Test Classique 1 (Statistiques et probabilités)

Mohamed Regragui

Ing2 – EPITA

Ce document est une version dactylographiée du cours de M. Mohamed Regragui, dispensé aux étudiants en deuxième année du cycle ingénieur de l'EPITA. Il a été réalisé par Rémi Berson et Benjamin Roux (respectivement SCIA et CSI 2014). Ce cours n'a pas vocation à être diffusé en dehors de l'EPITA.

Table des matières

1

Rappels de la loi Gamma et de la loi gaussienne

1.1 Loi gamma

Définition. On dit qu'une variable aléatoire positive X suit une loi gamma de paramètre r notée γ_r , si sa densité est donnée par :

$$f(X) = \frac{1}{\Gamma(r)} \cdot e^{-x} \cdot X^{r-1}, \quad \forall X > 0 \tag{1.1}$$

où la fonction $\Gamma(X)$ est définie par :

$$\Gamma(X) = \int_0^{+\infty} e^{-t} \cdot t^{X-1} dX, \quad \forall X > 0 \tag{1.2}$$

Propriété.

$$\Gamma(X+1) = x \cdot \Gamma(X) \tag{1.3}$$

$$\Gamma(1) = 1 \tag{1.4}$$

$$\Gamma(n+1) = n! \qquad \forall n \in \mathbb{N}^* \tag{1.5}$$

$$\lim_{X \to 0^+} \Gamma(X) = +\infty \tag{1.6}$$

$$\Gamma(k + \frac{1}{2}) = \frac{1 \times 3 \times 5 \times \cdots \times (2k-1)}{2^k} \cdot \Gamma(\frac{1}{2}) \qquad \forall n \in \mathbb{N}^* \tag{1.7}$$

$$\Gamma(\frac{1}{2}) = \sqrt{\pi} \tag{1.8}$$

Remarque. La fonction gamma est généralement perçue comme un prolongement de la factorielle à l'ensemble des nombres complexes (excepté les entiers négatifs ou nuls).

Espérance de la loi γ_r

$$\begin{aligned} E(X) &= \int_0^{+\infty} X \cdot f(x) \cdot dx \\ &= \frac{1}{\Gamma(r)} \int_0^{+\infty} X^r \cdot e^{-x} \cdot dx \\ &= \frac{\Gamma(r+1)}{\Gamma(r)} \\ E(X) &= r \end{aligned} \tag{1.9}$$

Variance de la loi γ_r

$$V(X) = E(X^2) - E^2(X)$$

Calculons $E(X^2)$:

$$\begin{aligned} E(X^2) &= \int_0^{+\infty} \cdot X^2 \cdot f(X) \cdot dX \\ &= \frac{1}{\Gamma(r)} \cdot \int_0^{+\infty} X^{r+1} e^{-X} dX \\ &= \frac{\Gamma(r+2)}{\Gamma(r)} \\ &= \frac{\Gamma(r+2)}{\Gamma(r+1)} \cdot \frac{\Gamma(r+1)}{\Gamma(r)} \\ &= (r+1) \cdot r \end{aligned}$$

On obtient donc :

$$V(X) = r \tag{1.10}$$

1.2 LOI DE LAPLACE-GAUSS (LOI NORMALE)

Définition. On dit qu'une variable aléatoire X suit une loi gaussienne Laplace–Gauss de paramètres (m, σ), si sa densité est :

$$f(X) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot (\frac{X-m}{\sigma})^2} \tag{1.11}$$

avec,

$$m = E(X) \qquad \text{(espérance)}$$
$$\sigma = \sqrt{V(x)} \qquad \text{(écart-type)}$$

Soit $U = \frac{X-m}{\sigma}$ une variable normale centrée et réduite. On peut dire que U suit une loi $L.G\ (0,1)$ ($m = 0$ et $\sigma = 1$). On dit qu'elle est centrée car la courbe de la Gaussienne est centrée sur l'axe des abscisse et réduite parce $f(X)$ est compris entre 0 et 1. De cette manière, on pourra utiliser une table (voir Annexe) contenant les valeurs de la gaussienne, pour toutes les variables aléatoires suivant une loi $L.G$ à condition de la réduire et de la centrer.

La fonction de répartition est la suivante :

$$F(u) = P[U < u] = \int_{\mathbb{R}} f(t) \cdot dt$$

et sa densité est :

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

Propriété. *$V(U) = 1$ avec U qui suit une loi $L.G\ (0,1)$*

Preuve.

$$\begin{aligned} V(U) &= E(U^2) - \cancel{E^2(U)} \\ &= E(U^2) \\ &= \int_{-\infty}^{+\infty} \frac{u^2}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{u^2}{2}} \cdot du \\ &= \frac{2}{\sqrt{2 \cdot \pi}} \cdot \int_{0}^{+\infty} u^2 \cdot e^{-\frac{u^2}{2}} \cdot du \qquad \text{(Car la fonction est paire)} \end{aligned}$$

Posons,

$$\begin{aligned} t &= \frac{u^2}{2} \\ \Rightarrow dt &= u \cdot du \\ \Rightarrow du &= \frac{dt}{u} = \frac{dt}{\sqrt{2 \cdot t}} \end{aligned}$$

En appliquant le changement de variable on obtient,

$$
\begin{aligned}
V(U) &= \frac{2}{\sqrt{2\cdot\pi}} \cdot \int_0^{+\infty} 2\cdot t\cdot e^{-t}\,\frac{dt}{\sqrt{2\cdot t}} \\
&= \frac{2}{\sqrt{\pi}} \cdot \int_0^{+\infty} t^{\frac{1}{2}}\cdot e^{-t}\cdot dt \\
&= \frac{2}{\sqrt{2\cdot\pi}} \cdot \int_0^{+\infty} t^{2}\cdot e^{-t}\cdot dt \\
&= \frac{\cancel{2}}{\sqrt{\pi}} \cdot \frac{1}{\cancel{2}} \cdot \Gamma(\frac{1}{2}) \\
&= \frac{\Gamma(\frac{1}{2})}{\sqrt{\pi}} \\
&= \frac{\cancel{\sqrt{\pi}}}{\cancel{\sqrt{\pi}}} \\
V(U) &= 1
\end{aligned}
$$

1.2.1 *Calcul des moments*

$$
\begin{aligned}
\mu_k &= E(U^k) \\
&= \int_{-\infty}^{+\infty} U^k \cdot f(u)\cdot du \\
&= \int_{-\infty}^{+\infty} U^k \frac{1}{\sqrt{2\pi}} e^{-\frac{U^2}{2}} du \qquad\qquad U \nearrow L.G\,(0,1)
\end{aligned}
$$

Calculons les moments dans le cas général. Observons le comportement de μ_k en fonction de la parité de k

***k* impair**

$$\mu_k = \int_{-\infty}^{+\infty} U^k \cdot \frac{1}{\sqrt{2\cdot\pi}} \cdot e^{-\frac{U^2}{2}} \cdot du = 0$$

Car $U^k \frac{1}{\sqrt{2\pi}} e^{-\frac{U^2}{2}}$ est impaire si k est impair et on l'intègre sur le domaine $]-\infty,+\infty[$.

k **pair**

$$\begin{aligned}\mu_{2k} &= E(U^{2k}) \\ &= \frac{1}{\sqrt{2\cdot\pi}}\cdot\int_{-\infty}^{+\infty} U^{2k}\cdot e^{-\frac{U^2}{2}}\cdot dU \\ &= \frac{2}{\sqrt{2\cdot\pi}}\cdot\int_{0}^{+\infty} U^{2k}\cdot e^{-\frac{U^2}{2}}\cdot dU\end{aligned}$$

Posons,

$$\begin{aligned} t &= \frac{U^2}{2} \\ \Rightarrow dt &= U\cdot dU \\ \Rightarrow U &= \sqrt{2\cdot t}\end{aligned}$$

On effectue le changement de variable,

$$\begin{aligned}\mu_{2k} &= \frac{2}{\sqrt{2\cdot\pi}}\cdot\int_{0}^{+\infty}(2\cdot t)^k\cdot e^{-t}\cdot\frac{dt}{\sqrt{2\cdot t}} \\ &= \frac{2^k}{\sqrt{\pi}}\cdot\int_0^{+\infty} t^{k-\frac{1}{2}}\cdot e^{-t}\cdot dt \\ &= \frac{2^k}{\sqrt{\pi}}\cdot\Gamma(k+\frac{1}{2}) \\ &= \frac{\cancel{2^k}}{\cancel{\sqrt{\pi}}}\cdot\frac{1\times 3\times 5\times\cdots\times(2\cdot k-1)}{\cancel{2^k}}\cdot\cancel{\Gamma(\frac{1}{2})} \\ &= 1\times 3\times 5\times\cdots\times(2k-1) \\ &= \frac{(2k)!}{2\times 4\times 6\times 8\times\cdots\times 2k} \\ &= \frac{(2k!)}{2^k k!}\end{aligned}$$

Grâce à cette formule générale, on pourra par exemple calculer directement :

$$\begin{aligned}\mu_4 &= 3 \\ \mu_2 &= 1\end{aligned}$$

2

Fonctions caractéristiques

Définition. La *fonction caractéristique* d'une variable réelle X est la transformée de *Fourier* de sa loi de probabilité. On la note :

$$\varphi_X(t) = E[e^{itX}] \quad i \in \mathbb{C},\ i^2 = -1 \tag{2.1}$$

Cas continu

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itX} f(X) dX$$

Cas discret

$$\varphi_X(t) = \sum_k e^{itk} P[X = k]$$

Propriété.

$$\varphi_{\lambda X}(t) = \varphi_X(\lambda t), \quad \forall \lambda \in \mathbb{C} \tag{2.2}$$

$$\varphi_{X+a}(t) = e^{ita} \cdot \phi_X(t), \quad \forall a \in \mathbb{C} \tag{2.3}$$

Conséquence. Si on pose $U = \frac{X-m}{\sigma}$, $\quad \lambda = \frac{X}{\sigma}$ et $a = -m$, on obtient :

$$\varphi_U(t) = e^{-\frac{itm}{\sigma}} \cdot \varphi_X(\frac{t}{\sigma})$$

Si $\frac{t}{\sigma} = v$ alors :

$$\varphi_X(v) = e^{ivm} \cdot \phi_U(\sigma v) \tag{2.4}$$

Proposition.

$$\varphi_X(0) = 1 \tag{2.5}$$

$$\underbrace{\varphi_X^{(k)}(0)}_{\text{Dérivée d'ordre } k} = i^k \cdot E[X^k] \tag{2.6}$$

Preuve. Supposons X une variable aléatoire continue, l'expression de la fonction caractéristique est donc :

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itX} \cdot f(X) \cdot dX$$

On en déduit immédiatement,

$$\varphi_X(0) = \int_{\mathbb{R}} f(X) \cdot dX$$

$$\varphi_X(0) = 1 \qquad \text{Car l'intégrale d'une densité vaut toujours 1}$$

La proposition (2.5) est donc vérifiée.

Afin de démontrer la proposition (2.6) on se propose de dériver l'expression de la fonction caractéristique ci-dessus k fois par rapport à t.

En dérivant une première fois on obtient :

$$\begin{aligned}\varphi_X'(t) &= \int_{\mathbb{R}} \frac{\partial}{\partial t}(e^{itX} \cdot f(X)) \cdot dX \\ &= \int_{\mathbb{R}} iXe^{itX} \cdot f(X) \cdot dX\end{aligned}$$

Si on dérive k fois on obtient :

$$\begin{aligned}\varphi_X^{(k)}(t) &= \int_{\mathbb{R}} (iX)^k e^{itX} \cdot f(X) \cdot dX \\ \varphi_X^{(k)}(0) &= i^k \int_{\mathbb{R}} X^k \cdot f(X) \cdot dX \\ &= i^k E[X^k]\end{aligned}$$

On en déduit la formule de *Mac-Laurin* :

$$\varphi_X(t) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} \cdot i^k \cdot E[X^k] \tag{2.7}$$

Remarque. La fonction caractéristique d'une somme de variables aléatoires indépendantes est égale au produit de leurs fonctions caractéristiques.

$$\begin{aligned}
\varphi_{X_1+X_2}(t) &= E[e^{it(X_1+X_2)}] \\
&= E[e^{itX_1}e^{itX_2}] \\
&= E[e^{itX_1}] \cdot E[e^{itX_2}] \quad \text{On peut séparer car } X_1 \text{ et } X_2 \text{ sont indépendantes} \\
&= \varphi_{X_1}(t) \cdot \varphi_{X_2}(t)
\end{aligned}$$

On peut généraliser cette formule :

$$\varphi_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^{n} \varphi_{X_i}(t) \tag{2.8}$$

3

Exercices sur les fonctions caractéristiques

Déterminer les fonctions caractéristiques dans les cas suivants :

1. *X suit une loi $B(p)$ (loi de Bernoulli) ;*
2. *X suit une loi $B(n, p)$ (loi Binomiale) ;*
3. *X suit une loi $P(\lambda)$ (loi de Poisson) ;*
4. *X suit une loi $U_{[-a,a]}$ (loi uniforme) ;*
5. *X suit une loi $L.G\,(0, 1)$ (loi gaussienne).*

3.1 Loi de Bernoulli

La variable aléatoire X suit une loi de Bernoulli, ce qui signifie :

$$f(X) = \begin{cases} 1 & \text{si un évènement } A \text{ se réalise avec une probabilité } p \\ 0 & \text{si un évènement } \bar{A} \text{ se réalise avec une probabilité } 1-p \end{cases}$$

On peut donc appliquer la définition dans le cas discret :

$$\begin{aligned} \phi_X(t) &= \sum_{k=0}^{1} e^{itk} \cdot P[X = k] \\ &= P[X = 0] + e^{it} \cdot P[X = 1] \\ &= 1 - p + p \cdot e^{it} \\ &= q + p \cdot e^{it} && \text{avec } q = 1 - p \end{aligned}$$

3.2 LOI BINOMIALE

La variable aléatoire X suit une loi Binomiale, ce qui signifie :

$$X = \sum_{i=1}^{n} X_i, \quad \text{somme de variables aléatoires indépendantes de } B(p)$$

On peut directement appliquer la remarque sur les fonctions caractéristiques de somme de variables aléatoires indépendantes :

$$\begin{aligned}
\phi_X(t) &= \phi_{\sum_{i=1}^{n} X_i}(t) \\
&= \prod_{i=1}^{n} \phi_{X_i}(t) \\
&= (q + p \cdot e^{it})^n
\end{aligned}$$

3.3 LOI DE POISSON

La variable aléatoire X suit une loi de Poisson, ce qui signifie :

$$P[X = k] = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \quad \forall k \in \mathbb{N}$$

On applique la définition discrète de la fonction caractéristique :

$$\begin{aligned}
\phi_X(t) &= \sum_{k=0}^{+\infty} e^{itk} \cdot P[X = k] \\
&= e^{-\lambda} \cdot \sum_{k=0}^{+\infty} e^{itk} \cdot \frac{\lambda^k}{k!}
\end{aligned}$$

Rappel : $\sum_0^{+\infty} \frac{X^k}{k!} = e^X$

$$\begin{aligned}
\phi_X(t) &= e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(e^{it}\lambda)^k}{k!} \\
&= e^{-\lambda} e^{\lambda e^{it}} \\
&= e^{\lambda(e^{it}-1)}
\end{aligned}$$

3.4 LOI UNIFORME

La variable aléatoire X suit une loi uniforme, ce qui signifie :

$$\text{Sa densité est } f(X) = \begin{cases} \frac{1}{2a} & \text{si } X \in [-a, a] \\ 0 & \text{sinon} \end{cases}$$

On applique la définition continue de la fonction caractéristique :

$$\begin{aligned} \phi_X(t) &= \int_{-a}^{a} e^{itX} \cdot \frac{1}{2a} \cdot dX \\ &= \frac{1}{2a} \cdot \int_{-a}^{a} e^{itX} \cdot dX \\ &= \frac{1}{2a} \cdot \left[\frac{e^{itX}}{it} \right]_{-a}^{a} \\ &= \frac{1}{2ait} \cdot (e^{ita} - e^{-ita}) \end{aligned}$$

Rappel 1 : $a - \bar{a} = 2i \times \mathfrak{Im}(a), \quad \forall a \in \mathbb{C}$
Rappel 2 : $re^{i\theta} = r \cdot (\cos(\theta) + i \sin(\theta))$

$$\begin{aligned} \phi_X(t) &= \frac{\cancel{2i} \cdot \sin(at)}{\cancel{2}a\cancel{i}t} \\ &= \frac{\sin(at)}{at} \end{aligned}$$

3.5 LOI GAUSSIENNE

La variable aléatoire X suit une loi gaussienne. On peut utiliser la formule de Mac-Laurin définie dans le cours :

$$\phi_X(t) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} \cdot i^k \cdot E[X^k]$$

On sait que $E[X^k] = 0$ si k est impair, donc on peut réduire la somme aux termes pairs :

$$\begin{aligned} \phi_X(t) &= \sum_{k=0}^{+\infty} \frac{t^{2k}}{\cancel{(2k)!}} \cdot \cancel{i^{2k}} \cdot \frac{\cancel{(2k)!}}{2^k k!} \\ &= \sum_{k=0}^{+\infty} \frac{(-\frac{t^2}{2})^k}{k!} \\ &= e^{-\frac{t^2}{2}} \end{aligned}$$

4

Exercices sur la loi normale

Nous rappelons la fonction de répartition de U lorsqu'elle suit une $L.G\ (0,1)$:

$$F(u) = P[U < u] = \int_{-\infty}^{u} f(t) \cdot dt, \quad \text{avec} \quad f(t) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot t^2}$$

Il est à noter que pour calculer les valeurs négatives de u on peut utiliser la formule suivant :

$$F(-a) = 1 - F(a)$$

Afin de calculer la probabilité d'un intervalle, rappelons qu'il convient d'utiliser la formule :

$$P(a < U < b) = F(b) - F(a)$$

Afin de réaliser les exercices suivants, il est nécessaire de savoir lire les valeurs de $L.G\ (0,1)$ dans la table fournie en annexe. Pour trouver la valeur de $F(a.bc)$, il suffit de lire la valeur à l'intersection de la ligne $[a.b]$ et de la colonne $[0.0c]$. Par exemple, pour calculer $F(0.93)$, regardons la valeur à l'intersection de la ligne $[0.9]$ et de la colonne $[0.03]$, pour obtenir 0.8238.

4.1 Exercice 1

Calculer les probabilités suivante en faisant usage de la table en annexe :

1. $P[U < 2.04]$

$$\begin{aligned} P[U < 2.04] &= F(2.04) \\ &= 0.9793 \end{aligned}$$

2. $P[U < -1.95]$

$$\begin{aligned} P[U < -1.95] &= 1 - F(1.95) \\ &= 0.0256 \end{aligned}$$

3. $P[0 < U < 2]$

$$\begin{aligned} P[0 < U < 2] &= F(2) - F(0) \\ &= 0.9772 - 0.5 \\ &= 0.4772 \end{aligned}$$

4. $P[-1 < U < 2]$

$$\begin{aligned} P[-1 < U < 2] &= F(2) - F(-1) \\ &= F(2) - 1 + F(1) \\ &= 0.8185 \end{aligned}$$

5. $P[-3 < U < -1]$

$$\begin{aligned} P[-3 < U < -1] &= F(-1) - F(-3) \\ &= 1 - F(1) - 1 + F(3) \\ &= F(3) - F(1) \\ &= 0.15735 \end{aligned}$$

6. $P[U > -2]$

$$\begin{aligned} P[U > -2] &= 1 - P[U \leq -2] \\ &= 1 - F(-2) \\ &= F(2) \\ &= 0.9772 \end{aligned}$$

7. $P[|U| < 2]$

$$\begin{aligned} P[|U| < 2] &= P[-2 < U < 2] \\ &= F(2) - F(-2) \\ &= 2 \times F(2) \\ &= 0.9544 \end{aligned}$$

8. $P[|U| > 1]$

$$\begin{aligned} P[|U| > 1] &= 1 - P[|U| \leq 1] \\ &= 1 - P[-1 < U < 1] \\ &= 1 - (F(1) - F(-1)) \\ &= 1 - (F(1) - 1 + F(1)) \\ &= 2 \times (1 - F(1)) \\ &= 0.3174 \end{aligned}$$

4.2 EXERCICE 2

Soit T suivant une loi $L.G(0,1)$. Déterminer $t \in \mathbb{R}$ dans chacun des cas suivants :

1. $P[T < t] = 0.8238$

$$\begin{aligned} & P[T < t] = 0.8238 \\ & \Rightarrow t = 0.93 && \text{en regardant dans la table directement} \end{aligned}$$

2. $P[T < t] = 0.1112$

$$\begin{aligned} & P[T < t] = 0.1112 \\ & \Rightarrow F(t) = 0.1112 \\ & \Rightarrow t < 0 \\ & \Rightarrow 1 - F(-t) = 0.1112 \\ & \Rightarrow F(-t) = 1 - 0.1112 \\ & \Rightarrow F(-t) = 0.8888 \\ & \Rightarrow -t = 1.22 \\ & \Rightarrow t = -1.22 \end{aligned}$$

3. $P[t < T < 1] = 0.6826$

$$\begin{aligned} & P[t < T < 1] = 0.6826 \\ & \Rightarrow F(1) - F(t) = 0.6826 \\ & \Rightarrow F(t) = F(1) - 0.6826 = 0.1587 \\ & \Rightarrow F(-t) = 1 - 0.1587 = 0.8413 \\ & \Rightarrow -t = 1 \\ & \Rightarrow t = -1 \end{aligned}$$

4. $P[0 < T < t] = 0.4878$

$$\begin{aligned} & P[0 < T < t] = 0.4878 \\ & \Rightarrow F(t) - F(0) = 0.4878 \\ & \Rightarrow F(t) = 0.5 + 0.4878 \\ & \Rightarrow t = 2.25 \end{aligned}$$

5. $P[|T| < t] = 0.95$

$$\begin{aligned} & P[|T| < t] = 0.95 \\ & \Rightarrow F(t) - F(-t) = 0.95 \\ & \Rightarrow 2 \times F(t) - 1 = 0.95 \\ & \Rightarrow F(t) = \frac{1.95}{2} = 0.975 \\ & \Rightarrow t = 1.96 \end{aligned}$$

4.3 EXERCICE 3

En une minute, une machine fabrique 3 pièces. Pendant un réglage, elle est arrêtée. La durée D en minutes d'un réglage suit une loi $L.G(20,3)$.

1. Quel est le nombre moyen de pièces non-fabriquées à cause d'un réglage ?

   Soit D la durée qui suit une loi $L.G(20,3)$. On a donc :

   $$\begin{cases} m = E(D) = 20 \\ \sigma = \sqrt{V(D)} = 3 \end{cases}$$

   Soit X la variable aléatoire représentant le nombre de pièces non-fabriquées à cause du réglage. On a :

   $$X = 3 \times D \Rightarrow E(X) = 3 \times E(D) = 3 \times 20 = 60$$

2. Quelle est la probabilité que le réglage entraîne une perte supérieure à 78 pièces ?

   On cherche $P[X > 78] = 1 - P[X \leq 78]$. On a,

   $$\begin{aligned} V(X) &= V(3 \times D) \\ &= 9 \times V(D) \qquad\qquad \text{car,} \quad V(\alpha \times X) = \alpha^2 \times V(X) \\ &= 9 \times 9 \\ &= 81 \end{aligned}$$

   d'où $\sigma(X) = 9$.
   On en déduit donc que X suit une loi $L.G(m = 60, \sigma = 9)$. Si on centre-réduit on obtient :

   $$P[X > 78] = 1 - P\left[\frac{X-m}{\sigma} \leq \frac{78-m}{\sigma}\right]$$

   avec, $U = \frac{X-60}{9}$ une variable aléatoire centrée et réduite (c'est-à-dire qu'elle suite une loi $L.G(0,1)$). On peut donc utiliser la table en annexe pour calculer la probabilité $P[X > 78]$:

   $$\begin{aligned} P[X > 78] &= 1 - P\left[\frac{X-60}{9} \leq \frac{78-60}{9}\right] \\ &= 1 - F(2) \\ &= 1 - 0.9772 \\ &= 0.0228 \end{aligned}$$

4.4 EXERCICE 4

Une usine fabrique des billes dont le diamètre, D, suit une loi continue de densité :

$$f(X) = \frac{1}{\sqrt{8\pi}} e^{-\left(\frac{X-100}{2\sqrt{2}}\right)^2}$$

1. Quelle est la probabilité pour une bille quelconque d'avoir un diamètre dans $[95 \text{ mm}, 105 \text{ mm}]$?

   Rappel Si X suit une loi $L.G(m, \sigma)$, alors sa densité est :

   $$g(X) = \frac{1}{2\sqrt{2\pi}} e^{-\left(\frac{1}{2}\left(\frac{X-m}{\sigma}\right)^2\right)}$$

   Or on remarque que,

   $$f(X) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-100}{2}\right)^2}$$

   On en déduit que D suit une loi $L.G(m = 100, \sigma = 2)$. Afin de répondre à la question, on cherche à calculer $P[95 \leq D \leq 105]$. Pour cela, on va centrer et réduire notre variable aléatoire afin de pouvoir utiliser la table fournie en annexe :

   $$P[95 \leq D \leq 105] = P\left[\frac{-5}{2} \leq \frac{D-100}{2} \leq \frac{5}{2}\right]$$

   $U = \frac{D-100}{2}$ suit bien une loi $L.G(0, 1)$. On peut donc calculer sa valeur avec la table en annexe :

   $$\begin{aligned} P[95 \leq D \leq 105] &= F\left(\frac{5}{2}\right) - F\left(\frac{-5}{2}\right) \\ &= 2 \times F\left(\frac{5}{2}\right) - 1 \\ &= 0.9876 \end{aligned}$$

2. Trouver l'intervalle centré sur $E(D)$ contenant 82% de la production.

   On cherche $I = [a, b]$ tel que $P[D \in I] = 0.82$

   Soit r le rayon de I (c'est à dire la *distance* entre $E(D)$ et a). On a :

   $$\begin{cases} b = E(D) + r \\ a = E(D) - r \end{cases}$$

   On cherche r tel que $P[|D - E(D)| \leq r] = 82\%$. On remarque que la variable aléatoire $|D - E(D)|$ est centrée, mais pas réduite. On la réduit :

$$P[|D-E(D)| \le r] = P\left[\frac{|D-E(D)|}{\sigma} \le \frac{r}{\sigma}\right]$$

$U = \frac{D-E(D)}{\sigma}$ suit une loi $L.G(0,1)$. Ce qui nous permet d'utiliser la table en annexe pour calculer la probabilité $P[|D-E(D)| \le r]$:

$$\begin{aligned}
& P\left[|U| \le \frac{r}{\sigma}\right] = 0.82 \\
& \Rightarrow F\left(\frac{r}{\sigma}\right) - F\left(-\frac{r}{\sigma}\right) = 0.82 \\
& \Rightarrow 2 \times F\left(\frac{r}{\sigma}\right) - 1 = 0.82 \\
& \Rightarrow F\left(\frac{r}{\sigma}\right) = \frac{1.82}{2} = 0.91 \\
& \Rightarrow \frac{r}{\sigma} = 1.34 && \text{, cf. table en annexe} \\
& \Rightarrow r = 2.68mm
\end{aligned}$$

On conclue donc que $I = [97.32, 102.68]$.

3. Un premier contrôle permet de répartir la production en 2 lots :
   - L_1 correspond à $D \notin [95, 105]$;
   - L_2 correspond à $D \in [95, 105]$.

   Quelle est la probabilité pour qu'une bille soit telle que son diamètre $D \in [95, 102]$?

   a. Quand elle appartient au lot L_1 ?

   C'est impossible car $[95, 102] \subset [95, 105]$ donc $P[D \in [95,\ 102] \mid D \notin [95,\ 105]]$ est absurde. On en déduit $P[D \in [95, 102] \mid L_1] = P[\emptyset] = 0$.

   b. Quand elle appartient au lot L_2 ?

   On cherche à calculer $P[95 \le D \le 102 \mid L_2]$.

   Rappel

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

   donc,

$$P[95 \le D \le 102|L_2] = \frac{P[(95 \le D \le 102) \cap L_2]}{P[L_2]}$$

   Nous avons déjà calculé le dénominateur dans une question précédente, il nous reste donc le numérateur. Pour cela, centrons et réduisons la variable aléatoire D :

$$\begin{aligned} P[95 \le D \le 102] &= P\left[\frac{95-100}{2} \le \frac{D-100}{2} \le \frac{102-100}{2}\right] \\ &= P\left[-\frac{5}{2} \le U \le 1\right] \end{aligned}$$

avec, $U = \frac{D-100}{2}$ une variable aléatoire centrée et réduite suivant une loi $L.G(0,1)$.

$$\begin{aligned} P[95 \le D \le 102] &= F(1) - F\left(-\frac{5}{2}\right) \\ &= F(1) + F\left(\frac{5}{2}\right) - 1 \\ &= 0.8351 \end{aligned}$$

finalement on a,

$$P[95 \le D \le 102 \mid L_2] = \frac{0.8351}{0.9876} = 0.8456$$

4. On extrait de la production totale, un échantillon de taille $n = 16$. Soit Δ le diamètre moyen de l'échantillon. Quelle est la probabilité $P[\Delta > 99]$?

Définition. Un *échantillon* D de taille n est une suite $(D_1, \ldots, D_n)$ de variables aléatoires indépendantes et de même loi D. On appelle *diamètre moyen de l'échantillon* la moyenne pondérée définie comme suit :

$$\Delta = \frac{\sum_{i=1}^{n} D_i}{n}$$

On cherche $P[\Delta > 99])1 - P[\Delta \le 99]$. Commençons par déterminer $E(\Delta)$ et $V(\Delta)$.

$$\begin{aligned} E(\Delta) &= \frac{\sum_{i=1}^{n} E(D_i)}{n} \\ &= \frac{n \times E(D)}{n} \\ &= 100 \\ V(\Delta) &= \frac{\sum_{i=1}^{n} V(D_i)}{n^2} && \text{car les } D_i \text{ sont indépendants} \\ &= \frac{n \times \sigma^2}{n^2} \\ &= \frac{\sigma^2}{n} \end{aligned}$$

Dans notre cas, $\sigma' = \sqrt{V(\Delta)} = \frac{\sigma}{\sqrt{n}} = \frac{1}{2}$. On en déduit donc que Δ suit une loi $L.G(100, \frac{1}{2})$. Nous pouvons maintenant centrer et réduire Δ afin de calculer $P[\Delta > 99]$:

$$\begin{aligned}
P[\Delta > 99] &= 1 - P[\Delta \leq 99] \\
&= 1 - P[2 \times (\Delta - 100) \leq 2 \times (99 - 100)] \\
&= 1 - F(-2) \\
&= F(2) \\
&= 0.9772
\end{aligned}$$

5

Convergence des suites de variables aléatoires

5.1 convergence en probabilité

Définition. Soit X_n une suite de variables aléatoires. On dit que la suite X_n converge en probabilité vers la constante a si et seulement si $\forall\ \varepsilon > 0$ et $\eta > 0$, $\exists\ n_0 \in \mathbb{N}^*$ tel que $\forall n > n_0$:

$$\P(|X_n - a| > \varepsilon) < \eta \quad \Leftrightarrow \quad \lim_{n \to \infty} P(|X_n - a| > \varepsilon) = 0$$

Cette définition est équivalente à :

$$\lim_{n \to \infty} P(|X_n - a| \leq \varepsilon) = 1$$

Notation. Si une suite X_n *converge en probabilité* vers a lorsque $n \to +\infty$ on note :

$$X_n \xrightarrow[n \to +\infty]{P} a$$

On définit alors la convergence en probabilité vers une variable aléatoire X comme la convergence vers 0 de la suite $X_n - X$. C'est-à-dire $\forall \varepsilon > 0$ et $\eta > 0$, $\exists\ n_0 \in \mathbb{N}^*$ tel que $\forall n > n_0$:

$$P(|X_n - x| > \varepsilon) < \eta \quad \Leftrightarrow \quad \lim_{n \to +\infty} P(|X_n - X| > \varepsilon) = 0$$

Notation. Si une suite X_n *converge en probabilité* vers variable aléatoire X lorsque $n \to +\infty$ on note :

$$X_n \xrightarrow[n\to+\infty]{P} X$$

Inégalité de Chebyshev

$$\boxed{\forall \varepsilon > 0, \quad P(|X - E(X)| > \varepsilon) < \frac{V(X)}{\varepsilon^2}}$$

Lorsque $\begin{cases} E(X_n) \xrightarrow[n\to+\infty]{} a \\ V(X_n) \xrightarrow[n\to+\infty]{} 0 \end{cases}$, alors $X_n \xrightarrow[n\to+\infty]{P} a$, car $\forall \varepsilon > 0$,

$$P(|X_n - \underbrace{E(X_n)}_{\xrightarrow[n\to+\infty]{} a}| > \varepsilon) < \underbrace{\frac{V(X_n)}{\varepsilon^2}}_{\xrightarrow[n\to+\infty]{} 0}$$

$$\Rightarrow P(|X_n - a| > \varepsilon) \xrightarrow[n\to+\infty]{} 0$$

$$\Rightarrow \boxed{X_n \xrightarrow[n\to+\infty]{P} a}$$

5.2 CONVERGENCE EN MOYENNE D'ORDRE p

On suppose que $E(|X_n - X|^p)$ existe, $p \in \mathbb{N}^*$.

Définition. On dit que la suite X_n converge en moyenne d'ordre p vers X si et seulement si $E(|X_n - X|^p) \xrightarrow[n\to+\infty]{} 0$.

Remarque. La plus utilisée est la convergence en moyenne quadratique : $p = 2$.

$$E(|X_n - X|^2) \xrightarrow[n\to+\infty]{} 0$$

Notation. Si une suite X_n *converge en moyenne d'ordre quadratique* vers une variable aléatoire X lorsque $n \to +\infty$ on note :

$$X_n \xrightarrow[n\to+\infty]{\text{m.q.}} X$$

5.3 CONVERGENCE EN LOI

Elle permet d'approximer la fonction de répartition de X_n par celle de X.

Définition. La suite X_n converge en loi vers la variable aléatoire X, si *en tout point de continuité* de $F(x)$ (fonction de répartition de X), la suite $F_n(x)$ (fonction de répartition de la suite X_n) converge vers $F(x)$:

$$\lim_{n\to+\infty} F_n(x) = F(x), \quad \forall X \text{ point de continuité de } F(x)$$

Rappel $F(x) = P[X < x]$ est la fonction de répartition de X.

Notation. Si une suite X_n *converge en loi* vers une variable aléatoire X lorsque $n \to +\infty$ on note :

$$X_n \xrightarrow[n\to+\infty]{\mathscr{L}} X$$

Remarque. Cette définition est équivalente à la convergence des fonctions caractéristiques :

$$\lim_{n\to+\infty} \underbrace{\varphi_{X_n}(t)}_{\substack{\text{fonction}\\ \text{caractéristique}\\ \text{de } X_n}} = \underbrace{\varphi_X(t)}_{\substack{\text{fonction}\\ \text{caractéristique}\\ \text{de } X}}$$

6

Approximation des lois discrètes par la loi gaussienne

6.1 Convergence en loi de la binomiale vers la loi de Laplace-Gauss

Théorème. (*Moivre-Laplace*)

X_n étant une suite de variables binomiales $\mathscr{B}(n,p)$ alors,

$$\frac{X_n - np}{\sqrt{npq}} \xrightarrow[n\to+\infty]{\mathscr{L}} L.G(0,1) \quad \text{avec } q = 1-p$$

Preuve. La fonction caractéristique de la loi $\mathscr{B}(n,p)$:

$$\varphi_{X_n}(t) = (pe^{it} + 1 - p)^n \quad \text{voir exercice 1}$$

or,

$$\varphi_{\frac{X-m}{\sigma}} = e^{-\frac{itm}{\sigma}} \cdot \varphi_X\left(\frac{t}{\sigma}\right)$$

donc la fonction caractéristique de $Y_n = \frac{X_n - np}{\sqrt{npq}}$ est :

$$\varphi_{Y_n}(t) = (pe^{\frac{it}{\sqrt{npq}}} + 1 - p)^n \cdot e^{-\frac{itnp}{\sqrt{npq}}} \text{, avec } \begin{cases} m = np \\ \sigma = \sqrt{npq} \end{cases}$$

on applique un logarithme népérien des deux côtés de l'égalité :

$$\ln\left(\varphi_{Y_n}(t)\right) = n \cdot \ln\left(p\left(e^{\frac{it}{\sqrt{npq}}} - 1\right) + 1\right) - \frac{itnp}{\sqrt{npq}}$$

Rappel. développement limité à l'ordre 2 lorsque x est au voisinage de 0 :

$$e^x \simeq 1 + \frac{x}{1!} + \frac{x^2}{2!}$$
$$\ln(1+x) \simeq x - \frac{x^2}{2}$$

On applique un double développement limité :

$$\begin{aligned}
\ln\left(\varphi_{Y_n}(t)\right) &\simeq n \cdot \ln\left(1 + p\left(\frac{it}{\sqrt{npq}} - \frac{t^2}{2npq}\right)\right) - \frac{itnp}{\sqrt{npq}} \\
&\simeq n\left(\frac{pit}{\sqrt{npq}} - \frac{\cancel{p}t^2}{2n\cancel{p}q} + \frac{p^2t^{\cancel{2}}}{2n\cancel{p}q}\right) - \frac{itnp}{\sqrt{npq}} \\
&\simeq \frac{npi\cancel{t}}{\cancel{\sqrt{}}npq} - \frac{t^2}{2q} + \frac{pt^2}{2q} - \frac{itn\cancel{p}}{\cancel{\sqrt{}}npq} \\
&\simeq \frac{t^2}{2q} \cdot \underbrace{(p-1)}_{-q} \\
&\simeq -\frac{\cancel{q}t^2}{2\cancel{q}} \\
\ln\left(\varphi_{Y_n}(t)\right) &\simeq -\frac{t^2}{2}
\end{aligned}$$

On applique une exponentielle des deux côtés de l'équation :

$$\varphi_{Y_n} \simeq e^{-\frac{t^2}{2}}, \quad \text{fonction caractéristique de } L.G(0,1)$$

Au final on a bien :

$$\boxed{\frac{X_n - np}{\sqrt{npq}} \xrightarrow[n\to+\infty]{\mathscr{L}} L.G(0,1)}$$

Remarque. Lorsque n est suffisamment grand, on peut approximer la loi $\mathscr{B}(n,p)$ par la loi de Gauss. Cependant, il convient d'effectuer ce que l'on appelle la *correction de continuité* :

$$P[X = x] \simeq P\left[\frac{x - np - \frac{1}{2}}{\sqrt{npq}} < U < \frac{x - np + \frac{1}{2}}{\sqrt{npq}}\right]$$
$$P[X \leq x] \simeq P\left[U < \frac{x - np + \frac{1}{2}}{\sqrt{npq}}\right]$$

avec, $U = \frac{X-n}{\sqrt{npq}} \simeq L.G(0.1)$.

Exemple

Soit X une variable aléatoire suivant une loi $\mathscr{B}(n,p)$ avec $\begin{cases} n = 40 \\ p = 0.3 \end{cases}$

on a donc $\begin{cases} np = 12 \\ npq = 8,4 \end{cases}$

- $P[X = 11]$:
  - valeur exacte : $P[X = 11] = 0.1319$;
  - formule d'approximation :

$$\begin{aligned} P[X = 11] &\simeq P\left[\frac{10.5 - 12}{\sqrt{8.4}} < U < \frac{11.5 - 12}{\sqrt{8.4}}\right] \\ &\simeq P[-0.52 < U < -0.17] \\ &\simeq 0.131 \end{aligned}$$

  Soit une erreur de moins de 1%.

- $P[X \leq 11]$:
  - valeur exacte : $P[X \leq 11] = 0.4406$;
  - formule d'approximation :

$$\begin{aligned} P[X \leq 11] &\simeq P\left[U \leq \frac{11.5 - 12}{\sqrt{8.4}}\right] \\ &\simeq 0.4325 \end{aligned}$$

  - valeur sans correction de continuité :

$$\begin{aligned} P[X \leq 11] &\simeq P\left[U \leq \frac{-1}{\sqrt{8.4}}\right] \\ &\simeq 0.3632 \end{aligned}$$

  Résultat très imprécis.

6.2 CONVERGENCE EN LOI DE LA LOI DE POISSON VERS GAUSS

Théorème. Soit X_n une famille de variables de Poisson $\mathcal{P}(\lambda)$ alors si $\lambda \to +\infty$:

$$\frac{X_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow[\lambda \to +\infty]{\mathscr{L}} L.G(0,1)$$

Preuve. La fonction de caractéristique de la variable de $\mathcal{P}(\lambda)$ est :

$$\varphi_{X_\lambda}(t) = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda e^{it} - \lambda}$$

La fonction caractéristique de $Y_\lambda = \frac{X_\lambda - \lambda}{\sqrt{\lambda}}$ est :

$$\begin{aligned}
\varphi_{Y_\lambda}(t) &= e^{\lambda(e^{\frac{it}{\sqrt{\lambda}}} - 1)} \cdot e^{\frac{-it\lambda}{\sqrt{\lambda}}} \\
&= e^{\lambda e^{\frac{it}{\sqrt{\lambda}}} - \lambda - it\sqrt{\lambda}} \\
&\simeq e^{\lambda(1 + \frac{it}{\sqrt{\lambda}} - \frac{t^2}{2\lambda}) - \lambda - it\sqrt{\lambda}} && \text{développement limité} \\
&\simeq e^{\frac{-t^2}{2}} && \text{fonction caractéristique de } L.G(0,1)
\end{aligned}$$

6.3 THÉORÈME CENTRAL-LIMITE

Théorème. Soit (X_n) une suite de variables aléatoires *indépendantes* et *de même loi*, d'espérance μ et d'écart type σ, alors,

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n \to +\infty]{Loi} L.G(0,1)$$

Preuve.

$$\frac{X_1 + X_2 + \ldots + X_n + n\mu}{\sigma\sqrt{n}} = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma\sqrt{n}} \right) = S_n$$

et,

$$\varphi_{S_n}(t) = \prod_{i=1}^{n} \varphi_{\frac{X_i - \mu}{\sigma\sqrt{n}}}(t) \quad \text{(car les } X_i \text{ sont indépendantes.)}$$

Rappel. Formule de Maclaurin :

$$\varphi_{\frac{X_j-\mu}{\sigma\sqrt{n}}}(t) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} i^k E\left(\left(\frac{X_j-\mu}{\sigma\sqrt{n}}\right)^k\right)$$

Or,

$$\begin{aligned}
E\left(\frac{X_i-\mu}{\sigma\sqrt{n}}\right) &= \frac{1}{\sigma\sqrt{n}} \cdot E\left(X_i-\mu\right) \\
&= \frac{E(X_i)-\mu}{\sigma\sqrt{n}} \\
&= 0 \\
V\left(\frac{X_i-\mu}{\sigma\sqrt{n}}\right) &= E\left(\left(\frac{X_i-\mu}{\sigma\sqrt{n}}\right)^2\right) \cancel{- E^2\left(\frac{X_i-\mu}{\sigma\sqrt{n}}\right)} \\
&= \frac{1}{\sigma^2 n} \cdot V\left(X_i-\mu\right) \\
&= \frac{V(X_i)}{\sigma^2 n} \\
&= \frac{\cancel{\sigma^2}}{\cancel{\sigma^2} n} \\
&= \frac{1}{n}
\end{aligned}$$

Maclaurin à l'ordre 2 nous donne :

$$\varphi_{\frac{X_i-\mu}{\sigma\sqrt{n}}}(t) \approx 1 - \frac{t^2}{2n}$$

donc,

$$\begin{aligned}
\varphi_{S_n}(t) &\approx \prod_{i=1}^{n}\left(1-\frac{t^2}{2n}\right) \\
&\approx \left(1-\frac{t^2}{2n}\right)^n
\end{aligned}$$

Rappel.

$$\left(1+\frac{X}{n}\right)^n \xrightarrow[n\to+\infty]{} e^X$$

donc,

$$\lim_{n\to+\infty} \varphi_{S_n}(t) = e^{-\frac{t^2}{2}}$$

qui est la fonction caractéristique de $L.G(0,1)$. Conclusion :

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to+\infty]{Loi} L.G(0,1)$$

6.4 EXERCICE 1

Énoncé Soit la suite X_n de variables aléatoires dont la densité est :

$$f_n(x) = \frac{ne^{-nx}}{(1+e^{-nx})^2}$$

Question Montrer que $X_n \xrightarrow[n\to+\infty]{P} 0$.

Solution On chercher à savoir si $\forall\varepsilon > 0$, $P[|X_n| > \varepsilon] \xrightarrow[n\to+\infty]{} 0$.
On a $P[|X_n| > \varepsilon] = 1 - P[|X_n| \leq \varepsilon]$.

Rappel.

$$P[a < X < b] = \int_a^b f(x)dx = F(b) - F(a)$$

Ici on a :

$$\begin{aligned}
P[|X_n| > \varepsilon] &= 1 - \int_{-\varepsilon}^{\varepsilon} f_n(x)dx \\
&= 1 - \int_{-\varepsilon}^{\varepsilon} \frac{ne^{-nx}}{(1+e^{-nx})^2}dx \\
&= 1 - \left[\frac{1}{1+e^{-nx}}\right]_{-\varepsilon}^{\varepsilon} \\
&= 1 - \underbrace{\frac{1}{1+e^{-n\varepsilon}}}_{\xrightarrow[n\to+\infty]{} 1} + \underbrace{\frac{1}{1+e^{n\varepsilon}}}_{\xrightarrow[n\to+\infty]{} 0}
\end{aligned}$$

Donc :

$$P[|X_n| > \varepsilon] \xrightarrow[n\to+\infty]{} 0 \quad \Leftrightarrow \quad \boxed{X_n \xrightarrow[n\to+\infty]{P} 0}$$

6.5 EXERCICE 2

Énoncé Soit X une variable aléatoire continue de densité :

$$g(x) = e^{-x-e^{-x}}$$

Questions

1. Déterminer la fonction de répartition G de X.
2. Déterminer la fonction de répartition H de $Z = e^{-X}$, ainsi que sa densité.
3. Calculer $E(Z)$ et $V(Z)$.
4. Soit $(Z_1, \ldots, Z_n)$ un échantillon de Z et $\overline{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$:
   a. montrer que $\overline{Z}_n \xrightarrow[n\to+\infty]{P} 1$;
   b. montrer que $\overline{Z}_n \xrightarrow[n\to+\infty]{m.q} 1$.

Solution

1.

$$\begin{aligned} G(x) &= P[X < x] \\ &= \int_{-\infty}^{x} g(t)dt \\ &= \int_{-\infty}^{x} e^{-t-e^{-t}} dt \\ &= \left[-e^{-e^{-t}}\right]_{-\infty}^{x} \\ G(x) &= \boxed{e^{-e^{-x}}} \end{aligned}$$

2. Fonction de répartition de $Z = e^{-X}$:

$$\begin{aligned} H(z) &= P[Z < z] \qquad\qquad \forall z > 0 \\ &= P[e^{-X} < \ln(z)] \\ &= P[-X < \ln(z)] \\ &= P[X > -\ln(z)] \\ &= 1 - G(-\ln(z)) \end{aligned}$$

Or on connait $G(x)$, donc :

$$\boxed{H(z) = 1 - e^{-z}, \quad \forall z > 0}$$

Enfin, sa densité $h(z)$ est :

$$\boxed{h(z) = H'(z) = e^{-z}, \quad \forall z > 0}$$

3. L'espérance de Z est :

$$\begin{aligned} E(Z) &= \int_0^{+\infty} h(z)dz \\ &= \left[-ze^{-z}\right]_0^{+\infty} + \int_0^{+\infty} e^{-z}dz \\ &= -\left[e^{-z}\right]_0^{+\infty} \\ &= 1 \end{aligned}$$

$$\boxed{E(Z) = 1}$$

La variance de Z est :

$$V(Z) = E(Z^2) - E^2(Z)$$

avec,

$$\begin{aligned} E(z^2) &= \int_0^{+\infty} z^2 h(z)dz \\ &= \int_0^{+\infty} z^2 e^{-z}dz \\ &= [-z^2e^{-z}]_0^{+\infty} + \int_0^{+\infty} 2ze^{-z}dz \qquad \text{par partie} \\ &= 2 \cdot \int_0^{+\infty} ze{-}zdz \\ &= 2 \cdot E(z) \\ E(z^2) &= 2 \end{aligned}$$

donc,

$$\boxed{V(Z) = 1}$$

4. Soit $(Z_1, \ldots, Z_n)$ un échantillon de Z, c'est-à-dire que Z_i est une suite de variables aléatoires indépendantes de même loi que Z.

$$\overline{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i \quad \text{moyenne empirique}$$

   a. On cherche à montrer que : $\overline{Z_n} \xrightarrow[n\to+\infty]{P} 1 = E(Z)$. Pour cela, on utilise l'inégalité de Chebyshev (5.1) :

$$\begin{aligned}E(\overline{Z}_n) &= \frac{1}{n}\sum_{i=1}^{n} E(Z_i)\\ &= \frac{1}{n}\sum_{i=1}^{n} 1\\ &= \frac{\cancel{n}}{\cancel{n}}\\ &= 1\end{aligned}$$

de même,

$$\begin{aligned}V(\overline{Z}_n) &= \frac{1}{n^2}\cdot V\left(\sum_{i=1}^{n} Z_i\right)\\ &= \frac{1}{n^2}\cdot\sum_{i=1}^{n} V(Z_i) \qquad \text{car les variables sont indépendantes}\\ &= \frac{1}{n^2}\cdot\sum_{i=1}^{n} 1\\ &= \frac{\cancel{n}}{n^{\cancel{2}}}\\ &= \frac{1}{n}\end{aligned}$$

donc,

$$V(\overline{Z}_n) = \frac{1}{n} \xrightarrow[n\to+\infty]{} 0$$

et $\forall\varepsilon > 0$,

$$\begin{aligned}P(|\overline{Z}_n - 1| \geq \varepsilon) &< \underbrace{\frac{1}{n\varepsilon^2}}_{\xrightarrow[n\to+\infty]{} 0}\\ &\Rightarrow \lim_{n\to+\infty} P(|\overline{Z_n} - 1| > \varepsilon) = 0\\ &\Rightarrow \boxed{\overline{Z}_n \xrightarrow[n\to+\infty]{P} 1 = E(Z)}\end{aligned}$$

b. On cherche à montrer que : $\overline{Z}_n \xrightarrow[n\to+\infty]{m.q} 1 = E(Z)$, donc que $E(|\overline{Z}_n - 1|^2) \xrightarrow[n\to+\infty]{} 0$.

On remarque que :

$$E(|\overline{Z}_n - 1|^2) = E(|\overline{Z}_n - E(\overline{Z}_n)|^2) = V(\overline{Z}_n)$$

Rappel. Par définition :

$$V(X) = E((X - E(X))^2))$$

Mais en pratique on utilise plus souvent la formule König-Huyghens :

$$V(X) = E(X^2) - E(X)^2$$

On conclue immédiatement que,

$$\begin{aligned} E(|\overline{Z}_n - E(\overline{Z}_n)|^2) &= V(\overline{Z}_n) \\ &= \frac{1}{n} \xrightarrow[n\to+\infty]{} 0 \\ &\Rightarrow \boxed{\overline{Z}_n \xrightarrow[n\to+\infty]{m.q} 1} \end{aligned}$$

6.6 EXERCICE 3

Énoncé Soit X une variable aléatoire de loi gamma γ_p

Questions

1. Calculer la fonction caractéristique de X ;
2. en déduire celle de $Y_p = \frac{X-p}{\sqrt{p}}$;
3. montrer que $\frac{X-p}{\sqrt{p}} \xrightarrow[p\to+\infty]{\mathscr{L}} L.G(0,1)$.

Rappel. On rappelle la densité de la loi γ_p :

$$f(X) = \frac{1}{\Gamma(p)} \cdot X^{p-1} \cdot e^{-X} \quad \forall X > 0$$

Solution

1. La fonction caractéristique de X est :

$$\begin{aligned}
\varphi_X(t) &= \int_{\mathbb{R}} e^{itx} f(x) dx \\
&= \int_0^{+\infty} e^{itx} \frac{1}{\Gamma(p)} \cdot x^{p-1} \cdot e^{-x} dx \\
&= \frac{1}{\Gamma(p)} \cdot \int_0^{+\infty} e^{(it-1)x} x^{p-1} dx
\end{aligned}$$

Posons : $I_{p-1} = \int_0^{+\infty} e^{(it-1)x} x^{p-1} dx$

Calculons les premiers termes de la suite pour essayer de faire apparaitre une relation de récurrence :

$$\begin{aligned}
I_0 &= \int_0^{+\infty} e^{(it-1)x} dx \\
&= \left[\frac{e^{(it-1)x}}{it-1} \right]_0^{+\infty}
\end{aligned}$$

or,

$$e^{(it-1)x} = e^{itx} e^{-x} \xrightarrow[x \to \infty]{} 0$$

car $|e^{itx}| = 1$.

$$\boxed{I_0 = \frac{-1}{it-1}}$$

Ensuite, en intégrant par partie avec $\begin{cases} v = x^{p-1} \to v' = (p-1)x^{p-2} \\ u' = e^{(it-1)x} \to u = \frac{e^{(it-1)x}}{it-1} \end{cases}$, on obtient :

$$\begin{aligned}
I_{p-1} &= \int_0^{+\infty} e^{(it-1)x} x^{p-1} dx \\
&= \underbrace{\cancel{\left[\frac{e^{(it-1)x}}{it-1} x^{p-1} \right]_0^{+\infty}}}_{\xrightarrow[x \to +\infty]{} 0} - \frac{p-1}{it-1} \int_0^{+\infty} e^{(it-1)x} x^{p-1} dx \\
&= -\frac{p-1}{it-1} \cdot I_{p-2} \quad \forall p \geq 2
\end{aligned}$$

Nous avons donc notre relation de récurrence :

$$\boxed{\begin{cases} I_0 = \frac{-1}{it-1} \\ I_{p-1} = -\frac{p-1}{it-1} \cdot I_{p-2} \end{cases}}$$

Ainsi par récurrence on obtient :

$$
\begin{aligned}
I_{p-1} &= -\frac{(p-1)}{it-1} \cdot \cancel{I_{p-2}} \\
\cancel{I_{p-2}} &= -\frac{(p-2)}{it-1} \cdot \cancel{I_{p-3}} \\
\cancel{I_{p-3}} &= -\frac{(p-4)}{it-1} \cdot \cancel{I_{p-4}} \\
&\dots \\
\cancel{I_2} &= -\frac{(2)}{it-1} \cdot \cancel{I_1} \\
\cancel{I_1} &= -\frac{(1)}{it-1} \cdot I_0
\end{aligned}
$$

En multipliant membre à membre toutes les équations, on obtient :

$$
\boxed{I_{p-1} = \frac{(-1)^{p-1}(p-1)! \cdot I_0}{(it-1)^{p-1}} = \frac{(-1)^p (p-1)!}{(it-1)^p} \quad \forall p \geq 1}
$$

Finalement,

$$
\begin{aligned}
\varphi_X(t) &= \frac{1}{\Gamma(p)} \cdot I_{p-1} \\
&= \frac{1}{\cancel{(p-1)!}} \cdot \frac{(-1)^p \cancel{(p-1)!}}{(it-1)^p} \\
&= (1-it)^{-p}
\end{aligned}
$$

$$
\boxed{\varphi_X(t) = (1-it)^{-p}}
$$

2. On cherche la fonction caractéristique de $Y_p = \frac{X-p}{\sqrt{p}}$ (centrée réduite).

Rappel.

$$
\varphi_{\frac{X-m}{\sigma}}(t) = e^{-\frac{itm}{\sigma}} \cdot \varphi_X\left(\frac{t}{\sigma}\right)
$$

$$
\varphi_{\frac{X-p}{\sqrt{p}}}(t) = e^{-\frac{itp}{\sqrt{p}}}(1 - \frac{it}{\sqrt{p}})^{-p}
$$

$$
\varphi_{Y_p}(t) = e^{-it\sqrt{p}}(1 - \frac{it}{\sqrt{p}})^{-p}
$$

3. On veut montrer qu'il s'agit d'une gaussienne. On applique un logarithme népérien sur chacun des membres de l'équation :

$$\begin{aligned}\ln\left(\varphi_{Y_p}(t)\right) &= \ln\left(e^{-it\sqrt{p}}(1-\frac{it}{\sqrt{p}})^{-p}\right)\\ &= \ln\left(e^{-it\sqrt{p}}\right)+\ln\left(\left(1-\frac{it}{\sqrt{p}}\right)^{-p}\right)\\ &= -it\sqrt{p}-p\ln\left(1-\frac{it}{\sqrt{p}}\right)\end{aligned}$$

On va effectuer un développement limité du logarithme népérien à l'ordre 2 afin de simplifier l'expression, puis appliquer une exponentielle sur les deux membres de l'équation afin de retrouver $\varphi_{Y_p}(t)$.

Rappel. $\ln(1+x) \simeq x - \frac{x^2}{2}$ à l'ordre 2.

Ainsi,

$$\begin{aligned}\ln\varphi_{Y_p}(t) &\simeq -it\sqrt{p}-p(-\frac{it}{\sqrt{p}}+\frac{t^2}{2p})\\ &\simeq \cancel{-it\sqrt{p}}+\frac{it\cancel{p}}{\cancel{\sqrt{p}}}-\frac{t^2}{2}\end{aligned}$$

Enfin, en repassant à l'exponentielle on obtient,

$$\varphi_{Y_p}(t) = e^{-\frac{t^2}{2}} \quad \text{fonction caractéristique de } L.G(0,1)$$

Conclusion,

$$\boxed{Y_p \xrightarrow[p\to+\infty]{\mathscr{L}} L.G(0,1)}$$

6.7 EXERCICE 4

Énoncé Le nombre de pannes par mois, sur une certaine machine suit une loi de Poisson $\mathcal{P}(3)$. Un atelier fonctionne avec 12 machines de ce type. Les machines sont indépendantes.

Questions

1. En un mois, quelle est la probabilité de constater plus de 42 pannes dans cet atelier ?
2. En un mois, quelle est la probabilité de constater entre 36 et 45 pannes dans cet atelier ?

Solution

1. Soit X_i la variable aléatoire qui représente le nombre de pannes en un mois de la machine i $(1 \leq i \leq 12)$. X_i suit une loi de Poisson $\mathcal{P}(\lambda = 3)$.
   Soit $S_n = \sum_{i=1}^{n} X_i$, avec X_i variables indépendantes (dans notre cas $n = 12$).

   Rappel. Soit X_λ une suite de variables aléatoires de Poisson $\mathcal{P}(\lambda)$:

   $$\frac{X_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow[\lambda \to +\infty]{\mathscr{L}} L.G(0,1)$$

   En pratique, on obtient une bonne approximation pour $\lambda \geq 18$.

   S_n est une somme de variables aléatoires indépendantes de Poisson. S_n suit une loi de Poisson $\mathcal{P}(n\lambda = \lambda')$.

   $$\lambda' = 12 \times 3 = 36$$

   Donc S_n suite une loi de Poisson $\mathcal{P}(36)$. On cherche la probabilité :

   $$P(S_n > 42)$$

   On se propose de centrer-réduire S_n, puis d'utiliser le théorème rappelé ci-dessus afin de pouvoir utiliser la table de valeurs d'une loi gaussienne centrée-réduite (fournie en annexe) :

   $$\begin{aligned} P(S_n > 42) &= 1 - P(S_n \leq 42) \\ &= 1 - P\left(\frac{S_{12} - 36}{6} \leq \frac{42 - 36}{6}\right) \end{aligned}$$

   Posons $U = \frac{S_{12} - 36}{6} \simeq L.G(0,1)$

   $$\begin{aligned} P(S_n > 42) &\simeq 1 - P[U \leq 1] \\ &\simeq 1 - F(1) \\ &\simeq -0.1587 \end{aligned}$$

2. $$\begin{aligned} P(36 \leq S_n \leq 45) &\simeq P\left(0 \leq U \leq \frac{45 - 36}{6}\right) \\ &\simeq P\left(0 \leq U \leq \frac{3}{2}\right) \\ &\simeq F(1.5) - F(0) \\ &\simeq 0.4332 \end{aligned}$$

6.8 EXERCICE 5

Énoncé Une usine fabrique des pièces dont 3% ont des défauts. Soit X la variable aléatoire représentant le nombre de pièces défectueuses. X suit une loi binomiale $\mathcal{B}(n,p)$ avec $p = 3$ (n dépendra de la question).

Questions

1. On prélève 1000 pièces au hasard, quelle est la probabilité d'avoir plus de 50 pièces défectueuses ?
2. Calculer $P(20 \leq X \leq 40)$.
3. On veut 1950 pièces sans défaut. On en prélève par prudence 2000 au hasard. Quelle est la probabilité d'avoir suffisamment de pièces en bon état ?

Solution

1.

> **Rappel.** Théorème de Moivre-Laplace :
>
> $$\mathcal{B}(n,p) \xrightarrow[n\to+\infty]{Loi} L.G(np, \sqrt{npq})$$

$$\begin{aligned} P(X > 50] &= 1 - P(X \leq 50) \\ &\approx 1 - P\left(\frac{X-np}{\sqrt{npq}} \leq \frac{50-np+0.5}{\sqrt{npq}}\right) \end{aligned}$$

On rajoute 0.5 pour effectuer une correction de continuité. On a $U = \frac{X-np}{\sqrt{npq}}$ qui suit une loi $L.G(0,1)$ donc on peut utiliser la table de valeurs en annexe :

$$\boxed{P(X > 50) \approx 1 - F(3.8) \approx 0}$$

2.

$$P(20 \leq X \leq 40) \approx P\left(\frac{20-np-0.5}{\sqrt{npq}} \leq \frac{X-np}{\sqrt{npq}} \leq \frac{40-np+0.5}{\sqrt{npq}}\right)$$

avec $U = \frac{X-np}{\sqrt{npq}} \approx L.G(0,1)$, donc

$$\begin{aligned} P(20 \leq X \leq 40) &\approx F(1.94) - F(-1.94) \\ &\approx 2 \cdot F(1.94) - 1 \\ &\approx 0.9476 \end{aligned}$$

3. Ici on a les paramètres $p = 3$ et $n = 2000$. On rappelle que X suit une loi binomiale $\mathcal{B}(n,p) \approx L.G(0,1)$.

$$\begin{aligned} P(X \leq 50) &\approx P\left(\frac{X-np}{\sqrt{npq}} \leq \frac{50-np+0.5}{\sqrt{npq}}\right) \\ &\approx P\left(U \leq \frac{50-60+0.5}{\sqrt{npq}}\right) \\ &\approx 0.1056 \end{aligned}$$

6.9 EXERCICE 6

Énoncé Soit,

$$U_n = \sum_{i=1}^{n} X_i^2$$

Où les X_i suivent une loi $L.G(0,1)$ et sont indépendantes. U_n est une variable χ-2 de degré de liberté n.

$$\chi_n = \|X\|^2 \quad \text{avec, } X = (X_1, X_2, \cdots, X_n) \quad \text{vecteur gaussien}$$

Questions Montrer que :

$$\frac{U_n - n}{\sqrt{2n}} \xrightarrow[n\to+\infty]{Loi} L.G(0,1)$$

Solution Il faut montrer que : $\begin{cases} E(U_n) = n \\ \sigma(U_n) = \sqrt{2n} \end{cases}$

$$\begin{aligned} E(U_n) &= \sum_{i=1}^{n} E(X_i^2) \\ &= \sum i = 1n1 \\ &= n \end{aligned}$$

$$\begin{aligned} \sigma(U_n) &= \sigma\left(\sum_{i=1}^{n} X_i^2\right) \\ &= \sqrt{V\left(\sum_{i=1}^{n} X_i^2\right)} \\ &= \sqrt{\sum_{i=1}^{n} V(X_i^2)} \\ &= \sqrt{\sum_{i=1}^{n}\left(E(X_i^4) - E^2(X_i^2)\right)} \end{aligned}$$

Rappel. Si X suit une loi gaussienne $L.G(0,1)$, alors grâce à la formule des moments on a :

$$E(X^{2k}) = \frac{(2k)!}{2^k k!}$$

On obtient donc :

$$
\begin{aligned}
E(X_i^4) &= \frac{4!}{4 \cdot 2!} = 3 \\
V(X_i^2) &= E(X_i^4) - E^2(X_i^2) = 3 - 1 = 2 \\
V(U_n) &= \sum_{i=1}^{n} 2 = 2n
\end{aligned}
$$

d'où,

$$\boxed{\sigma(U_n) = \sqrt{2n}}$$

conclusion, grâce au *théorème de central-limite* on obtient :

$$\frac{U_n - n}{\sqrt{2n}} \xrightarrow[n \to +\infty]{Loi} L.G(0,1)$$

7

Couples de variables aléatoires

7.1 Cas discret

Définition. On appelle loi conjointe, la loi du couple (X, Y), où X et Y sont des variables aléatoires discrètes, définie par :

$$P\Big((X = x_i) \cap (Y = y_j)\Big) = p_{ij} \quad \text{(Loi de l'intersection)}$$

X \ Y	y_1	$\cdots$	y_j	$\cdots$	y_q	
x_1			$\vdots$			$p_{1\cdot}$
$\vdots$			$\vdots$			
x_i	$\cdots$	$\cdots$	p_{ij}	$\cdots$	$\cdots$	$p_{i\cdot}$
$\vdots$			$\vdots$			
x_p			$\vdots$			$p_{p\cdot}$
	$p_{\cdot 1}$		$p_{\cdot j}$		$p_{\cdot q}$	

$$\sum_{i,j} P_{ij} = 1$$

Définition. On appelle *lois marginales*, les lois de probabilité de X et Y prises séparément :

Loi marginale de X

$$P[X = x_i] = \sum_{j=1}^{q} P_{ij} = P_{i.}$$

Loi marginale de Y

$$P[Y = y_j] = \sum_{i=1}^{q} P_{ij} = P_{.j}$$

Définition. *Lois conditionnelles* :
Si $Y = y_j$

$$P(X = x_i \setminus Y = y_j) = \frac{P_{ij}}{P_{.j}} = \frac{P\Big((X = x_i) \cap (Y = y_j)\Big)}{P(Y = y_j)}$$

Si $X = y_i$

$$P(Y = y_j \setminus X = x_i) = \frac{P_{ij}}{P_{i.}} = \frac{P\Big((X = x_i) \cap (Y = y_j)\Big)}{P_{i.}}$$

Définition. Les variables aléatoires X et Y sont indépendantes si et seulement si :

$$P_{ij} = P_{i.} \times P_{.j} \tag{7.1}$$

Définition.

$$E(X \cdot Y) = \sum_{i,j} x_i y_j P_{ij} = \sum_{i,j} x_i y_j P\Big((X = x_i) \cap (Y = y_j)\Big)$$

Définition. On note $Cov(X,Y) = E(X \cdot Y) - E(X)E(Y)$ la covariance de couple (X,Y). On note le coefficient de corrélation linéaire entre X et Y :

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_x \sigma_y} = \frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$$

7.2 CAS CONTINU

Définition. Soient X et Y deux variables aléatoires continues. La loi du couple (X,Y) est définie par la densité :

$$f(x,y) \ : \begin{cases} (i) & f(x,y) \geq 0 \\ (ii) & \int\int_{\mathbb{R}^2} f(x,y)dxdy = 1 \end{cases}$$

Définition. la fonction de répartition $F(x,y)$ du couple (X,Y) est définie par :

$$\begin{aligned} F(x,y) &= P[X < x \text{ et } Y < y] \\ &= \int_{-\infty}^{x} \int_{-\infty}^{y} f(x,y) \cdot dx \cdot dy \end{aligned}$$

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$$

Définition. *Lois marginales* :
Loi marginale de X

$$g(x) = \int_{\mathbb{R}} f(x,y)dy$$

Loi marginale de Y

$$h(x) = \int_{\mathbb{R}} f(x,y)dx$$

Définition. *Lois conditionnelles* :
Loi conditionnelle de X

$$f_{X/Y}(x \ / \ Y = y) = \frac{f(x,y)}{h(y)}$$

Loi conditionnelle de X

$$f_{X/Y}(y \ / \ X = x) = \frac{f(x,y)}{g(y)}$$

Définition.

$$E(X \cdot Y) = \int\int_{\mathbb{R}^2} x \cdot y \cdot f(x,y) \cdot dx \cdot dy$$

$$Cov(X,Y) = E(X \cdot Y) - E(X)E(Y)$$

$$\boxed{\rho(X,Y) = \frac{Cov(X,Y)}{\partial x \partial y}}$$, coefficient de corrélation linéaire.

Remarque. Le produit scalaire défini sur l'espace des variables aléatoires :

$$\begin{aligned}
\langle X,Y\rangle &= E(X \cdot Y) \\
\|X\| &= \sqrt{E(X^2)} \\
\sigma &= \sqrt{V(X)} \\
&= \sqrt{E\left((X - E(X))^2\right)} \\
&= \|X - E(X)\| && \text{norme de } X - E(X) \\
Cov(X,Y) &= \langle X - E(X), Y - E(Y)\rangle && \text{produit scalaire des variables centrées} \\
\rho(X,Y) &= \frac{Cov(X,Y)}{\sigma_x \sigma_y} \\
&= \frac{\langle X - E(X), Y - E(Y)\rangle}{\|X - E(X)\|\|Y - E(X)\|} \\
&= \cos\left(X - E(\widehat{X), Y} - E(Y)\right)
\end{aligned}$$

On a donc toujours :

$$\boxed{|\rho| \leq 1}$$

8

Exercices sur les couples de variables aléatoires

8.1 Exercice 1

Énoncé Soit (X, Y) un vecteur aléatoire. Soit la densité du couple définie par :

$$f(x,y) = \begin{cases} 2(e-1)^{-1}xe^y & \text{si, } \begin{array}{l} 0 \leq x \leq 1 \\ 0 \leq y \leq 1 \end{array} \\ 0 & \text{sinon} \end{cases}$$

Questions

1. Montrer que $f(x,y)$ est une densité sur $\mathbb{R}^2$.
2. Calculer les densités marginales de X et Y.
3. X et Y sont-elles indépendantes ?

Solution

1.

Rappel. f est une densité sur $\mathbb{R}^2$ si :

a. $f(x,y) \geq 0$

b. $\iint_{\mathbb{R}^2} f(x,y)dxdy = 1$

Ici on a,

a. $f(x,y) \geq 0 \quad \forall\ (x,y) \in \mathbb{R}^2$

b.

$$\begin{aligned}
\iint_{\mathbb{R}^2} f(x,y)dxdy &= \int_0^1 \int_0^1 2(e-1)^{-1} xe^y dxdy \\
&= 2(e-1)^{-1} \int_0^1 xdx \int_0^1 e^y dy \\
&= 2(e-1)^{-1} \left[\frac{x^2}{2}\right]_0^1 [e^y]_0^1 \\
&= 2(e-1)^{-1} \frac{1}{2}(e-1) \\
&= 1
\end{aligned}$$

2. Densité marginale de X,

$$\begin{aligned}
g(x) &= \int_0^1 f(x,y)dy \\
&= 2(e-1)^{-1} x \int_0^1 e^y dy \\
&= 2(e-1)^{-1} x [e^y]_0^1 \\
&= 2(e-1)^{-1} x (e-1) \\
&= 2x
\end{aligned}$$

Densité marginale de Y,

$$\begin{aligned}
h(y) &= \int_0^1 f(x,y)dx \\
&= 2(e-1)^{-1} e^y \int_0^1 xdx \\
&= 2(e-1)^{-1} e^y \left[\frac{x^2}{2}\right]_0^1 \\
&= (e-1)^{-1} e^y
\end{aligned}$$

3. On souhaite montrer que $f(x,y) = g(x) \cdot h(x)$,

$$\begin{aligned}
g(x) \cdot h(x) &= 2x(e-1)^{-1} e^y \\
&= f(x,y)
\end{aligned}$$

8.2 EXERCICE 2

Énoncé Soit X et Y deux variables aléatoires. Soit f la densité du couple définie par :

$$f(x,y) = \begin{cases} k(x+y) & \text{si, } (x,y) \in D \\ 0 & \text{sinon} \end{cases}$$

Le domaine D est défini de la manière suivante :

$$D = \left\{ (x,y) \in \mathbb{R}^2 \;/\; y \geq 0 \;\wedge\; y \leq x \;\wedge\; x+y \leq 2 \right\}$$

Questions

1. Déterminer k ;
2. déterminer les densités marginales ;
3. les variables X et Y sont-elles indépendantes ?

Solution

1. Commençons par nous représenter le domaine D graphiquement.
   Si $f(x,y)$ est une densité alors on a :
   a. $k(x+y) \geq 0$
   b. $\iint_{\mathbb{R}^2} k(x+y)dxdy = 1$
   Déterminons k tel que ces conditions soient respectées :

$$\begin{aligned}
\iint_D f(x,y)dx,dy &= 1 \\
k\left[\int_{x=0}^{x=1}\left(\int_{y=0}^{y=x}(x+y)dy\right)dx + \int_{X=1}^{X=2}\left(\int_{y=0}^{y=2-x}(x+y)dy\right)\right] &= 1 \\
k\left[\int_0^1\left[xy+\frac{y^2}{2}\right]_{y=0}^{y=x}dx + \int_1^2\left[xy+\frac{y^2}{2}\right]_{y=0}^{y=2-x}dx\right] &= 1 \\
k\left[\int_0^1 \frac{3}{2}x^2 + \int_1^2\left(x(2-x)+\frac{(2-x)^2}{2}\right)dx\right] &= 1 \\
k\left[\frac{3}{2}\left[\frac{x^3}{3}\right]_0^1 + \int_1^2\left(\cancel{2x}-x^2+2-\cancel{2x}+\frac{x^2}{2}\right)dx\right] &= 1 \\
k\left[\frac{1}{2}+\int_1^2\left(2-\frac{1}{2}x^2\right)dx\right] &= 1 \\
k\left(\frac{1}{2}+2-\frac{1}{2}\left[\frac{x^3}{3}\right]_1^2\right) &= 1 \\
k\left(\frac{1}{2}+2-\frac{1}{2}\left(\frac{8}{3}-\frac{1}{3}\right)\right) &= 1 \\
k\left(\frac{5}{2}-\frac{4}{3}+\frac{1}{6}\right) &= 1 \\
k\left(\frac{15-8+1}{6}\right) &= 1 \\
k\frac{4}{3} &= 1
\end{aligned}$$

Au final on obtient,

$$\boxed{k = \frac{3}{4}}$$

2. **Densité de** X. Il est nécessaire de faire deux cas :
   a. Premier cas ($0 \leq x \leq 1$) :

$$\begin{aligned} g(x) &= \frac{3}{4}\int_{y-0}^{y=x}(x+y)dy \\ &= \frac{3}{4}\left(x^2+\left[\frac{y^2}{2}\right]_0^x\right) \\ &= \frac{3}{4}\left(x^2+\frac{x^2}{2}\right) \\ &= \frac{3}{4}\cdot\frac{3}{2}x^2 \\ &= \frac{9}{8}x^2 \end{aligned}$$

   b. Deuxième cas ($1 \leq x \leq 2$) :

$$\begin{aligned} g(x) &= \frac{3}{4}\int_{y=0}^{y=2-x}(x+y)dy \\ &= \frac{3}{4}\left(x(2-x)+\left[\frac{y^2}{2}\right]_0^{2-x}\right) \\ &= \frac{3}{4}\left(2x-x^2+\frac{1}{2}(2-x)^2\right) \\ &= \frac{3}{4}\left(2x-x^2+2-2x+\frac{x^2}{2}\right) \\ &= \frac{3}{4}\left(2-\frac{x^2}{2}\right) \end{aligned}$$

Densité de Y.

$$\begin{aligned} h(y) &= \int_{x=y}^{x=2-y} f(x,y)dx \\ &= \frac{3}{4}\int_{x=y}^{x=2-y}(x+y)dx \\ &= \frac{3}{4}\left(\left[\frac{x^2}{2}\right]_y^{2-y}+y(2-2y)\right) \\ &= \frac{3}{4}\left(\frac{(2-y)^2}{2}-\frac{y^2}{2}+2y-2y^2\right) \\ &= \frac{3}{4}\left(2-2y+\frac{y^2}{2}-\frac{y^2}{2}+2y-2y^2\right) \\ &= \frac{3}{4}(2-2y^2) \\ &= \frac{3}{2}(1-y^2) \end{aligned}$$

3. X et Y ne sont pas indépendantes car $\boxed{f(x,y) \neq g(x) \cdot h(y)}$.

8.3 EXERCICE 3

Énoncé Soient n contrats d'assurance automobiles identiques selon lesquels, moyennant le versement d'une prime p de la part de l'assuré, l'assureur s'engage à le dédommager pour les accidents de l'année à venir.

En supposant que les risques d'accidents sont identiques pour tous les assurés, la loi de probabilité de X est la même pour tous les assurés. X représentant la somme que l'assureur devra lui verser. D'après les statistiques des années passées :

$$E(X) = a \quad \sigma(X) = \sigma \quad \text{connus.}$$

Questions Calculer la valeur de la prime p que doit demander l'assureur afin de pouvoir tenir ses engagements avec une probabilité : 99%. Voici les valeurs nécessaires à l'application numérique :

$$\begin{cases} n = 20000 \\ a = 1200 \\ \sigma = 600 \end{cases}$$

Solution Soit $S_n = \sum_{i=1}^{n} X_i$ la somme totale. On cherche p tel que :

$$P(S_n \leq np) = 0.99$$

On a,

$$\begin{aligned} E(S_n) &= na \\ V(S_n) &= \sum V(X_i) = n\sigma^2 \\ P\left(\frac{S_n - na}{\sigma\sqrt{n}} \leq \frac{np - na}{\sigma\sqrt{n}}\right) &= 99\% \end{aligned}$$

Si on pose $U_n = \frac{S_n - na}{\sigma\sqrt{n}} \approx L.G(0,1)$, on obtient :

$$P\left(U_n \leq \frac{n}{\sigma\sqrt{n}}(p-a)\right) = 0.99$$

En utilisant la table on a $F(t_0) = 0.99 \Rightarrow t_0 = 2.33$, car :

$$\begin{aligned} \frac{\sqrt{n}}{\sigma}(p-a) &= t_0 \\ p &= \frac{\sigma}{\sqrt{n}} t_0 + a \end{aligned}$$

ANNEXE A

Loi normale réduite

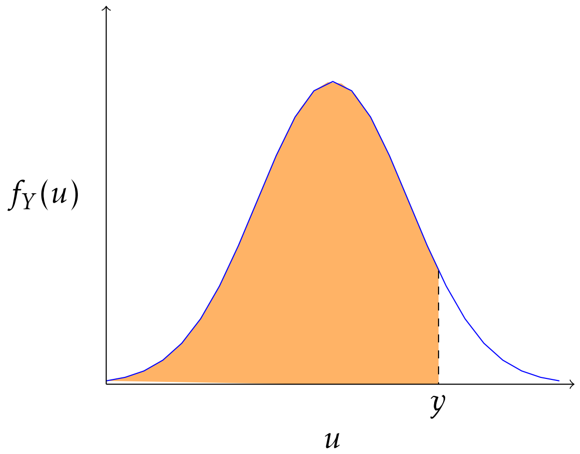

Figure A.1 – Fonction de répartition de la loi normale réduite (Probabilité de trouver une valeur inférieure à u.

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5190	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7157	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7969	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8513	0.8554	0.8577	0.8529	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9215	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9492	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

Table A.1 – Table de valeurs de la loi normale réduite pour u positif.

u	3.0	3.1	3.2	3.3	3.4	3.5	3.8	4.0	4.5
$F(u)$	0.9986	0.9990	0.9993	0.9995	0.99966	0.99976	0.999928	0.999968	0.9999997

Table A.2 – Table pour les grandes valeurs de u.