Probabilités (LINGE1113)

Université catholique de Louvain

Examen du jeudi 29 août 2024 – série bleue

NOM, Prénom : ....

NOMA : ....

Consignes générales

- Veuillez écrire votre nom, prénom et NOMA sur cette page-ci et sur les pages 2 et 3.
- L'examen consiste en deux parties :
  - 2 questions ouvertes (pondération : 4/20), à répondre sur ce document-ci ;
  - 15 questions à choix multiples sur les exercices (pondération : 16/20), *à répondre sur la grille à lecture optique*—voir les consignes à la fin de ce questionnaire.
- L'examen est à cours fermé.
- L'usage du formulaire—sans annotations—et d'une calculatrice est autorisé.
- L'examen prend **2 heures**.

Consignes pour la partie QCM

- Les QCM 1 à 3 concernent la partie équivalente au test de mars. La note totale pour ces 3 questions sera comparée avec la note au test ramenée sur 4 points. Le maximum entre ces deux notes sera retenu comme note finale pour ces 3 questions.
- Pour les QCM 1 à 11, exactement une des cinq alternatives proposées est correcte. Pondération : 4/3 pour une réponse correcte aux questions 1 à 3 et 5/4 pour une réponse correcte aux questions 4 à 11 ; réponse vide ou fausse, 0.
- Pour les QCM 12 à 15, exactement une des dix alternatives proposées est correcte, et aucune alternative n'est correcte pour deux questions ou plus. Pondération : réponse correcte, 1/2 ; réponse vide ou fausse, 0.

Bon travail !

NOM, Prénom : ....

NOMA : ....

Question théorique 1.

Si c est une constante et Y une variable aléatoire, prouver que

$$\sigma_{cY} = |c|\,\sigma_Y.$$

Réponse.

NOM, Prénom : . . . .

NOMA : . . . .

Question théorique 2.

Si X et Y sont deux variables aléatoires et si $\mu_X = E[X]$ et $\mu_Y = E[Y]$, prouver que

$$\mathrm{Cov}(X,Y) = E[XY] - \mu_X \mu_Y.$$

Réponse.

Questions à choix multiples

Réponses à donner sur la grille à lecture optique.

Les questions 1 à 3 concernent la partie équivalente au test de mars.

1. Pour me rendre au bureau, j'ai le choix entre le train et la voiture. Par défaut, je préfère le train. Si le train a beaucoup de retard, ce qui arrive avec une probabilité de 10%, je prends la voiture. Si, en plus, il y a beaucoup d'embouteillages sur l'autoroute, ce qui arrive avec une probabilité de 30%, je ne vais pas au bureau du tout et je travaille de chez moi. Les retards de train sont indépendants des embouteillages sur l'autoroute. Sachant que je suis au bureau, quelle est la probabilité que j'y sois allé en train ? [Indication : on cherche une probabilité conditionnelle de la forme $P(\overline{A} \mid \overline{A} \cup \overline{B})$, où A et B sont deux événements indépendants, et donc $\overline{A}$ et $\overline{B}$ aussi.]

A. 0.9 **B.** 0.9278 **C.** 0.97 **D.** 0.7216 **E.** 0.7778

Résolution. Par simplicité de notation, on va inverser A et $\overline{A}$ et de même B et $\overline{B}$ par rapport à l'indication. Soit A l'événement que le train n'a *pas* beaucoup de retard et B l'événement qu'il n'y *pas* beaucoup d'embouteillages sur l'autoroute. Alors A et B sont indépendants et $P(A) = 1 - 0.1 = 0.9$ et $P(B) = 1 - 0.3 = 0.7$. On trouve

$$P(A \mid A \cup B) = \frac{P[A \cap (A \cup B)]}{P(A \cup B)} = \frac{P(A)}{P(A \cup B)}.$$

Par la loi additive et l'indépendance de A et B,

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A)\,P(B) \\ &= 0.9 + 0.7 - 0.9 \cdot 0.7 = 0.97. \end{aligned}$$

On conclut que

$$P(A \mid A \cup B) = \frac{0.9}{0.97} = \mathbf{0.9278}.$$

2. Dans un jeu de cartes, on distingue 4 « couleurs » ou « enseignes » : cœur ♡, carreau ♢, trèfle ♣ et pique ♠. On prend 4 cartes, une de chaque couleur, donc 4 cartes au total, dont on tire des cartes aléatoirement, une par une, *avec* remise. On définit la variable aléatoire X à valeurs dans $\{1,2,3\}$ comme suit :

1. Si la couleur de la 2e carte est égale à celle de la 1re, on définit $X = 1$. Exemple : ♡♡.
2. Sinon, les deux premières cartes sont différentes et on tire une 3e carte. Si sa couleur est égale à celle de la 1re ou 2e carte, on met $X = 2$. Exemples : ♡♠♡ ou ♡♠♠.
3. Sinon, les 3 cartes tirées sont toutes différentes et on met $X = 3$. Exemple : ♡♠♣.

Calculer $E(X)$. [Indication : la distribution de X n'est pas égale à une des distributions usuelles. On peut trouver $P(X = k)$ pour $k = 1,2,3$ par un raisonnement direct.]

A. $17/8$	**B.** $11/8$	**C.** 2	**D.** $9/8$	**E.** $15/8$

Résolution. On trouve

$$\begin{aligned} P(X=1) &= \frac{1}{4}, \\ P(X=2) &= \frac{3\cdot 2}{4^2} = \frac{3}{8}, \\ P(X=3) &= \frac{3\cdot 2}{4^2} = \frac{3}{8}. \end{aligned}$$

[Contrôle : la somme des probabilités vaut bien 1.] On obtient

$$E(X) = 1\cdot P(X=1) + 2\cdot P(X=2) + 3\cdot P(X=3) = \frac{1}{4} + \frac{3}{4} + \frac{9}{8} = \mathbf{\frac{17}{8}}.$$

3. On modélise l'arrivée des messages sur ma téléphone par une distribution de Poisson. Si, en moyenne, je reçois un nouveau message tous les 15 minutes, quelle est la probabilité que, lors d'une heure donnée, je reçois au moins 3 messages ?

A. 0.0803	**B.** 0.9084	**C.** 0.5665	**D.** 0.1954	**E.** 0.7619

Résolution. La distribution du nombre Y de messages reçus par heure est Poisson avec paramètre $\lambda = 4$. On obtient

$$\begin{aligned} P(Y \geqslant 3) &= 1 - P(Y=0) - P(Y=1) - P(Y=2) \\ &= 1 - \mathrm{e}^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2}\right) \\ &= \mathbf{0.7619}. \end{aligned}$$

4. En musique, la fréquence de la note de *La* est de 440 Hz. Une chanteuse s'entraîne à améliorer son intonation. Quand elle chante un *La*, l'accordeur (un petit appareil pour mesurer la fréquence d'un son) indique qu'elle le fait à une fréquence de 440 Hz en moyenne mais avec un écart-type de 1 Hz. Dès que la fréquence est en-dessous de 436 Hz ou au-dessus de 444 Hz, la différence devient audible et la note chantée n'est plus appréciée comme bien accordée. Sur base de ces informations, et sans aucune hypothèse supplémentaire sur la distribution de la fréquence des notes chantées, quel est l'énoncé correct le plus précis qu'on sache faire sur la probabilité p qu'une note de *La* chantée soit bien accordée ?

A. $p \geqslant \frac{1}{2}$ **B.** $p \geqslant \frac{3}{4}$ **C.** $p \geqslant \frac{8}{9}$ **D.** $p \geqslant \frac{1}{4}$ **E.** $p \geqslant \frac{15}{16}$

Résolution. Soit Y la fréquence d'une note chantée. Il est donné que

$$\begin{aligned} \mu &= E(Y) = 440, \\ \sigma^2 &= V(Y) = 1^2. \end{aligned}$$

Grâce à l'inégalité de Chebyshev,

$$p = P(436 \leqslant Y \leqslant 440) = P(\mu - 4\sigma \leqslant Y \leqslant \mu + 4\sigma) \geqslant 1 - \frac{1}{4^2} = \mathbf{\frac{15}{16}}.$$

5. Le rendement annuel d'un actif financier est modélisé par une distribution normale avec une espérance de $\mu = 0.02$ et une variance de $\sigma^2 = (0.10)^2$. On suppose que les rendements annuels de plusieurs années consécutives sont indépendantes. Quelle est la probabilité qu'après 4 ans, le rendement annuel moyen soit positif ?

A. 0.3446 **B.** 0.4207 **C.** 0.7881 **D.** 0.5793 **E.** 0.6554

Résolution. On cherche $P\left(\overline{Y} > 0\right)$ où $\overline{Y} = (Y_1 + \cdots + Y_4)/4$, les variables aléatoires $Y_1, \ldots, Y_4$ étant indépendantes et de distribution identique $N(0.02, (0.10)^2)$. La distribution de $\overline{Y}$ est alors $N(0.02, (0.10)^2/4)$, et on trouve

$$
\begin{aligned}
P\left(\overline{Y} > 0\right) &= P\left(\frac{\overline{Y} - 0.02}{\sqrt{(0.10)^2/4}} > \frac{0 - 0.02}{\sqrt{(0.10)^2/4}}\right) \\
&= 1 - \Phi\left(\frac{-0.02}{0.10/2}\right) = 1 - \Phi(-0.40) && \text{standardisation} \\
&= \Phi(0.40) = 1 - (1 - \Phi(0.40)) && \text{symétrie} \\
&= 1 - 0.3446 = \mathbf{0.6554}. && \text{table}
\end{aligned}
$$

6. Une des variables utilisées pour mesurer la qualité d'une équipe de football est le pourcentage de passes réussis (*pass accuracy*, %). Si le pourcentage de passes réussis de l'équipe de l'Espagne est de 85 % en général, quelle est la probabilité (approximative) que, lors d'un match donné, au moins 350 des 400 passes de cette équipe soient réussis ?

A. 0.49 **B.** 0.4349 **C.** 0.125 **D.** 0.0808 **E.** 0.0918

Résolution. On cherche $P(Y \geqslant 350)$ où la distribution de Y est Binomiale avec paramètres $n = 400$ et $p = 0.85$. Grâce au théorème central limite avec correction de continuité, on trouve

$$
\begin{aligned}
P(Y \geqslant 350) &= P(Y \geqslant 349.5) && \text{correction de continuité} \\
&= P\left(\frac{Y - np}{\sqrt{np(1-p)}} \geqslant \frac{349.5 - np}{\sqrt{np(1-p)}} \right) \\
&\approx 1 - \Phi\left(\frac{349.5 - np}{\sqrt{np(1-p)}} \right) && \text{théorème central limite} \\
&= 1 - \Phi(1.3303) = \mathbf{0.0918}. && \text{table}
\end{aligned}
$$

Contrôle : via le logiciel R on trouve que $P(Y \geqslant 350) = 0.0894$. Si on applique l'approximation normale sans correction de continuité, on trouve 0.0808, ce qui est donc plus éloigné de la vraie probabilité.

7. La durée de mon entraînement de course à pied est une variable aléatoire exponentielle X avec paramètre $\beta = 1$. Sachant la durée $X = x > 0$ de mon entraînement, mon temps sur une distance donnée est une variable aléatoire Y de distribution exponentielle avec paramètre $\beta = 1/x$. Calculer $P(Y \leqslant 2)$. [Indication : calculer la densité jointe de (X, Y) en $(x, y) \in]0, +\infty[^2$ à partir de la densité marginale de X en x et la densité conditionnelle de Y en y sachant $X = x$; ensuite calculer la densité marginale de Y en y via l'intégrale de la densité jointe par rapport à x entre 0 et $+\infty$; finalement calculer la fonction de répartition de Y. Une intégrale utile à savoir est $\int_0^{+\infty} x \, \mathrm{e}^{-ax} \, \mathrm{d}x = a^{-2}$ pour $a > 0$. La fonction primitive de $(1+t)^{-2}$ est $-(1+t)^{-1}$.]

A. $2/3$ **B.** $1/2$ **C.** $4/5$ **D.** 1 **E.** $3/4$

Résolution. Il est donné que

$$
\begin{aligned}
f_X(x) &= \mathrm{e}^{-x}, && x > 0, \\
f_{Y|X}(y|x) &= x \, \mathrm{e}^{-xy}, && x, y > 0.
\end{aligned}
$$

La densité jointe de (X, Y) est donc

$$
f_{X,Y}(x, y) = f_X(x) \, f_{Y|X}(y|x) = x \, \mathrm{e}^{-x(1+y)}, \qquad x, y > 0.
$$

On obtient la densité marginale de Y par intégration sur $x > 0$:

$$
f_Y(y) = \int_0^{+\infty} x \, \mathrm{e}^{-x(1+y)} \, \mathrm{d}x = \frac{1}{(1+y)^2}, \qquad y > 0.
$$

La probabilité cherchée est donc

$$P(Y \leqslant 2) = \int_0^2 (1+y)^{-2}\,\mathrm{d}y = \left[(1+y)^{-1}\right]_{y=0}^2 = \frac{1}{1+0} - \frac{1}{1+2} = \mathbf{\frac{2}{3}}.$$

8. Le rendement logarithmique du prix d'un actif financier est défini comme $Y = \ln(X_1/x_0)$, où x_0 est le prix actuel (connu, observé) et X_1 le prix futur (inconnu, aléatoire). On en déduit que le prix futur vaut $X_1 = x_0 \exp(Y)$. Un modèle commun pour le rendement logarithmique Y est la distribution normale $N(\mu, \sigma^2)$, dont les paramètres μ et σ^2 varient avec l'actif financier en question. Sous ces hypothèses, quelle est le prix futur attendu, $E(X_1)$? [Indication : la fonction génératrice de moments d'une variable aléatoire Z de distribution $N(0,1)$ est $m_Z(t) = E[\exp(tZ)] = \exp(t^2/2)$ pour $t \in \mathbb{R}$.]

A. $x_0 \exp(\mu - \sigma^2/2)$ **B.** $x_0 \exp(\mu + \sigma/2)$ **C.** $x_0 \exp(\mu + \sigma^2/2)$ **D.** $x_0 \exp(\mu - \sigma/2)$ **E.** $x_0 \exp(\mu)$

Résolution. On écrit $Y = \mu + \sigma Z$ où la distribution de $Z = (Y-\mu)/\sigma$ est $N(0,1)$. On obtient

$$\begin{aligned}
E[X] &= E[x_0 \exp(Y)] && \text{par définition} \\
&= E[x_0 \exp(\mu + \sigma Z)] && \text{réduction} \\
&= E[x_0 \exp(\mu)\exp(\sigma Z)] && \text{fonction exponentielle} \\
&= x_0 \exp(\mu)\, E[\exp(\sigma Z)] && \text{linéarité de l'espérance} \\
&= x_0 \exp(\mu)\exp(\sigma^2/2) && \text{mgf de la } N(0,1) \\
&= \boldsymbol{x_0 \exp(\mu + \sigma^2/2)} && \text{fonction exponentielle.}
\end{aligned}$$

9. Les rendements futurs X et Y de deux actifs financiers sont supposés d'avoir les propriétés suivantes :

$$\begin{aligned} E(X) &= 0.05, & V(X) &= (0.10)^2, \\ E(Y) &= 0.03, & V(Y) &= (0.06)^2, \\ \rho = \operatorname{corr}(X,Y) &= -0.5. \end{aligned}$$

Le rendement d'une portefeuille financière composée de ces deux actifs est de

$$Z = aX + (1-a)Y,$$

où $0 \leqslant a \leqslant 1$ est la proportion du capital allouée au premier actif et $(1-a)$ celle allouée au deuxième actif. Le critère moyenne-variance introduit par H. Markowitz pour valoriser la portefeuille est définie par

$$\operatorname{MV}(Z) = E(Z) - \frac{b}{2} V(Z),$$

où $b > 0$ indique l'aversion pour le risque (plus b est grand, plus on évite le risque et on préfère une variance petite). Calculer $\operatorname{MV}(Z)$ si $a = 1/2$ et $b = 1$.

A. 0.04 **B.** 0.0381 **C.** 0.03755 **D.** 0.03905 **E.** 0.0383

Résolution. Comme $a = 1/2$, on a $Z = (X+Y)/2$. Le rendement attendu est de

$$E(Z) = [E(X) + E(Y)]/2 = (0.05 + 0.03)/2 = 0.04.$$

La covariance entre X et Y est de

$$\operatorname{Cov}(X,Y) = \rho\sqrt{V(X)\,V(Y)} = -0.5 \cdot 0.10 \cdot 0.06 = -0.003,$$

ce qui donne

$$\begin{aligned} V(Z) &= (1/2)^2\,V(X) + (1/2)^2\,V(Y) + 2(1/2)(1/2)\operatorname{Cov}(X,Y) \\ &= \frac{1}{4}\big((0.10)^2 + (0.06)^2 - 0.006\big) \\ &= \frac{1}{4}(0.01 + 0.0036 - 0.006) = 0.0019. \end{aligned}$$

On obtient

$$\mathrm{MV}(Z) = 0.04 - \frac{1}{2} \cdot 0.0019 = \mathbf{0.03905}.$$

10. Si on lance un ballon de foot à partir du sol à un angle α (entre $0°$ et $90° = \pi/2$) à une vitesse v_0 (en m/s), alors le ballon touchera la pelouse de nouveau à une distance (m) de $h(\alpha) = (v_0^2/g)\sin(2\alpha)$, où $g = 9.81$ m/s^2 est la pesanteur terrestre. Si l'angle α est la réalisation d'une variable aléatoire A qui est uniformément distribuée entre 0 et $\pi/2$, alors quelle est la distance attendue $E[h(A)]$? [Indication : la fonction primitive de $\sin$ est $-\cos$.]

A. $\dfrac{\pi v_0^2}{g}$ **B.** $\dfrac{v_0^2}{\pi g}$ **C.** $\dfrac{\pi v_0^2}{2g}$ **D.** $\dfrac{v_0^2}{g}$ **E.** $\dfrac{2v_0^2}{\pi g}$

Résolution. Par linéarité de l'espérance,

$$E[h(A)] = E[(v_0^2/g)\sin(2A)] = (v_0^2/g)\,E[\sin(2A)].$$

Comme A est uniformément distribuée entre 0 et $\pi/2$, sa densité est

$$f(\alpha) = \begin{cases} \dfrac{1}{\pi/2 - 0} = \dfrac{2}{\pi} & \text{si } 0 \leqslant \alpha \leqslant \pi/2, \\ 0 & \text{sinon.} \end{cases}$$

On trouve

$$\begin{aligned}
E[\sin(2A)] &= \int_{-\infty}^{+\infty} \sin(2\alpha)\, f(\alpha)\, \mathrm{d}\alpha && \text{définition de l'espérance} \\
&= \frac{2}{\pi} \int_0^{\pi/2} \sin(2\alpha)\, \mathrm{d}\alpha && \text{densité } f \\
&= \frac{1}{\pi} \int_0^{\pi} \sin(\beta)\, \mathrm{d}\beta && \text{substitution } 2\alpha = \beta \\
&= \frac{1}{\pi} [-\cos(\beta)]_{\beta=0}^{\pi} && \text{primitive de } \sin \\
&= \frac{1}{\pi} [-\cos\pi - (-\cos 0)] = \frac{1}{\pi}[-(-1) - (-1)] = \frac{2}{\pi}.
\end{aligned}$$

La distance attendue est donc

$$E[h(A)] = \boldsymbol{\frac{2v_0^2}{\pi g}}.$$

11. Afin de se qualifier pour un tournoi d'athlétisme, un sauteur en hauteur doit sauter plus haut qu'une certaine hauteur minimale. L'athlète dispose de trois tentatives, qui sont réalisées indépendamment l'une de l'autre. Quel doit être le taux de réussite minimal d'un saut individuel pour que l'athlète ait au moins 80% de chances de se qualifier ? [Indication : considérer l'événement que l'athlète ne parvient pas à se qualifier, c'est-à-dire, que ces trois tentatives sont toutes des échecs. La probabilité de cet événement doit être inférieure ou égale à 20%. Exprimer la probabilité de cet événement en fonction de la probabilité de succès d'un saut individuel.]

A. 0.2667 **B.** 0.5848 **C.** 0.7333 **D.** 0.9283 **E.** 0.4152

Résolution. Soit p la probabilité de succès d'un saut. L'athlète ne se qualifie pas après trois échecs consécutifs, ce qui se passe avec une probabilité de $(1-p)^3$. Il faut donc que $(1-p)^3$ soit inférieur à $1-0.80 = 0.20$. On trouve

$$\begin{aligned} (1-p)^3 \leqslant 0.20 &\iff 1-p \leqslant (0.20)^{1/3} \\ &\iff p \geqslant 1-(0.20)^{1/3} = 1-0.5848 = \mathbf{0.4152}. \end{aligned}$$

On vérifie que la probabilité que l'athlète se qualifie est alors de $1-(1-p)^3 = 0.8000036$.

12–15. Ci-dessous, 4 variables aléatoires sont décrites. Pour chaque variable, indiquez la loi exacte ou approchée la plus appropriée parmi les 10 distributions données. Une distribution ne peut être choisie qu'une seule fois.

Les variables :

12. Le nombre de cartes à couleur trèfle (♣) si on tire 4 cartes sans remise d'un jeu de cartes contenant 13 cartes de chaque couleur (52 cartes au total).

13. L'angle (en degrés) auquel un ballon de foot est lancé à partir du sol, si l'angle est de 28° en moyenne avec un écart-type de $(8\sqrt{3})^\circ$.

14. Le nombre de passes ratés par une équipe de football lors d'un quart d'heure d'un match s'il y a 52 passes tentés par l'équipe et si la probabilité de réussite d'un pass est de 75%.

15. La distribution de la somme de 4 variables aléatoires indépendantes, si la distribution de chaque variable est normale avec une espérance de 13 et un écart-type de 1.

Les distributions :

A. Normale($\mu = 52$, $\sigma^2 = 4$)

B. Bernoulli($p = 1/4$)

C. Exponentielle($\beta = 4$)

D. Uniforme($\theta_1 = 4$, $\theta_2 = 52$)

E. Poisson($\lambda = 4$)

F. Binomiale($n = 52$, $p = 1/4$)

G. Géométrique($p = 1/4$)

H. Hypergéométrique($N = 52$, $M = 13$, $n = 4$)

I. Normale($\mu = 13$, $\sigma^2 = 4$)

J. Binomiale Négative($r = 13$, $p = 1/4$)

Résolution. **12.** H ; **13.** D ; **14.** F ; **15.** A.