

B. LOIS USUELLES CONTINUES

I. Loi normale

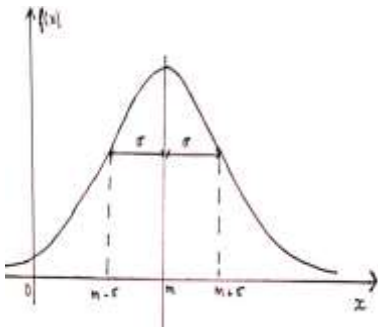
1. Définition

Une variable aléatoire continue X suit une loi normale de paramètres m et σ^2 notée $N(m, \sigma^2)$ où $m \in \mathbb{R}$ et $\sigma > 0$, si et seulement si sa densité de probabilité est la fonction f définie sur \mathbb{R} par :

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

Cette densité forme une courbe en cloche. Elle admet la droite $x = m$ comme axe de symétrie et deux points d'inflexion d'abscisse $x = m \pm \sigma$

m et σ représentent respectivement la moyenne et l'écart-type de la v.a X .

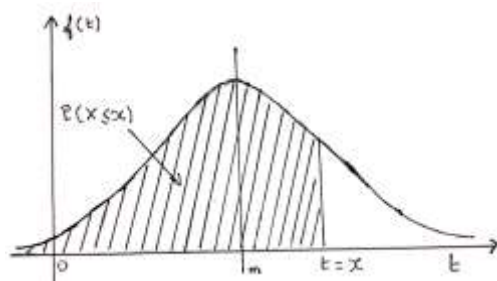


2. Fonction de répartition

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(t-m)^2} dt$$

i. Graphiquement

C'est l'aire délimitée par la courbe en cloche, l'axe des abscisses et à gauche de la droite $t = x$



ii. Analytiquement

Cette intégrale ne peut être calculée directement. Pour ce faire on doit passer par le cas particulier où $m=0$ et $\sigma = 1$, on parle alors de la loi normale centrée réduite.

3. Propriété des 3 σ

On montre que :

- $P[m - \sigma < X < m + \sigma] = 0.6826 \Leftrightarrow 68\%$ des valeurs sont dans $[m - \sigma ; m + \sigma]$
 - $P[m - 2\sigma < X < m + 2\sigma] = 0.9545 \Leftrightarrow 95\%$ des valeurs sont dans $[m - 2\sigma ; m + 2\sigma]$
 - $P[m - 3\sigma < X < m + 3\sigma] = 0.9973 \Leftrightarrow 99,7\%$ des valeurs sont dans $[m - 3\sigma ; m + 3\sigma]$
- voir démonstration page 9.

II Loi normale centrée réduite $N(0; 1)$

1. Définition :

Une v.a.r continue Z suit une loi normale centrée réduite si et seulement si sa densité de probabilité est la fonction f définie sur \mathbb{R} par :

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad \forall z \in \mathbb{R}$$

Cette densité forme une courbe en cloche. Elle admet la droite $x = 0$ comme axe de symétrie et deux points d'inflexion d'abscisse $x = \pm 1$

2. Espérance et Variance

On montre que si Z une variable aléatoire de loi $N(0,1)$ alors $E(Z)=0$ et $V(Z)=1$.

Démonstration

$$\text{On a } E(Z) = \int_{-\infty}^{\infty} z f(z) dz = \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 0 \quad (\text{intégrale sur } \mathbb{R} \text{ d'une fonction impaire})$$

$$\text{et } E(Z^2) = \int_{-\infty}^{\infty} z^2 f(z) dz = \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1 \quad (\text{à admettre})$$

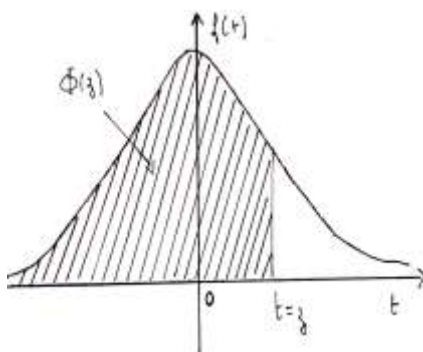
$$\text{alors } V(Z) = E(Z^2) - E^2(Z) = 1$$

3. La fonction de répartition

$$F_Z(z) = P(Z \leq z) = \int_{-\infty}^z f(t) dt = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

i. Graphiquement :

C'est l'aire délimitée par la courbe en cloche, l'axe des abscisses et par la droite $t = z$.
La fonction de répartition de la loi normale centrée réduite se note généralement Φ .



$$P(Z \leq z) = \Phi(z)$$

ii. Analytiquement

Cette intégrale a été calculée par interpolation et donnée sous forme de table statistique

Ses valeurs peuvent se lire sur une table donnée en annexe.

La table ne donne que les valeurs de $P(Z \leq z) = \Phi(z)$ pour z positif. Pour les valeurs de z qui sont négatives, on les déduit par symétrie.

iii. Propriétés

$$\Phi(0) = P(Z \leq 0) = \frac{1}{2} = P(Z \geq 0)$$

$\Phi(-z) = 1 - \Phi(z)$ par symétrie de la densité $f(z)$ par rapport à la droite $z=0$

$$\Phi(z) \geq \frac{1}{2} \Leftrightarrow z \geq 0 \text{ et } \Phi(z) \leq \frac{1}{2} \Leftrightarrow z \leq 0$$

4. Lecture de la table $N(0,1)$

a. Lecture directe : z est donné, lire $\alpha = \Phi(z)$

- si $0 \leq z < 4$ lire $\Phi(z)$ sur la table

Exemple : $\Phi(1.24) = ?$

$z=1,24$ s'écrit comme $1,2 + 0,04$ alors $\Phi(1.24)$ se trouve à l'intersection de la ligne 1,2 et de la colonne 0,04. On lit $\Phi(1.24) = 0,8925$

- si $z \geq 4$ alors $\Phi(z) = 1$

- si $z < 0$ on utilise la propriété de symétrie soit $\Phi(z) = 1 - \Phi(-z)$

Exemple : $\Phi(-1.24) = ?$

On a $\Phi(-1.24) = 1 - \Phi(1.24) = 1 - 0,8925 = 0,1075$

b. Lecture inverse : $\Phi(z) = \alpha$ est donné, lire $z = \Phi^{-1}(\alpha)$ le quantile d'ordre α , noté aussi q_α

- si $\Phi(z) = \alpha \geq \frac{1}{2}$ lire $z = ?$

Exemple1 : $\Phi(z) = 0.67 \Rightarrow z = ?$

On voit que $\Phi(z) = \alpha = 0.67$ est au croisement de la ligne 0.4 et de la colonne 0.04 alors $z = 0.44$

Exemple2 : $\Phi(z) = 0.8930 \Rightarrow z = ?$

On remarque que 0.8930 n'existe pas sur la table, on prendra la valeur la plus proche de $\Phi(z)$. Par encadrement, on lit sur la table que

$$0.8925 < 0.8930 < 0.8944$$

qui correspond à $\Phi(1.24) < \Phi(z) < \Phi(1.25)$.

$\Phi(1.24) = 0.8925$ est la valeur la plus proche 0.8930 proche, par conséquent on prendra $\Phi(z) = \Phi(1.24)$ donc $z = 1.24$

- si $\Phi(z) \leq \frac{1}{2}$ lire z . Utiliser la relation de symétrie

Exemple : Trouver z tel que $\Phi(z)=0.25$

On a $\Phi(z) \leq \frac{1}{2} \Rightarrow z < 0 \Rightarrow -z > 0$. On calculera donc $\Phi(-z)$

$$\Phi(-z) = 1 - \Phi(z) = 1 - 0.25 = 0.75 \Rightarrow -z = 0.67 \Rightarrow z = -0.67$$

5. Théorème

Si la variable aléatoire X suit une loi normale $N(m, \sigma^2)$, alors la variable aléatoire

- $Z = \frac{X-m}{\sigma}$ suit la loi normale centrée réduite $N(0,1)$.
- $P(X \leq x) = \Phi(z)$
- $P(X \geq x) = \Phi(-z)$ tel $z = \frac{x-m}{\sigma}$

Démonstration

i) A admettre

ii) On a $P(X \leq x) = P(X-m \leq x-m) = P\left(\frac{X-m}{\sigma} \leq \frac{x-m}{\sigma}\right)$ car $\sigma > 0$

En posant $Z = \frac{X-m}{\sigma}$ et $z = \frac{x-m}{\sigma}$, de plus d'après i) on sait que $Z \rightarrow N(0,1)$

Alors on a bien $P(X \leq x) = P(Z \leq z) = \Phi(z)$

iii) $P(X \geq x) = 1 - P(X \leq x) = 1 - \Phi(z) = \Phi(-z) \Leftrightarrow P(X \geq x) = \Phi(-z)$

Exercice

- Démontrer la propriété des 3σ de la loi $N(m, \sigma^2)$ citée précédemment.
- Montrer que si X est une variable aléatoire de loi $N(m, \sigma^2)$ alors $E(X) = m$ et $V(X) = \sigma^2$

Solution

1. Montrons que Si $X \rightarrow N(m, \sigma^2) \Rightarrow P[m - \sigma < X < m + \sigma] = 0.6826$?

En effet,

D'après le théorème, on sait que si $X \rightarrow N(m, \sigma^2)$ alors $Z = \frac{X-m}{\sigma} \rightarrow N(0,1)$

$$\text{On a } P[m - \sigma < X < m + \sigma] = P\left(\frac{(m-\sigma)-m}{\sigma} \leq \frac{X-m}{\sigma} \leq \frac{(m+\sigma)-m}{\sigma}\right) = P(-1 \leq Z \leq 1) =$$

$$P(Z \leq 1) - P(Z \leq -1) = \Phi(1) - \Phi(-1) = \Phi(1) - [1 - \Phi(1)] = 2 \times \Phi(1) - 1 = 2 \times 0.8413 - 1 = 0.6826$$

Remarque les deux autres propriétés se démontrent de la même façon.

2. Toujours d'après le théorème, si $X \rightarrow N(m, \sigma^2)$ alors $Z = \frac{X-m}{\sigma} \rightarrow N(0,1)$

De plus on sait que si $Z \rightarrow N(0,1)$ alors $E(Z)=0$ et $V(Z)=1$.

$$\text{On a } Z = \frac{X-m}{\sigma} \Leftrightarrow X = m + \sigma Z \Rightarrow E(X) = m + \sigma E(Z) = m \text{ et } V(X) = \sigma^2 V(Z) = \sigma^2$$

Exercice 10 de la série de TD

Dans une population humaine, la mesure de la quantité d'urée dans le sang en (mg/100ml) a donné une moyenne de 27mg/100ml et une variance de 25 (mg/100ml)². On suppose que cette variable est de loi Normale. On prend un individu au hasard.

1. Quelle est la probabilité que sa quantité d'urée soit inférieure à 24,5 ?
2. Déterminer la limite x_0 , appelé seuil pathologique, tel qu'il y ait 2,5% d'individus dont la quantité d'urée est supérieure à cette valeur x_0 .

solution

X = la quantité d'urée dans le sang, $E(X) = 27$ et $V(X) = 25$

1. $P(X < 24.5)$?

On sait que si $X \rightarrow N(m, \sigma^2)$ alors $Z = \frac{X - m}{\sigma} \rightarrow N(0, 1)$ et $P(X \leq x) = \Phi(z)$ où $z = \frac{x - m}{\sigma}$

$$P(X < 24.5) = P(X \leq 24.5) = \Phi\left(\frac{24.5 - 27}{\sqrt{25}}\right) = \Phi(-0.5) = 1 - \Phi(0.5) = 1 - 0.6915 = 0.3085$$

2. $x_0 = ?$ tel $P(X > x_0) = 0.025$

$$\text{On a } P(X > x_0) = 1 - P(X \leq x_0) = 0.025 \Rightarrow P(X \leq x_0) = 1 - 0.025 = 0.975$$

de plus on sait que $P(X \leq x_0) = \Phi(z_0) = 0.975 \Rightarrow z_0 = 1.96$ avec $z_0 = \frac{x_0 - m}{\sigma}$

$$\text{donc } x_0 = m + \sigma \times z_0 = 27 + 5 \times 1.96 = 36.8$$

On conclut que le seuil pathologique est de 36.8 mg/100ml

III - Approximation d'une loi binomiale par une loi Normale

Théorème

Pour n suffisamment grand, et p ni trop grand, ni trop petit, la loi binomiale peut être approximée par la loi normale

Dans la pratique on doit avoir $n > 30$, $n \times p \geq 5$ et $n \times q \geq 5$.

Dans ce cas la loi $\mathcal{B}(n, p) \underset{\infty}{\approx} N(m, \sigma^2)$ avec $m = n \times p$ et $\sigma^2 = n \times p \times q$

Démonstration

La démonstration découle du théorème central limite. Voir page 15

Remarque :

Dans certains cas limites, la loi binomiale pourrait être approximée à la fois par la loi de Poisson et la loi normale. On préférera l'approximation par la loi normale, car il n'y aura qu'une seule probabilité à calculer.

exemple

On lance deux dés 200 fois. Quelle est la probabilité d'avoir au moins 10 fois le double as ?

Soit X = le nombre de fois où on obtient le double as en 200 jets de 2 dés

$$X \rightarrow \mathcal{B}(n, p) \text{ avec } n = 200 \text{ et } p = \frac{1}{36} \quad p = P(1, 1)$$

$$P(X \geq 10) = 1 - P(X < 10) = 1 - \sum_{k=0}^9 P(X = k) = 1 - \sum_{k=0}^9 C_{200}^k \left(\frac{1}{36}\right)^k \left(\frac{35}{36}\right)^{200-k}$$

ces calculs sont très ardues donc on approxime cette loi.

- on a $n = 200 > 50$, $p = \frac{1}{36} < 0.1$ et $np = 5.56 < 10$

donc on peut approximer cette loi $\mathcal{B}(200, \frac{1}{36})$ par une loi $P(5,56)$

$$\text{et alors } P(X \geq 10) = 1 - P(X < 10) = 1 - \sum_{k=0}^9 e^{-5.56} \frac{(5.56)^k}{k!}$$

Là on aurait 10 calculs de probabilités à faire.

- mais on a aussi on a $n = 200 > 30$, $np = 5.56 > 5$ et $nq = 194.44 > 5$

donc on peut approximer cette loi $\mathcal{B}(n, p)$ par une loi $N(m, \sigma^2)$ avec $m = np = 5.56$ et $\sigma^2 = npq = 5.4$

$$\text{donc } P(X \geq 10) = 1 - P(X \leq 10) = 1 - \Phi\left(\frac{10 - 5.56}{\sqrt{5.4}}\right) = 1 - \Phi(1.91) = 1 - 0.9719 = 0.0281$$

Exercice 12 de la série de TD

On veut connaître la proportion, p , d'individus d'une population très grande, résistant à un certain antibiotique. Sur un échantillon de 1000 individus, on a observé que le nombre moyen d'individus résistant à l'antibiotique est 62,5.

Soit X le nombre d'individus résistant à l'antibiotique parmi les 1000 observés.

1. Donner la loi de la v.a X et donner une estimation de la proportion p .
2. Par quelle loi peut-on approximer la loi de X ? Calculer alors $P(X > 30)$ et $P(30 < X < 40)$.

Solution

X le nombre d'individus résistant à l'antibiotique parmi les 1000 observés. $E(X) = 62.5$

1. Loi de la v.a X ?

$X \rightarrow \mathcal{B}(n, p)$ avec $n = 1000$ et $p = ?$ (voir corrigé de l'ex 12 de la série de TD)

On sait que si $X \rightarrow \mathcal{B}(n, p)$ alors $E(X) = n \times p \Leftrightarrow p = \frac{E(X)}{n} = \frac{62.5}{1000} = 0.0625$.

On peut donc estimer p à 6.25%

2. On a $n = 1000 > 30$, $n \times p = 62.5 > 5$ et $n \times q = n - n \times p = 937.5 > 5$ alors on peut approximer cette loi binomiale par une loi $N(m, \sigma^2)$ avec $m = n \times p = 62.5$ et $\sigma^2 = n \times p \times q = 58.59$

$P(X > 30) = ?$

Si $X \rightarrow N(m, \sigma^2)$ alors $Z = \frac{X - m}{\sigma} \rightarrow N(0, 1)$ et $P(X \geq x) = \Phi(-z)$ où $z = \frac{x - m}{\sigma}$

$$\text{Alors } P(X > 30) = \Phi\left(-\frac{30 - 62.5}{\sqrt{58.59}}\right) = \Phi(4.24) = 1$$

$P(30 < X < 40) = ?$

$$\begin{aligned} P(30 < X < 40) &= P(X \leq 40) - P(X \leq 30) = \Phi\left(\frac{40 - 62.5}{\sqrt{58.59}}\right) - 0 \quad (\text{car } P(X \leq 30) = 1 - P(X > 30) = 0) \\ &= \Phi(-2.94) = 1 - \Phi(2.94) = 1 - 0.9984 = 1.6 \cdot 10^{-3} \end{aligned}$$

IV. Propriétés de la loi Normale

1. Théorème 1:

Soit $(X_i)_{i=1,\dots,n}$ une suite de variables aléatoires **indépendantes** de lois $N(m_i, \sigma_i^2)$ et a et b deux réels quelconques. On montre que :

1. $X = \sum_{i=1}^n X_i \rightarrow N(m, \sigma^2)$ avec $m = \sum_{i=1}^n m_i$ et $\sigma^2 = \sum_{i=1}^n \sigma_i^2$
2. $Y = aX + b \rightarrow N(am + b, a^2 \sigma^2)$

2. Théorème 2:

Soit $(X_i)_{i=1,\dots,n}$ une suite de variables aléatoires indépendantes identiquement distribuées(i.i.d) de loi $N(m, \sigma^2)$. On montre que :

- a. la variable aléatoire $X = \sum_{i=1}^n X_i \rightarrow N(nm, n\sigma^2)$
- b. $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \rightarrow N(m, \frac{\sigma^2}{n})$ \bar{X} est la moyenne échantillonnale
- c. $Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{X} - m)}{\sigma} \rightarrow N(0,1)$

3. Théorème Central Limite (T.C.L)

Soit $(X_i)_{i=1,\dots,n}$ une suite de variables aléatoires indépendantes identiquement distribuées(i.i.d) de moyenne $E(X_i) = m$ et $V(X_i) = \sigma^2 \quad \forall i = 1, 2, \dots, n$.

On montre qu' à partir de $n > 30$:

- a. la variable aléatoire $X = \sum_{i=1}^n X_i \rightarrow N(nm, n\sigma^2)$
- b. $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \rightarrow N(m, \frac{\sigma^2}{n})$ et $Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{X} - m)}{\sigma} \rightarrow N(0,1)$

V. Lois dérivées de la loi normale.

Parfois, on est amené à utiliser d'autres lois que la loi normale, entre autres la loi du khi-deux et la loi de Student très utiles en statistique inférentielle. Ces lois dépendent d'un paramètre ν entier, appelé degré de liberté (d.d.l)

De même que pour la loi normale $N(0,1)$, on disposera de tables statistiques pour ces lois.

1. La loi du Khi-deux (χ^2)

a. Définition

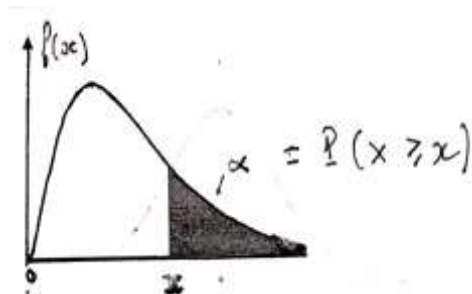
Soient X_1, X_2, \dots, X_n des variables aléatoires indépendantes de même loi $N(0,1)$.

Par définition la v.a $X = \sum_{i=1}^n X_i^2$ suit une loi du khi-deux à n degrés de liberté.

On note $X \rightarrow \chi^2_{(n)}$

Remarques :

- Si X_1 une variable aléatoire indépendante de loi $N(0,1)$ alors par définition la variable aléatoire $X = X_1^2$ suit une loi du khi-deux à 1 degré de liberté. On note $X \rightarrow \chi^2_{(1)}$
- Une variable de loi du χ^2 est toujours positive, donc sa densité n'est pas symétrique
- La loi du χ^2 admet une densité difficile à retenir qui nous est pas très utile car la loi est tabulée.
- La table donne la valeur $x = F^{-1}(1-\alpha)$ telle que $P(X > x) = \alpha$ qui le quantile d'ordre $(1-\alpha)$



b. Espérance et Variance

si la v.a $X \rightarrow \chi^2_{(\nu)}$ alors $E(X) = \nu$ et $V(X) = 2\nu$

c. Approximation par la loi normale

Si $n > 30$ et $X \rightarrow \chi^2_{(\nu)}$, on peut montrer d'après le TCL que $X \rightarrow N(\nu, 2\nu)$

d. Exemples de lecture de la table

Soit $X \rightarrow \chi^2_{(20)}$, trouver x tel que $P(X > x) = 0.10$

On croise la ligne $\nu=20$ et la colonne $\alpha = 0.10$, on lit $x = 28.412$

- Soit $X \rightarrow \chi^2_{(10)}$, trouver x tel que $P(X \leq x) = 0.05$

$$P(X \leq x) = 0.05 \Rightarrow P(X > x) = 0.95$$

On croise la ligne $\nu=10$ et la colonne $\alpha = 0.95$, on lit $x=3.940$

- Soit $X \rightarrow \chi^2_{(40)}$, trouver x tel que $P(X \leq x) = 0.05$

$\nu = 40 > 30$ alors on approxime la loi du $\chi^2_{(\nu)}$ par la loi $N(\nu, 2\nu)$

On considère que $X \rightarrow N(\nu, 2\nu)$

$$P(X \leq x) = \Phi\left(\frac{x-\nu}{\sqrt{2\nu}}\right) = \Phi(z) = 0.05, \quad z = \frac{x-\nu}{\sqrt{2\nu}} \Rightarrow x = \nu + \sqrt{2\nu} \times z$$

On a $\Phi(z) \leq \frac{1}{2} \Rightarrow z < 0 \Rightarrow -z > 0$. On calculera donc $\Phi(-z)$

$$\Phi(-z) = 1 - \Phi(z) = 1 - 0.05 = 0.95 \text{ n'existe pas sur la table,}$$

On a $0.9495 < 0.95 < 0.9505$, l'écart est le même de chaque côté, on approchera $\Phi(z)$ par l'une ou l'autre des deux valeurs.

$$\text{On prend } \Phi(-z) = 0.9495 = \Phi(1.64) \Rightarrow -z = 1.64 \Rightarrow z = -1.64$$

$$\Rightarrow x = \nu + \sqrt{2\nu} \times z = x = 40 + \sqrt{2 \times 40} \times (-1.64) = 25.33$$

e. Théorème

Soient X_1 et X_2 deux variables aléatoires indépendantes telles que

$X_1 \rightarrow \chi^2_{(\gamma_1)}$ et $X_2 \rightarrow \chi^2_{(\gamma_2)}$ alors on montre que $X_1 + X_2 \rightarrow \chi^2_{(\gamma_1 + \gamma_2)}$

Exercice 14 de la série

Soit T le taux sanguin (en u/ml) d'une substance. On note T_M ce taux chez les malades atteints de la maladie (notés:M) et T_{NM} chez les non malades (notés:NM). On connaît les lois de T_M et T_{NM} . Soit Y la v.a représentant la différence du taux sanguin de cette substance entre les malades et les non malades. $Y = T_M - T_{NM}$.

Malades M	Distribution de $T_M : N(\mu_M, \sigma_M^2)$	$\mu_M = 110$	$\sigma_M = 12$
Non Malades NM	Distribution de $T_{NM} : N(\mu_{NM}, \sigma_{NM}^2)$	$\mu_{NM} = 100$	$\sigma_{NM} = 10$

➤ **QCM 1** la v.a Y est de loi

- A. $N(\mu_M - \mu_{NM}, \sigma_M^2 - \sigma_{NM}^2)$ B. $N(\mu_M - \mu_{NM}, \sigma_M^2 + \sigma_{NM}^2)$ C. Khi-2 à 4ddl
D. Khi-2 à 2ddl E. une autre loi

➤ **QCM 2 :**

- A. $P(Y=10) = 0,5$ B. $P(Y < 15) = 0,3745$ C. $P(Y \leq 15) = 0,4129$
D. $P(Y \leq 15) = 0,5871$ E. $P(Y \leq 15) = 0,6255$

Solution

QCM1 réponse B

$$T_M \rightarrow N(\mu_M, \sigma_M^2), T_{NM} \rightarrow N(\mu_{NM}, \sigma_{NM}^2) \Rightarrow Y = T_M - T_{NM} \rightarrow N(m, \sigma^2)$$

les variables T_M et T_{NM} sont indépendantes alors

$$m = E(Y) = E(T_M - T_{NM}) = E(T_M) - E(T_{NM}) = \mu_M - \mu_{NM}$$

$$\sigma^2 = V(Y) = V(T_M - T_{NM}) = V(T_M) + V(T_{NM}) = \sigma_M^2 + \sigma_{NM}^2$$

QCM 2 réponse E

A faux car Y est une variable continue alors $P(Y=y) = 0 \quad \forall y \in \mathbb{R}$, ce qui est juste c'est $P(Y \leq 10) = 0,5$ car $E(Y) = 10$

$P(Y \leq 15) = ?$

$$Y \rightarrow N(m, \sigma^2) \quad m = \mu_M - \mu_{NM} = 110 - 100 = 10, \quad \sigma^2 = \sigma_M^2 + \sigma_{NM}^2 = 12^2 + 10^2 = 244$$

$$P(Y \leq 15) = \Phi\left(\frac{15-10}{\sqrt{244}}\right) = \Phi(0.32) = 0.6255$$

Exercice 15 de la série

On considère une variable X_1 distribuée selon une loi normale d'espérance 0 et de variance 4, et une variable X_2 distribuée selon une loi normale centrée réduite; X_1 et X_2 sont indépendantes

➤ QCM 1

- A. $\text{var}(X_1 + X_2) = 9$
- B. (donc $X_1/2$ et X_2 aussi), $\text{var}(X_1/2 + X_2) = 2$
- C. $(X_1/2 + X_2)$ suit une loi Normale centrée réduite.
- D. $X_1 + X_2$ suit une loi normale d'espérance 0 et de variance 4
- E. $X_1 + X_2$ suit une loi normale d'espérance 0 et de variance 5

➤ QCM 2:

- A. $E(X_1^2) = 0$ B. $E(X_1^2) = 4$ C. $E(X_1^2 + X_2^2) = 4$
- D. $X_1^2 + X_2^2$ est distribuée selon un χ^2 à 2 degrés de liberté
- E. $X_1^2/4 + X_2^2$ est distribuée selon un χ^2 à 2 degrés de liberté

Solution**QCM1 réponses B et E**

$X_1 \rightarrow N(0,4)$, $X_2 \rightarrow N(0,1)$, X_1 et X_2 sont indépendantes

- A. $V(X_1 + X_2) = V(X_1) + V(X_2) = 4 + 1 = 5$ alors A est faux
- B. $V\left(\frac{X_1}{2} + X_2\right) = \left(\frac{1}{2}\right)^2 V(X_1) + V(X_2) = 2$ alors B est juste
- C. X_1 et X_2 sont de loi normales alors $\frac{X_1}{2} + X_2$ est de loi normale
 $E\left(\frac{X_1}{2} + X_2\right) = \frac{1}{2}E(X_1) + E(X_2) = 0$ et $V\left(\frac{X_1}{2} + X_2\right) = 2$ alors C est faux
- D. $V(X_1 + X_2) = V(X_1) + V(X_2) = 5$ alors D est faux
- E. X_1 et X_2 sont de loi normales alors $X_1 + X_2$ est de loi normale
 $E(X_1 + X_2) = E(X_1) + E(X_2) = 0 + 0 = 0$ et $V(X_1 + X_2) = 5$ alors E est juste

QCM 2 réponse E

$V(X_1) = E(X_1^2) - E^2(X_1) \Rightarrow E(X_1^2) = V(X_1) + E^2(X_1) = 4 + 0 = 4$ A faux, B juste, C faux
D est faux car X_1 n'est pas une $N(0,1)$

E. $\frac{X_1}{2} \rightarrow N(0,1)$ alors $\left(\frac{X_1}{2}\right)^2 \rightarrow \chi_{(1)}^2$, de même $X_2 \rightarrow N(0,1)$ alors $X_2^2 \rightarrow \chi_{(1)}^2$, de plus

X_1 et X_2 sont indépendantes, on conclut que $\left(\frac{X_1}{2}\right)^2 + X_2^2 \rightarrow \chi_{(1+1)}^2 = \chi_{(2)}^2$ alors E est juste

Exercice 17 de la série

Dans une certaine pathologie, la durée d'hospitalisation est distribuée selon un χ^2 à 18 degrés de liberté.

- A. L'écart-type de la durée d'hospitalisation est 6 jours
- B. 50% des sujets ont une durée d'hospitalisation supérieure à 18 jours
- C. Plus de 5% des patients restent hospitalisés au moins 4 semaines
- D. Plus de 10% des patients restent hospitalisés au moins 4 semaines
- E. Moins de 25% des patients restent hospitalisés entre 3 et 4 semaines

solution A, C et E

Soit X : la durée d'hospitalisation, $X \rightarrow \chi_{(\nu)}^2$, $\nu=18$

- A. On sait que $V(X) = 2\nu = 36 \Rightarrow \sigma(X) = \sqrt{36} = 6$ A vraie
- B. $P(X < 18) = 0.05$, or $E(X) = \nu = 18$, la loi du $\chi^2_{(\nu)}$ n'est pas symétrique B fausse
- C. 4 semaines = 28 jours. La table donne $P(X > 28.87) = 0.05$,
 $28 < 28.87$ alors $P(X > 28) > P(X > 28.87)$ C vraie
- D. La table donne $P(X > 25.99) = 0.1$, $28 > 25.99$ alors $P(X > 28) < P(X > 25.99)$ D fausse
- E. La table donne $P(X > 20.60) = 0.3$ et $P(X > 28.87) = 0.05 \Rightarrow$
 $P(20.60 < X < 28.87) = 0.3 - 0.05 = 0.25$
 comme $[21, 28] \subset [20.60, 28.87] \Rightarrow P(21 < X < 28) < P(20.60 < X < 28.87)$ E vraie

2. Loi de Student

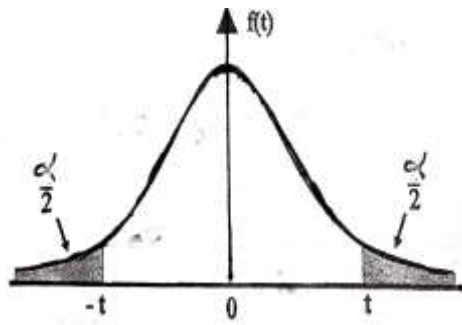
a. définition

Soient X et Y deux variables aléatoires indépendantes de lois respectives $N(0,1)$ et $\chi^2_{(\nu)}$.

Posons
$$T = \frac{X}{\sqrt{\frac{Y}{\nu}}} = \sqrt{\nu} \frac{X}{\sqrt{Y}}$$

Par définition la variable aléatoire T suit une loi de Student à ν degrés de liberté

On note $T \rightarrow St_{(\nu)}$ ou bien $T \rightarrow t_{(\nu)}$



b. propriétés

- La densité de la variable T est en forme de « cloche », mais plus aplatie que celle d'une loi normale de mêmes paramètres. Elle est symétrique par rapport à l'axe $t=0$, on retrouve donc les propriétés de symétrie de la loi $N(0,1)$ soit $P(T < -t) = P(T > t)$
- La loi de Student admet une densité difficile à retenir qui nous est pas très utile car la loi est tabulée.
- La table donne la valeur positive t telle que $P(|T| > t) = \alpha = P(T < -t) + P(T > t) = \frac{\alpha}{2} + \frac{\alpha}{2}$

c. Espérance et Variance

si la v.a $T \rightarrow St_{(\nu)}$ alors $E(T) = 0$ et $V(T) = \frac{\nu}{\nu - 2} \quad \forall \nu > 2$

d. Approximation par la loi normale

Si $n > 30$ et $T \rightarrow St_{(\nu)}$ alors on peut l'approcher par $T \rightarrow N(0, 1)$

e. Exemples de lecture de la table

- Soit $T \rightarrow St_{(20)}$, trouver t tel que $P(|T| > t) = 0.05$

On croise la ligne $\nu = 20$ et la colonne $\alpha = 0.05$, on lit $t = 2.086$

- Soit $T \rightarrow St_{(10)}$, trouver t tel que $P(T > t) = 0.1$

La table donne $P(|T| > t) = \alpha = P(T < -t) + P(T > t) = \frac{\alpha}{2} + \frac{\alpha}{2}$; $P(T > t) = \frac{\alpha}{2} = 0.1$ donc $\alpha = 0.20$

On croise la ligne $\nu = 10$ et la colonne $\alpha = 0.2$, on lit $t = 1.372$

- Soit $T \rightarrow St_{(15)}$, trouver t tel que $P(T < -t) = 0.01$

Idem $P(T < -t) = \frac{\alpha}{2} = 0.01$ donc $\alpha = 0.02$

On croise la ligne $\nu = 15$ et la colonne $\alpha = 0.02$, on lit $t = 2.602$ alors $-t = -2.602$

- Soit $T \rightarrow St_{(35)}$, trouver t tel que $P(T < t) = 0.025$

$\nu = 35 > 30$ alors $T \rightarrow N(0, 1)$

$P(T < t) = 0.025 = \Phi(t) < 0.025$ donc $\Phi(-t) = 1 - \Phi(t) = 0.975 \Rightarrow -t = 1.96 \Rightarrow t = -1.96$