EISTI
ING1-MG : Probabilités 17-18

TD 6-7: Lois usuelles (Corrigé)

Exercice 1. Un étudiant passant un oral de mathématique choisit dans une urne 5 enveloppes contenant chacune une question à traiter. Il y a dans l'urne, au totale, 12 questions de probabilités et 8 questions d'analyse. On appelle X le nombre de questions de probabilités qu'il a tirée.

a) Donner la loi de X, $\mathbb{E}(X)$ et $\mathbb{V}(X)$

b) Reprendre le problème dans le cas où il y a dans l'urne 60 questions de probabilité et 40 questions d'analyse.

Solution :

Le choix des questions correspond ici à un tirage sans remise de $n = 5$ questions parmi $N = 12 + 8 = 20$ questions comportant:

- $N_1 = 12$ questions de probabilité et $N_2 = 8$ questions d'analyse.

- Il y a donc une proportion $p = 12/20 = 0,6$ de questions de probabilité.

Si l'on qualifie le choix d'une question de probabilité de "succès" et celui d'une question d'analyse "d'échec" alors :

la v.a. X qui compte le nombre de succès parmi les n épreuves suit $\mathcal{H}(N, n, p)$ (tirage sans remise)

$$\mathbb{P}(X = 2) = \frac{C_{12}^2 C_8^3}{C_{20}^5} = \frac{66 \times 56}{15 \times 504} = 0,2384.$$

$E(X) = \frac{nN_1}{N} = np = 5 \times 0,6 = 3$

$$\begin{aligned} \mathbb{V}(X) &= = n\frac{N_1}{N} \times \frac{N_2}{N} \times \frac{N-n}{N-1} \\ &= npq\frac{N-n}{N-1} = 5 \times 0,6 \times 0,4\frac{15}{19} = \frac{77}{323} = 0,9474. \end{aligned}$$

b) Avec 60 questions de probs et 40 d'analyse on a $N = 100$ et $p = 0,6 = \frac{60}{100}$ reste inchangée. Alors $X \hookrightarrow \mathcal{H}(100, 5; 0,6)$.

- Comme $N/n = 20$ est important (supérieur à 10), on peut approximer cette loi par $\mathcal{B}(n = 5; p = 0,6)$.

- par approximation:

$$\mathbb{P}(X = 2) = C_3^2(0,6)^2(0,4)^2 = 0,2304,$$

$$\mathbb{E}(X) = np = 3 \qquad \text{et} \quad \mathbb{V}(X) = npq = 5 \times 0,6 \times 0,4 = 1,2.$$

-Par la loi hypèrgéométrique

$$\mathbb{P}(X = 2) = \frac{C_{60}^2 C_{40}^3}{C_{100}^5} = 0,2323$$

et $V(X) = npq\frac{N-n}{N-1} = 5 \times 0,6 \times 0,4\frac{95}{99} = 1,152$.

L'erreur relative commise est d'environs 0.19% pour la probabilité, et de

$$1,2 - 1,152 = 4.8\%$$

pour la variance.

Exercice 2. On tire avec remise une carte d'un jeu de 52 cartes, on note Y la v.a.r. égale au nombre de tirages effectués lorsque l'on obtient, pour la 4^e fois, un roi.

Donner la loi de Y, sa fonction génératrice et en déduire $\mathbb{E}(Y)$ et $V(Y)$.

Utiliser les deux formules suivantes:

$$\forall q \in [0,1], \quad \forall r \in \mathbb{N}^*, \quad \sum_{k=r-1}^{\infty} C_k^{r-1} q^{k-r+1} = \frac{1}{(1-q)^r} \quad \text{et} \quad \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}.$$

Solution :

Définition 1. La v.a. T_r qui compte le nombre de répétition nécessaires pour avoir r succès (ou le temps d'attente d'avoir r premier succès) suit une loi de Pascal de paramètres r et p. On note

$$T_r \hookrightarrow \mathcal{P}(r,p).$$

Loi de Y:

Y est la v.a.r. égale au rang d'arrivé du 4-ème roi. Donc Y suit une loi de Pascal $\mathcal{P}(r=4, p=\frac{1}{13})$ avec

$Y(\Omega) = \{4, 5, \dots, \}$ il faut au moins 4 tirages pour obtenir 4 fois roi.

L'événement $(Y = k)$ est réalisé lorsqu'au cours des $k-1$ premiers tirages, roi est arrivé $4-1$ fois et que le k-ème tirage amène le 4ème roi.

Il y a C_{k-1}^{4-1} façons de choisir les rangs d'apparition des $4-1$ premiers rois au cours de $k-1$ premiers tirages.

D'où

$$\mathbb{P}(Y = k) = C_{k-1}^{r-1} p^{r-1} q^{k-r} p = C_{k-1}^{3} q^{k-4} p^4 \quad \text{avec} \quad C_Y = \{4, 5, \dots, \}$$

Montrons que $\sum_{k=r}^{\infty} \mathbb{P}(Y = k) = 1$. Observons d'abord que:

$$\sum_{k=r}^{\infty} \mathbb{P}(Y = k) = \sum_{k=r}^{\infty} C_{k-1}^{r-1} p^r q^{k-r}$$

Changeons d'indice en posant $k-1 = h$ et mettons p^r en facteur, on obtient :

$$\sum_{k=r}^{\infty} \mathbb{P}(Y = k) = p^r \sum_{h=r-1}^{\infty} C_h^{r-1} q^{h-r+1} \tag{1}$$

Utilisons maintenant la formule suivante

$$\forall q \in]0,1[, \quad \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$

Pour évaluer le second membre de (1).

En dérivant cette égalité r-fois, nous obtenons :

$$\begin{aligned}
\sum_{k=0}^{\infty} k(k-1)\dots(k-r+1)q^{k-r} &= \frac{1.2.\dots.(r-1)}{(1-q)^r} \\
\sum_{k=r}^{\infty} k(k-1)\dots(k-r+1)q^{k-r} &= \frac{1.2.\dots.(r-1)}{(1-q)^r} \\
& \text{on change } k \to k+1 \\
\sum_{k=r-1}^{\infty} k(k-1)\dots(k-r+2)q^{k-r+1} &= \frac{1.2.\dots.(r-1)}{(1-q)^r}
\end{aligned}$$

d'où:

$$\begin{aligned}
\frac{1}{(1-q)^r} = \sum_{k=r-1}^{\infty} \frac{k(k-1)\dots(k-r+2)}{1.2.\dots.(r-1)} q^{k-r+1} &= \sum_{k=r-1}^{\infty} \frac{k!}{(k-r+1)!(r-1)!} q^{k-r+1} \\
&= \sum_{k=r-1}^{\infty} C_k^{r-1} q^{k-r+1}
\end{aligned}$$

Il vient donc de (1) que

$$\begin{aligned}
\sum_{k=r}^{\infty} \mathbb{P}(Y=k) &= \sum_{k=r}^{\infty} C_{k-1}^{r-1} p^r q^{k-r} \\
& \text{Changeons d'indice en posant } k-1=h \text{ et mettons } p^r \text{ en facteur, on obtient :} \\
\sum_{k=r}^{\infty} \mathbb{P}(Y=k) &= p^r \sum_{h=r-1}^{\infty} C_h^{r-1} q^{h-r+1} \\
&= \frac{p^r}{(1-q)^r} = 1.
\end{aligned}$$

-Fonction génératrice:

$$\begin{aligned}
G_X(t) &= \mathbb{E}(t^X) = \sum_{k=r}^{\infty} C_{k-1}^{r-1} p^r q^{k-r} t^k = \left(\frac{p}{q}\right)^r \sum_{k=r}^{\infty} C_{k-1}^{r-1} (qt)^k \\
&= \left(\frac{p}{q}\right)^r \sum_{k=r-1}^{\infty} C_k^{r-1} (qt)^{k+1-r} (qt)^r \\
&= p^r t^r \sum_{k=r-1}^{\infty} C_k^{r-1} (qt)^{k+1-r} \\
&= \frac{(pt)^r}{(1-qt)^r} = \left(\frac{pt}{1-qt}\right)^r.
\end{aligned}$$

$$G'_X(t) = \left(\frac{pt}{1-qt}\right)^{r-1} . \frac{rp}{(1-qt)^2} \qquad \Rightarrow \quad G'_X(1) = \frac{r}{p}.$$

$$\begin{aligned} G''_X(t) &= r(r-1)\left(\frac{pt}{1-qt}\right)^{r-2}.\frac{p^2}{(1-qt)^4}+r.\left(\frac{pt}{1-qt}\right)^{r-1}.\frac{2pq}{1-qt)^3} \\ &\Rightarrow\ G''_X(1)=\frac{r(r-1)+2rq}{p^2} \end{aligned}$$

d'où

$$\mathbb{E}(Y)=\frac{r}{p}=4\times 13=52$$

$$V(Y)=\frac{r(r-1)+2rq}{p^2}+\frac{r}{p}-\frac{r^2}{p^2}=\frac{rq}{p^2}=\frac{4(1-\frac{1}{13})}{(\frac{1}{13})^2}=624.$$

Exercice 3. Soient X_n des variables aléatoires i.i.d (indépendantes identiquement distribuées) suivant une loi de Bernoulli de paramètres p. On pose

$$Y_n=X_nX_{n+1}\quad \text{et}\quad U_n=Y_1+\ldots+Y_n.$$

1) Quelle est la loi de Y_n ?
2) Les Y_n sont-elles deux à deux indépendantes ?
3) Calculer l'espérance et la variance de U_n.

Solution :
1) Support de Y_n est $D_Y=\{0,1\}$. Donc

$$\mathbb{P}(X_nX_{n+1}=k)=0\quad \text{si}\quad k\notin\{0,1\}$$

l'événement $\{k=1\}$ est équivaut $\{X_n=1,X_{n+1}=1\}$, donc

$$\mathbb{P}(X_nX_{n+1}=1)=\mathbb{P}(X_n=1)\times\mathbb{P}(X_{n+1}=1)=p^2.$$

$$Y_n\hookrightarrow\mathcal{B}(p^2)$$

2) Non car cela entraînerait que $\mathbb{E}(Y_1Y_2)=\mathbb{E}(Y_1)\mathbb{E}(Y_2)$. Or $\mathbb{E}(Y_1)\mathbb{E}(Y_2)=p^4$ et

$$\mathbb{E}(Y_1Y_2)=\mathbb{E}(X_1X_2X_2X_3)=\mathbb{E}(X_1X_2^2X_3)=\mathbb{E}(X_1)\mathbb{E}(X_2^2)\times\mathbb{E}(X_3)=p^3$$

car X_2^2 a même loi que X_2 et les trios variables sont indépendantes dans leur ensemble (résultat du cours).
3) $\mathbb{E}(U_n)=np^2$
Variance: On écrit:

$$V(U_n)=\mathbb{E}(U_n^2)-(\mathbb{E}(U_n))^2=\mathbb{E}(U_n^2)-n^2p^4$$

Pour calculer $\mathbb{E}(U_n^2)$, on écrit:

$$U_n^2=\sum_{i=1}^n\sum_{j=1}^n Y_iY_j=\sum_{i=1}^n Y_i^2+2\sum_{i<j}Y_iY_j.$$

On peut donc distinguer:

a) Soit i et j sont séparés d'au moins 2 ($j = i + k, k \geq 2$). Dans ce cas

$$\mathbb{E}(Y_i Y_j) = \mathbb{E}(Y_i)\mathbb{E}(Y_j) = p^4$$

i.e. Y_i et Y_{i+k} sont indépendantes (pour $k \geq 2$)

b) Soit i et j sont espacés d'un ($j = i + 1$), on utilise le calcule de la question 2, on trouve

$$\mathbb{E}(Y_i Y_{i+1}) = p^3.$$

c) Soit i et j sont égaux. On a alors $\mathbb{E}(Y_i^2) = p^2$.

Reste à compter les occurrences de tout ces termes.

- Pour les termes diagonaux, on a n.
- La sous-diagonale est $n - 1$.
- De plus, le reste est

$$\sum_{i=1}^{n-2} i = (n-1)(n-2)/2.$$

Ainsi

$$V(U_n) = np^2 + 2[(n-1)(n-2)/2p^4 + (n-1)p^3] - n^2p^4 = np^2 + 2(n-1)p^3 + (2-3n)p^4$$

Exercice 4. I) Soit X une variable aléatoire continue qui suit une loi exponentielle de paramètres $\theta > 0$ et $\nu \in \mathbb{R}$. Déterminer :

- a) La fonction génératrice

-b) La moyenne μ

-c) La variance σ^2

II)

Un sous-marin nucléaire voyage en plongée mais doit néanmoins faire surface pour renouveler son atmosphère. La durée d'une plongée en jours suit une loi exponentielle. En dépouillant tous les livres de bord, on constate que 88% des plongées ont duré plus de six jours.

a) Donner la fonction de répartition de la variable aléatoire X représentant le temps de plongée du sous-marin.

b) Calculer la probabilité pour qu'une plongée dure plus d'une semaine.

c) Sachant que le sous-marin évolue immergé depuis une semaine, calculer la probabilité pour que la plongée dépasse dix jours.

Solution :

I a)

$$C_X = [\nu, +\infty[, \qquad f_X(x) = \theta e^{-\theta(x-\nu)} \mathbf{1}_{C_X}(x).$$

$$
\begin{aligned}
M_X(t) = \mathbb{E}[e^{tX}] &= \int_{\nu}^{+\infty} e^{xt} f_X(x) dx \\
&= \theta \int_{\nu}^{\infty} e^{xt} e^{-\theta(x-\nu)} dx \\
& \quad \text{posons} \quad y = x - \nu \quad dy = dx \\
&= \theta \int_{0}^{\infty} e^{\nu t} e^{ty - y\theta} dy = \theta e^{\nu t} \int_{0}^{\infty} e^{-y(\theta - t)} dy \\
&= \theta e^{\nu t} \left[\frac{e^{-y(\theta-t)}}{t-\theta} \right]_0^{\infty} \\
&= \frac{\theta e^{\nu t}}{\theta - t} \mathbf{1}_{\{t<\theta\}}.
\end{aligned}
$$

b1) Moyenne de X

$\mu_X = \frac{dM}{dt}_{|t=0}$.

Or

$$
\begin{aligned}
M_X'(t) &= \frac{\nu \theta e^{\nu t}(\theta - t) + \theta e^{\nu t}}{(\theta - t)^2} = \frac{\theta[\nu(\theta - t)e^{\nu t} + e^{\nu t}]}{(\theta - t)^2} \\
\mu_X &= \frac{dM}{dt}_{|t=0} = \frac{\nu\theta + 1}{\theta} = \nu + \frac{1}{\theta}.
\end{aligned}
$$

b2) Variance de X

$$
\begin{aligned}
M_X''(t) &= \frac{\theta\left[\nu^2 e^{\nu t}(\theta - t) - \nu e^{\nu t} + \nu e^{\nu t}\right](\theta - t)^2 + 2(\theta - t)\theta\left[\nu e^{\nu t}(\theta - t) + e^{\nu t}\right]}{(\theta - t)^4} \\
&= \frac{\theta(\theta - t)^2 \nu^2 e^{\nu t}(\theta - t) + 2(\theta - t)\left[\theta\nu e^{\nu t}(\theta - t) + \theta e^{\nu t}\right]}{(\theta - t)^4} \\
M_X''(0) &= \frac{\theta^4\nu^2 + 2\theta^2[\theta\nu + 1]}{\theta^4} = \nu^2 + \frac{2\nu}{\theta} + \frac{2}{\theta^2}
\end{aligned}
$$

Donc

$$
\begin{aligned}
\sigma_X^2 &= \mathbb{E}(X^2) - \mu_X^2 = \frac{d^2M}{dt^2}_{|t=0} - \mu_X^2 \\
&= \nu^2 + \frac{2\nu}{\theta} + \frac{2}{\theta^2} - \nu^2 - \frac{2\nu}{\theta} - \frac{1}{\theta^2} \\
&= \frac{1}{\theta^2} \qquad \text{(indépendant de } \nu).
\end{aligned}
$$

II) a) Soit T=la durée d'une plongée en jours:

On a

$$
f_T(t) = \theta e^{-\theta t} \mathbf{1}_{t \geq 0}
$$

$$
F_T(t) = \begin{cases} 1 - e^{-\theta t} & \text{si} \quad t \geq 0 \\ 0 & \text{sinon} \end{cases}
$$

Détermination de θ

$$\begin{aligned}
\mathbb{P}(T > 6) &= 1 - F_T(6) = 0,88 \\
&\Leftrightarrow e^{-6\theta} = 0,88 \\
\theta &= -\frac{\ln(0,88)}{6} = 0,02
\end{aligned}$$

b)

$$\mathbb{P}(T > 7) = 1 - F_T(7) == e^{-7\times 0,02} = 0,87.$$

c)

$$\begin{aligned}
\mathbb{P}(T > 10|T > 7) &= \frac{\mathbb{P}\left[(T > 10) \cap (T > 7)\right]}{\mathbb{P}(T > 7)} \\
&= \frac{\mathbb{P}\left[T > 10\right]}{\mathbb{P}(T > 7)} \\
&= \frac{e^{-10\theta}}{e^{-7\theta}} = e^{-3\times 0,02} = 0,94.
\end{aligned}$$

Exercice 5. Soient X_1 et X_2 deux variables aléatoires indépendantes suivant une loi exponentielle de paramètres respectifs λ_1 et λ_2. On pose $Y = \min(X_1, X_2)$.

1. Pour tout réel y, calculer $\mathbb{P}(Y > y)$. En déduire que Y suit une loi exponentielle de paramètre $\lambda_1 + \lambda_2$.
2. Deux guichets sont ouverts à une banque. Le temps de service au premier guichet (resp. au deuxième) suit une loi exponentielle de moyenne 20 min (resp. 30 min). Deux clients rentrent simultanément, l'un choisit le guichet 1 et l'autre le guichet 2. En moyenne, après combien de temps sort le premier?
3. En moyenne, après combien de temps sort le dernier?

Solution :

$$X_i \hookrightarrow \mathcal{E}(\lambda_i) \Rightarrow f_{X_i}(x) = \lambda_i e^{-\lambda_i x}\mathbf{1}_{[0,\ \infty[}(x)$$

1. On va commencer par calculer $P(X_1 > y)$.

   Si $y \leq 0$, alors $P(X_1 > y) = 1$ (car X_1 est à valeurs dans $\mathbb{R}_+$).

   Sinon, pour $y > 0$, on a

   $$\mathbb{P}(X_1 > y) = \int_y^{+\infty} \lambda_1 e^{-\lambda_1 y} dy = e^{-\lambda_1 y}.$$

   De même,

   $$P(X_2 > y) = e^{-\lambda_2 y}.$$

   Y est à valeurs positives, donc $P(Y > y) = 1$ si $y \leq 0$. Si $y > 0$, alors

   $$P(Y > y) = P(\{X_1 > y\} \cap \{X_2 > y\}) = P(X_1 > y)P(X_2 > y) = e^{-(\lambda_1+\lambda_2)y},$$

   où on a utilisé l'indépendance de X_1 et X_2. Ainsi,

$$F_Y(y) = \begin{cases} 0 & si \quad y \leq 0 \\ 1 - e^{-(\lambda_1+\lambda_2)y} & si \quad y > 0 \end{cases}$$

On reconnait la fonction de répartition d'une loi exponentielle de paramètre $\lambda_1 + \lambda_2$. Comme la fonction de répartition caractérise la loi, on en déduit que Y suit une loi exponentielle de paramètre $\lambda_1 + \lambda_2$.

2. On cherche $E(Y)$, avec $\lambda_1 = 1/20$ et $\lambda_2 = 1/30$. L'espérance de Y est donc

$$E(Y) = \frac{1}{\lambda_1 + \lambda_2} = \frac{60}{5} = 12.$$

3. Si on pose $X = \max(Y_1, Y_2)$, on cherche l'espérance de X. Il serait possible de procéder comme précédemment, en cherchant la fonction de répartition de X. Mais comme on cherche son espérance, il y a un raisonnement plus facile. En effet, il est facile de remarquer que

$$X_1 + X_2 = X + Y$$

(la somme de deux nombres est égale à la somme de leur minimum et de leur maximum). Prenant l'espérance, on trouve

$$E(X) = E(X_1) + E(X_2) - E(Y) = 20 + 30 - 12 = 38.$$

Exercice 6. On considère une v.a. X de loi normale $\mathcal{N}(0, 1)$.
1) Montrer que pour tout $n \in \mathbb{N}$, on a: $\mathbb{E}(X^{n+2}) = (n+1)\mathbb{E}(X^n)$
2) Que vaut $\mathbb{E}(X^2)$? Déduire de ce résultat de la question précédente la valeur $\mathbb{E}(X^4)$
2) Que vaut $\mathbb{E}(X^3)$?

Solution :

Soit φ la densité de X donnée par

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \ \mathbf{1}_{x \in \mathbb{R}}$$

1) On a d'après le théorème de Transfère que, $\forall n \in \mathbb{N}$:

$$\begin{aligned} \mathbb{E}(X^{n+2}) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{n+2} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{n+1} \left[-e^{-\frac{x^2}{2}} \right]' dx \\ &= \frac{1}{\sqrt{2\pi}} \left(\left[-x^{n+1} e^{-\frac{x^2}{2}} \right]_{-\infty}^{\infty} + (n+1) \int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} dx \right) \\ &= 0 + (n+1)\mathbb{E}(X^n) \end{aligned}$$

2) $\mathbb{E}(X^2) = 1.\mathbb{E}(1) = 1$
$\mathbb{E}(X^4) = (2+1).\mathbb{E}(X^{2+2}) = (3)\mathbb{E}(X^2) = 3$
3) $\mathbb{E}(X^3) = 1.\mathbb{E}(X^{1+2}) = 2\mathbb{E}(X) = 0$
Ainsi:
Si X est centrée alors

$$\forall n \in \mathbb{N}, \quad \mathbb{E}(X^{2n+1}) = 0$$

Par récurrence, si $\mathbb{E}(X^{2n+1}) = 0$, alors

$$\mathbb{E}(X^{2n+3}) = \mathbb{E}(X^{2n+1+2}) = (2n+2)\mathbb{E}(X^{2n+1} = 0$$

Exercice 7. Dans une école d'ingénieurs, on suppose que le résultat (sur 100) à un examen d'analyse est aléatoire suivant une loi normale $\mathcal{N}(70, 14^2)$. On désire diviser en quatre classes (A, B, C, D) les résultats des étudiants qui se présentent au prochain examen en supposant que les nombres réels positifs C_1, C_2, C_3:

A="Elèves ayant un résultat supérieur ou égal à C_1"

B="Elèves ayant un résultat inférieur à C_1 mais supérieur ou égal à C_2"

C="Elèves ayant un résultat inférieur à C_2 mais supérieur ou égal à C_3"

D="Elèves ayant un résultat inférieur à C_3"

Déterminer les constantes C_1, C_2, C_3 telles que les classes A, B, C et D contiennent respectivement $10\%, 30\%, 45\%$ et 15% des étudiants.

Solution :

On a d'après l'énoncé que:

$$\begin{array}{rcl}
\mathbb{P}(A) & = & \mathbb{P}(X \geq C_1) = 0.1 \\
\mathbb{P}(B) & = & \mathbb{P}(C_2 \leq X \leq C_1) = 0.3 \\
\mathbb{P}(C) & = & \mathbb{P}(C_3 \leq X < C_2) = 0.45 \\
\mathbb{P}(D) & = & \mathbb{P}(X < C_3) = 0.15
\end{array}$$

$$\begin{array}{rcl}
\mathbb{P}(X \geq C_1) & = & \mathbb{P}(\frac{X-70}{14} \geq \frac{C_1-70}{14}) = 0.1 \\
& \Leftrightarrow & \mathbb{P}(Z \geq C_1') = 0.1 \quad \text{où} \quad Z \hookrightarrow \mathcal{N}(0,1), \quad C_1' = \dfrac{C_1-70}{14} \\
& \Leftrightarrow & \mathbb{P}(Z \leq C_1') = 0.9
\end{array}$$

D'après la table 2, on obtint $C_1' = 1.2816$ et donc $C_1 = 87,95$.

$$\begin{array}{rcl}
\mathbb{P}(C_2 \leq X \leq C_1) & = & \mathbb{P}(\frac{C_2-70}{14} \leq \frac{X-70}{14} \leq \frac{C_1-70}{14}) = 0.3 \\
& \Leftrightarrow & \mathbb{P}(C_2' \leq Z \leq C_1') = 0.3 \quad \text{où} \quad Z \hookrightarrow \mathcal{N}(0,1), \quad C_1' = \dfrac{C_1-70}{14}, \ C_2' = \dfrac{C_2-70}{14}, \\
& \Leftrightarrow & F_Z(C_1') - F_Z(C_2') = 0.3 \Leftrightarrow 0.9 - F_Z(C_2') = 0.3 \Leftrightarrow F_Z(C_2') = 0.6
\end{array}$$

D'après la table 1, on obtint $C_2' = 0.253$ et donc $C_2 = 73,54$.

Exercice 8. Une machine fabrique des vis ayant comme diamètre une variable aléatoire normale d'espérance 10mm et d'écart-type 1 mm. Une autre machine fabrique des écrous ayant comme diamètre une variable aléatoire normale d'espérance 11 mm et d'écart-type 0, 5 mm. En choisissant au hasard une vis et un écrou, quelle est la probabilité pour que la vis rentre dans l'écrou ?

Solution :

Soit X la v.a. =diamètre des vis, $X \hookrightarrow \mathcal{N}(10, 1)$
Soit Y la v.a. =diamètre des écrous, $Y \hookrightarrow \mathcal{N}(11, (0.5)^2)$
Notons $Z = Y - X$,

$$\mathbb{E}(Z) = \mathbb{E}(Y) - \mathbb{E}(X) = 1$$

$$V(Z) = V(Y) + V(X) = 1 + (0.5)^2 = 1.25$$

Donc

$$Z \hookrightarrow \mathcal{N}(1, 1.25)$$

On cherche

$$\begin{aligned}
\mathbb{P}(Z \geq 0) &= \mathbb{P}\left(\frac{Z-1}{\sqrt{1.25}} \geq \frac{-1}{\sqrt{1.25}}\right) \\
&= \mathbb{P}(T \geq -0.8944) \quad \text{où} \quad T \hookrightarrow \mathcal{N}(0,1) \\
&= \mathbb{P}(T \leq -0.8944) = 0.81327
\end{aligned}$$

Exercice 9. Dans le cadre d'une enquête hospitalière on suppose connaître pour chaque sujet la cause de son décès :

1) décès lié au cancer des bronches,
2) décès lié à toute cause (accident, autre maladie, ete, ...)

Pour les sujets de la première catégorie, on admet que la distribution du délai de survie X (exprimé en mois) suit une loi Lognormale d'espérance μ_X et de variance σ_X^2, c'est à dire que l'on peut trouver des constantes a, x_0, b, telles que la variable :

$$Y = a \ln(X - x_0) + b$$

suivre une loi normale d'espérance μ_Y et de variance σ_Y^2.

a) Calculer μ_X en fonction de a, x_0, b, μ_Y et σ_Y

b) Pour étudier la moyenne des délais de survie sur un groupe de n sujets, revient-il au même de considérer les moyennes empiriques $\overline{X}$ et $\overline{Y}$?

c) On admet maintenant que la transformation $Y = \ln X$ est telle que Y suit la loi $\mathcal{N}(1.8\,;1)$.

Avant la fin de l'enquête on désire étudier le délai moyen de survie observé sur les 16 premiers sujets, tous décédés.

Quelle est la probabilité pour que $\overline{Y} > 2.8$?

Solution :

$$X \hookrightarrow LogN(\mu_X,\ \sigma_X^2).$$

a) **Moyenne**

On a

$$\begin{aligned}
Y &= a\ln(X - x_0) + b \quad \Rightarrow \quad e^Y = e^{a\ln(X-x_0)+b} = (X - x_0)^a e^b \\
&\Rightarrow \quad (X - x_0)^a = e^{Y-b} \\
&\text{soit} \quad X = e^{\frac{Y-b}{a}} + x_0.
\end{aligned}$$

On a ainsi
$$X=\phi(Y)\quad \text{où} \quad \phi(y)=e^{\frac{y-b}{a}}+x_0.$$
Donc la moyenne de la survie de la première catégorie
$$\mu_X=\mathbb{E}(X)=\mathbb{E}(\phi(Y))$$
et d'après le théorème de transfère, on a
$$\mu_X=\mathbb{E}(\phi(Y))=\int_{-\infty}^{\infty}\phi(y)f_Y(y)dy.$$

Or
$$Y\hookrightarrow\mathcal{N}(\mu_Y,\ \sigma_Y^2)\quad\Rightarrow\quad f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_Y}e^{-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2}.$$

Notons $\mu_Y=\mu$ et $\sigma_Y=\sigma$, il vient donc que

$$\begin{aligned}
\mu_X &= \int_{-\infty}^{\infty}\frac{\left(e^{\frac{y-b}{a}}+x_0\right)}{\sqrt{2\pi}\sigma_Y}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy\\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{\frac{y-b}{a}-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy+x_0\\
&= \frac{e^{-b/a}}{\sqrt{2\pi}\sigma_Y}\int_{-\infty}^{\infty}e^{\frac{y}{a}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy+x_0\\
\text{on pose}\quad z&=\frac{y-\mu}{\sigma}\quad\Leftrightarrow\quad y=\sigma z+\mu\quad\text{et}\quad dy=\sigma dz\\
\mu_X &= x_0+\frac{e^{-b/a}e^{\mu/a}}{\sigma}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}\left[(z-\frac{\sigma}{a})^2-\frac{\sigma^2}{a^2}\right]}dz\\
&= x+e^{-\frac{b-\mu}{a}+\frac{\sigma^2}{2a^2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}t^2}dt\quad\text{après changement de variables}\\
&= x+e^{-\frac{b-\mu_Y}{a}+\frac{\sigma^2}{2a^2}}.
\end{aligned}$$

D'où
$$\mu_X=x+e^{-\frac{b-\mu_Y}{a}+\frac{\sigma_Y^2}{2a^2}}.$$

b) On a
$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\quad\text{et}\quad\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i.$$
or
$$\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i=\frac{1}{n}\sum_{i=1}^{n}[a\ln(X_i-x_0)+b]=b+\frac{1}{n}a\sum_{i=1}^{n}[\ln(X_i-x_0).$$
$$\Rightarrow\qquad\overline{X}\neq\overline{Y}$$

Il n'y a pas donc équivalence entre $\overline{X}$ qui est la moyenne arithmétique (empirique) et $\overline{Y}$ qui retrouver la "moyenne géométrique" ou "logarithmique" pondérée par a. Ceci ne donne pas la même chose.

c) Soit

$$Y \hookrightarrow \mathcal{N}(1,8\,;1) \qquad Y = \ln X$$

On cherche à calculer $\mathbb{P}(\overline{Y} > 2,8)$.

Or $\overline{Y} = \frac{1}{16}\sum_{i=1}^{16} Y_i$ avec

$$\mathbb{E}(\overline{Y}) = \mu_{Y_1} = 1,8 \qquad \mathbb{V}(Y_i) = \frac{\sigma_{Y_1}^2}{16} = \frac{1}{16}.$$

En notant

$$Z = \frac{\overline{Y} - 1,8}{1/\sqrt{16}} = 4(\overline{Y} - 1,8) \quad \hookrightarrow \mathcal{N}(0,1),$$

on obtient

$$\begin{aligned}
\mathbb{P}(\overline{Y} > 2,8) &= 1 - \mathbb{P}(\overline{Y} \leq 2,8) \\
&= 1 - \mathbb{P}(Z \leq 4(2,8 - 1,8)) = 1 - \Phi(4) = 1 - 0,999985 = 0,000014 = 0,1 \times 10^{-4}.
\end{aligned}$$

La valeur dans la table se trouve entre $\Phi(3,891)$ et $\Phi(4,417)$.

Exercice 10. (**Loi gamma et khi-deux**).

1.a) Montrez que pour tout $a > 0, b > 0$, la fonction

$$f_X : \mathbb{R} \to \mathbb{R}_+, \quad x \mapsto \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \mathbf{1}_{\mathbb{R}_+^*}. \tag{2}$$

est une densité de probabilité.

Pour une v.a. réelle X, on notera $X \hookrightarrow \gamma(a,b)$ lorsque f_X est définie par (2).

Pour la suite on suppose que $X \hookrightarrow \gamma(a,b)$.

1.b) Déterminer la fonction génératrice des moments M_X de X et en déduire que :

$$\mathbb{E}(X) = \mu_X = \frac{a}{b} \quad \text{et} \quad \sigma_X^2 = \frac{a}{b^2}.$$

2.a) Soient $X_1 \hookrightarrow \gamma(a_1, b)$ et $X_2 \hookrightarrow \gamma(a_2, b)$. Montrez que $X_1 + X_2$ suit une loi $\gamma(a_1 + a_2, b)$.

3.a) Soit $Y \sim \mathcal{N}(0,1)$. Montrez que

$$Y^2 \hookrightarrow \Gamma(\tfrac{1}{2})\pi^{-\frac{1}{2}}\gamma(\tfrac{1}{2}, \tfrac{1}{2}).$$

En déduire la valeur de $\Gamma(\frac{1}{2})$.

3.b) Soit $n \in \mathbb{N}^*$. Soient $X_i, i \in \{1, \ldots, n\}$ i.i.d t.q. $X_1 \sim \mathcal{N}(0,1)$. Montrez que

$$\sum_{i=1}^{n} X_i^2 \sim \gamma(\tfrac{n}{2}, \tfrac{1}{2}).$$

C'est-à-dire une loi de χ^2 (prononcer khi carré, voire khi-deux) à n degrés de liberté.

Solution :

1.a) Rappelons que $\Gamma(a) = \int_0^{+\infty} t^{a-1}e^{-t}\,dt$, pour tout $a \in \mathbb{C}$ t.q. $\mathrm{Re}(a) > 0$. La fonction $f_{X_{a,b}}$ est bien une probabilité car pour tout $x \in \mathbb{R}$, $f_{X_{a,b}}(x) \geq 0$ et $\int_{\mathbb{R}} f_{X_{a,b}}(x)\,dx = 1$. En effet

$$\begin{aligned}
\int_{\mathbb{R}} f_{X_{a,b}}(x)\,dx &= \frac{b^a}{\Gamma(a)}\int_0^{\infty} x^{a-1}e^{-bx}dx \\
& \quad y = bx \Rightarrow \; dy = bdx \\
&= \frac{b^a}{\Gamma(a)}\int_0^{\infty} (\frac{y}{b})^{a-1}e^{-y}\frac{dy}{b} \\
&= \frac{1}{\Gamma(a)}\int_0^{\infty} y^{a-1}e^{-y}dy = 1
\end{aligned}$$

1.b) **Méthode directe pour calculer espérance et variance.** Soit $n \in \mathbb{N}$.

$$\begin{aligned}
E[X_{a,b}^n] &= \frac{b^a}{\Gamma(a)}\int_0^{\infty} x^{n+a-1}e^{-bx}dx \\
& \quad bx = t \\
&= \frac{1}{\Gamma(a)b^n}\int_0^{\infty} t^{n+a-1}e^{-t}dt \\
&= \frac{\Gamma(n+a)}{\Gamma(a)b^n}
\end{aligned}$$

d'où pour tout $n \in \mathbb{N}^*$,

$$\begin{aligned}
\mathbb{E}[X_{a,b}^n] &= \frac{\Gamma(n+a)}{\Gamma(a)b^n} \times 1 = \frac{1}{b^n}\prod_{i=1}^{n}\frac{\Gamma((i-1+a)+1)}{\Gamma(i-1+a)} \\
&= \frac{\prod_{i=1}^{n}(i-1+a)}{b^n} = \frac{a(a+1)\dots(a+n-1)}{b^n}.
\end{aligned}$$

Cela donne $\mu_{X_{a,b}} = \frac{\Gamma(1+a)}{\Gamma(a)}\frac{1}{b} = \frac{a}{b}$ et $\sigma^2_{X_{a,b}} = \frac{\Gamma(2+a)}{\Gamma(a)}\frac{1}{b^2} - \frac{a^2}{b^2} = \frac{(1+a)a-a^2}{b^2} = \frac{a}{b^2}$.

calcul de M_X:

$$M(u) = \mathbb{E}(e^{uX}) = \frac{b}{\Gamma(a)}\int_0^{\infty} e^{-x(b-u)}\;(bx)^{a-1}dx$$

cette intégrale est convergente si et seulement si $b-u > 0$. Plaçons nous dans ce cas et faisons le changement de variable.

$$x(b-u) = t \Rightarrow x = \frac{t}{b-u} \Rightarrow bx = \frac{bt}{b-u},$$

on obtient

$$\begin{aligned}
M(u) &= \frac{b}{\Gamma(a)}\int_0^{\infty} e^{-t}\left(\frac{bt}{b-u}\right)^{a-1}\frac{dt}{b-u} \\
&= \left(\frac{b}{b-u}\right)^{a}\frac{1}{\Gamma(a)}\int_0^{\infty} e^{-t}t^{a-1}dt \\
&= \left(\frac{b}{b-u}\right)^{a}.
\end{aligned}$$

Calcul des moments en utilisant la fonction génératrice
Pour $k \geq 1$: $M^{(k)}(0) = \mathbb{E}(X^k)$.

$$M(u) = (g(u))^a \to M(u)' = ag(u)'(g(u))^{a-1}$$
$$M(u)' = \frac{ab^a}{(b-u)^{a+1}} \to M'(0) = \frac{a}{b}$$

$$M'(u) = ab^a h(u)^{a+1} \to M''(u) = \frac{ab^a(a+1)}{(b-u)^{a+2}} \to M''(0) = \frac{a(a+1)}{b^2} \to V(X) = \frac{a}{b^2}.$$

2.a) Pour X_1 et X_2 indépendantes, on a:

$$M_{X_1+X_2}(u) = \left(\frac{b}{b-u}\right)^{a_1}\left(\frac{b}{b-u}\right)^{a_2} = \left(\frac{b}{b-u}\right)^{a_1+a_2},$$

$$X_1 + X_2 \hookrightarrow \gamma(a_1+a_2, b).$$

3.a) Soit $z \in \mathbb{R}$.
Si $z \leq 0$ alors $F_{Y^2}(z) = 0$.
Si $z > 0$, $F_{Y^2}(z) = \mathbb{P}(-\sqrt{z} < Y < \sqrt{z}) = 2\int_0^{\sqrt{z}} f_Y(t)\,dt$, d'où

$$\begin{aligned} f_{Y^2}(z) &= 2 \quad \frac{1}{2\sqrt{z}} f_Y(\sqrt{z}) \\ &= \frac{1}{\sqrt{z}\sqrt{2\pi}} e^{-z/2}\mathbf{1}_{\{z\geq 0\}}. \\ &= \frac{\sqrt{2}\Gamma(1/2)}{\sqrt{2\pi}} \times \frac{1}{\sqrt{2}\Gamma(1/2)\sqrt{z}} e^{-z/2}\mathbf{1}_{\{z\geq 0\}}. \end{aligned}$$

Donc
$$Y^2 \sim \frac{\Gamma(\frac{1}{2})}{\sqrt{\pi}}\gamma(\tfrac{1}{2}, \tfrac{1}{2}),$$

donc en prenant l'espérance on trouve

$$1 = \mathbb{E}[Y^2] = \frac{\Gamma(\frac{1}{2})}{\sqrt{2\pi}}\mathbb{E}[\gamma(\tfrac{1}{2}, \tfrac{1}{2})],$$

d'où $\Gamma(\frac{1}{2}) = \sqrt{2\pi}$.

3.b) Par récurrence,
$$\sum_{i=1}^n X_i^2 \sim \gamma(\sum_{i=1}^n a_i, b).$$
Comme $X_i \hookrightarrow \mathcal{N}(0,1)$ d'où $X_i^2 \hookrightarrow \gamma(\frac{1}{2}, \frac{1}{2})$. Donc on obtient bien $\sum_{i=1}^n X_i^2 \hookrightarrow \gamma(n/2; \frac{1}{2})$.

Transformée d'une v.a. Rappeler le théorème du cours

Exercice 11. Déterminer les caractériqtiques de la variable aléatoire : $Y = \varphi(X)$ où $X \sim \mathcal{U}([0,1])$ et $\varphi(x) = e^x$.
Il faut préciser : le support, la densité de probabilité, la fonction de répartition, l'espérance et la variance de Y.

Solution :
$Y = \varphi(X)$ où $X \sim \mathcal{U}([0,1])$ et $\varphi(x) = e^x$, $\psi(x) = \varphi^{-1}(x) = \ln(x)$.
Support de $X = C_X = [0,1] \Longrightarrow C_Y = \varphi(C_X) = \exp([0,1]) = [1,e]$.
Densité de probabilité :

$$f_Y(y) = \begin{cases} f_X(\psi(y))|\psi'(y)| = 1 \times \frac{1}{y} = \frac{1}{y} & si \quad y \in [1,e] \\ 0 & sinon \end{cases}$$

Fonction de répartition :

$$F_y(y) = \begin{cases} 0 & \text{si} \quad y < 1 \\ \ln(y) & \text{si} \quad y \in [1,e] \\ 1 & \text{si} \quad y > e \end{cases}$$

Espérance : $E(Y) = E(e^X) = \displaystyle\int_{-\infty}^{+\infty} e^x f_x(x)dx = \int_0^1 e^x dx = e - 1.$

Variance de : $E(Y^2) = \displaystyle\int_{-\infty}^{+\infty} (e^x)^2 f_x(x)dx = \int_0^1 e^{2x} dx = \frac{1}{2}(e^2 - 1).$

Donc $Var(Y) = E(Y^2) - E(Y)^2 = \frac{1}{2}(e^2-1) - (e-1)^2 = \frac{-e^2+4e-3}{2}$.

Exercice 12. Soit X une variable aléatoire réelle, de densité f.
1) Déterminer la densité g de l.a. $Y = e^X$
2) Donner la forme de f_Y lorsque $\mathcal{L}(Y) \equiv \mathcal{N}(\mu, \sigma^2)$ $(\mu \in \mathbb{R}, \sigma > 0)$

Solution :
$Y = \varphi(X) = e^X$ est une v.a. réelle à valeurs strictement positives.
- L'application $\varphi : \mathbb{R} \longrightarrow \mathbb{R}_*^+$ est bijective et $\psi(y) = \varphi^{-1}(y) = \ln y$.
- Le support de Y est $C_Y = \varphi(C_X) = \exp(\mathbb{R}) =]0, +\infty[$.
- Ainsi,

$$\frac{d\varphi^{-1}(y)}{dy} = \frac{1}{y},$$

- la densité g est donnée par

$$g(y) = f_Y(y) = \frac{1}{y} f(\ln y) \mathbf{1}_{]0,+\infty[}(y).$$

2) $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$, alors $Y = e^X \hookrightarrow \log N(\mu, \sigma^2)$ où $y > 0$, qu'on appelle loi log-normale ayant pour densité

$$g(y) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{y}` \exp\left(-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right) \mathbf{1}_{]0,+\infty[}(y)$$

En raison de son symétrie (FiG. 5.9), l loi log-normle est fort utilisée pour modéliser des variables aléatoires continues ne prennant que des valeurs positives et ayant tendance à présenter des valeurs très élevées avec une faible probabilité (débit de rivières présentant de courtes périodes de crue, concentration en polluants avec présence de pics de pollution, etc...)

Exercice 13. Soit X une variable aléatoire réelle, de densité f.
1) Déterminer la densité g de l.a. $Y = \frac{1}{X}$
2) Donner la forme de f_Y lorsque $\mathcal{L}(Y) \equiv \mathcal{N}(0,1)$.

Solution :
1) On a $Y = \varphi(X)$,

$$\varphi : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : \quad x \mapsto \frac{1}{x}.$$

En toute rigueur, il faudrait utiliser le Théorème du cours que pour $\mathbb{R} \setminus \{0\}$. Cependant, la v.a. X étant absolument continue, on a $\mathbb{P}(X = 0) = 0$, de sorte que l'on peut négliger l'origine. Dans ce cas, une application de ce théorème donne

$$g(y) = \frac{1}{y^2} f(\frac{1}{y}) \quad (y \neq 0).$$

2)

$$g(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{y^2} \exp\left(-\frac{1}{2y^2}\right) \quad (y \neq 0)$$

On constante que

Remarque. Dans le cas où l'application φ n'est pas bijective, il n'y a pas de méthode générale pour déterminer la densité de $\varphi \circ X$, mais, dans la plus part des cas particuliers qui se présentent, il est possible de trouver une solution adéquate.

Exercice 14. Soit X une v.a. réelle, absolument continue, de densité f.
1) Déterminer la densité g de $Y = |X|$.
2) Donner la forme g lorsque X est paire, en particulier si $\mathcal{L}(X) \equiv \mathcal{N}(0,1)$.

Solution :

L'application $\varphi : \mathbb{R} \longrightarrow \mathbb{R}^+$, $x \mapsto |x|$ n'est pas bijective. On adopter alors la méthode suivante:

On commence par calculer la fonction de répartition F_Y de Y:

$$F_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ \mathbb{P}(Y \leq y) = \mathbb{P}(-y \leq X \leq y) & \quad y > 0. \end{cases}$$

Rappel: Soit X une v.a. admet une fonction de répartition F_X et une fonction de masse $p(x) = \mathbb{P}(X = x)$. Alors F_X est continue sur tout $\mathbb{R} \Leftrightarrow$ l'ensemble $\{x \in \mathbb{R} \quad | \; p(x) = \mathbb{P}(X = x) > 0\} = \emptyset$; on dit alors que la loi de probabilité de X est diffuse.

Comme la loi de probabilité de X est diffuse, i.e. l'ensemble $\{x \in \mathbb{R} \quad | \; p(x) = \mathbb{P}(X = x)\} = \emptyset$, alors F_X est continue sur tout $\mathbb{R}$, on peut alors écrire:

$$F_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ F_X(y) - F_X(-y) & y > 0. \end{cases}$$

On en déduit la densité f_Y de Y par dérivation:

$$f_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ f_X(y) - f_X(-y) & y > 0. \end{cases}$$

2) X est paire, on a :

$$f_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ 2f_X(y) & y > 0. \end{cases}$$

Si $\mathcal{L}(X) \equiv \mathcal{N}(0,1)$.

$$f_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ \frac{2}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} & y > 0. \end{cases}$$

Exercice 15. Soit X une v.a. réelle, absolument continue, de densité f.

1) Déterminer la densité g de $Y = X^2$.

2) Donner la forme g lorsque X est paire, en particulier si $\mathcal{L}(X) \equiv \mathcal{N}(0,1)$.

Solution :

L'application $\varphi : \mathbb{R} \longrightarrow \mathbb{R}^+$, $x \mapsto x^2$ n'est pas bijective.

On commence par calculer la fonction de répartition F_Y de Y:

$$F_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ \mathbb{P}(Y \leq y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) & y > 0. \end{cases}$$

Comme la loi de probabilité de X est diffuse, i.e. l'ensemble $\{x \in \mathbb{R} \quad | \, p(x) = \mathbb{P}(X = x)\} = \emptyset$, alors F_X est continue sur tout $\mathbb{R}$, on peut alors écrire:

$$F_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ F_X(\sqrt{y}) - F_X(-\sqrt{y}) & y > 0. \end{cases}$$

Par dérivation

$$f_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ \frac{1}{2\sqrt{y}} \left[f_X(\sqrt{y}) - f_X(-\sqrt{y}) \right] & y > 0. \end{cases}$$

Si X est paire, on a

$$f_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ \frac{1}{\sqrt{y}} f_X(\sqrt{y}) & y > 0. \end{cases}$$

Si $X \equiv \mathcal{N}(0,1)$, on a

$$f_Y(y) = \begin{cases} 0 & \text{pour} \quad y \leq 0; \\ \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} e^{-\frac{y^2}{2}} & \quad y > 0. \end{cases}$$

C'est la loi de chi-deux à un degré de liberté.

On peut aussi utiliser la méthode suivante pour trouver la densité d'une $Y = \varphi(X)$

Rappel: Théorème. Une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R})$ admet la densité f si cette fonction est borélienne positive sur $\mathbb{R}$ et vérifié

$$\mathbb{P}(A) = \int_A f(x)dx = \int f(x)\mathbf{1}_A(x)dx \quad \forall \quad A \in \mathcal{B}(\mathbb{R})$$

Exercice 16. Étant donnée une variable aléatoire X de loi uniforme sur $]0,1[$, donner la loi de la variable $Y = \lambda X + \mu$ pour λ et μ deux réels.

Solution :

On écrira, pour $a < b$,

$$\mathbb{P}(a \leq \lambda X + \mu \leq b)$$

comme l'intégrale entre a et b d'une fonction de densité. On utilisera pour cela le changement de variable $y = \lambda x + \mu$.

On pose $Y = \lambda X + \mu$ (Y est une variable aléatoire). On suppose $\lambda > 0$ (sinon il faut juste faire attention aux bornes des intégrales). Par définition de la densité de Y , notée f_Y, on a

$$\mathbb{P}(Y \in [a,b]) = \int_a^b f_Y(y)dy$$

Par ailleurs

$$\begin{aligned} \mathbb{P}(Y \in [a,b]) &= \mathbb{P}(\lambda X + \mu \in [a,b]) \\ &= \mathbb{P}\left(X \in \left[\frac{a-\mu}{\lambda}, \ \frac{b-\mu}{\lambda}\right]\right) \\ &= \int_{\frac{a-\mu}{\lambda}}^{\frac{b-\mu}{\lambda}} f_X(x)dx \\ &= \int_a^b \frac{1}{\lambda} f_X\left(\frac{y-\lambda}{\lambda}\right) dy. \end{aligned}$$

On a utilisé un changement de variable dans la dernière égalité. Cette égalité est vraie pour tous a et b. En comparant avec la définition de f_Y, on voit donc que

$$f_Y(y) = \frac{1}{\lambda} f_X\left(\frac{y-\lambda}{\lambda}\right).$$

Donc $f_Y(y) = 0$ si $y < \mu$, $f_Y(y) = 0$ si $y > \mu + \lambda$, $f_Y(y) = \frac{1}{\lambda}$ sinon.
Y est donc une v.a. de loi uniforme sur l'intervalle $[\mu, \mu + \lambda]$.

Exercice 17. Étant donnée une variable aléatoire X de loi gaussienne centrée réduite, i.e. de loi de densité

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \ \mathbf{1}_{x \in \mathbb{R}}$$

donner la loi de $Y = m + \sigma X$, pour m réel et σ réel strictement positif.

Solution :

On écrira, pour $a < b$,

$$\mathbb{P}(a \leq m + \sigma X \leq b)$$

comme l'intégrale entre a et b d'une fonction de densité. On utilisera pour cela le changement de variable $y = m + \sigma$.

C'est encore le même principe; on trouve que $Y = m + \sigma X$ est une v.a. de loi $\mathcal{N}(m, \sigma^2)$.