CURSUL 9 (6 decembrie 2021)

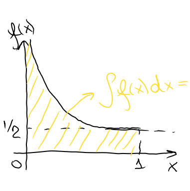
Analogie v.a. discrete VS v.a. continue

$\mathbb{P}(X=x)$ fct. de masă	$\mathbb{P}(X=x)=0$
$\mathbb{P}(X\in A)=\sum_{x\in A}\mathbb{P}(X=x)$	$\mathbb{P}(X\in A)=\int_A f(x)\,dx$
$\sum$	$\int$
$f(x)$ (PMF) fct. de masă	$f(x)\,dx$ densitatea

Obs: Densitatea de repartiție nu este o probabilitate

Exp: (O densitate poate lua valori oricât de mari)

V.a. X cu densitatea $f(x)=\begin{cases}\frac{1}{2\sqrt{x}}, & 0<x\le 1\\ 0, & \text{altfel}\end{cases}$



f este densitate

a) $f(x)\ge 0$

b) $\int_{-\infty}^{\infty} f(x)dx=\int_0^1\frac{1}{2\sqrt{x}}dx=$

$=(\sqrt{x})\Big|_0^1=1$

Pentru $x\to 0$, $f(x)\to+\infty$

Obs: Suportul v.a. X este $\{x\mid f(x)$

Funcția de rep. pt o v.a X este $F:\mathbb{R}\to[0,1]$

$F(x)=\mathbb{P}(X\le x)$, $\forall x\in\mathbb{R}$

$=\mathbb{P}(X\in\underbrace{(-\infty,x]}_{A})$

$F(x)=\mathbb{P}(X\in(-\infty,x])$ X v.a continuă cu densitatea f

$=\int_{(-\infty,x]}f(t)dt=\int_{-\infty}^{x}f(t)dt$, $\forall x\in\mathbb{R}$

Proprietățile funcției de repartiție sunt:

a) F este crescătoare

b) F este continuă la dreapta

c) $\lim_{x\to-\infty}F(x)=0$ și $\lim_{x\to\infty}F(x)=1$

Obs: În cazul în care X este o v.a discretă cu fct. de masă $f(x)=\mathbb{P}(X=x)$ atunci

$F(x)=\mathbb{P}(X=x)=\sum_{\substack{y\le x\\ y\in\operatorname{Im}X}}f(y)$

În cazul în care X este o v.a. continuă cu densitatea f

$F(x)=\int_{-\infty}^{x}f(t)dt$

Din teorema fundamentală a analizei dacă f este continuă în x_0 atunci F este derivabilă în x_0 și $F'(x_0)=f(x_0)$

Dacă f este continuă atunci $F'(x)=f(x)$, $\forall x$.

Dacă știm F și vrem să găsim f atunci folosim $F'(x)=f(x)$, $\forall x$

Dacă știm f și vrem să găsim F $\Rightarrow$

$F(x)=\int_{-\infty}^{x}f(t)dt$

Exp: Fie X o v.a continuă cu densitatea f def prin $f(x)=\frac{e^x}{(1+e^x)^2}$, $x\in\mathbb{R}$ (logistic)

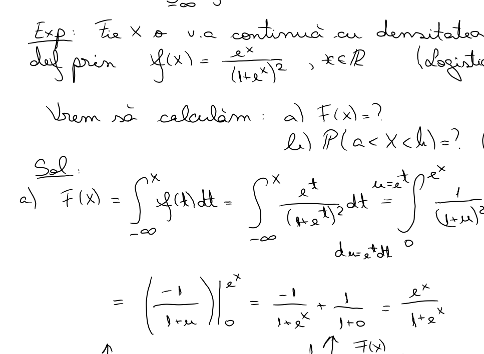
Vrem să calculăm: a) $F(x)=?$

b) $\mathbb{P}(a<X<b)=?$ $\left(\begin{matrix}a=-1\\ b=2\end{matrix}\right)$

Sol:

a) $F(x)=\int_{-\infty}^{x}f(t)dt=\int_{-\infty}^{x}\frac{e^t}{(1+e^t)^2}dt\overset{u=e^t}{\underset{du=e^tdt}{=}}\int_0^{e^x}\frac{1}{(1+u)^2}du=$

$=\left(\frac{-1}{1+u}\right)\Big|_0^{e^x}=\frac{-1}{1+e^x}+\frac{1}{1+0}=\frac{e^x}{1+e^x}$



b) $\mathbb{P}(a<X<b)=\mathbb{P}(X<b)-\mathbb{P}(X<a)$

$=\mathbb{P}(X\le b)-\mathbb{P}(X\le a)$

$=F(b)-F(a)=$

$=\frac{e^b}{1+e^b}-\frac{e^a}{1+e^a}$

$\underset{X\in(a,b)}{\mathbb{P}(a<X<b)}=\int_a^b f(x)dx=\int_a^b\frac{e^x}{(1+e^x)^2}dx$ s.v. $u=1+e^x$, $du=e^xdx$

$=\int_{1+e^a}^{1+e^b}\frac{1}{u^2}du$

Media și momentele v.a. continue

Reamintim că dacă X este o v.a discretă atunci

a) $E[X]=\sum_x xf(x)=\sum_x x\,\mathbb{P}(X=x)$

(atunci când $\sum_x x\mathbb{P}(X=x)<\infty$)

b) $E[X^k]=\sum_x x^k\,\mathbb{P}(X=x)=\sum_x x^kf(x)$ (moment de ordin k)

c) $E[(X-a)^k]=\sum_x(x-a)^k\mathbb{P}(X=x)=$

$=\sum_x(x-a)^kf(x)$ (moment centrat în a de ordin k)

d) $a=E[X]$ atunci

$E[(X-E[X])^k]=\sum_x(x-E[x])^k\mathbb{P}(X=x)$

(moment centrat de ordin k) $=\sum_x(x-E[x])^kf(x)$

e) $E[g(X)]=\sum_x g(x)\,\mathbb{P}(X=x)=\sum_x g(x)f(x)$

(media unei funcții de v.a. X)

Def: Fie X o v.a. continuă cu densitatea de repartiție f. Media v.a X este definită prin:

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

dacă $E[X]=\int_{-\infty}^{\infty}|x|f(x)dx<\infty$. În caz contrar spunem că media nu există.

În mod similar avem:

a) Momentul de ordin k al v.a. X este:

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

b) Mom. centrat în a de ordin k al v.a X este:

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

c) Mom. centrat de ordin k ($a=E[X]$) al v.a. X

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

k=2 avem varianța:

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

$=E[X^2]-E[X]^2$

d) Media unei funcții de v.a X:

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

Proprietățile mediei și varianței de la v.a discrete rămân valabile și în cazul v.a. continue:

i) Dacă $X=c$, c o const. atunci $E[X]=c$, $Var(X)=0$

ii) Dacă $X\ge 0$ atunci $E[X]\ge 0$

iii) Dacă $X\ge Y$ atunci $E[X]\ge E[Y]$

iv) $E[aX+bY]=aE[X]+bE[Y]$, $\forall a,b\in\mathbb{R}$

v) $Var(X)\ge 0$

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

vii) Dacă X și Y sunt indep. $E[XY]=E[X]E[Y]$

viii) Dacă X și Y sunt indep. $Var(X+Y)=Var(X)+Var(Y)$

Exemple de repartiții continue

1) Repartiția uniformă

Def: O v.a U este repartizată uniform pe intervalul $[a,b]$ dacă densitatea de repartiție $f(x)$ este constantă pe $[a,b]$

$f(x)=c$, $x\in[a,b]$

Cum f este dens. de rep $\Rightarrow c\ge 0$ și $\int_a^b f(x)dx=1$

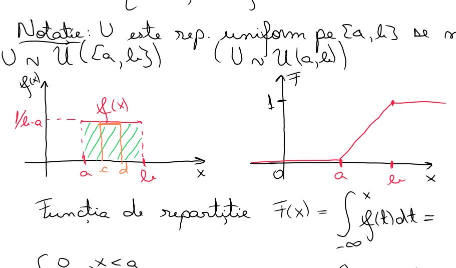
$\left(\int_{-\infty}^{+\infty}f(x)dx=\int_a^b f(x)dx\right)$

$\int_a^b c\,dx=1\Rightarrow$ $c=1/b-a$

Dens. rep. uniform pe $[a,b]$

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

Notație: U este rep. uniform pe $[a,b]$ se notează $U\sim\mathcal{U}([a,b])$ ($U\sim\mathcal{U}(a,b)$)



Funcția de repartiție $F(x)=\int_{-\infty}^{x}f(t)dt=$

$=\begin{cases}0, & x<a\\ \int_a^x f(t)dt, & a\le x\le b\\ \underbrace{\int_a^b f(t)dt}_{1}, & x>b\end{cases}\Rightarrow F(x)=\begin{cases}0, & x<0\\ \int_a^x\frac{1}{b-a}dt=\frac{x-a}{b-a}, & a\le x\le b\\ 1, & x>b\end{cases}$

Dacă $(c,d)\subseteq(a,b)$ și $U\sim\mathcal{U}([a,b])$ atunci

$\mathbb{P}(U\in(c,d))=\int_c^d f(x)dx=\int_c^d\frac{1}{b-a}dx=\frac{d-c}{b-a}$

prob. ca U să se afle într-un subinterval este proporțională cu lungimea intervalului.

Dacă $U\sim\mathcal{U}([0,1])$ atunci $V=a+(b-a)U\sim\mathcal{U}([a,b])$

translatare scalare trans. de locație și scală

$F_V(x)=\mathbb{P}(V\le x)$

$=\mathbb{P}(a+(b-a)U\le x)$

$=\mathbb{P}\left(U\le\frac{x-a}{b-a}\right)=\frac{x-a}{b-a}$, $x\in(a,b)$

Obs: Dacă $U\sim\mathcal{U}([0,1])$ atunci $f_U(x)=1$, $\forall x\in[0,1]$,

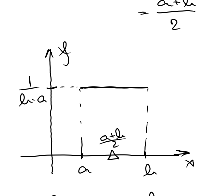
iar $F_U(x)=\begin{cases}0, & x<0\\ x, & x\in[0,1]\\ 1, & x>1\end{cases}$

Media lui $V\sim\mathcal{U}[a,b]$:

$E[V]=\int_{-\infty}^{+\infty}xf_V(x)dx=\int_{-\infty}^{+\infty}x\frac{1}{b-a}\mathbb{1}_{[a,b]}(x)dx=$

$=\int_a^b\frac{x}{b-a}dx=\frac{1}{b-a}x^2/2\Big|_a^b=\frac{1}{2(b-a)}(b^2-a^2)=$

$=\frac{a+b}{2}$



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

$E[V]=a+(b-a)E[U]$

$E[U]=\int_0^1x\,dx=1/2$

$E[V]=a+\frac{b-a}{2}=\frac{a+b}{2}$

Varianța lui $V\sim\mathcal{U}[a,b]$

$Var(V)=E[V^2]-E[V]^2$

$E[V]^2=\int_{-\infty}^{+\infty}x^2f_V(x)dx=\int_{-\infty}^{+\infty}x^2\frac{1}{b-a}\mathbb{1}_{[a,b]}(x)dx$

$=\frac{1}{b-a}\int_a^b x^2dx=\frac{1}{3(b-a)}(b^3-a^3)=$

$=\frac{(b-a)(b^2+ab+a^2)}{3(b-a)}=\frac{b^2+ab+a^2}{3}$

$Var(V)=\frac{b^2+ab+a^2}{3}-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}$

Folosind transformarea $V=a+(b-a)U$, $U\sim\mathcal{U}[0,1]$

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

$=(b-a)^2Var(U)$

$Var(U)=E[U^2]-E[U]^2=\int_0^1x^2dx-1/4=\frac{1}{3}-\frac{1}{4}=\frac{1}{12}$

Temă: Calculați momentele de ordin superior!

Universalitatea repartiției uniforme

(T) Fie X o v.a. cu funcția de repartiție F continuă și strict crescătoare (este bijectivă). Atunci:

a) Dacă $U\sim\mathcal{U}([0,1])$ atunci $F^{-1}(U)$ are aceeași repartiție ca X

b) $F(X)\sim\mathcal{U}([0,1])$

Dem: a) Cum f este continuă și strict crescătoare $\Rightarrow$ F este bijectivă deci există F^{-1} (tot crescătoare)

Fie $Y=F^{-1}(U)$

$\mathbb{P}(Y\le x)=\mathbb{P}(F^{-1}(U)\le x)=\mathbb{P}(U\le F(x))=F(x)$

Ev. $\{F^{-1}(U)\le x\}=\{U\le F(x)\}$ (aplic $F\uparrow$)

Astfel $F^{-1}(U)$ și X sunt repartizate la fel

b) Fie $Y=F(X)$, $F:\mathbb{R}\to[0,1]\Rightarrow Y\in(0,1)$

$\mathbb{P}(Y\le y)=\mathbb{P}(F(X)\le y)$

$=\mathbb{P}(F^{-1}(F(X))\le F^{-1}(y))=\mathbb{P}(X\le F^{-1}(y))$ aplic $F^{-1}(\uparrow)$

$=F(F^{-1}(y))=y$, $y\in(0,1)$

Pt $y<0$ avem $\mathbb{P}(Y\le y)=0$; $y>1$ avem $\mathbb{P}(Y\le y)=1$ $\Rightarrow Y\sim\mathcal{U}([0,1])$

Obs: Rezultatul teoremei de universalitate este valabil în general:

$F^{-1}(u)=\inf\{x\mid F(x)\ge u\}$ ← Funcție cuantilă

Dacă F este continuă și crescătoare $\Rightarrow F^{-1}$ este chiar inversa lui F

Exp: (Repartiția logistică)

$f(x)=\frac{e^x}{(1+e^x)^2}$, $x\in\mathbb{R}$, $F(x)=\frac{e^x}{1+e^x}$, $x\in\mathbb{R}$

Dacă $U\sim\mathcal{U}(0,1)$

$F^{-1}(u)=?$ Rezolvăm ecuația $F(x)=u$

$\Rightarrow\frac{e^x}{1+e^x}=u$, $u\in(0,1)\Leftrightarrow e^x=\frac{u}{1-u}\Rightarrow\ldots$

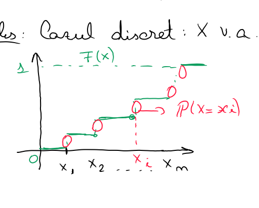
$F^{-1}(u)=\ln\left(\frac{u}{1-u}\right)$

Atunci $F^{-1}(U)=\ln\left(\frac{U}{1-U}\right)$ este repartizată logistic

Dacă vreau să simulăm n obs. dintr-o rep. logistică

- generăm n obs. uniforme pe (0,1) (runif(n))
- calculăm $\ln\left(\frac{U}{1-U}\right)$ (histograma)

Obs: Cazul discret: X v.a. discretă cu fct. de rep. F



$X\sim f$

$X(\omega)$ – o observație

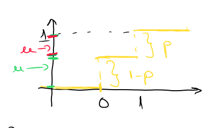
$\omega_1,\omega_2,\ldots\omega_n$

$\underbrace{X(\omega_1),X(\omega_2),\ldots X(\omega_n)}_{n\text{ obs.}}$

Exp: $X\sim\mathcal{B}(p)$, $X\in\{0,1\}$

$\mathbb{P}(X=1)=p$

$F(x)=\begin{cases}0, & x<0\\ 1-p, & 0\le x<1\\ 1, & x\ge 1\end{cases}$



$F^{-1}(u)=\inf\{x\mid F(x)\ge u\}$

$F^{-1}(u)=\begin{cases}0, & u\le 1-p\\ 1, & u>1-p\end{cases}$

Generăm $U\sim\mathcal{U}(0,1)$ și considerăm $F^{-1}(U)=\begin{cases}0, & U\le 1-p\\ 1, & U>1-p\end{cases}$

sau

$F^{-1}(U)=\begin{cases}0, & p\le 1-U\\ 1, & p>1-U\end{cases}$

Pt a genera o observație Bernoulli de param. p:

- generăm $U\sim\mathcal{U}([0,1])$
- dacă $U<p$ atunci $X=1$ altfel $X=0$

2) Repartiția exponențială

Def: O variabilă aleatoare X este rep. exponențial de param. $\lambda>0$, notăm $X\sim Exp(\lambda)$ dacă densitatea de repartiție a lui X este

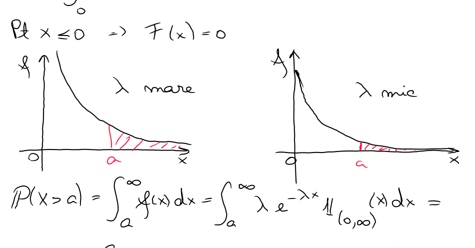
$\boxed{f(x)=\lambda e^{-\lambda x},\ x>0}$

Funcția de repartiție este:

$F(x)=\int_{-\infty}^{x}f(t)dt=\int_{-\infty}^{+\infty}\lambda e^{-\lambda t}\mathbb{1}_{(0,\infty)}(t)dt=$

$=\int_0^x\lambda e^{-\lambda t}dt=(-e^{-\lambda t})\Big|_0^x=1-e^{-\lambda x}$, $x>0$

Pt $x\le 0\Rightarrow F(x)=0$



$\mathbb{P}(X>a)=\int_a^{\infty}f(x)dx=\int_a^{\infty}\lambda e^{-\lambda x}\mathbb{1}_{(0,\infty)}(x)dx=$

$=\int_a^{\infty}\lambda e^{-\lambda x}dx=e^{-\lambda a}$ ($\mathbb{P}(X>a)$ s.n. funcția de supraviețuire)

Probabilitatea ca v.a X să depășească pragul a scade exponențial.

Media v.a $Exp(\lambda)$:

Dacă $X\sim Exp(\lambda)$ atunci

$E[X]=\int_{-\infty}^{+\infty}xf(x)dx=\int_{-\infty}^{+\infty}x\lambda e^{-\lambda x}\mathbb{1}_{(0,\infty)}(x)dx=$

$=\int_0^{\infty}x\lambda e^{-\lambda x}dx=\int_0^{\infty}x(-e^{-\lambda x})'dx=$

$=-xe^{-\lambda x}\Big|_0^{\infty}-\int_0^{\infty}x'(-e^{-\lambda x})dx=$

$=0+\int_0^{\infty}e^{-\lambda x}dx=\frac{1}{\lambda}$

Varianța v.a. $Exp(\lambda)$

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

$E[X^2]=\int_{-\infty}^{+\infty}x^2f(x)dx=\int_0^{\infty}x^2\overbrace{\lambda e^{-\lambda x}}^{(-e^{-\lambda x})'}dx=$

$=\underbrace{-x^2e^{-\lambda x}\Big|_0^{\infty}}_{=0}+\int_0^{\infty}2xe^{-\lambda x}dx=$

$=\frac{2}{\lambda}\underbrace{\int_0^{\infty}x\lambda e^{-\lambda x}dx}_{E[X]}=\frac{2}{\lambda^2}$

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

(P) (Lipsa de memorie)

Spunem că v.a X verifică prop. lipsei de memorie dacă

$\mathbb{P}(X\ge s+t\mid X\ge s)=\mathbb{P}(X\ge t)$, $\forall s,t\ge 0$

a) V.a. repartizate $Exp(\lambda)$ admit proprietatea lipsei de memorie

b) Dacă X este o v.a. continuă, pozitivă care admite prop lipsei de memorie atunci X este repartizată exponențial.

Dem: a)

$\mathbb{P}(X\ge s+t\mid X\ge s)=\frac{\mathbb{P}(X\ge s+t,X\ge s)}{\mathbb{P}(X\ge s)}=$

$=\frac{\mathbb{P}(X\ge s+t)}{1-F(s)}=\frac{1-F(s+t)}{1-F(s)}=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}=$

$=e^{-\lambda t}=\mathbb{P}(X\ge t)$

b) temă? (Ec. funcțională Cauchy: $h(s+t)=h(s)h(t)$, $\forall s,t>0$) atunci $h(s)=e^{-s\cdot k}$