Probabilités (LINGE1113)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Examen du vendredi 17 août 2017 – série bleu

Consignes générales

- L'examen consiste de deux parties :
  - les 19 questions à choix multiples (QCM) de ce document-ci – pondération : 16/20 ;
  - les 2 questions ouvertes sur l'autre partie de l'examen – pondération : 4/20.
- L'examen est à livres fermés. L'usage du formulaire est autorisé.
- L'examen prend 3 heures.

Consignes pour la partie QCM

- Pour les QCM 1 à 14, exactement une des cinq alternatives proposées est correcte. Pondération : réponse correcte, +1 ; réponse vide, +0.2 ; réponse fausse, 0.
- Pour les QCM 15 à 19, exactement une des dix alternatives proposées est correcte, et aucune alternative n'est correcte pour deux questions ou plus. Pondération : réponse correcte, +0.4 ; réponse vide ou fausse, 0.
- Les QCM 1 à 4 concernent la partie équivalente au test de mars. La note totale pour ces 4 questions sera comparée avec la note au test ramenée sur 4 points ; le maximum entre ces deux notes sera retenue comme note finale pour ces 4 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 4 concernent la partie équivalente au test de mars.

1. On place un rat dans une plateforme expérimentale. Il se trouve face à 4 chemins dont un seul lui permet de sortir du labyrinthe, où il trouvera un morceau de fromage, tandis que les autres chemins sont sans issue. À chaque essai infructueux, on le replace à l'endroit initial. En supposant que le rat n'est pas doué d'apprentissage, quelle est la probabilité qu'exactement 5 essais lui soient nécessaires pour sortir du labyrinthe ?

A. 0.200 **B.** 0.059 **C.** 0.333 **D.** 0.250 **E.** 0.079

Résolution. On définit la variable aléatoire Y : le nombre d'essais jusqu'à ce que le rat prenne le bon chemin. Y est une variable aléatoire où on compte le nombre de tentatives pour arriver au premier succès, le succès arrivant lorsque le rat prend le chemin vers la sortie. Donc $Y \sim \text{Geom}(p = 0.25)$, où p est la probabilité de succès sur une tentative. Donc,

$$\Pr(Y = 5) = p(1-p)^{5-1} = 0.079$$

2. Un clavier de 12 touches (comprenant les neuf chiffres de 1 à 9 et les trois lettres A, B, C) permet de composer le code d'entrée d'un immeuble. Ce code, de longueur six, est composé d'un chiffre suivi de deux lettres *distinctes* puis de trois chiffres. Dans les chiffres, les répétitions sont bien autorisés. Par exemple, le code « 1AB111 » est autorisé, tandis que le code « 1AA234 » ne l'est pas. Le codes « 0AB123 » (contenant le chiffre 0) et « 1CD234 » (contenant une lettre non égale à A, B, ou C) ne sont pas autorisés non plus. Combien de codes différents peut-on former ?

A. 19683 **B.** 18144 **C.** 59049 **D.** 39366 **E.** 27216

Résolution. Concernant les chiffres, les répétitions sont autorisées, il y a donc 9 possibilités à chaque fois. Concernant les lettres, l'ordre a une importance, nous avons donc $P_2^3 = 6$ possibilités. Il y a donc $9 \times 6 \times 9 \times 9 \times 9 = 39366$ codes possibles.

3. Dans un paquet de 150 pralines, 80 sont au chocolat au lait, 40 au chocolat blanc et 30 au chocolat noir. Selon la vendeuse, ces dernières sont composées à moitié de pralines alcoolisées. De plus, elle t'a également informé qu'il y a 20% de pralines au chocolat au lait qui contiennent de l'alcool et que les pralines blanches ne sont que 5% à être alcoolisées. Si un ami vient de manger une des pralines alcoolisées, quelle est la probabilité qu'elle n'était *pas* noire ?

A. $\frac{6}{11}$ **B.** $\frac{1}{5}$ **C.** $\frac{4}{5}$ **D.** $\frac{5}{11}$ **E.** $\frac{1}{10}$

Résolution. Si B = Chocolat blanc, L = Chocolat au lait, N = Chocolat noir et A = Praline alcoolisée, alors, par le théorème de Bayes, nous avons :

$$\begin{aligned}
P(N|A) &= \frac{P(A|N)P(N)}{P(A|B)P(B)+P(A|N)P(N)+P(A|L)P(L)} \\
&= \frac{1/2\times 30/150}{5/100\times 40/150+1/2\times 30/150+20/100\times 80/150} \\
&= \frac{15/150}{200/15000+1500/15000+1600/15000} \\
&= \frac{1/10}{33/150} = \frac{15}{33} \\
P(\overline{N}|A) &= 1-P(N|A) \\
&= \frac{18}{33} = \frac{6}{11}
\end{aligned}$$

4. Vous invitez deux amis, Pierre et Paul, à passer la soirée avec vous. La probabilité que Pierre vienne mais que Paul ne vienne pas est de 0.3. La probabilité que Pierre vienne est de 0.65, et la probabilité qu'au moins un des deux soit disponible est de 0.8. Quelle est la probabilité que Paul vienne ?

A. 0.85 **B.** 0.35 **C.** 0.50 **D.** 0.30 **E.** 0.15

Résolution. Puisque $P(A\backslash B) = 0.3$, $P(A) = 0.65$ et $P(A\cup B) = 0.8$, on trouve que $P(A\cap B) = P(A) - P(A\backslash B) = 0.35$, et donc $P(B) = P(A\cup B) - P(A) + P(A\cap B) = 0.8 - 0.65 + 0.35 = 0.5$.
Ou plus simple : $P(B) = P(A\cup B) - P(A\setminus B) = 0.8 - 0.3 = 0.5$.

5. Soit une machine qui remplit des paquets de bonbons telle que le contenu d'un paquet, exprimé en grammes, suit une loi normale, $\mathcal{N}(\mu, \sigma^2)$. La variance dépend des réglages de la machine. Quelle valeur maximale de la variance peux-tu donner pour que la quantité contenue dans le paquet ne soit éloignée de plus de 10 grammes de la moyenne que dans 0.5% des cas ? [Indice : si X est le contenu d'un paquet, exprimer l'événement d'intérêt $|X-\mu| > 10$ en termes de la version standardisée de la variable X.]

A. 0.71 **B.** 12.66 **C.** 0.50 **D.** 3.56 **E.** Manque d'info

Résolution. Soit X la quantité de bonbons fournie par la machine. On sait par l'énoncé que $X \sim \mathcal{N}(\mu, \sigma^2)$. Posons $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$. On cherche la valeur maximale de la variance telle que :

$$\begin{aligned}
\Pr(|X-\mu| \le 10) \ge 0.995 &\iff \Pr(\mu - 10 \le X \le \mu + 10) \ge 0.995 \\
&\iff \Pr\left(\frac{\mu - 10 - \mu}{\sigma} \le Z \le \frac{\mu + 10 - \mu}{\sigma}\right) \ge 0.995 \\
&\iff \Pr\left(\frac{-10}{\sigma} \le Z \le \frac{10}{\sigma}\right) \ge 0.995 \\
&\iff 1 - 2\Pr\left(Z > \frac{10}{\sigma}\right) \ge 0.995 \\
&\iff \Pr\left(Z > \frac{10}{\sigma}\right) \le 0.0025
\end{aligned}$$

En utilisant la table du formulaire pour Z, on obtient que $\frac{10}{\sigma} \ge 2.81$. D'où l'écart-type $\sigma \le \frac{10}{2.81} = 3.56$ et la variance $\sigma^2 \le 12.66$.

6. La distribution du nombre de voyages effectués par les employés d'une entreprise pendant l'été est la suivante :

y	0	1	2	3	4
$P(Y=y)$	0.32	0.37	0.21	0.07	0.03

Quelle est la variance de cette variable ?

A. 2.32 **B.** 1.12 **C.** 1.0656 **D.** 0.2434 **E.** -1.011

Résolution. $E(Y)=\sum_y y\,P(Y=y)=0.37+0.42+0.21+0.12=1.12$ et $E(Y^2)=\sum_y y^2\,P(Y=y)=0.37+0.84+0.63+0.48=2.32$. On trouve donc $V(Y)=E(Y^2)-(E(Y))^2=2.32-1.12^2=1.0656$.

7. Soit Y une variable aléatoire de distribution khi-carrée de paramètre $\nu=3$. Que vaut l'espérance de Y^3 ?

A. 27 **B.** 35 **C.** 1 **D.** 105 **E.** 15

Résolution. Puisque $Y\sim\chi^2_3$, sa fonction génératrice des moments est $m(t)=(1-2t)^{-3/2}$. Par conséquent, $m'''(t)=3\times5\times7\times(1-2t)^{-9/2}$, et donc $E(Y^3)=m'''(0)=105$.

8. Soient Y_1 et Y_2 deux variables aléatoires dont la fonction de densité jointe est donnée par $f(y_1,y_2)=\frac{1}{y_1}$ si $0\le y_2\le y_1\le 1$ et 0 sinon. Que vaut la corrélation entre ces deux variables ? [Indices : $E(Y_1)=1/2$, $E(Y_2)=1/4$, $E(Y_1^2)=1/3$, et $E(Y_2^2)=1/9$.]

A. 0.1667 **B.** 0.3158 **C.** 0.0417 **D.** 0.6547 **E.** 10.2857

Résolution. On trouve d'abord

$$\begin{aligned}
E(Y_1Y_2)&=\int_0^1\int_0^{y_1} y_2\,\mathrm{d}y_2\,\mathrm{d}y_1=\int_0^1 (y_1^2/2)dy_1=1/6,\\
V(Y_1)&=E(Y_1^2)-(E(Y_1))^2=1/12,\\
V(Y_2)&=E(Y_2^2)-(E(Y_2))^2=7/144.
\end{aligned}$$

Par conséquent,

$$\begin{aligned}
\mathrm{Cov}(Y_1,Y_2)&=E(Y_1Y_2)-E(Y_1)E(Y_2)=1/24,\\
\mathrm{Corr}(Y_1,Y_2)&=\frac{\mathrm{Cov}(Y_1,Y_2)}{\sqrt{(}V(Y_1)V(Y_2))}=\sqrt{3/7}=0.6547.
\end{aligned}$$

9. Soient Y_1,Y_2 deux variables aléatoires dont la fonction de densité jointe est

$$f(y_1,y_2)=\begin{cases}4y_1y_2, & 0\le y_1\le 1, 0\le y_2\le 1,\\ 0, & \text{sinon.}\end{cases}$$

Quelle est la probabilité que la somme de Y_1 et Y_2 soit plus petite que 1 ?

A. 1 **B.** $(1-y_1)^2$ **C.** $\frac{1}{12}$ **D.** $(1-y_2)^2$ **E.** $\frac{1}{6}$

Résolution.

$$\begin{aligned}
\Pr(Y_1+Y_2 \le 1) &= \int_0^1 \int_0^{1-y_1} 4y_1y_2 \mathrm{d}y_2\mathrm{d}y_1 \\
&= \int_0^1 2y_1(1-y_1)^2 \mathrm{d}y_1 \\
&= \int_0^1 2\left(y_1^3-2y_1^2+y_1\right)\mathrm{d}y_1 \\
&= 2\left[\frac{y_1^4}{4}-\frac{2y_1^3}{3}+\frac{y_1^2}{2}\right]_0^1 = \frac{1}{6}
\end{aligned}$$

10. La distribution de la variable aléatoire discrète Y est donnée par le tableau suivant :

y	-2	1	2	100
$P(Y=y)$	a	$\frac{1}{6}$	$\frac{1}{6}$	b

Notons $F(y)$, la fonction de répartition de Y. Sachant que $1-F(1)=\frac{2}{6}$, laquelle des propositions ci-dessous est exacte ?

A. $b=0$ et $F(-5)=0$ **B.** Manque d'info **C.** $a=b=\frac{1}{3}$ **D.** $a=\frac{4}{6}$ et $b=0$ **E.** $F(-2)=\frac{1}{2}$ et $b=\frac{1}{6}$

Résolution. Sachant que $\sum_{y\in\mathscr{Y}} P(Y=y)=1$ où $\mathscr{Y}=\{-2;1;2;100\}$, $F(y)=\Pr(Y\le y)$ et $1-F(y)=\Pr(Y>y)$, nous avons $a=F(-2)=\frac{1}{2}$ et $b=\frac{1}{6}$.

11. Le gestionnaire d'un café situé dans un aéroport est intéressé par les montants dépensés en boissons par les passagers fréquentant cet aéroport. Le service qualité de l'aéroport lui communique les deux résultats suivants :

— l'espérance du montant dépensé par un seul passager est de 5.80 euro ;
— la variance de ce montant est de 11.56 euro².

Un jour, 100 passagers sont choisis aléatoirement et la moyenne arithmétique des montants qu'ils ont dépensé en boissons est calculée. Que vaut (approximativement) la probabilité que cette moyenne arithmétique soit supérieure à 6.50 euro ?

A. ≈ 0.3520 **B.** ≈ 0.6480 **C.** ≈ 0.2709 **D.** ≈ 0.0197 **E.** < 0.001

Résolution. Soit Y le montant dépensé par un passager et $\overline{Y}$ la moyenne arithmétique des montants dépensés par 100 passagers choisis au hasard. Grâce au théorème central limite, la loi approchée de $\overline{Y}$ est $\mathscr{N}(\mu,\sigma^2/100)$ où $\mu=5.80$ et $\sigma^2=11.56$. Par conséquent, si Z est une variable aléatoire $N(0,1)$, on a

$$P(\overline{Y}>6.50)\approx P\left(Z>\frac{6.50-\mu}{\sigma/10}\right)=P(Z>2.0588)=0.01976.$$

La table de la distribution normale donne plutôt $1-\Phi(2.06)=0.0197$.

12. Parmi vos connaissances, le nombre moyen de jours de vacances à l'étranger pendant l'été est de 7.4, avec une variance de 12.25. Quelle est l'information la plus précise que l'on puisse fournir à propos de la probabilité que le nombre de jours de vacances à l'étranger pendant l'été soit au moins de 14 ?

A. ≥ 0.4697 **B.** ≤ 3.4450 **C.** ≤ 0.1406 **D.** ≤ 0.2812 **E.** ≥ 0.2348

Résolution. Soit Y le nombre de jours de vacances à l'étranger pendant l'été. Grâce au théorème de Tchebysheff, on sait que $P(Y \geq 14) = P(Y \geq 7.4 + 3.5k) \leq 1/k^2$ avec $k = 1.89$. La probabilité demandée est donc ≤ 0.2812.

13. Assis sur un banc autour du lac de Louvain-la-Neuve, tu regardes le nombre de chiens passer. D'expérience, tu sais qu'il y en a en moyenne 2 par heure. Cette fois-ci, tu es déjà là depuis une heure sans qu'un seul chien ne soit passé. Quelle est la probabilité que tu doives attendre encore au moins une heure de plus avant de voir le premier chien ?

A. 0.6321 **B.** 0.1353 **C.** 0.5 **D.** 0.6065 **E.** 0.3679

Résolution. Le temps (h) jusqu'au premier chien qui passe est variable aléatoire, disons T, de distribution exponentielle avec moyenne $\beta = E(T) = 0.5$. Grâce à la propriété de « perte de mémoire », on trouve que $P(T > 1) = \exp(-\beta) = 0.6065$.

14. Assis sur un banc autour du lac de Louvain-la-Neuve, tu regardes le nombre de chiens passer. Si en moyenne il y en a 2 par heure, quelle est la probabilité que tu en voies au moins 3 en restant une heure sur le banc ?

A. 0.3233 **B.** 0.1429 **C.** 0.2500 **D.** 0.1804 **E.** 0.6766

Résolution. Notons X, la variable aléatoire représentant le nombre de chiens qui sont passés devant ton banc. Celui-ci suit une distribution Poisson de paramètre λ soit $\Pr(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$. Par l'énoncé, on sait que $\lambda = E(X) = V(X) = 2$. De là, on calcule la probabilité demandée :

$$\begin{aligned}
\Pr(X \geq 3) &= 1 - \Pr(X \leq 2) \\
&= 1 - \Pr(X = 2) - \Pr(X = 1) - \Pr(X = 0) \\
&= 1 - e^{-2}\frac{2^2}{2!} - e^{-2}\frac{2^1}{1!} - e^{-2}\frac{2^0}{0!} \\
&= 1 - 0.2706706 - 0.2706706 - 0.1353353 \\
&= 0.3233235
\end{aligned}$$

15–19. Ci-dessous, 5 variables aléatoires sont décrites. Pour chacune de ces variables, il faut indiquer, parmi les 10 distributions données, la loi exacte ou approchée la plus appropriée. Une distribution ne peut être choisie qu'une seule fois.

Les variables :

15. Si, parmi quatre chocolats, il y en a un seul qui contient de l'alcool, et si on choisit un des quatre chocolats au hasard, le nombre de chocolats (0 ou 1) qu'on mange contenant de l'alcool.

16. Si, dans un aéroport, un passager sur quatre s'offre un sandwich avant d'embarquer, le nombre de passagers dans un avion de 100 passagers qui achètent un sandwich.

17. Le temps en minutes jusqu'au départ de votre avion, si le pilote vient d'annoncer que les préparations prendront au maximum 25 minutes.

18. Le contenu en grammes d'un paquet de bonbons, si, en moyenne, le contenu est de 100 grammes.

19. Si, malgré toutes les informations de sécurité, 25% des passagers portent un couteau de poche dans leur bagage à main, le nombre de passagers qui amènent par les contrôles de sécurité jusqu'au premier qui est attrapé avec un tel couteau sur soi.

Les distributions :

A. Binomiale($n = 100, p = 0.25$)

B. Binomiale Négative($r = 100, p = 0.25$)

C. Bernoulli($p = 0.25$)

D. Khi-carré($\nu = 25$)

E. Uniforme($\theta_1 = 0, \theta_2 = 25$)

F. Exponentielle($\beta = 0.25$)

G. Normale($\mu = 100, \sigma = (2.5)^2$)

H. Hypergéométrique($M = 100, N = 25, n = 25$)

I. Poisson($\lambda = 25$)

J. Géométrique($p = 0.25$)

Résolution

15. C
16. A
17. E
18. G
19. J