Contrôle

Les documents sont interdits, la calculatrice est autorisée.
Le barème est donné à titre indicatif.

Exercice 1 : Combinaison de lois normales (6 points)

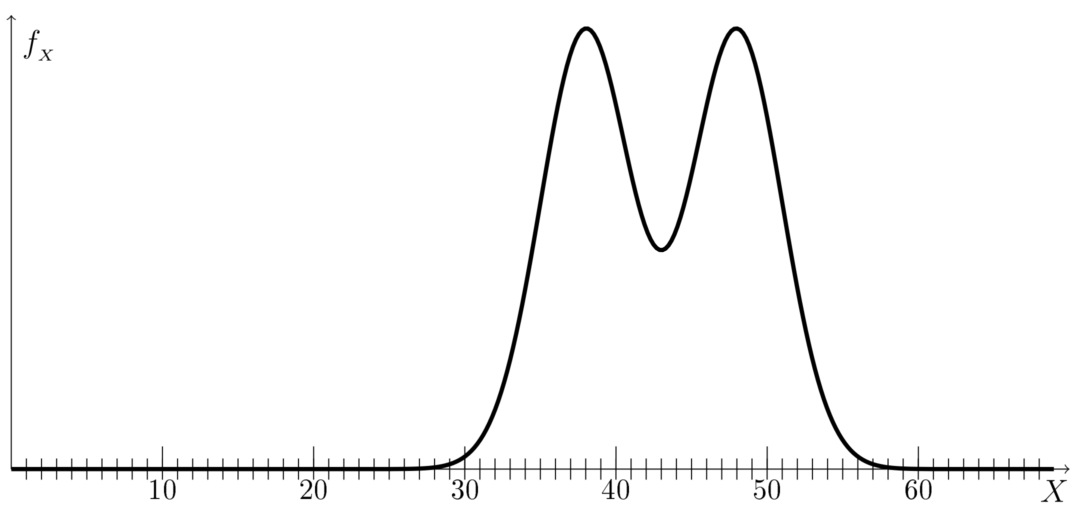
Une étude s'est intéressée à la quantité maximale de fumées toxiques (d'une certaine substance) que peuvent supporter les individus. Chez les femmes, ce seuil est modélisé par une variable X_f suivant la loi normale $\mathcal{N}(38, 9)$, et chez les hommes par X_h suivant la loi normale $\mathcal{N}(48, 9)$.

On prend un individu au hasard dans la population et on note X le seuil d'enfumage toxique qu'il peut supporter.

1. Justifier que pour tout x,

$$\mathbf{P}(X \leqslant x) = \frac{1}{2}\mathbf{P}(X_f \leqslant x) + \frac{1}{2}\mathbf{P}(X_h \leqslant x)$$

2. En déduire l'expression de la densité f_X de X.
   Nous la représentons ci-dessous.



3. Estimer graphiquement l'espérance et l'écart-type de X.
4. On observe une quantité d'enfumage de 36,7. On souhaite déterminer la part d'individus en danger.
   Calculer à 10^{-2} près les probabilités $\mathbf{P}(X_f \leqslant 36{,}7)$, $\mathbf{P}(X_h \leqslant 36{,}7)$ et $\mathbf{P}(X \leqslant 36{,}7)$.
5. En déduire simplement la quantité d'enfumage insupportable pour 83% de la population.

Exercice 2 : Produit de Pareto

On considère la loi de Pareto de densité définie sur $[1, +\infty[$ par

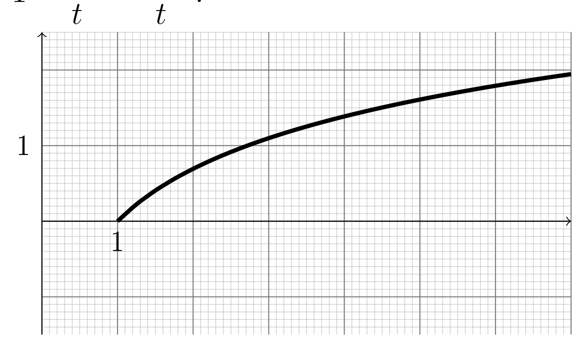
$$f(x) = \frac{1}{x^2}.$$

1. **Étude de la loi (5 points)**
   Soit X une variable aléatoire suivant la loi de Pareto.
   (a) Calculer l'espérance de X.
   (b) Déterminer la fonction de répartition de X.
   (c) Trouver la médiane de X, c'est-à-dire la valeur de m telle que $\mathbf{P}(X \leqslant m) = \frac{1}{2}$.
   (d) Soit $Z = \ln(X)$. Montrer que Z suit la loi exponentielle $\mathcal{E}(1)$.

2. **Couple de variables (5 points)**
   Soient X et Y deux variables indépendantes suivant toutes les deux la loi de Pareto.
   (a) Soit $P = XY$. Montrer que la fonction de répartition de P est définie par

$$\forall t \geqslant 1, \quad F_{P}(t) = 1 - \frac{1}{t} - \frac{\ln(t)}{t}.$$

   (b) Nous donnons ci-contre le graphe de $t \mapsto \ln(t)$ pour $t \in [1, 7]$.
   Estimer la médiane de P, en utilisant ce graphe ou une calculatrice.



3. **Théorème central limite (4 points)**
   Soit maintenant $n \geqslant 30$ et $X_1, \ldots, X_n$ des variables indépendantes suivant la loi de Pareto.
   On pose $P_n = \prod_{i=1}^{n} X_i$.
   (a) Peut-on appliquer le théorème central limite aux variables X_i ?
   (b) Écrire le théorème central limite pour une somme de n variables indépendantes Z_i de même loi $\mathcal{E}(1)$.
   (c) Déterminer ainsi la valeur de a pour laquelle $\mathbf{P}\left(\sum_{i=1}^{n} Z_i \leqslant a\right) \approx \frac{1}{2}$.
   (d) En déduire, à l'aide de la question 1-d, une estimation de la médiane de P_n.

Formulaire de probabilités

Probabilités élémentaires

• Axiomes : $(\Omega, \mathbf{P})$ espace probabilisé si $\mathbf{P}(\Omega) = 1$ et pour tous ensembles disjoints A_n,
$$\mathbf{P}(\cup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} \mathbf{P}(A_n)$$
• Union : $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$

• Indépendance : $\mathbf{P}(A \cap B) = \mathbf{P}(A) \times \mathbf{P}(B)$

• Proba conditionnelles :
$$\mathbf{P}_B(A) = \mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$
• Formules de Bayes : $\mathbf{P}(B|A) = \dfrac{\mathbf{P}(B)\mathbf{P}(A|B)}{\mathbf{P}(A)}$

et $\mathbf{P}(A) = \mathbf{P}(A|B)\mathbf{P}(B) + \mathbf{P}(A|\overline{B})\mathbf{P}(\overline{B})$

Densité, fonction de répartition

• Variable aléatoire : $X : \Omega \to \mathbf{R}$

• Cas des variables continues

Densité f_X de X : $\mathbf{P}(X \in [a, b]) = \int_a^b f_X$

Fonction de répartition : $F_X(x) = \mathbf{P}(X \leqslant x)$

Déduire la densité de la f° de répart. : $f_X = F'_X$

Espérance, Variance

• Espérance : $\mathbf{E}(X) = \sum_i x_i \mathbf{P}_X(x_i)$ (discret)

$\mathbf{E}(X) = \int_{-\infty}^{+\infty} x f_X(x) \mathrm{d}x$ (continu)

• Variance : $\mathbf{V}(X) = \sum_i (x_i - \mathbf{E}(X))^2 \mathbf{P}_X(x_i)$

$\mathbf{V}(X) = \int_{-\infty}^{+\infty} (x - \mathbf{E}(X))^2 f_X(x) \mathrm{d}x$

• Écart-type : $\sigma(X) = \sqrt{\mathbf{V}(X)}$

• Propriétés : $\mathbf{E}(X + \alpha Y) = \mathbf{E}(X) + \alpha \mathbf{E}(Y)$

$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$

$\mathbf{V}(\alpha X + \beta) = \alpha^2 \mathbf{V}(X)$

• X, Y indépendantes $\implies$ $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$

et $\mathbf{V}(X + Y) = \mathbf{V}(X) + \mathbf{V}(Y)$

Couple de variables

X, Y variables indépendantes de densités f_X, f_Y

$\mathcal{A} \subset \mathbf{R}^2$.

Alors $\mathbf{P}((X, Y) \in \mathcal{A}) = \iint_{\mathcal{A}} f_X(x) f_Y(y) \mathrm{d}y \mathrm{d}x$.

Représenter $\mathcal{A}$ et le paramétrer pour Fubini.

Lois usuelles

• Bernoulli : $X \sim \mathcal{B}(p)$

$\mathbf{P}(X = 1) = p, \quad \mathbf{P}(X = 0) = 1 - p$

$\mathbf{E}(X) = p \qquad \mathbf{V}(X) = p(1 - p)$

• Binomiale : $X \sim \mathcal{B}(n, p)$

$\mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k},\ k \in [\![0, n]\!]$

$\mathbf{E}(X) = np \qquad \mathbf{V}(X) = np(1 - p)$

• Géométrique : $X \sim \mathcal{G}(p)$

$\mathbf{P}(X = k) = p(1 - p)^{k-1}, k \in \mathbf{N}^*$

$\mathbf{E}(X) = \frac{1}{p} \qquad \mathbf{V}(X) = \frac{1-p}{p^2}$

• Uniforme : $X \sim \mathcal{U}([a, b])$

$f_X(x) = \frac{1}{b-a}, \quad x \in [a, b].$

$\mathbf{E}(X) = \frac{a+b}{2} \qquad \mathbf{V}(X) = \frac{(b-a)^2}{12}$

• Exponentielle : $X \sim \mathcal{E}(\lambda)$

$f_X(x) = \lambda \mathrm{e}^{-\lambda x}, \quad x \in \mathbf{R}_+$

$\mathbf{E}(X) = \frac{1}{\lambda} \qquad \mathbf{V}(X) = \frac{1}{\lambda^2}$

• Loi normale : $X \sim \mathcal{N}(m, \sigma^2)$

$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbf{R}$

$\mathbf{E}(X) = m \qquad \mathbf{V}(X) = \sigma^2.$

$X \sim \mathcal{N}(m, \sigma^2) \implies \frac{X-m}{\sigma} \sim \mathcal{N}(0, 1)$

$X_1 \sim \mathcal{N}(m_1, \sigma_1^2)$, $X_2 \sim \mathcal{N}(m_2, \sigma_2^2)$ indép.

$\implies X_1 + X_2 \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

Théorèmes limites

• Loi forte : (X_k) iid d'espérance m

alors $\frac{1}{n} \sum_{k=1}^n X_k \to m$ presque sûrement

• TCL : (X_k) iid d'espérance m, variance σ^2

alors $\dfrac{\sum_{k=1}^n X_k - nm}{\sigma\sqrt{n}} \to_{n\infty,\ \text{en loi}} \mathcal{N}(0, 1)$

• En pratique : pour $n \geqslant 30$, on a à peu près
$$\frac{\sum_1^n X_k - nm}{\sigma\sqrt{n}} \sim \mathcal{N}(0, 1) \qquad \sum_{k=1}^n X_k \sim \mathcal{N}(nm, n\sigma^2)$$
$$\frac{\frac{1}{n}\sum_1^n X_k - m}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$$

Quelques valeurs de la fonction de répartition de la loi normale $\mathcal{N}(0, 1)$:

x	0	0,25	0,43	0,61	0,84	1,08	1,30	1,65	1,96	2,33	2,58	3,0
$F_{\mathcal{N}}(x)$	0,5	0,60	0,66	0,73	0,80	0,86	0,90	0,95	0,975	0,99	0,995	0,999

$$F_{\mathcal{N}}(-a) = 1 - F_{\mathcal{N}}(a) \qquad \mathbf{P}_{\mathcal{N}}([-a, a]) = 2F_{\mathcal{N}}(a) - 1$$