\mu}{\sigma\sqrt{n}}\right) = P\left(Z_n \le \frac{c-n\mu}{\sigma\sqrt{n}}\right) =$$

$$= \Phi\left(\frac{c-n\mu}{\sigma\sqrt{n}}\right)$$

Aproximarea de Moivre-Laplace a binomului

$S_n \sim B(n,p)$, $S_n = X_1+\dots+X_n$, $X_i \sim B(p)$ $\perp\!\!\!\perp$

$\mu = E[X_i] = p$

$\sigma^2 = Var(X_i) = p(1-p)$

$$P(k \le S_n \le l) = \Phi\left(\frac{l+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)$$

cu precizarea că $P(S_n = k) \simeq P(k-\frac{1}{2} \le S_n \le k+\frac{1}{2})$

statistică

STATISTICĂ

Media

$$\overline{X}_n = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$\overline{T}_n^{n}$

$\leftarrow$ media empirică / a eșantionului

Varianța

$$V_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X}_n)^2$$ $\leftarrow$ varianța empirică

$$S_n = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X}_n)^2$$ $\leftarrow$ varianța eșantionului

Momente de ordin r

$$M_r = \frac{1}{n}\sum_{i=1}^{n} x_i^r$$ $\leftarrow$ moment empiric de ord. r

Momente centrate de ordin r

$$M_r' = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X}_n)^r$$ $\leftarrow$ moment empiric centrat de ord. r

Covarianța

$$C_n = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X}_n)(Y_i - \overline{Y}_n)$$ $\leftarrow$ covarianța empirică

$$\rho_n = \frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)(Y_i - \overline{Y}_n)}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 \sum_{i=1}^{n}(Y_i - \overline{Y}_n)^2}}$$ $\leftarrow$ coeficientul de corelație liniară empiric

POPULAȚIA NORMALĂ

$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$

$\overline{X}_n \sim N(\mu, \sigma^2/n)$

$\overline{X}_n \perp\!\!\!\perp S_n^2$

$$\frac{(n-1)S_n^2}{\sigma_n^2} \sim \chi^2(n-1)$$

$$\frac{\overline{X}_n - \mu}{S_n/\sqrt{n}} \sim t(n-1)$$

$$\frac{S_{n_1}^2}{S_{n_2}^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1-1, n_2-1)$$

$$Var(S_n^2) = \frac{2\sigma^4}{n-1}$$

Repartitia hi-patrat

$Z \sim \chi^2(v)$

$\hookrightarrow$ gr. de libertate , $Z = Z_1^2 + \dots + Z_n^2$

cu $Z_1, \dots, Z_n$ iid $\sim N(0,1)$

$E[Z] = v$

$Var(Z) = 2v$

$Z_1 \sim \chi^2(v_1) \perp\!\!\!\perp Z_2 \sim \chi^2(v_2) \Rightarrow Z_1 + Z_2 \sim \chi^2(v_1 + v_2)$

Repartitia t-Student

$$Z \sim N(0,1) \perp\!\!\!\perp Q \sim \chi^2(v) \Rightarrow T = \frac{Z}{\sqrt{\frac{Q}{v}}} \sim t(v)$$

Repartitia Fisher-Snedecor

$$U \sim \chi^2(v_1) \perp\!\!\!\perp V \sim \chi^2(v_2) \Rightarrow F = \frac{U/v_1}{V/v_2} \sim F(v_1, v_2)$$

$$1/F \sim F(v_2, v_1)$$

ESTIMATORI

Nedeplasarea

$b_\theta(\hat{\theta}_n) = E_\theta[\hat{\theta}_n] - \theta$ ← deplasarea estimatorului

$\hat{\theta}_n$ nedeplasat $\Leftrightarrow b_\theta(\hat{\theta}_n) = 0 \Rightarrow E_\theta(\hat{\theta}_n) = \theta, \ \forall \theta$

$g: \mathbb{R} \to \mathbb{R}$ $\quad g(\hat{\theta}_n)$ nu e nedeplasat pt. $g(\theta)$

Consistența

$\hat{\theta}_n$ consistent $\Leftrightarrow \hat{\theta}_n \xrightarrow{P} \theta$

$\overline{X}_n \xrightarrow{P} \mu$

$S_n \xrightarrow{P} \sigma^2$

Estimare punctuală

$X_1, X_2, \dots, X_n \sim f_\theta, \ \theta \in \Theta \subseteq \mathbb{R}^k$

$\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \dots, X_n)$ ← estimator punctual

↳ statistică

Eroarea pătratică medie

$X_1, X_2, \dots, X_n \sim f_\theta$

$MSE_\theta(\hat{\theta}_n) = E_\theta[(\hat{\theta}_n - \theta)^2]$

$= Var_\theta(\hat{\theta}_n) - b_\theta(\hat{\theta}_n)$ ← eroarea pătratică medie

$MSE_\theta(\hat{\theta}_n) \xrightarrow[n \to \infty]{} 0 \Rightarrow \hat{\theta}_n$ consistent

CONSTRUIREA ESTIMATORILOR

Metoda momentelor

$X=(X_1,\dots,X_n)\sim f_\theta$

$$E[X^j]=\int_{-\infty}^{+\infty} x^j f_\theta(x)\,dx\ ,\ j=\overline{1,k}$$

momente de ord j

Rez. sist.:
$$\begin{cases}\frac{1}{n}\sum_{i=1}^{n} X_i = E[X^1]=g_1(\theta_1,\dots,\theta_k)\\ \frac{1}{n}\sum_{i=1}^{n} X_i = E[X^2]=g_2(\theta_1,\dots,\theta_k)\\ \vdots\\ \frac{1}{n}\sum_{i=1}^{n} X_i = E[X^k]=g_k(\theta_1,\dots,\theta_k)\end{cases}$$

$\Rightarrow$ estimatorul

Metoda verosimilității maxime

$X_1,\dots,X_n\sim f_\theta(x)$ $\quad f_\theta(x_1,\dots,x_n)=\prod_{i=1}^{n} f_\theta(x_i)$

↳ densitatea(masa) comună

$$L(\theta\,|\,x_1,x_2,\dots,x_n)=f_\theta(x_1,\dots,x_n)=\prod_{i=1}^{n} f_\theta(x_i)$$

funcția de verosimilitate

$$\ell(\theta\,|\,x_1,x_2,\dots,x_n)=\log L(\theta\,|\,x_1,\dots,x_n)$$

logaritmul fct. de verosimilitate

$$\hat{\theta}_n=\arg\max_{\theta\in\Theta} L(\theta\,|\,x)=\arg\max_{\theta\in\Theta} \ell(\theta\,|\,x)$$

estimator de verosimilitate maximă

$\frac{\partial}{\partial\theta_i}L(\theta\,|\,x)=0$, $i=\overline{1,k}$ $\xrightarrow{\text{rez. sist}}$ posibili candidați pt. estimatorul de verosimilitate maximă (MLE)

· Pop normală $\Rightarrow$ $\overline{X}_n$ MLE pt. μ

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x}_n)^2 \quad \text{MLE pt. } \sigma^2$$

MLE nu este unic

INTERVALE DE ÎNCREDERE

$X_1,\dots,X_n \sim f_\theta$, $\alpha\in(0,1)$, $A_\alpha, B_\alpha : \mathbb{R}^n \to \mathbb{R}$

$A_\alpha(x_1,\dots,x_n) \le B_\alpha(x_1,\dots,x_n) \quad \forall (x_1,\dots,x_n)\in\mathbb{R}^n$

$[A_\alpha(x_1,\dots,x_n), B_\alpha(x_1,\dots,x_n)] = IC^{1-\alpha}(\theta)$

cu $P([A_\alpha(x_1,\dots,x_n), B_\alpha(x_1,\dots,x_n)] \ni \theta) \ge 1-\alpha$

Interval de încredere cu coeficientul de încredere $1-\alpha$

Metoda pivotului

$g : \mathbb{R}\times\Theta \to \mathbb{R}$ $\leftarrow$ Funcție pivot

repartiția $g(x_1,\dots,x_n,\theta)$ nu depinde de θ

$u_1 \le g(x_1,\dots,x_n) \le u_2 \xRightarrow{\text{rez}} a(x_1,\dots,x_n) \le \theta \le b(x_1,\dots,x_n)$

$\forall u_1,u_2\in\mathbb{R} \quad u_1\le u_2$

$x_p = F^{-1}(p) = \inf\{x \mid F(x)\ge p\}$, $p\in(0,1)$ $\leftarrow$ cuantilă de ord p

$P(X\le x_p) = p \qquad P(X\ge x_p) = (1-p)$

Intervale de încredere pt. populația normală

$X_1, \dots, X_n \sim N(\mu, \sigma^2)$

μ necunoscut
σ^2 cunoscut $\Rightarrow IC^{1-\alpha}(\mu) = [\bar{X}_n - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X}_n + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}]$

↳ cuantilă de ord $1-\alpha/2$ din $N(0,1)$

μ necunoscut
σ^2 necunoscut $\Rightarrow IC^{1-\alpha}(\mu) = [\bar{X}_n - t_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}]$

↳ cuantilă de ord $1-\alpha/2$ din $t(v)$