CURSUL 13 (3 ianuarie 2022)

Media condiționată (cont)

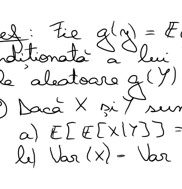
Fie X o v.a. continuă și A un eveniment a.î. $\mathbb{P}(A)>0$. Definim media condiționată a lui X la A, $\mathbb{E}[X|A]$, prin:

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

În particular, dacă $A=\{Y=y\}$, unde Y este o v.a. atunci

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

Formula probab. totale aplicată mediei: Fie $A_1, A_2, \dots A_m$ o partiție a lui Ω cu $\mathbb{P}(A_i)>0, \forall i$ și X o v.a. cont.

Ω 

$$\mathbb{E}[X] = \sum_{i=1}^{m} \mathbb{E}[X|A_i]\,\mathbb{P}(A_i)$$

FPT: $f_X(x) = \sum_{i=1}^{m} f_{X|A_i}(x)\,\mathbb{P}(A_i)$ (înmulțim cu x și integrăm)

Dacă Y este o v.a. continuă atunci

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} \mathbb{E}[X|Y=y]\, f_Y(y)\,dy$$

Def: Fie $g(y) = \mathbb{E}[X|Y=y]$. Atunci media condiționată a lui X la Y, $\mathbb{E}[X|Y]$, este variabila aleatoare $g(Y)$.

(P) Dacă X și Y sunt două v.a. atunci

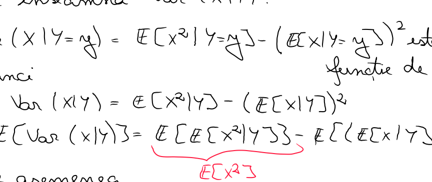
a) $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$

b) $Var(X) = Var(\mathbb{E}[X|Y]) + \mathbb{E}[Var(X|Y)]$

Interpretarea geometrică a mediei condiționate:

$\mathbb{E}[X|Y] = g(Y)$ - este acea funcție de Y care minimizează eroarea pătratică medie: $\mathbb{E}[(X-g(Y))^2]$

$$\mathbb{E}[X|Y] = \underset{h}{\operatorname{argmin}}\ \mathbb{E}[(X-h(Y))^2]$$



Exp: (Regresie liniară simplă)

Y - variabilă răspuns (dependentă)

X - variabilă explicativă / covariabilă / variabilă predictor

$$\mathbb{E}[Y|X] = a + bX$$

Relația de mai sus este echivalentă cu

$$Y = a + bX + \varepsilon$$

ε este termen eroare care verifică $\mathbb{E}[\varepsilon|X]=0$.

De ce este pct b) adevărat?

Ce înseamnă $Var(X|Y)$?

$Var(X|Y=y) = \mathbb{E}[X^2|Y=y] - (\mathbb{E}[X|Y=y])^2$ este o funcție de y

atunci

$$Var(X|Y) = \mathbb{E}[X^2|Y] - (\mathbb{E}[X|Y])^2$$

$$\mathbb{E}[Var(X|Y)] = \underbrace{\mathbb{E}[\mathbb{E}[X^2|Y]]}_{\mathbb{E}[X^2]} - \mathbb{E}[(\mathbb{E}[X|Y])^2] \quad (1)$$

De asemenea

$$Var(\mathbb{E}[X|Y]) = \mathbb{E}[\mathbb{E}[X|Y]^2] - \mathbb{E}^2[\mathbb{E}[X|Y]]$$
$$= \mathbb{E}[\mathbb{E}[X|Y]^2] - \mathbb{E}^2[X] \quad (2)$$

Adunând (1) cu (2) avem:

$$\mathbb{E}[Var(X|Y)] + Var(\mathbb{E}[X|Y]) = (\mathbb{E}[X^2] - \mathbb{E}[\mathbb{E}[X|Y]^2]) + (\mathbb{E}[\mathbb{E}[X|Y]^2] - \mathbb{E}[X]^2)$$
$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = Var(X)$$

Exp: În magazin intră N clienți, N este o v.a. $\mathbb{E}[N]<\infty$. Fiecare client cheltuie $X_1, X_2, \dots$ sume de bani. Vom presupune că $X_1, X_2, \dots$ sunt v.a independente de medie μ și varianță σ^2 și că N este indep. de $X_1, X_2, \dots$.

Vrem să determinăm suma totală medie cheltuită în magazin.

$T = \sum_{i=1}^{N} X_i$ - suma totală cheltuită

Ziua i: ω_i atunci $T(\omega_1) = \sum_{i=1}^{N(\omega_1)} X_i(\omega_1)$

$N(\omega_1) = 10$, $N(\omega_2) = 20$

$$\mathbb{E}[T] = \mathbb{E}\left[\sum_{i=1}^{N} X_i\right] = \mathbb{E}[\mathbb{E}[T|N]]$$

Cum $\mathbb{E}[\mathbb{E}[T|N]] = \sum_n \mathbb{E}[T|N=n]\,\mathbb{P}(N=n)$

iar $\mathbb{E}[T|N=n] = \mathbb{E}\left[\sum_{i=1}^{N} X_i \,|\, N=n\right] =$

$$= \mathbb{E}\left[\sum_{i=1}^{n} X_i \,|\, N=n\right] = \sum_{i=1}^{n} \mathbb{E}[X_i|N=n] =$$

$$= \sum_{i=1}^{n} \mathbb{E}[X_i] = n\mu$$

Găsim că

$$\mathbb{E}[T] = \mathbb{E}[\mathbb{E}[T|N]] = \sum_n n\mu\,\mathbb{P}(N=n)$$
$$= \mu \sum_n n\,\mathbb{P}(N=n) = \mu\,\mathbb{E}[N]$$

astfel

$$\mathbb{E}[T] = \mathbb{E}[X_1]\,\mathbb{E}[N]$$

Pentru varianță găsim că:

$$Var(T) = \mathbb{E}[N]\,Var(X_1) + \underbrace{(\mathbb{E}[X_1])^2\,Var(N)}_{\text{apare datorită lui } N}$$

(avem mai multă variabilitate)

Covarianța și corelația v.a.

Def: Fie X și Y două v.a. Se numește covarianța dintre X și Y

$$\boxed{Cov(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]}$$

Obs: Intuitiv covarianța a două v.a măsoară tendința celor două v.a. de a crește simultan sau de a scădea simultan.

În particular, pt $X=Y$ avem:

$$Cov(X,X) = \mathbb{E}[(X-\mathbb{E}[X])^2] = Var(X)$$

Obs: Avem că

$$\boxed{Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]}$$
$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Def: Spunem că două v.a X și Y sunt necorelate dacă $Cov(X,Y)=0$.

Obs: Dacă $X \perp Y$ (independente) atunci

$$\mathbb{E}[XY] = \mathbb{E}[X]\cdot\mathbb{E}[Y]$$

prin urmare

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

Astfel, independența implică necorelare.

Exp: (Necorelarea nu implică independența)

$$\left.\begin{array}{l} X \sim N(0,1) \\ Y = X^2 \end{array}\right\} \Rightarrow \begin{array}{l} \mathbb{E}[X]\cdot\mathbb{E}[Y] = 0 \\ \mathbb{E}[XY] = \mathbb{E}[X^3] = \int x^3 \varphi(x)\,dx = 0 \end{array}$$

Astfel, $Cov(X,Y)=0 \Rightarrow X$ și Y sunt necorelate.

Dar X și Y nu sunt independente!

(P) Covarianța verifică urm. proprietăți:

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

b) $Cov(X,a) = 0$, a constantă

c) $Cov(X,Y) = Cov(Y,X)$ (simetrie)

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

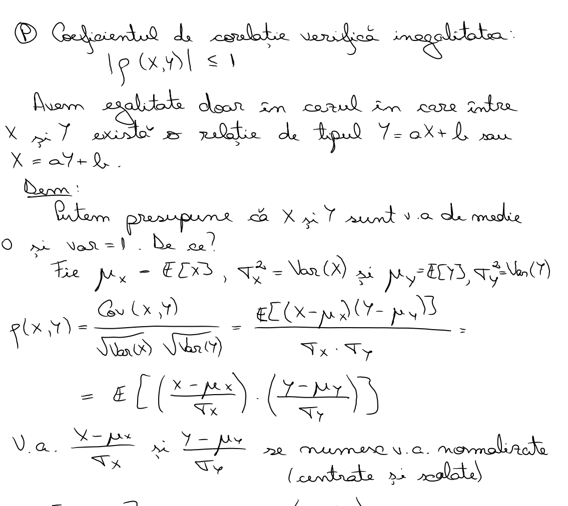
e) $Cov(X+Y, Z) = Cov(X,Z) + Cov(Y,Z)$

f) $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

g) $Var(X_1+X_2+\dots+X_n) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i,X_j)$

Def: Se numește corelația dintre X și Y, sau coeficientul de corelație dintre X și Y,

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



(P) Coeficientul de corelație verifică inegalitatea:

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

Avem egalitate doar în cazul în care între X și Y există o relație de tipul $Y = aX+b$ sau $X = aY+b$.

Dem:

Putem presupune că X și Y sunt v.a de medie 0 și var $=1$. De ce?

Fie $\mu_X = \mathbb{E}[X]$, $\sigma_X^2 = Var(X)$ și $\mu_Y = \mathbb{E}[Y]$, $\sigma_Y^2 = Var(Y)$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X\cdot\sigma_Y} =$$
$$= \mathbb{E}\left[\left(\frac{X-\mu_X}{\sigma_X}\right)\cdot\left(\frac{Y-\mu_Y}{\sigma_Y}\right)\right]$$

V.a. $\frac{X-\mu_X}{\sigma_X}$ și $\frac{Y-\mu_Y}{\sigma_Y}$ se numesc v.a. normalizate (centrate și scalate)

$$\mathbb{E}\left[\frac{X-\mu_X}{\sigma_X}\right] = 0 \quad \text{și} \quad Var\left(\frac{X-\mu_X}{\sigma_X}\right) = 1$$

Astfel, dacă presupunem că X și Y sunt de medie 0 și de varianță 1 avem:

$$\rho(X,Y) = \mathbb{E}[XY]$$

Avem: $\mathbb{E}[(X+\lambda Y)^2] \ge 0, \forall \lambda \in \mathbb{R}$

$$\lambda^2\,\mathbb{E}[Y^2] + 2\lambda\,\mathbb{E}[XY] + \mathbb{E}[X^2] \ge 0, \forall \lambda$$

$$\Delta = 4\mathbb{E}[XY]^2 - 4\mathbb{E}[X^2]\,\mathbb{E}[Y^2] \le 0 \quad \text{(semnul funcției de gr. 2)}$$

$\Rightarrow \mathbb{E}[XY]^2 \le \mathbb{E}[X^2]\cdot\mathbb{E}[Y^2]$ (Ineg Cauchy-Schwarz)

În cazul nostru, $\mathbb{E}[X^2] = \mathbb{E}[Y^2] = 1$ și ineg devine $\mathbb{E}[XY]^2 \le 1 \Rightarrow |\rho(X,Y)| \le 1$

Pp. că $\rho(X,Y) = 1 \Rightarrow \Delta = 0$ $(\Delta = 4\rho(X,Y)^2 - 4)$

$$\Rightarrow \lambda^2\underbrace{\mathbb{E}[Y^2]}_{=1} + 2\lambda\underbrace{\mathbb{E}[XY]}_{\rho(X,Y)=1} + \underbrace{\mathbb{E}[X^2]}_{=1} = \lambda^2 + 2\lambda + 1 = (\lambda+1)^2$$

$$\mathbb{E}[(X+\lambda Y)^2] = (\lambda+1)^2 \ge 0$$

Pentru $\lambda = -1$ avem că $\mathbb{E}[(X-Y)^2] = 0 \Rightarrow \mathbb{P}(X=Y) = 1$

adică

$$\mathbb{P}\left(\frac{X-\mu_X}{\sigma_X} = \frac{Y-\mu_Y}{\sigma_Y}\right) = 1$$

$$\Rightarrow Y = \underbrace{\frac{\sigma_Y}{\sigma_X} X}_{aX} + \underbrace{\frac{\sigma_X\mu_Y - \sigma_Y\mu_X}{\sigma_X}}_{b}$$

Inegalități și teoreme limită

Dacă nu putem evalua de manieră exactă o probabilitate sau medie sau expresie ce implică probabilități atunci putem apela la urm. strategii:

a) simulare

b) mărginirea prob. prin inegalități

c) aproximarea la limită

(T) (Ineg. Cauchy-Schwarz)

Pt orice două v.a. X și Y cu $Var(X)<\infty$, $Var(Y)<\infty$ avem

$$|\mathbb{E}[XY]| \le \sqrt{\mathbb{E}[X^2]\,\mathbb{E}[Y^2]}$$

Versiuni:

a) $\left(\sum_{i=1}^{n} a_i b_i\right)^2 \le \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right)$

Dacă $X \sim \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ 1/n & 1/n & \dots & 1/n \end{pmatrix}$, $Y \sim \begin{pmatrix} b_1 & b_2 & \dots & b_n \\ 1/n & 1/n & \dots & 1/n \end{pmatrix}$ indep

b) $|\langle X,Y\rangle| \le \|X\|\cdot\|Y\|$

$\langle X,Y\rangle = \mathbb{E}[X\cdot Y]$ atunci $\|X\|^2 = \mathbb{E}[X^2]$, $\|Y\|^2 = \mathbb{E}[Y^2]$

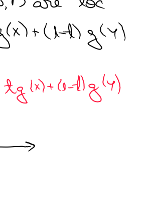
Inegalitatea lui Jensen

$g \nearrow$ convexă $(g''(x) \ge 0)$

$g(x) = x^2$

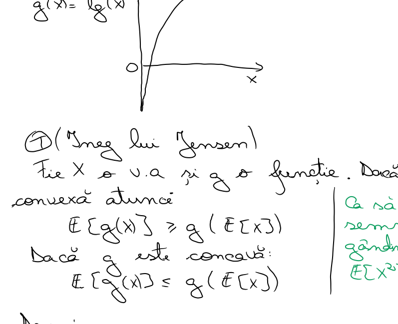
$g''(x) = 2 > 0$

$g(x) = |x|$

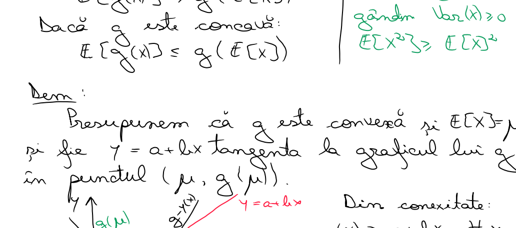


$\forall x, y$, $\forall t \in [0,1]$

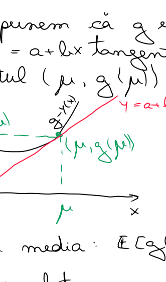
$$g(tx+(1-t)y) \le t g(x) + (1-t) g(y)$$



g este concavă $(g''<0)$: $\forall x, y$ și $\forall t \in [0,1)$ are loc $g(tx+(1-t)y) \ge t g(x) + (1-t) g(y)$



$g(x) = \lg(x)$



(T) (Ineg lui Jensen)

Fie X o v.a și g o funcție. Dacă g este convexă atunci

$$\mathbb{E}[g(X)] \ge g(\mathbb{E}[X])$$

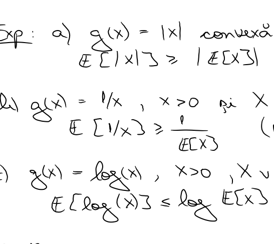
Dacă g este concavă:

$$\mathbb{E}[g(X)] \le g(\mathbb{E}[X])$$

Ca să reținem semnul ne gândim $Var(X) \ge 0$: $\mathbb{E}[X^2] \ge \mathbb{E}[X]^2$

Dem:

Presupunem că g este convexă și $\mathbb{E}[X] = \mu$ și fie $Y = a+bx$ tangenta la graficul lui g în punctul $(\mu, g(\mu))$.



Din convexitate: $g(x) \ge a+bx$, $\forall x$ atunci $g(X) \ge a+bX$

Luând media: $\mathbb{E}[g(X)] \ge \mathbb{E}[a+bX]$ (monotonia)

Din liniaritate:

$$\mathbb{E}[a+bX] = a + b\,\mathbb{E}[X] = a + b\mu = g(\mu) = g(\mathbb{E}[X])$$

Astfel:

$$\mathbb{E}[g(X)] \ge g(\mathbb{E}[X])$$

Să pp. că avem egalitate: $\mathbb{E}[g(X)] = g(\mathbb{E}[X])$

și fie $\left.\begin{array}{l} Y = g(X) - a - bX \ge 0 \\ \mathbb{E}[Y] = 0 \end{array}\right\} \Rightarrow \mathbb{P}(Y=0) = 1$

$\Rightarrow \mathbb{P}(g(X) = a+bX) = 1 \Rightarrow g$ este afină $(g(x) = a+bx)$

Exp: a) $g(x) = |x|$ convexă

$$\mathbb{E}[|X|] \ge |\mathbb{E}[X]|$$

b) $g(x) = 1/x$, $x>0$ și X v.a. pozitivă

$$\mathbb{E}[1/X] \ge \frac{1}{\mathbb{E}[X]} \quad (1/x \text{ convexă})$$

c) $g(x) = \log(x)$, $x>0$, X v.a. pozitivă

$$\mathbb{E}[\log(X)] \le \log \mathbb{E}[X] \quad (\log(x) \text{ concavă})$$

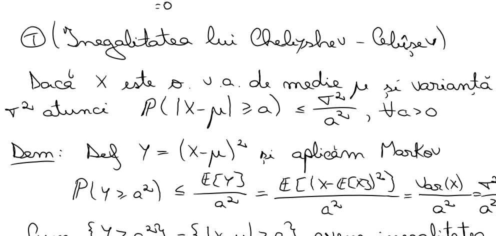
(T) Inegalitatea lui Markov

Dacă X este o v.a. pozitivă atunci

$$\mathbb{P}(X \ge a) \le \frac{\mathbb{E}[X]}{a}, \forall a>0$$

Dem: Definim v.a. Y_a pt $a>0$ fixat astfel

$$Y_a = \begin{cases} 0, & X < a \\ a, & X \ge a \end{cases}$$



Avem

$Y_a \le X$ ← aplicăm media

$$\left.\begin{array}{l} \mathbb{E}[Y_a] \le \mathbb{E}[X] \\ 0\cdot\mathbb{P}(Y_a=0) + a\,\mathbb{P}(Y_a=a) = a\cdot\mathbb{P}(X\ge a) \end{array}\right\} \Rightarrow a\,\mathbb{P}(X\ge a) \le \mathbb{E}[X] \;|:a$$

Exp: $X \sim \mathcal{U}[0,4] \Rightarrow \mathbb{E}[X] = 2$

$$\mathbb{P}(X \ge a) \le \frac{\mathbb{E}[X]}{a} = \frac{2}{a}$$

$a=1 \Rightarrow \underbrace{\mathbb{P}(X\ge 1)}_{3/4} \le 2$

$a=2 \Rightarrow \underbrace{\mathbb{P}(X\ge 2)}_{=1/2} \le 1$

$a=4 \Rightarrow \underbrace{\mathbb{P}(X\ge 4)}_{=0} \le \frac{1}{2} = 0.5$

(T) (Inegalitatea lui Chebyshev - Cebîșev)

Dacă X este o v.a. de medie μ și varianță σ^2 atunci $\mathbb{P}(|X-\mu| \ge a) \le \frac{\sigma^2}{a^2}, \forall a>0$

Dem: Def $Y = (X-\mu)^2$ și aplicăm Markov

$$\mathbb{P}(Y \ge a^2) \le \frac{\mathbb{E}[Y]}{a^2} = \frac{\mathbb{E}[(X-\mathbb{E}[X])^2]}{a^2} = \frac{Var(X)}{a^2} = \frac{\sigma^2}{a^2}$$

Cum $\{Y \ge a^2\} = \{|X-\mu| \ge a\}$ avem inegalitatea dorită.

Obs: Dacă $a = K\sigma$ atunci

$$\mathbb{P}(|X-\mu| \ge K\sigma) \le \frac{\sigma^2}{(K\sigma)^2} = \frac{1}{K^2}$$

(T) (Inegalitatea lui Chernoff)

Fie X o v.a, $a>0$ și $t>0$. Atunci

$$\mathbb{P}(X \ge a) \le \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$$

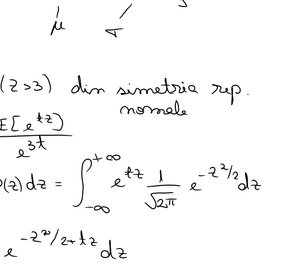
Dem: $g(x) = e^{tx}$ cresc, bijectivă

$$\mathbb{P}(X\ge a) = \mathbb{P}(g(X) \ge g(a)) = \mathbb{P}(e^{tX} \ge e^{ta}) \overset{\text{Ineg Markov}}{\le} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$$

Exp: Fie $Z \sim N(0,1)$ și vrem să $\mathbb{P}(|Z|>3) \le ?$ folosind Markov, Chebyshev și Chernoff.

a) Markov

$\mathbb{P}(|Z|>3) \approx 0.003$



$$\mathbb{P}(|Z|>3) \le \frac{\mathbb{E}[|Z|]}{3}$$

$$\mathbb{E}[|Z|] = \int_{\mathbb{R}} |z|\varphi(z)\,dz =$$
$$= \int_{\mathbb{R}} |z| \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz = 2\int_0^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz =$$
$$= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} z e^{-z^2/2}\,dz = \frac{2}{\sqrt{2\pi}} \left(-e^{-z^2/2}\right)\Big|_0^{\infty} =$$
$$= \sqrt{\frac{2}{\pi}} \Rightarrow \frac{\mathbb{E}[|Z|]}{3} = \frac{1}{3}\sqrt{\frac{2}{\pi}} \approx 0.027$$

b) Chebyshev:

$$\mathbb{P}(|Z|>3) = \mathbb{P}(|\underset{\mu}{Z-0}| > 3\cdot\underset{\sigma}{1}) \le \frac{1}{9} \approx 0.11$$

c) Chernoff

$$\mathbb{P}(|Z|>3) = 2\,\mathbb{P}(Z>3) \text{ din simetria rep. normale}$$
$$\le 2\frac{\mathbb{E}[e^{tZ}]}{e^{3t}}$$

$$\mathbb{E}[e^{tZ}] = \int_{\mathbb{R}} e^{tz}\varphi(z)\,dz = \int_{-\infty}^{+\infty} e^{tz}\frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz$$
$$= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2+tz}\,dz$$
$$= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-1/2(z^2-2tz+t^2-t^2)}\,dz$$
$$= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}}\cdot e^{t^2/2}\,dz$$
$$= e^{t^2/2}\underbrace{\int \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}}\,dz}_{\text{densitatea } N(t,1) \;=1}$$
$$= e^{t^2/2}$$

$$\mathbb{P}(|Z|>3) = 2\cdot\frac{e^{t^2/2}}{e^{3t}} = 2e^{t^2/2-3t}, \forall t$$

Minimul se atinge pt $t=3$

$$\mathbb{P}(|Z|>3) \le 2\cdot e^{-9/2} \approx 0.022$$