Option économique

MATHEMATIQUES

8 Janvier 2024

Problème 1

Correction proposée par Frédéric Gaunard

Partie 1 : Une loi de probabilité à deux paramètres

Toutes les variables aléatoires intervenant dans cet exercice sont supposées définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

Soient p, q deux réels de $]0; 1[$ tels que $p + q = 1$ et $n \in \mathbb{N}^*$.

On dispose d'une urne contenant des boules rouges en proportion p et des boules vertes en proportion q . On effectue dans cette urne une suite de tirages d'une boule avec remise jusqu'à ce que l'on ait pioché n boules rouges.

On suppose que les résultats des différents tirages sont indépendants.

Pour tout i de $\mathbb{N}^*$, on note R_i (resp. V_i) l'événement : " le i -ième tirage amène une boule rouge (resp. verte)". Ainsi, on a : $\forall i \in \mathbb{N}^*, P(R_i) = p$ et $P(V_i) = q$.

On note Z la variable aléatoire qui prend la valeur du nombre de boules vertes obtenues avant l'apparition de la n -ième boule rouge. On admet qu'on définit bien une variable aléatoire.

1. On peut n'avoir obtenu aucune boule verte (c'est à dire les n boules rouges consécutivement) comme on peut en avoir un nombre arbitrairement grand avant d'avoir eu toutes nos boules rouges donc $Z(\Omega) = \mathbb{N}$.

2. (a) L'évènement $[Z = n]$ signifie qu'on a obtenu n boules vertes avant l'apparition de la première boule rouge

- Si $n = 0$, on a : $[Z = 0] = R_1$ donc $P([Z = 0]) = p = pq^0$.

- Si $n > 0$, on a : $[Z = n] = V_1 \cap \dots \cap V_n \cap R_{n+1}$

ce qui donne par indépendance des tirages :

$$P([Z = n]) = P(V_1 \cap \dots \cap V_n \cap R_{n+1}) = P(V_1) \times \dots \times P(V_n) \times P(R_{n+1}) = pq^n.$$

Ainsi, on a bien : $\forall n \in \mathbb{N}, P([Z = n]) = pq^n$.

(b) i. On a $Z(\Omega) = \mathbb{N}$ donc $(Z + 1)(\Omega) = \mathbb{N}^*$.

Pour tout $n \in \mathbb{N}^*$, on a : $P(Z + 1 = n) = P(Z = n - 1) = pq^{n-1}$ d'après la question précédente, on reconnaît alors la formule d'une loi géométrique donc $Z + 1$ suit une loi géométrique de paramètre p .

ii. On en déduit d'après le cours que : $E(Z + 1) = \frac{1}{p}$.

Ainsi, linéarité de l'espérance, Z admet une espérance et on a

$$E(Z + 1) = E(Z) + 1 = \frac{1}{p} \text{ donc } E(Z) = \frac{1}{p} - 1 = \frac{1 - p}{p} \text{ soit } E(Z) = \frac{q}{p}.$$

iii. De même : $V(Z + 1) = \frac{q}{p^2}$ puis par propriété de la variance, Z admet une variance et on a :

$$V(Z + 1) = V(Z) \text{ soit } \boxed{V(Z) = \frac{q}{p^2}.}$$

3. (a) On reconnaît pour X_1 le temps d'attente du premier succès (obtenir la boule rouge) lors de répétitions identiques et indépendantes du même schéma de Bernoulli. Ainsi, X_1 suit donc une loi géométrique de paramètre p .
- (b) Observons que si $X_1(\Omega) = \mathbb{N}^*$, alors $X_2(\Omega) = \llbracket 2, +\infty \llbracket$. Soit donc $(k, j) \in X_1(\Omega) \times X_2(\Omega)$. Premièrement, si $j \leq k$, alors $P([X_1 = k] \cap [X_2 = j]) = 0$ car la seconde boule rouge arrive nécessairement après la première. Si $j > k$, alors

$$\begin{aligned} P([X_1 = k] \cap [X_2 = j]) &= P(V_1 \cap \dots \cap V_{k-1} \cap R_k \cap V_{k+1} \cap \dots \cap V_{j-1} \cap R_j) \\ &= \prod_{i=1}^{k-1} P(V_i) \times P(R_k) \times \prod_{i=k+1}^{j-1} P(V_i) \times P(R_j) \\ &= p^2 q^{j-2} \end{aligned}$$

par indépendance des tirages

Au final

$$\boxed{P([X_1 = k] \cap [X_2 = j]) = \begin{cases} 0, & \text{si } j \leq k \\ p^2 q^{j-2}, & \text{si } j > k \end{cases}}$$

(c) Soit $j \geq 2$. Par la formule des probabilités totales appliquée au s.c.e $\{[X_1 = k] : k \in \mathbb{N}^*\}$, on a

$$\begin{aligned} P(X_2 = j) &= \sum_{k=1}^{+\infty} P([X_1 = k] \cap [X_2 = j]) \\ &= \sum_{k=1}^{j-1} P([X_1 = k] \cap [X_2 = j]) = \sum_{k=1}^{j-1} p^2 q^{j-2} \\ &= \boxed{(j-1)p^2 q^{j-2}} \end{aligned}$$

(d) On sait que X_2 admet une espérance si et seulement si la série $\sum jP(X_2 = j)$ converge. Or,

$$jP(X_2 = j) = j(j-1)p^2 q^{j-2}$$

et on reconnaît le multiple du terme général d'une série géométrique dérivée d'ordre 2, de raison q (avec $0 < q < 1$) donc convergente. Ainsi, X_2 admet une espérance et

$$E(X_2) = p^2 \sum_{j=2}^{+\infty} j(j-1)q^{j-2} = p^2 \times \frac{2}{(1-q)^3} = \boxed{\frac{2}{p}}$$

(e) Comme Z représente ici le nombre de boules vertes tirées avant d'obtenir deux rouges, on a $Z = X_2 - 2$. Par linéarité de l'espérance, Z admet une espérance et

$$E(Z) = E(X_2) - 2 = \frac{2}{p} - 2 = \frac{2-2p}{p} = \frac{2(1-p)}{p} = \frac{2q}{p}.$$

4. La variable nR compte le nombre de boules rouges obtenues. On pioche donc jusqu'à en avoir n , en augmentant à chaque fois le nombre t de tirages de 1. On récupère alors le nombre de vertes en soustrayant n au nombre de tirages.

```

def simul_BN(n, q):
    nR=0
    t=1
    while nR<n :
        if rd.random()>q:
            nR=nR+1
        t=t+1
    return (t-n)

```

Partie 2 : Un calcul statistique

Soit $N \in \mathbb{N}$. On rappelle que, si $(x_i)_{i \in [1, N]}$ et $(y_i)_{i \in [1, N]}$ sont deux séries statistiques,

- on désigne par $\bar{x}$ (resp. $\bar{y}$) la moyenne empirique associée à (x_i) (resp. à (y_i));
- σ_x et σ_y désignent les écarts-types empiriques des séries (x_i) et (y_i) .
- $\text{cov}(x, y)$ la covariance empirique de la série statistique double (x_i, y_i) ;
- $\rho(x, y)$ le coefficient de corrélation linéaire empirique de cette même série statistique double.

5. D'après le cours,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \sigma_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

et

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Par König-Huyguens,

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}$$

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \text{cov}(x, x)$$

6. On applique simplement la formule ci-dessous avec la commande `np.mean()` qui renvoie la moyenne des éléments d'une liste en argument.

```

def covariance(x, y):
    prod=x*y
    return (np.mean(prod)-np.mean(x)*np.mean(y))

```

7. (a) La fonction `simul_BN()` permet de simuler chacune des variables X_i .

```

def sample_TW(q, N):
    X_1=np.zeros(N)
    X_2=np.zeros(N)
    X_3=np.zeros(N)
    for i in range(N):
        X_1[i]=simul_BN(1, p)
        X_2[i]=simul_BN(1, p)
        X_3[i]=simul_BN(1, p)
    T=[X_1[i]+ X_2[i] for i in range(N)]
    W=[X_2[i]+ X_3[i] for i in range(N)]
    return ([T,W])

```

- (b) La formule pour r calcule le coefficient de corrélation linéaire des séries statistiques (t_i) et (w_i) obtenues par simulation (pour un échantillon de taille 1000) avec cinq valeurs différentes de q entre 0.1 et 0.9 . On voit que le résultat affiché semble toujours proche de 1/2 et ne dépendrait donc pas de la valeur de q .
8. (a) On a par double linéarité de la covariance :

$$\begin{aligned} \text{Cov}(T, W) &= \text{Cov}(X_1 + X_2, X_2 + X_3) \\ &= \underbrace{\text{Cov}(X_1, X_2)} + \underbrace{\text{Cov}(X_1, X_3)}_{=0 \text{ par indép.}} + \underbrace{\text{Cov}(X_2, X_2)}_{=V(X_2)} + \underbrace{\text{Cov}(X_2, X_3)}_{=0 \text{ par indép.}} \\ &= V(X_2) \end{aligned}$$

Comme $\text{Cov}(T, W) = V(X_2) \neq 0$, on en déduit que les variables T et W ne sont pas indépendantes.

(b) On a :

$$\begin{aligned} V(T) &= V(X_1 + X_2) \\ &= V(X_1) + V(X_2) \quad \text{par indépendance} \\ &= 2V(X_2) \quad \text{même loi} \end{aligned}$$

De même, on a : $V(W) = 2V(X_2)$

D'autre part, on a :

$$\rho(T, W) = \frac{\text{Cov}(T, W)}{\sqrt{V(T)}\sqrt{V(W)}} = \frac{V(X_2)}{\sqrt{V(T)}\sqrt{V(W)}} = \frac{V(X_2)}{\sqrt{2V(X_2)}\sqrt{2V(X_2)}} = \frac{V(X_2)}{\sqrt{4V(X_2)}} \quad \text{soit}$$

Notre conjecture de la question 4(b)ii est bien vérifiée.

Partie 3 : Loi du maximum

9. • Si $i < j$, l'évènement $([U = i] \cap [V = j])$ est impossible car le maximum de X et Y ne peut pas être strictement inférieur au minimum de X et Y .

Donc $P([U = i] \cap [V = j]) = 0$ si $i < j$.

- Si $i = j$, le maximum et le minimum sont égaux à i , c'est à dire que X et Y sont tous les deux égaux à i donc on a : $([U = i] \cap [V = i]) = ([X = i] \cap [Y = i])$. On a alors :

$$\begin{aligned} P([U = i] \cap [V = i]) &= P([X = i] \cap [Y = i]) \\ &= P([X = i]) \times P([Y = i]) \quad \text{par indépendance} \\ &= pq^i \times pq^i \quad \text{loi binomiale négative} \\ &= p^2q^{2i} \\ &-11/17 \end{aligned}$$

Ainsi, $P([U = i] \cap [V = i]) = p^2q^{2i}$ si $i = j$

- Si $i > j$, l'une des deux variables est égale à i (le maximum) et l'autre à j (le minimum) donc on a : $([U = i] \cap [V = i]) = (([X = i] \cap [Y = j]) \cup ([X = j] \cap [Y = i]))$. On a alors :

$$\begin{aligned} P([U = i] \cap [V = j]) &= P(([X = i] \cap [Y = j]) \cup ([X = j] \cap [Y = i])) \\ &= P([X = i] \cap [Y = j]) + P([X = j] \cap [Y = i]) \quad \text{par incompatibilité} \\ &= P([X = i]) \times P([Y = j]) + P([X = j]) \times P([Y = i]) \quad \text{par indépendance} \\ &= pq^i \times pq^j + pq^j \times pq^i \quad \text{loi binomiale négative} \\ &= 2p^2q^{i+j} \end{aligned}$$

Ainsi,

$$\span style="border: 1px solid black; padding: 2px;"> $P([U = i] \cap [V = j]) = 2p^2q^{i+j}$ si $i > j$$$

10. (a) D'après la FPT avec le SCE $([U = i])_{i \in \mathbb{N}}$, on a :

$$\begin{aligned} P([V = j]) &= \sum_{i=0}^{+\infty} P([U = i] \cap [V = j]) \\ &= \sum_{i=0}^{j-1} \underbrace{P([U = i] \cap [V = j])}_{=0 \text{ car } i < j} + \sum_{i=j}^{+\infty} P([U = i] \cap [V = j]) \\ &= \sum_{i=j}^{+\infty} P([U = i] \cap [V = j]) \end{aligned}$$

Ainsi, $\forall j \in \mathbb{N}, P([V = j]) = \sum_{i=j}^{+\infty} P([U = i] \cap [V = j]).$

(b) On a alors :

$$\begin{aligned} P([V = j]) &= \sum_{i=j}^{+\infty} P([U = i] \cap [V = j]) \\ &= P([U = j] \cap [V = j]) + \sum_{i=j+1}^{+\infty} P([U = i] \cap [V = j]) \\ &= p^2 q^{2j} + \sum_{i=j+1}^{+\infty} 2p^2 q^{i+j} \\ &= p^2 q^{2j} + 2p^2 q^j \sum_{i=j+1}^{+\infty} q^i \\ &= p^2 q^{2j} + 2p^2 q^j \sum_{k=0}^{+\infty} q^{k+j+1} \quad \text{changement d'indice } k = i - (j + 1) \\ &= p^2 q^{2j} + 2p^2 q^j q^{j+1} \sum_{k=0}^{+\infty} q^k \\ &= p^2 q^{2j} + 2p^2 q^{2j+1} \frac{1}{1-q} \\ &= p^2 q^{2j} + 2p q^{2j+1} \\ &= p q^{2j} (p + 2q) \\ &= p q^{2j} (1 - q + 2q) \\ &= p(1 + q) q^{2j} \end{aligned}$$

$$\forall j \in \mathbb{N}, P([V = j]) = p(1 + q) q^{2j}.$$

(c) Il s'agit de montrer que $P([V = j]) = \tilde{p}(1 - \tilde{p})^j$ avec $\tilde{p} = (1 - q^2)$ soit $P([V = j]) = (1 - q^2) (q^2)^j$.

Or, on a

$$P([V = j]) = p(1 + q) q^{2j} = (1 - q)(1 + q) (q^2)^j = (1 - q^2) (q^2)^j$$

Ainsi, V suit une loi binomiale négative de paramètre $(1 - q^2)$.

(d) On utilise la formule de la relation $(\star)$ en remplaçant p par $(1 - q^2)$ et donc q par $1 - (1 - q^2) = q^2$,

ce qui donne : $E(V) = \frac{q^2}{1 - q^2}.$

11. (a) Si $U = \max(X, Y) = X$, alors $V = \min(X, Y) = Y$ et donc $U + V = X + Y$
 Si $U = \max(X, Y) = Y$, alors $V = \min(X, Y) = X$ et donc $U + V = X + Y$.

Ce sont les deux seules possibilités donc dans tous les cas, on a $\boxed{U + V = X + Y}$.

(b)

$$E(U) = E(X + Y - V) = E(X) + E(Y) - E(V) \quad \text{par linéarité de l'espérance}$$

$$\begin{aligned} &= \frac{q}{p} + \frac{q}{p} - \frac{q^2}{1 - q^2} \\ &= \frac{2q}{(1 - q)} - \frac{q^2}{1 - q^2} = \frac{2q(1 + q)}{(1 - q)(1 + q)} - \frac{q^2}{1 - q^2} = \frac{2q + 2q^2 - q^2}{1 - q^2} = \frac{2q + q^2}{1 - q^2} \end{aligned}$$

Ainsi, $\boxed{E(U) = \frac{2q + q^2}{1 - q^2}}$.

Partie 4 : Partie entière d'une variable à densité

12. (a) Un simple calcul donner $\lim_{x \rightarrow 0^+} f(x) = \lambda$ et $\lim_{x \rightarrow 0^-} f(x) = 0 \neq \lambda$. Ainsi f n'est pas continue sur $\mathbb{R}$.

- (b) La fonction f est clairement positive sur $\mathbb{R}$, continue sur $\mathbb{R}$ sauf en 0 et comme f est nulle sur $\mathbb{R}_-^*$,

$\int_{-\infty}^0 f(t) dt$ converge et vaut 0. Par ailleurs, pour $A > 0$:

$$\int_0^A f(t) dt = \int_0^A \lambda e^{-\lambda t} dt = [-e^{-\lambda t}]_0^A = 1 - e^{-\lambda A}.$$

Or $\lim_{A \rightarrow +\infty} e^{-\lambda A} = 0$, donc $\int_0^{+\infty} f(t) dt$ converge vers 1, et de même $\int_{-\infty}^{+\infty} f(t) dt$ converge vers 1. Ainsi

$\boxed{f \text{ est bien une densité de probabilité}}$.

13. (a) On a clairement, pour tout $x < 0$, $\int_{-\infty}^x f(t) dt = 0$, donc $F_{S_1}(x) = 0$.

Si $x \geq 0$,

$$\int_{-\infty}^x f(t) dt = \int_{-\infty}^0 f(t) dt + \int_0^x f(t) dt = \int_0^x \lambda e^{-\lambda t} dt = [-e^{-\lambda t}]_0^x = 1 - e^{-\lambda x}$$

On a : $\boxed{F_{S_1}(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{sinon.} \end{cases}}$.

- (b) Pour tout entier k , calculer $P(T_1 = k)$ puis en déduire que T_1 suit une loi binomiale négative de paramètre $p_\lambda = 1 - e^{-\lambda}$.

On remarquera que T_2 suit la même loi que T_1 .

On a $T_1(\Omega) = \mathbb{N}$. Soit $k \in \mathbb{N}$. On a :

$$\begin{aligned} P(T_1 = k) &= P(k \leq S_1 < k + 1) \\ &= F_{S_1}(k + 1) - F_{S_1}(k) \\ &= 1 - e^{-\lambda(k+1)} - 1 + e^{-\lambda k} \\ &= e^{-\lambda k} (1 - e^{-\lambda}) \\ &= (e^{-\lambda})^k (1 - e^{-\lambda}) \\ &= q_\lambda^k p_\lambda \end{aligned}$$

avec $p_\lambda = 1 - e^{-\lambda}$ et $q_\lambda = 1 - p_\lambda = e^{-\lambda}$.

Ainsi, T_1 suit une loi binomiale négative de paramètre $p_\lambda = 1 - e^{-\lambda}$

14. (a) Comme S_1 et S_2 sont indépendantes, on en déduit que T_1 et T_2 sont indépendantes par le lemme des coalitions.

(b) On note $q_\lambda = 1 - p_\lambda$. Montrer que : $P(T_1 = T_2) = \sum_{k=0}^{+\infty} p_\lambda^2 q_\lambda^{2k}$.

On a $T_1(\Omega) = \mathbb{N}$. D'après la FPT avec le SCE $(T_1 = k)_{k \in \mathbb{N}}$, on a :

$$\begin{aligned} P(T_1 = T_2) &= \sum_{k=0}^{+\infty} P((T_1 = T_2) \cap (T_1 = k)) \\ &= \sum_{k=0}^{+\infty} P((T_2 = k) \cap (T_1 = k)) \\ &= \sum_{k=0}^{+\infty} P(T_2 = k) P(T_1 = k) \quad \text{par indépendance} \\ &= \sum_{k=0}^{+\infty} q_\lambda^k p_\lambda q_\lambda^k p_\lambda \\ &= \sum_{k=0}^{+\infty} p_\lambda^2 q_\lambda^{2k} \end{aligned}$$

Ainsi, $P(T_1 = T_2) = \sum_{k=0}^{+\infty} p_\lambda^2 q_\lambda^{2k}$.

(c) Calculer alors $P(T_1 = T_2)$ en fonction de p_λ et q_λ puis vérifier que $P(T_1 = T_2) = \frac{1 - e^{-\lambda}}{1 + e^{-\lambda}}$. On a alors :

$$\begin{aligned} P(T_1 = T_2) &= \sum_{k=0}^{+\infty} p_\lambda^2 q_\lambda^{2k} \\ &= p_\lambda^2 \sum_{k=0}^{+\infty} (q_\lambda^2)^k \\ &= p_\lambda^2 \frac{1}{1 - q_\lambda^2} \quad \text{car } 0 < q_\lambda^2 = e^{-2\lambda} < 1 \\ &= p_\lambda^2 \frac{1}{(1 - q_\lambda)(1 + q_\lambda)} \\ &= \frac{p_\lambda^2}{p_\lambda(1 + q_\lambda)} = \frac{p_\lambda}{1 + q_\lambda} = \frac{1 - e^{-\lambda}}{1 + e^{-\lambda}} \end{aligned}$$

Ainsi, $P(T_1 = T_2) = \frac{1 - e^{-\lambda}}{1 + e^{-\lambda}}$.