LOIS À DENSITÉ (PARTIE 3) : LES LOIS NORMALES DÉMONSTRATIONS / RÉPONSES

PROPRIÉTÉ

Soit X une v.a.r. suivant la loi normale centrée réduite. Alors : $\mathrm{E}(\mathrm{X})=0$; $\mathrm{V}(\mathrm{X})=1$; $\sigma(\mathrm{X})=1$.

Démonstrations : $\mathrm{V}(\mathrm{X})=1$ est admis, donc $\sigma(\mathrm{X})=1$ également. Démontrons que $\mathrm{E}(\mathrm{X})=0$:

Par définition, $\mathrm{E}(\mathrm{X})=\int_{-\infty}^{+\infty} t\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t$, autrement dit :

$$\mathrm{E}(\mathrm{X})=\lim_{y\to-\infty}\int_y^0 t\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t+\lim_{x\to+\infty}\int_0^x t\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

• Soit x un réel positif : $\int_0^x t\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t=\frac{1}{\sqrt{2\pi}}\int_0^x t\,\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t$.

Or $t\,\mathrm{e}^{-\frac{t^2}{2}}$ est de la forme $-u'\mathrm{e}^u$ donc une primitive de $t\,\mathrm{e}^{-\frac{t^2}{2}}$ est $-\mathrm{e}^{-\frac{t^2}{2}}$:

$$\frac{1}{\sqrt{2\pi}}\int_0^x t\,\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t=\frac{1}{\sqrt{2\pi}}\left[-\mathrm{e}^{-\frac{t^2}{2}}\right]_0^x=\frac{1}{\sqrt{2\pi}}\left(-\mathrm{e}^{-\frac{x^2}{2}}+\mathrm{e}^0\right)=\frac{1}{\sqrt{2\pi}}\left(1-\mathrm{e}^{-\frac{x^2}{2}}\right).$$

Par composition de limites : $\lim_{x\to-\infty}\mathrm{e}^{-\frac{x^2}{2}}=0$ donc $\lim_{x\to+\infty}\frac{1}{\sqrt{2\pi}}\left(1-\mathrm{e}^{-\frac{x^2}{2}}\right)=\frac{1}{\sqrt{2\pi}}$

d'où $\lim_{x\to+\infty}\int_0^x t\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t=\frac{1}{\sqrt{2\pi}}$.

• De même, on montrerait que $\lim_{y\to-\infty}\int_y^0 t\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t=-\frac{1}{\sqrt{2\pi}}$.

• Par somme de limites : $\mathrm{E}(\mathrm{X})=0$.

PROPRIÉTÉS.

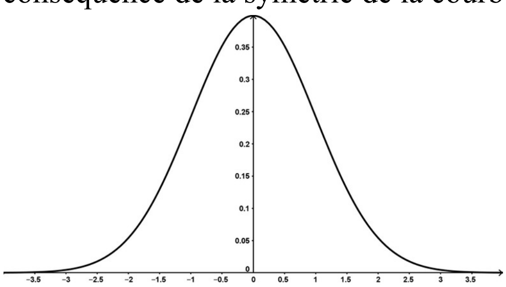
Soit T une variable aléatoire suivant la loi normale centrée réduite $\mathcal{N}(0;1)$.

• Pour tout réel $u>0$: $\mathrm{p}(\mathrm{T}\leqslant-u)=\mathrm{p}(\mathrm{T}\geqslant u)$ et $\mathrm{p}(-u\leqslant\mathrm{T}\leqslant u)=2\,\mathrm{p}(\mathrm{T}\leqslant u)-1$.

• $\mathrm{p}(\mathrm{T}\geqslant0)=\mathrm{p}(\mathrm{T}\leqslant0)=\frac{1}{2}$.

Démonstration : • $\mathrm{p}(\mathrm{T}\leqslant-u)=\mathrm{p}(\mathrm{T}\geqslant u)$ est une conséquence de la symétrie de la courbe, tout comme $\mathrm{p}(\mathrm{T}\geqslant0)=\mathrm{p}(\mathrm{T}\leqslant0)=\frac{1}{2}$.

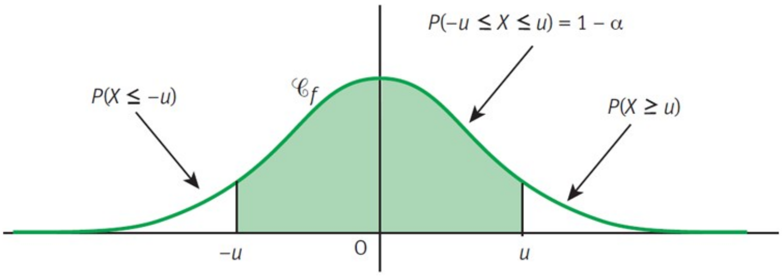
•
$$\begin{aligned}\mathrm{p}(-u\leqslant\mathrm{T}\leqslant u)&=\mathrm{p}(\mathrm{T}\leqslant u)-\mathrm{p}(\mathrm{T}<-u)\\&=\mathrm{p}(\mathrm{T}\leqslant u)-\mathrm{p}(\mathrm{T}>u)\\&=\mathrm{p}(\mathrm{T}\leqslant u)-(1-\mathrm{p}(\mathrm{T}\leqslant u))\\&=2\,\mathrm{p}(\mathrm{T}\leqslant u)-1\end{aligned}$$



ROC

PROPRIÉTÉ.

Soit T une variable aléatoire suivant la loi normale centrée réduite $\mathcal{N}(0;1)$.
Pour tout $\alpha \in]0;1[$, il existe un unique réel $u_\alpha > 0$ tel que $p(-u_\alpha \leqslant T \leqslant u_\alpha) = 1-\alpha$.



Démonstration :

Soit T une variable aléatoire suivant la loi normale centrée réduite $\mathcal{N}(0;1)$. Soit $\alpha \in]0;1[$.

1. Démontrer que pour tout réel x strictement positif :
$p(-x \leqslant T \leqslant x) = 2F(x)$ où F est la primitive de f qui s'annule en 0.

2. Démontrer que : $0 < \frac{1-\alpha}{2} < \frac{1}{2}$.

3. a) On admet que F est continue et strictement croissante sur $[0;+\infty[$.
Démontrer qu'il existe un unique réel strictement positif, noté u_α, tel que : $F(u_\alpha) = \frac{1-\alpha}{2}$.

b) En déduire qu'il existe un unique réel $u_\alpha > 0$ tel que $p(-u_\alpha \leqslant T \leqslant u_\alpha) = 1-\alpha$.

Correction : **1.** Pour tout réel $x > 0$:
$p(-x \leqslant T \leqslant x) = 2p(0 \leqslant T \leqslant x)$ par symétrie de la densité f de la loi normale centrée réduite

$$= 2\int_0^x f(t)\,dt = 2F(x)$$ où F est la primitive de f qui s'annule en 0.

2. $0 < \alpha < 1 \Rightarrow 0 > -\alpha > -1 \Rightarrow 1 > 1-\alpha > 0 \Rightarrow \frac{1}{2} > \frac{1-\alpha}{2} > 0$.

3. a) $F(0) = 0$ et $\lim_{x \to +\infty} F(x) = \frac{1}{2}$ (aire sous la « courbe en cloche » en partant de l'origine).

On a alors :

x	0 $\qquad\qquad\qquad\qquad$ $+\infty$
F	$0 \nearrow \frac{1}{2}$

Or, F est continue et $0 < \frac{1-\alpha}{2} < \frac{1}{2}$ donc d'après le théorème des valeurs intermédiaires, il existe un unique réel strictement positif, noté u_α, tel que : $F(u_\alpha) = \frac{1-\alpha}{2}$.

b) $p(-x \leqslant T \leqslant x) = 1-\alpha \Leftrightarrow 2F(x) = 1-\alpha \Leftrightarrow F(x) = \frac{1-\alpha}{2}$.

D'après les questions précédentes, il existe donc un unique réel $u_\alpha > 0$ tel que $p(-u_\alpha \leqslant T \leqslant u_\alpha) = 1-\alpha$.

III. Avec la calculatrice

Exercice III.1 : soit X une v.a.r. suivant la loi normale $\mathcal{N}\left(46{,}75\,;6{,}23^2\right)$. Avec la calculatrice :

1. $p(40\leqslant X\leqslant 60)\approx 0{,}84398$

2. $p(X\leqslant 60)\approx 0{,}98328$ directement avec calculatrice (-1E99 jusqu'à 60)
ou $p(X\leqslant 60)=p(X\leqslant 46{,}75)+p(46{,}75\leqslant X\leqslant 60)=0{,}5+p(46{,}75\leqslant X\leqslant 60)\approx 0{,}5+0{,}48328=0{,}98328$

3. $p(X>40)\approx 0{,}86069$ directement avec calculatrice (40 jusqu'à 1E99)
ou $p(X>40)=p(40\leqslant X\leqslant 46{,}75)+p(X>46{,}75)=p(40\leqslant X\leqslant 46{,}75)+0{,}5\approx 0{,}36069+0{,}5=0{,}86069$

4. $p(X\geqslant 60)\approx 0{,}016718$ directement avec calculatrice (60 jusqu'à 1E99)
ou $p(X\geqslant 60)=p(X>60)=1-p(X\leqslant 60)=1-(p(X<46{,}75)+p(46{,}75\leqslant X\leqslant 60))=1-(0{,}5+p(46{,}75\leqslant X\leqslant 60))$
$=0{,}5-p(46{,}75\leqslant X\leqslant 60)\approx 0{,}5-0{,}48328=0{,}01672$

Exercice III.2 : soit X une v.a.r. suivant la loi normale $\mathcal{N}\left(46{,}75\,;6{,}23^2\right)$.
Déterminer une valeur approchée du réel k tel que :

1. $p(X\leqslant k)=0{,}75$ donne $k\approx 50{,}952$.

2. $p(X<k)=0{,}95 \Leftrightarrow p(X\leqslant k)=0{,}95$ et donne $k\approx 56{,}997$.

3. $p(X\geqslant k)=0{,}95 \Leftrightarrow 1-p(X\leqslant k)=0{,}95 \Leftrightarrow p(X\leqslant k)=0{,}05$ donne $k\approx 36{,}502$.
On peut aussi utiliser (sur CASIO) le *Tail: Right* et trouver directement le résultat.

Exercice III.3 : **voir fichier Geogebra** « Comparaison majoration erreur Uspensky – Esseen.ggb »

1. $0{,}588\geqslant 0{,}9568\left(p^2+(1-p)^2\right) \Leftrightarrow 2p^2-2p+1\leqslant\dfrac{0{,}588}{0{,}9568}$.

Or, on montre facilement (poly. de degré 2) que :

x	0		$\frac{1}{2}$		1
$f(x)$	1	$\searrow$	$\frac{1}{2}$	$\nearrow$	1

$\dfrac{0{,}588}{0{,}9568}\approx 0{,}61$ donc d'après le TVI, il existe 2 solutions α_1 et α_2 à l'équation $2p^2-2p+1=\dfrac{0{,}588}{0{,}9568}$ et :
$0{,}588\geqslant 0{,}9568\left(p^2+(1-p)^2\right)$ sur $[\alpha_1;\alpha_2]$ donc « Asof est meilleur que Uspensky sur $[\alpha_1;\alpha_2]$ ».
$0{,}588< 0{,}9568\left(p^2+(1-p)^2\right)$ sur $[0;\alpha_1[\cup]\alpha_2;1]$ donc « Uspensky est meilleur que Asof sur $[\alpha_1;\alpha_2]$ ».
Remarque : on trouve facilement, à la calculatrice, $\alpha_1\approx 0{,}26$ et $\alpha_2\approx 0{,}74$.

2. Il suffit de trouver à partir de quel coefficient *C* on a eu $\dfrac{0{,}588}{C}\geqslant 0{,}5$.

On trouve $C\leqslant\dfrac{0{,}588}{0{,}5}$ *ie* $C\leqslant 1{,}176$.

Donc jusqu'en 2009 le théorème d'Uspensky était toujours meilleur que celui de Berry-Esseen, avec Korolev et Shevtsova qui trouvent 1,0258.

IV. Le continu remplace le discret… mais attention !

1. Cela n'est pas très judicieux d'approcher $p(X \leqslant 45)$ par $p(Z \leqslant 45)$ car X prend ses valeurs dans ℕ, donc calculer $p(X \leqslant 45)$ signifie bien sûr $p(0 \leqslant X \leqslant 45)$: on devrait approcher $p(X \leqslant 45)$ par $p(0 \leqslant Z \leqslant 45)$.

2. Si n est grand, on approche X par Z avec $Z \sim \mathcal{N}(np\ ;\ \sqrt{np(1-p)}^2)$.

3. Avec $n=255$ et $p=0{,}02$: $np=5{,}1$ et $\sqrt{np(1-p)} \approx 2{,}2356$.

a) On peut utiliser directement la calculatrice (sur CASIO : menu STAT, puis DIST-BINM-Bcd) et trouver : $p(X \leqslant 5) \approx 0{,}59825$.

On peut aussi revenir à la définition de la loi binomiale et écrire :

$$\begin{aligned} p(X \leqslant 5) &= \sum_{k=0}^{5} \binom{255}{k} 0{,}02^k\, 0{,}98^{255-k} \\ &= \binom{255}{0} 0{,}98^{255} + \binom{255}{1} 0{,}02 \times 0{,}98^{254} + \binom{255}{2} 0{,}02^2\, 0{,}98^{253} + \binom{255}{3} 0{,}02^3\, 0{,}98^{252} + \binom{255}{4} 0{,}02^4\, 0{,}98^{251} + \binom{255}{5} 0{,}02^5\, 0{,}98^{250} \\ &\approx 0{,}59825. \end{aligned}$$

b) Avec la calculatrice (sur CASIO : menu STAT, puis DIST-NORM-Ncd) :
$p(Z \leqslant 5) \approx p(-10^{99} \leqslant Z \leqslant 5) \approx 0{,}48216$ et $p(0 \leqslant Z \leqslant 5) \approx 0{,}47089$.

c) Observation 1 : remplacer X par Z n'est déjà pas un bon choix… On obtient $\approx 0{,}471$ au lieu de $\approx 0{,}598$!
Tout d'abord, n n'est peut-être pas assez grand pour le p choisi… Le théorème de Moivre-Laplace parle de limite lorsque n tend vers $+\infty$.
Mais on pourrait corriger tout cela avec ce qu'on appelle la « correction de continuité ».
Observation 2 : l'écart entre $p(Z \leqslant 5)$ et $p(0 \leqslant Z \leqslant 5)$ est d'environ $0{,}01127$, soit 1,1 % environ…
Ce n'est pas négligeable !