

TD STATISTIQUE INFÉRENTIELLE
L3 Maths-Eco

Frédéric Lavancier

1 Compléments de probabilités

Exercice 1.

1. Soit X une variable aléatoire suivant une loi $\gamma(a, b)$ avec $a > 0$ et $b > 0$. Montrer que $X \in L^2$ et que $E(X) = a/b$, $V(X) = a/b^2$.
2. Soit X et Y deux variables aléatoires indépendantes avec $X \sim \gamma(a, b)$, $Y \sim \gamma(a', b)$, où $a > 0$, $a' > 0$, $b > 0$. Montrer que $X + Y \sim \gamma(a + a', b)$.
Indication : on pourra dans un premier temps montrer que le produit de convolution des densités est proportionnelle à $e^{-bx}x^{a+a'}1_{x>0}$, puis conclure.
3. Soit $Z \sim \mathcal{N}(0, 1)$, montrer que $Z^2 \sim \gamma(1/2, 1/2)$. On rappelle que $\Gamma(1/2) = \sqrt{\pi}$.
4. Dédurre des deux questions précédentes que si Z_1, \dots, Z_n sont indépendantes et identiquement distribuées selon une $\mathcal{N}(0, 1)$, alors $\sum_{i=1}^n Z_i^2 \sim \chi^2(n)$.
5. Soit X une variable aléatoire suivant une loi de Student à n degré de liberté, où $n \in \mathbb{N}^*$. Montrer que $X \in L^1$ ssi $n > 1$, que $X \in L^2$ ssi $n > 2$, et que pour $n > 1$, $E(X) = 0$.

Exercice 2. Soit $(U_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes de loi uniforme sur l'intervalle $[0, \theta]$, où $\theta > 0$. On pose pour $n \geq 1$, $M_n = \max_{1 \leq i \leq n} U_i$.

1. Montrer que $(M_n)_{n \geq 1}$ converge presque sûrement et déterminer sa limite. On pourra calculer $P(|M_n - \theta| > \epsilon)$ pour $\epsilon > 0$.
2. Etudier la convergence en loi de la suite $(n(\theta - M_n))_{n \geq 1}$.

Exercice 3. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes telle que pour tout $n \geq 1$, X_n suit la loi de Bernoulli de paramètre $1/\sqrt{n}$:

$$P(X_n = 1) = \frac{1}{\sqrt{n}}, \quad P(X_n = 0) = 1 - \frac{1}{\sqrt{n}}.$$

On note, pour tout $n \geq 1$, $S_n = \sum_{i=1}^n X_i$ et $Y_n = S_n/E(S_n)$.

1. Montrer que $V(S_n) \leq E(S_n)$.
2. En utilisant l'inégalité de Tchebychev, montrer que la suite $(Y_n)_{n \geq 1}$ converge vers 1 en probabilité.
On rappelle que, pour tout $n \geq 1$, $2\sqrt{n} - 2 \leq \sum_{i=1}^n \frac{1}{\sqrt{i}} \leq 2\sqrt{n}$.

Exercice 4. On effectue n séries de 400 tirages de pile ou face avec une pièce équilibrée. On observe les fréquences empiriques de pile F_1, \dots, F_n dans ces séries.

1. Quelle est (approximativement) la loi de probabilité du nombre N de ces fréquences $(F_i)_{1 \leq i \leq n}$ qui ne vérifient pas la condition $0.45 < F_i < 0.55$, lorsque $n = 20$?
2. Est-il plus probable que $N = 0$, que $N = 1$ ou que $N \geq 2$?

Exercice 5. Soit X_N une variable aléatoire de loi hypergéométrique de paramètre (N, m, n) . On rappelle que X_N représente le nombre de boules blanches obtenues lors d'un tirage sans remise de n boules hors d'une urne contenant m boules blanches et $N - m$ boules noires.

1. Vérifier que $P(X_N = k) = \frac{\binom{m}{k} \binom{N-m}{n-k}}{\binom{N}{n}} = \frac{\binom{n}{k} \binom{N-n}{m-k}}{\binom{N}{m}}$ pour $n - N + m \leq k \leq m$ et $0 \leq k \leq n$.
2. On suppose que le tirage est de taille n fixé et que m et N tendent vers $+\infty$ avec $\lim_{N \rightarrow +\infty} m/N = \theta \in [0, 1]$ (θ est la proportion limite du nombre de boules blanches dans l'urne). Montrer que la suite $(X_N)_{N \in \mathbb{N}^*}$ converge en loi vers la loi binomiale de paramètre (n, θ) .
3. Application aux sondages : on considère une population de N individus dont un nombre inconnu m ont l'intention de voter pour le candidat A lors d'une élection. On effectue un sondage sur 1000 personnes pour estimer la proportion de votes pour A dans la population totale. Donner la loi du nombre d'intentions de votes pour A dans l'échantillon, selon que le sondage a été effectué avec remise ou sans remise. Quelle approximation peut-on faire ?

Exercice 6. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles indépendantes et identiquement distribuées suivant la loi gaussienne $\mathcal{N}(0, 1)$. On note, pour tout $n \geq 1$,

$$S_n = \sum_{i=1}^n X_i, \quad M_n = \exp(S_n - n/2).$$

1. Calculer, pour $s \in \mathbb{R}$, $E(e^{sX_1})$.

2. Justifier la convergence presque sûre de $(S_n/n)_{n \geq 1}$ et préciser la limite.
3. En déduire que $(M_n)_{n \geq 1}$ convergence presque sûrement vers 0.
4. Montrer que la suite de terme général

$$Y_n = \frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n e^{X_i}}$$

converge presque sûrement et préciser la limite.

5. Calculer, pour tout $n \geq 1$, $E(M_n)$.
6. La convergence de $(M_n)_{n \geq 1}$ a-t-elle lieu dans L^1 ?

Exercice 7. Soit (X_n) une suite de variables aléatoires réelles indépendantes et identiquement distribuées. On suppose que X_1 est de carré intégrable et on note $m = E(X_1)$, $\sigma^2 = V(X_1)$, $\bar{X}_n = \sum_{i=1}^n X_i/n$, $S_n^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2/n$ et $T_n = \sqrt{n}(\bar{X}_n - m)$

1. Montrer que la suite de terme général (T_n, \bar{X}_n) converge en loi vers (G, m) où G suit la loi $\mathcal{N}(0, \sigma^2)$.
2. En déduire que la suite de terme général $U_n = \sqrt{n}((\bar{X}_n)^2 - m^2)$ (resp. $W_n = \sqrt{n}(\sqrt{n}\bar{X}_n - \sqrt{nm})$) converge en loi vers U (resp. W) de loi $\mathcal{N}(0, 4m^2\sigma^2)$ (resp. $\mathcal{N}(0, \sigma^2)$).
3. En utilisant la méthode delta, retrouver le résultat de la question précédente.
4. Montrer que la suite de terme général $Z_n = T_n/S_n$ converge en loi vers Z de loi $\mathcal{N}(0, 1)$.
5. Quelle est la loi de Z_n si l'on suppose que X_1 suit une loi normale ?

Exercice 8. Soit (X_1, X_2) un vecteur gaussien centré dans \mathbb{R}^2 et de matrice de variance-covariance.

$$\Sigma = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$$

1. Déterminer la densité du vecteur (X_1, X_2) .
2. Considérons la variable aléatoire réelle $Y = X_2 - aX_1$, où a est un nombre réel quelconque. Quelle est la loi de du couple (X_1, Y) ?
3. En déduire que X_1 et Y sont indépendantes si et seulement si $a = 1$.

Exercice 9. Soit (X, Y) un couple aléatoire dont la densité sur \mathbb{R}^2 est donnée par

$$f_{X,Y}(x, y) = \frac{1}{4\pi} e^{-\frac{1}{2}(\frac{1}{2}x^2 - xy + y^2)}, \quad x, y \in \mathbb{R}.$$

1. Déterminer la loi du couple (X, Y) .
2. Considérons la variable aléatoire U définie par $Y = aX + U$, où $a \in \mathbb{R}$. Montrer que U et X sont indépendantes si et seulement si $a = 1/2$.

Exercice 10. Soit X une variable aléatoire de loi $\mathcal{N}(0, 1)$ et T indépendante de X telle que $P(T = 1) = P(T = -1) = 1/2$.

1. Déterminer la loi de $Y = TX$.
2. Montrer que le vecteur aléatoire (X, Y) n'est pas gaussien.
Indication : calculer $P(X + Y = 0)$.
3. Calculer la covariance de X et de Y .
4. Montrer que X et Y ne sont pas indépendants.

Exercice 11. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles indépendantes et identiquement distribuées suivant la loi gaussienne $\mathcal{N}(0, 1)$. Considérons pour tout $n \geq 1$ la variable $B_n = \sum_{i=1}^n X_i$.

1. Déterminer la loi du vecteur aléatoire (B_1, \dots, B_n) .
2. Étant donné $m \in \{1, \dots, n - 1\}$, on pose $Y_m = B_m - \frac{m}{n}B_n$. Montrer que Y_m et B_n sont indépendantes.