Probabilités (LINGE1113)

Université catholique de Louvain

Examen du samedi 21 août 2021 – série bleue

Consignes

- L'examen est à livres fermés. L'usage du formulaire est autorisé.
- L'examen prend 2 heures.
- Pour les QCM 1 à 11, exactement une des huit alternatives proposées est correcte. Pondération : réponse correcte, +1 ; réponse vide ou fausse, 0.
- Pour les QCM 12 à 17, exactement une des douze alternatives proposées est correcte, et aucune alternative n'est correcte pour deux questions ou plus. Pondération : réponse correcte, +0.5 ; réponse vide ou fausse, 0.
- L'examen est donc sur 14 points.
- 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 3 points ; le maximum entre ces deux notes sera retenu comme note finale pour ces 3 questions.
- Veuillez utiliser la grille à lecture optique pour donner votre réponse.

Bonne chance !

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. Vous décidez de passer une journée à Pairi Daiza. Cependant, en fonction de leur humeur, certains animaux décident de ne pas se montrer et ne sont donc pas toujours visibles. Pour un jour donné, les gérants du parc estiment qu'il y a 63% de chance de voir le lion, 72% de chance de voir le rhinocéros, 59% de chance de voir la girafe et 40% de chance de voir le léopard. Le comportement de chaque animal étant indépendant de celui des autres, quelle est la probabilité que vous aperceviez effectivement au cours de votre visite à la fois le lion et la girafe ou à la fois le rhinocéros et le léopard ? *(Indice : le "ou" est ici non-exclusif : on s'intéresse au cas où au moins un des deux événements reliés par le "ou" se produit.)*

A. 0.541	**B.** 0.583	**C.** 0.485	**D.** 0.603
E. 0.736	**F.** 0.553	**G.** 0.767	**H.** 0.690

Résolution. Si l'événement "*Voir le lion*" (resp. le rhinocéros, la girafe et le léopard) est désigné par A (resp. B, C et D), alors, puisque les événements sont indépendants :

$$\begin{aligned} P((A\cap C)\cup(B\cap D)) &= P(A\cap C)+P(B\cap D)-P(A\cap C\cap B\cap D) \\ &= P(A)P(C)+P(B)P(D)-P(A)P(C)P(B)P(D) \\ &= 0.63\times 0.59+0.72\times 0.4-0.63\times 0.59\times 0.72\times 0.4 \\ &= 0.553 \end{aligned}$$

2. Si le nombre moyen d'accidents du travail sur une période de 10 jours dans une usine est de 2, et que les nombres d'accidents sont indépendants d'un jour à l'autre, quelle est la probabilité qu'au moins trois accidents du travail se produisent sur une période de 30 jours ?

A. 0.938	**B.** 0.6	**C.** 0.8488	**D.** 0.3233
E. 0.0892	**F.** 0.1428	**G.** 0.9558	**H.** 0.8772

Résolution. Le nombre Y d'accidents du travail sur une période de 30 jours suit une loi de Poisson de paramètre $\lambda = 6$. La probabilité recherchée est donc

$$\begin{aligned} P(Y\geqslant 3) &= 1-P(Y<3) \\ &= 1-P(Y=0)-P(Y=1)-P(Y=2) \\ &= 1-e^{-6}\frac{6^0}{0!}-e^{-6}\frac{6}{1!}-e^{-6}\frac{6^2}{2!} \\ &= 0.938. \end{aligned}$$

3. A l'UCLouvain, 5 postes administratifs distincts (requérant les mêmes compétences, mais impliquant des tâches différentes) sont à pourvoir. Parmi les personnes ayant postulé, 13 ont les compétences requises. Combien y a-t-il de façons possibles de remplir les 5 postes vacants ?

A. 120	**B.** 51891840	**C.** 65	**D.** 336
E. 1287	**F.** 525875	**G.** 371293	**H.** 154440

Résolution. On compte ici le nombre de façons possibles de sélectionner 5 personnes à partir d'un groupe de 13 personnes, en tenant compte de l'ordre puisque les postes sont bien différents (et donc assigner la personne 1 au poste A et la personne 2 au poste B est différent d'assigner la personne 2 au poste A et la personne 1 au poste B). Il s'agit d'une permutation : $P_5^{13} = \frac{13!}{8!} = 154440$.

4. Soit Y une variable aléatoire à perte de mémoire dont la fonction de densité est définie par :

$$f(y) = \begin{cases} 2e^{-2y}, & \text{si } 0 < y. \\ 0, & \text{sinon.} \end{cases}$$

Que vaut la variance de $2Y + 1$?

A. 9	**B.** 0	**C.** 2	**D.** 1.5
E. 16	**F.** 17	**G.** 1	**H.** 8

Résolution. D'après le formulaire, la fonction de densité de Y correspond à celle d'une loi exponentielle de paramètre $\frac{1}{2}$. On a donc $V(2Y+1) = 2^2 V(Y) = 2^2 \left(\frac{1}{2}\right)^2 = 1$.

5. Chaque vol d'une compagnie aérienne peu fiable a une probabilité de 0.3 d'avoir du retard. On choisit au hasard 200 vols parmi ceux prévus par cette compagnie dans les 6 prochains mois et on suppose que le fait qu'un vol ait du retard n'influence pas les autres vols. Quelle est la probabilité approchée (le plus précisément possible) qu'au moins 65 de ces vols aient du retard ?

A. 0.2810	**B.** 0.2451	**C.** 0.3314	**D.** 0.1977
E. 0.2206	**F.** 0.1566	**G.** 0.4562	**H.** 0.3910

Résolution. Soit Y le nombre de vols ayant du retard parmi les 200 vols sélectionnés : $Y \sim Bin(200, 0.3)$. Cette distribution peut être approchée par une distribution normale de moyenne $200 \times 0.3 = 60$ et de variance $200 \times 0.3 \times 0.7 = 42$: $N(60; 42)$. Par conséquent, en utilisant la correction de continuité, on trouve

$$P(Y \geqslant 65) = P(Y \geqslant 64.5) = P\left(\frac{Y-60}{\sqrt{42}} \geqslant \frac{64.5-60}{\sqrt{42}}\right) \approx P(Z \geqslant 0.69) = 0.2451.$$

Pour information, la probabilité exacte (calculée avec un logiciel) est 0.2421.

6. Soit X le nombre de personnes admises dans un hôpital chaque jour. Si on sait juste que la variable X est de moyenne 36 et de variance 11, quelle est la borne maximale la plus précise pour $P(X \geqslant 40)$?

A. 0.1708	**B.** 0.9931	**C.** 0.3438	**D.** 0.6875
E. 0.1139	**F.** 0.3125	**G.** 0.0068	**H.** 0.8292

Résolution. On sait que $E(X) = \mu = 36$ et que $V(X) = \sigma^2 = 11$. On utilise le théorème de Tchebysheff en se souvenant que, pour deux événements A et B, on a $P(A) \leqslant P(A \cup B)$:

$$P(X \geqslant 40) \leqslant P(X \geqslant 40) + P(X \leqslant 32) = P(|X - 36| \geqslant 4) = P(|X - \mu| \geqslant k\sigma) \leqslant \frac{1}{k^2},$$

où $4 = 40 - 36 = k\sqrt{11}$, et donc $k = \frac{4}{\sqrt{11}}$. On en déduit que $P(X \geqslant 40) \leqslant \frac{11}{4^2} = 0.6875$.

7. Soient deux variables aléatoires X et Y dont la fonction de densité jointe est donnée par :

$$f(x, y) = \begin{cases} 3x & \text{si } x - 1 \leqslant y \leqslant 1 - x \text{ et } 0 \leqslant x \leqslant 1 \\ 0 & \text{sinon.} \end{cases}$$

Calculer la probabilité $P(X \geqslant Y)$.
Indice : le calcul de la probabilité de l'événement complémentaire est plus rapide.

A. 0.875	**B.** 0.7358	**C.** 1	**D.** 0.5
E. 0	**F.** 0.125	**G.** 0.2642	**H.** 0.75

Résolution.
On définit la région D du plan dans laquelle la densité jointe est strictement positive :

$$D = \left\{(x, y) \in \mathbb{R}^2 \mid x - 1 \leqslant y \leqslant 1 - x,\ 0 \leqslant x \leqslant 1\right\}$$

Notons dans un premier temps la région correspondant à $Y > X$ par B, qui est donnée par :

$$B = \left\{(x, y) \in \mathbb{R}^2 \mid y > x\right\}$$

alors que la région du plan correspondant à $X \geqslant Y$ est notée A et est donnée par :

$$A = \left\{(x, y) \in \mathbb{R}^2 \mid x \geqslant y\right\}$$

Graphiquement, la région d'intérêt est donnée par la zone verte hachurée $A \cap D$. Cependant, nous voyons que la zone rouge $B \cap D$ est plus facile à calculer dans un premier temps.

La région hachurée rouge $B \cap D$ est donc donnée par :

$$B \cap D = \left\{(x, y) \in \mathbb{R}^2 \mid 0 \leqslant x \leqslant 1/2,\ x < y \leqslant 1 - x\right\}$$

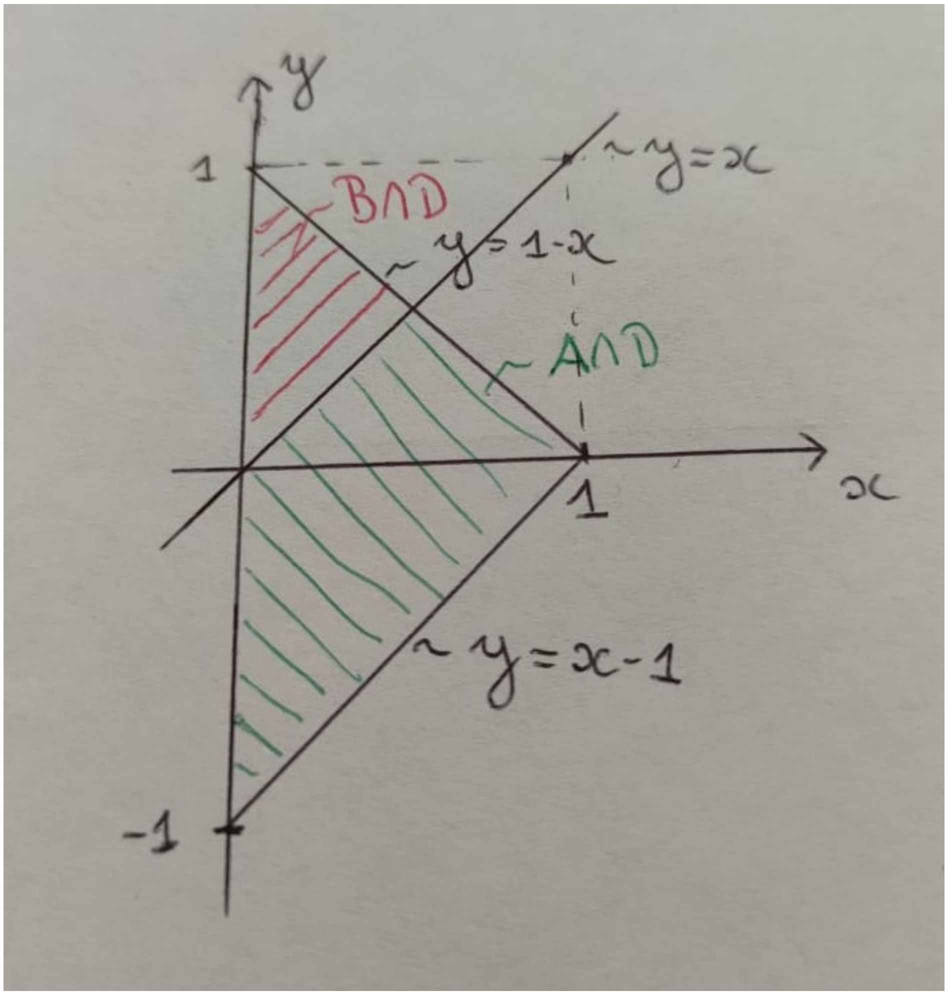

FIGURE 1 – Région d'intérêt

La probabilité étant la double intégrale de la densité jointe sur la région $B \cap D$, nous obtenons :

$$
\begin{aligned}
P(Y > X) &= \int_0^{1/2} \left(\int_x^{1-x} 3x\,dy \right) dx \\
&= \int_0^{1/2} 3x\,[y]_x^{1-x}\,dx \\
&= \int_0^{1/2} 3x\,(1-x-x)\,dx \\
&= \int_0^{1/2} \left(3x - 6x^2\right) dx \\
&= \left[\frac{3}{2}x^2 - 2x^3 \right]_0^{1/2} \\
&= \frac{3}{8} - \frac{2}{8} = \frac{1}{8}
\end{aligned}
$$

La probabilité recherchée étant $P(X \geqslant Y)$ (correspondant à la zone verte hachurée $A \cap D$), nous obtenons :

$$P(X \geqslant Y) = 1 - P(Y > X) = 1 - \frac{1}{8} = \frac{7}{8}.$$

8. Soient Z_1 et Z_2 deux variables aléatoires indépendantes telles que $V(Z_1) = 2$ et $V(Z_2) = 1/2$. On définit $A_1 = Z_1 + Z_2$ et $A_2 = Z_1 - Z_2$. Que vaut la covariance entre A_1 et A_2 ?

A. 0	**B.** 1.5	**C.** 2	**D.** -2.5
E. 2.5	**F.** -2	**G.** -1.5	**H.** Il manque de l'information

Résolution. On utilise une des formules pour la covariance (première ligne) puis une des formules de la variance (cinquième ligne) :

$$\begin{aligned}
Cov(Z_1+Z_2, Z_1-Z_2) &= E[(Z_1+Z_2)(Z_1-Z_2)] - E(Z_1+Z_2)E(Z_1-Z_2) \\
&= E\left[Z_1^2 - Z_1Z_2 + Z_1Z_2 - Z_2^2\right] - [E(Z_1)+E(Z_2)][E(Z_1)-E(Z_2)] \\
&= E(Z_1^2) - E(Z_2^2) - [E(Z_1)]^2 + E(Z_1)E(Z_2) - E(Z_1)E(Z_2) + [E(Z_2)]^2 \\
&= E(Z_1^2) - [E(Z_1)]^2 - \left(E(Z_2^2) - [E(Z_2)]^2\right) \\
&= V(Z_1) - V(Z_2) = 2 - 0.5 = 1.5.
\end{aligned}$$

9. Si le prix (en dollars) d'une action suit une loi normale de moyenne 60 et d'écart-type 0.8, quelle est la valeur sous laquelle le prix se situera dans 97.5% des cas ?

A. 60.8203	**B.** 61.773	**C.** 60.1308	**D.** 58.7456
E. 61.568	**F.** 60.3904	**G.** 58.432	**H.** 61.2544

Résolution. Soit Y le prix d'une action. $Y \sim \mathcal{N}(60, 0.8^2)$ et le prix q recherché est tel que :

$$P(Y \leqslant q) = 0.975 \Leftrightarrow P\left(Z \leqslant \frac{q-60}{0.8}\right) = 0.975 \Leftrightarrow P\left(Z > \frac{q-60}{0.8}\right) = 0.025$$

avec $Z \sim \mathcal{N}(0,1)$. Dans la table de la loi normale, on trouve $\frac{q-60}{0.8} = 1.96$, et donc $q = 60 + 0.8 \times 1.96 = 61.568$.

10. Soient deux variables aléatoires X et Y dont la distribution bivariée est définie par la fonction de probabilité jointe suivante :

$$p(x,y) = \begin{cases} \frac{(x+y)}{11} & \text{pour } (x,y) \in \{(0,1),(0,3),(1,0),(1,1),(1,3)\} \\ 0 & \text{sinon.} \end{cases}$$

Que vaut la fonction de probabilité marginale de X au point $x = 0$?

A. 0.3636	**B.** 0.0909	**C.** 0.6364	**D.** 0.2727
E. 0.5	**F.** 0	**G.** 0.5455	**H.** 1

Résolution. La fonction de densité marginale de X au point $x = 0$, $p_X(0)$, est obtenue comme suit :

$$p_X(0) = \sum_y p(0, y) = p(0,0) + p(0,1) + p(0,3) = \frac{0}{11} + \frac{1}{11} + \frac{3}{11} = \frac{4}{11} = 0.3636.$$

11. Soit Y une variable aléatoire de loi exponentielle de paramètre 3. Que vaut $V\left(Y^2\right)$, la variance de Y^2 ?
Indice : pour rappel, $\frac{d}{dx}x^\alpha = \alpha x^{\alpha-1}$.

A. 81 **B.** 36 **C.** 72 **D.** 1944
E. 18 **F.** 162 **G.** 9 **H.** 1620

Résolution. On sait que $V\left(Y^2\right) = E\left(Y^4\right) - \left(E\left(Y^2\right)\right)^2$ et que, puisque $Y \sim Exp(3)$, $E(Y) = 3$ et $V(Y) = 3^2$. On trouve donc $E\left(Y^2\right) = V(Y) + (E(Y))^2 = 3^2 + 3^2 = 18$. Il suffit donc de trouver $E\left(Y^4\right)$ en dérivant la fonction génératrice des moments de Y donnée par $m_Y(t) = (1-3t)^{-1}$:

$$\begin{aligned}
\frac{dm_Y(t)}{dt} &= 3(1-3t)^{-2} \\
\frac{d^2m_Y(t)}{dt^2} &= 18(1-3t)^{-3} \\
\frac{d^3m_Y(t)}{dt^3} &= 162(1-3t)^{-4} \\
\frac{d^4m_Y(t)}{dt^4} &= 1944(1-3t)^{-5}.
\end{aligned}$$

On trouve alors

$$E(Y^4) = \left.\frac{d^4m_Y(t)}{dt^4}\right|_{t=0} = 1944.$$

Finalement, $V\left(Y^2\right) = 1944 - 18^2 = 1620$.

12–17. Ci-dessous, 6 variables aléatoires sont décrites. Pour chacune de ces variables, il faut indiquer, parmi les 12 distributions données, la loi exacte ou approchée la plus appropriée. Une distribution ne peut être choisie qu'une seule fois.

Les variables :

12. Le poids total (en kg) de deux bébés à la naissance, si le poids moyen de chacun d'entre eux est de 2.5 kg, avec variance 0.1.

13. Le temps (en minutes) qu'un étudiant passe sur une question d'un quiz Moodle, si le test est paramétré pour ne laisser à l'étudiant que 10 minutes par question.

14. Le temps (en heures) entre l'arrivée de deux clients dans une boutique, si, en moyenne, 5 clients par heure arrivent dans la boutique.

15. Le nombre d'hommes parmi 5 personnes sélectionnées au hasard dans un groupe de 10 personnes, dont 2 hommes.

16. Le nombre de services qu'une personne apprenant à jouer au tennis doit tenter pour arriver à réussir un service.

17. Le nombre de réponses correctes d'un étudiant qui répond au hasard à 10 questions à choix multiples offrant chacune 5 propositions de réponse.

Les distributions :

A. Exponentielle($\beta = 0.2$)

B. Normale($\mu = 2$, $\sigma^2 = 1.6$)

C. Poisson($\lambda = 10$)

D. Exponentielle($\beta = 5$)

E. Géométrique($p = 0.2$)

F. Binomiale($n = 10$, $p = 0.2$)

G. Binomiale Négative($r = 10$, $p = 0.2$)

H. Normale($\mu = 5$, $\sigma^2 = 0.2$)

I. Uniforme($\theta_1 = 0.2$, $\theta_2 = 10$)

J. Géométrique($p = 1$)

K. Bernoulli($p = 0.2$)

L. Hypergéométrique($N = 10, M = 2, n = 5$)

Résolution

12. H
13. I
14. A
15. L
16. E
17. F