Probabilités (LINGE1113)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Examen du mardi 7 juin 2016 – série bleu

Consignes

— L'examen comporte les 20 questions à choix multiples de ce document-ci (pondération : 16/20 plus une question de bonus pour 1 point) ainsi que 2 questions ouvertes sur l'autre partie de l'examen (pondération : 4/20).

— Pour les questions à cinq alternatives, exactement une des cinq alternatives proposées est correcte. Pondération : réponse correcte, +1 ; réponse vide, +0.2 ; réponse fausse, 0. Pour la question de bonus, une réponse vide n'apporte aucun point.

— Pour les questions 16 à 20, 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.

— Veuillez utiliser la grille à lecture optique pour donner votre réponse.

— L'examen est à livres fermés. L'usage du formulaire est autorisé.

— L'examen prend 3 heures.

Bon travail !

Questions à choix multiples

Réponses à donner sur la grille à lecture optique.

1. Un athlète et son coach doivent se voir pour mettre au point une stratégie avant la compétition. Ils sélectionnent tous les deux aléatoirement et indépendamment l'un de l'autre un moment d'arrivée durant une même période d'une heure. Chacun est d'accord pour attendre l'autre maximum 10 minutes une fois arrivé au lieu de rencontre. Quelle est la probabilité que le coach et l'athlète se rencontrent ? [Indice : en ramenant les 60 minutes vers l'intervalle $[0,1]$, on cherche la probabilité que $-1/6 \leq Y_1 - Y_2 \leq 1/6$, où (Y_1, Y_2) est un couple de variables aléatoires dans le carré $[0,1]^2$ et avec une certaine distribution continue. Faites un dessin et évitez les doubles intégrales en faisant appel à un raisonnement géométrique.]

A. $\frac{1}{3}$ **B.** $\frac{1}{2}$ **C.** $\frac{11}{36}$ **D.** $\frac{4}{9}$ **E.** $\frac{1}{6}$

Résolution. Soit Y_1 le moment où le coach arrive et Y_2 le moment où l'athlète arrive. Y_1 et Y_2 sont uniformément distribuées sur l'intervalle $[0,1]$. Par indépendance entre Y_1 et Y_2, le vecteur (Y_1, Y_2) est uniformément distribué sur le carré unité. Il suffira donc de faire un calcul d'aire. En faisant un dessin, on voit que la zone d'intérêt se trouve entre les droites d'équation $y_2 = y_1 - \frac{1}{6}$ et $y_2 = y_1 + \frac{1}{6}$. La probabilité recherchée vaut donc $1 - 2 \cdot \frac{\frac{5}{6} \cdot \frac{5}{6}}{2} = \frac{11}{36}$ (en faisant 1 moins l'aire des deux parties extérieures à cette bande, qui sont deux triangles).

2. Lors du deuxième jour des Jeux Olympiques, il y a 6 matchs de badminton, 2 matchs de foot et 4 matchs de water-polo, tous diffusés à des heures différentes en Belgique. Vous décidez de regarder 3 matchs, choisis au hasard. Quelle est la probabilité que vous regardiez au moins 2 matchs de badminton ou au moins 1 match de foot ?

A. 0.6818 **B.** 0.7727 **C.** 0.8182 **D.** 0.7456 **E.** 0.9545

Résolution. Si $B =$ badminton et $F =$ foot, alors on cherche

$$\begin{aligned} P(\{B \geq 2\} \cup \{F \geq 1\}) &= P(B \geq 2) + P(F \geq 1) - P(B \geq 2 \cap F \geq 1) \\ &= P(B = 2) + P(B = 3) + 1 - P(F = 0) - P(B = 2, F = 1) \\ &= 1 + \frac{\binom{6}{2}\binom{6}{1} + \binom{6}{3} - \binom{10}{3} - \binom{6}{2}\binom{2}{1}}{\binom{12}{3}} \approx 0.8182 \end{aligned}$$

Solution alternative :

$$\begin{aligned} 1 - P(\{B \leq 1\} \cap \{F = 0\}) &= 1 - P(B = 0, F = 0) - P(B = 1, F = 0) \\ &= 1 - \frac{\binom{4}{3} + \binom{6}{1}\binom{4}{2}}{\binom{12}{3}} \end{aligned}$$

3. Vous êtes en train de regarder l'épreuve du lancer du poids féminin à la télévision. Le commentateur sportif annonce que, dans le passé, la distance moyenne atteinte par la prochaine concurrente a été de 19.98m, avec un écart-type de 0.15m. Sans faire d'autres hypothèses, quel est l'énoncé le plus précis que vous puissiez faire à propos de la probabilité que le poids de cette concurrente parcoure au moins 20.70m, la distance de la médaillée d'or en 2012 ?

A. ≥ 0.0217 **B.** ≤ 0.0434 **C.** ≤ 0.0217 **D.** ≤ 0.9566 **E.** ≥ 0.4783

Résolution. Par le théorème de Tchebysheff, $P(|Y-19.98| > 0.15k) \leq \frac{1}{k^2}$, donc $P(Y > 19.98 + 0.15k) \leq \frac{1}{k^2}$. Puisque $19.98 + 0.15k = 20.70$, on en déduit que $k = 4.8$, et donc $P(Y > 20.70) \leq 0.0434$.

4. Le diamètre d'un ballon de basketball utilisé aux Jeux Olympiques suit une distribution normale de moyenne 24 cm et d'écart-type 0.5 cm. Quel est en moyenne le volume (en cm^3) d'un de ces ballons de basketball ? [Indice : on rappelle que le volume d'une sphère de rayon R est donné par $\frac{4}{3}\pi R^3$. De plus, $\frac{d^2}{dx^2}\exp(ax+bx^2) = ((a+2bx)^2+2b)\exp(ax+bx^2)$.]

A. 7257.08 **B.** 57981.23 **C.** 7238.23 **D.** 7247.65 **E.** 58056.63

Résolution. Soit V le volume du ballon. On veut calculer $E(V) = \frac{4}{3}\pi E(R^3) = \frac{\pi}{6}E(D^3)$ où D représente le diamètre du ballon. Or, si $D \sim N(\mu, \sigma^2)$, alors la FGM de D est donnée par $m(t) = \exp(\mu t + \frac{t^2\sigma^2}{2})$ et donc $E(D^3) = \frac{d^3}{dt^3}m(t)|_{t=0} = \mu^3 + 3\mu\sigma^2$. Avec $\mu = 24$ et $\sigma = 0.5$, on obtient $E(D^3) = 13842$ et finalement $E(V) = 7247.65$.

5. Les organisateurs des Jeux Olympiques ont commandé 200 balles de hockey chez le distributeur I, qui promet un taux de produits défectueux de 5%. Malheureusement, le distributeur I n'a pas assez de balles en stock et n'en livre que 120 : les organisateurs doivent alors également commander 80 balles chez le distributeur II, dont le taux de produits défectueux est de 10%. Lors d'un match de hockey, une balle s'avère être défectueuse. Quelle est la probabilité qu'elle provienne du distributeur II ?

A. 0.6064 **B.** 0.5714 **C.** 0.2593 **D.** 0.3520 **E.** 0.3936

Résolution. Si D = défectueux, on calcule la probabilité qu'une balle vienne du distributeur II sachant qu'elle est défectueuse :

$$P(II \mid D) = \frac{P(D \mid II)P(II)}{P(D \mid II)P(II) + P(D \mid I)P(I)} = \frac{0.1 \times 0.4}{0.1 \times 0.4 + 0.05 \times 0.6} \approx 0.5714$$

6. Lors de ses entrainements, une lanceuse de javelot a remarqué que la distance moyenne de ses lancers est de 66.23m avec un écart-type de 2.05m. Quelle est la probabilité (approximative) que, lors d'une série de 20 lancers successifs, elle réalise une distance moyenne d'au moins 67m ?

A. 0.3536 **B.** 0.0930 **C.** 0.3756 **D.** 0.0465 **E.** 0.9535

Résolution. Grâce au théorème central limite, la distribution de $\bar{Y}_{20}$ peut être approchée par la distribution $N(66.23, (2.05)^2/20)$. On trouve que

$$\begin{aligned} P(\bar{Y}_{20} > 67) &\approx 1 - \Phi\left(\sqrt{20}\frac{67-66.23}{2.05}\right) \\ &= 1 - \Phi(1.679778) = 0.0465 \end{aligned}$$

7. La fonction de répartition du temps de course (en secondes) d'Usain Bolt sur 100m est donnée par

$$F(y) = \begin{cases} 0 & \text{si } y < 9.55 \\ 1000 \times (y - 9.55)^2/9 & \text{si } 9.55 \leq y < 9.58 \\ 1 - \dfrac{9}{10} \exp\{6 \times (9.58 - y)\} & \text{si } y \geq 9.58 \end{cases}$$

Quelle est la probabilité (à 10^{-2} près) qu'il ne batte ni son record du monde de 9.58s obtenu aux Championnats du monde de Berlin en 2009, ni son temps de 9.63s obtenu aux Jeux Olympiques de 2012 ?

A. 0.9 **B.** 0.67 **C.** 0.10 **D.** 0.33 **E.** 0.71

Résolution.

$$\begin{aligned} P(Y \geq 9.63, Y \geq 9.58) &= P(Y \geq 9.63) \\ &= 1 - F(9.63) = \frac{9}{10} \exp\{6 \times (9.58 - 9.63)\} \\ &= \frac{9}{10} \exp\{-0.3\} \\ &\simeq \frac{9}{10} \times 0.74 \simeq 0.67. \end{aligned}$$

8. Vous essayez d'acheter un billet pour la finale du 100 mètres des Jeux Olympiques par téléphone. Puisque c'est un événement fort demandé, les organisateurs ont communiqué que la probabilité de devoir faire au moins 3 appels avant d'être servi est de 81%. Quelle est la probabilité que vous ayez un billet lors du deuxième essai ?

A. 0.09 **B.** 0.1 **C.** 0.1539 **D.** 0.19 **E.** 0.0292

Résolution. Soit Y = nombre d'essais jusqu'au premier succès, $Y \sim \text{Geo}(p)$.

$$\begin{aligned} P(Y \geq 3) &= 1 - P(Y = 1) - P(Y = 2) = 1 - p - (1 - p) \times p \\ &= 1 - 2p + p^2 = (1 - p)^2 = 0.81, \end{aligned}$$

alors $p = 0.1$. La probabilité recherchée est $P(Y = 2) = (1 - p) \times p = 0.09$.

9. Un athlète qui s'entraine pour l'épreuve du 100 mètres papillon en natation a l'impression que ses performances lors des entrainements sont fort variables. Il aimerait pouvoir quantifier la variabilité dans le temps qu'il met habituellement pour parcourir cette distance. Selon son entraineur, dans 90% des cas, il met moins de 54.1625 secondes. Par contre, il n'arrive à passer sous la barre des 52.386 secondes que dans 20% des cas. Si l'on suppose que le temps qu'il met pour nager les 100 mètres suit une distribution normale, que vaut son écart-type ?

A. 1.78 **B.** 53.09 **C.** 0.84 **D.** 59.22 **E.** 53.27

Résolution. Puisque $X \sim N(\mu, \sigma^2)$, alors, d'une part, $P(X < 54.1625) = 0.9 \Leftrightarrow \frac{54.1625 - \mu}{\sigma} = 1.28$. D'autre part, $P(X < 52.386) = 0.2 \Leftrightarrow \frac{\mu - 52.386}{\sigma} = 0.84$. En résolvant le système formé de ces deux équations, on trouve $\mu = 53.09$ et $\sigma = 0.84$.

10. Lors du 20 kilomètres marche, les concurrents ne peuvent pas courir, c'est-à-dire qu'un pied au moins doit être en permanence en contact avec le sol sous peine de recevoir un avertissement. Au bout de 3 avertissements, un athlète est disqualifié. Soient X et Y les distances parcourues (en dizaines de kilomètres) par un athlète avant de recevoir respectivement le premier et le deuxième avertissement. La fonction de densité jointe de X et Y est donnée par $f(x,y) = \frac{3}{4}x$ si $0 \leq x \leq y \leq 2$ et $f(x,y) = 0$ sinon. On s'intéresse aux athlètes recevant leur deuxième avertissement exactement à la mi-course. Quelle distance (en kilomètres) ont parcouru en moyenne ces athlètes-là au moment de recevoir leur premier avertissement ?

A. 6.667 **B.** 5.333 **C.** 7.5 **D.** 5 **E.** 3.333

Résolution. On veut calculer $E(X|Y=1)$. La densité marginale de Y est donnée par $f_Y(y) = \int_0^y \frac{3}{4}x\,dx = \frac{3y^2}{8}$ si $0 \leq y \leq 2$. La densité conditionnelle de X sachant Y est alors donnée par $f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{3x/4}{3y^2/8} = \frac{2x}{y^2}$ pour $0 \leq x \leq y \leq 2$. Donc, $f(x|y=1) = 2x$ pour $0 \leq x \leq 1$ et $E(X|Y=1) = \int_0^1 2x^2\,dx = \frac{2}{3}$, ce qui correspond à $20/3 = 6.667$ kilomètres.

11. Le nombre de buts marqués par les Belges lors d'un match de foot aux Jeux Olympiques est en moyenne de 3 par heure. Lors du premier match contre les Français, les Belges sont menés 0-1 à l'issue de la première mi-temps. Un match de foot dure 90 minutes au total. Quelle est la probabilité que les Belges gagnent le match (sans passer par des penalties), si les Français ne marquent plus aucun but ?

A. 0.4422 **B.** 0.6575 **C.** 0.5714 **D.** 0.8946 **E.** 0.8009

Résolution. Soit Y le nombre de buts marqués par les Belges, $Y \sim \text{Pois}(\lambda)$, où $\lambda = 3 \times 0.75 = 2.25$ par 45 minutes. On cherche

$$P[Y \geq 2] = 1 - P[Y=0] - P[Y=1] = 1 - e^{-\lambda} - \lambda e^{-\lambda} = 0.6575.$$

12. La cérémonie d'ouverture des Jeux Olympiques 2016 va commencer dans 15 minutes dans le stade de Maracanã. Or, l'équipe chargée d'accueillir la flamme olympique et son porteur n'a toujours aucune nouvelle de ceux-ci, et le stress commence à monter. On suppose que le temps nécessaire pour que le porteur de la flamme arrive au stade suit une distribution exponentielle de moyenne 12 minutes. Quelle est la probabilité que la flamme olympique arrive au stade pendant les 5 premières minutes de la cérémonie d'ouverture ?

A. 0.1254 **B.** 0.3408 **C.** 0.2865 **D.** 0.8111 **E.** 0.0976

Résolution. Soit $X \sim \text{Exp}(12)$ le temps nécessaire au porteur de la flamme pour arriver au stade. Alors $P(15 < X < 20) = \int_{15}^{20} \frac{1}{12} e^{-x/12} dx = 0.0976$.

13. Après l'épreuve du 400 mètres, Kévin Borlée se rend auprès des journalistes pour répondre à leurs questions. Il y a 3 journalistes belges, 2 journalistes brésiliens et un journaliste anglais présents. Kévin décide de répondre aléatoirement à exactement 3 journalistes au total. Quelle est la probabilité que Kévin réponde à exactement un des trois journalistes belges sachant qu'il répondra à au moins un des deux journalistes brésiliens ?

A. $\frac{3}{4}$ **B.** $\frac{9}{16}$ **C.** $\frac{9}{20}$ **D.** $\frac{3}{10}$ **E.** $\frac{3}{8}$

Résolution. Soient Y_1 et Y_2 respectivement le nombre de journalistes belges et brésiliens auxquels Kévin répond. La fonction de probabilité jointe de Y_1 et Y_2 est donnée par $p(y_1, y_2) = \frac{C^3_{y_1} C^2_{y_2} C^1_{3-y_1-y_2}}{C^6_3}$ où $y_1 \in \{0,1,2,3\}$, $y_2 \in \{0,1,2\}$ et $0 \leq y_1 + y_2 \leq 3$. De plus, $p_2(y_2) = \frac{C^2_{y_2} C^4_{3-y_2}}{C^6_3}$. On trouve donc finalement que $P(Y_1 = 1|Y_2 \geq 1) = \frac{P(Y_1=1, Y_2 \geq 1)}{P(Y_2 \geq 1)} = \frac{p(1,1)+p(1,2)}{1-p_2(0)} = \frac{6/20+3/20}{1-4/20} = \frac{9}{16}$.
Résolution alternative : parmi les $\binom{6}{3} = 20$ triples de journalistes possibles, il y en a $\binom{4}{3} = 4$ sans aucun journaliste brésilien et donc $20-4 = 16$ avec au moins un journaliste brésilien. Parmi ces 16 triples, il y en a $3 \times 2 = 6$ qui sont composé d'un journaliste belge, brésilien et anglais et encore 3 qui sont composé d'un journaliste belge et les deux journalistes brésiliens. La probabilité demandée est donc de $(6+3)/16$.

14. La densité jointe de Y_1, le temps de course (en heures) en cyclisme, et de Y_2, le temps de course total (en heures) d'un athlète de triathlon, est

$$f(y_1, y_2) = \begin{cases} e^{-y_2} & \text{si } 0 \leq y_1 \leq y_2 \\ 0 & \text{sinon.} \end{cases}$$

Sachant que $E[Y_1] = 1$, $E\left[Y_1^2\right] = 2$, $E[Y_2] = 2$, $E\left[Y_2^2\right] = 6$ et

$$\int_0^\infty x^n e^{-x} dx = n!$$

pour tout n entier positif, quelle est la corrélation entre ces deux temps ?

A. $1/2$ **B.** $1/\sqrt{3}$ **C.** 0.1 **D.** $1/\sqrt{2}$ **E.** $\sqrt{2/3}$

Résolution.

$$\begin{aligned} E[Y_1 Y_2] &= \int_0^\infty \int_0^{y_2} y_1 y_2 e^{-y_2} dy_1 dy_2 = \int_0^\infty \left[\frac{y_1^2}{2}\right]_0^{y_2} y_2 e^{-y_2} dy_2 \\ &= \frac{1}{2}\int_0^\infty y_2^3 e^{-y_2} dy_2 = \frac{1}{2}\int_0^\infty x^3 e^{-x} dx \\ &= \frac{1}{2}\Gamma(4) = 3 \end{aligned}$$

Les densités marginales sont

$$f_{Y_1}(y_1) = e^{-y_1} \sim \text{Exp}(1)$$
$$\text{donc} \qquad E[Y_1] = 1 \qquad \text{et} \qquad \text{V}(Y_1) = E[Y_1^2] - (E[Y_1])^2 = 1,$$

et

$$f_{Y_2}(y_2) = y_2 e^{-y_2} \sim \Gamma(2, 1)$$
$$\text{donc} \qquad E[Y_2] = 2 \qquad \text{et} \qquad \text{V}(Y_2) = E[Y_2^2] - (E[Y_2])^2 = 2.$$

On conclut que

$$\text{Cov}(Y_1, Y_2) = 3 - 2 = 1 \qquad \text{et} \qquad \text{Cor}(Y_1, Y_2) = \frac{1}{\sqrt{2}}.$$

15. [Question bonus.] L'équipe marketing qui gère la boutique en ligne officielle des Jeux Olympiques aimerait davantage connaître les habitudes d'achats de ses clients les plus importants, selon leur sexe. Ils sont particulièrement intéressés par la variabilité dans les montants dépensés par les hommes.
Parmi tous les clients ayant effectué au moins un achat jusqu'à présent, 43% sont des femmes. Un client est considéré comme important s'il dépense au minimum 200 euros sur le site officiel. Jusqu'à présent, 27% des clients correspondent à cette description. Parmi les femmes, ce taux augmente à 39%. Par ailleurs, le montant moyen dépensé par les hommes est de 130 euros, et on peut supposer que la distribution des montants dépensés par les hommes suit une distribution normale.
Quelle est alors la variance du montant dépensé par un homme ?

A. 0.1795 **B.** 1111.11 **C.** 33.33 **D.** 5789.23 **E.** 76.09

Résolution. Soient F et A les événements 'être une femme' et 'être un client important', respectivement. On sait que $P(A) = 0.27$, $P(A|F) = 0.39$, $P(F) = 0.43$. De plus, on suppose que le montant dépensé par un homme est $Y \sim N(130, \sigma^2)$. Par la formule des probabilités totales, on trouve alors que $P(A|\overline{F}) = 0.1795$. Ce qui veut dire que $P(Y > 200) = P(Z > \frac{200-130}{\sigma}) = 0.1795$. On en déduit que $\frac{200-130}{\sigma} = 0.92$, d'où $\sigma = 76.0870$, et donc $\sigma^2 = 5789.2250$.

16–20. 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 :

16. Le nombre de lancers dont une lanceuse de javelot a besoin afin de dépasser les 67m.
17. La durée, en minutes, jusqu'au premier but dans un tournoi de foot.
18. Le nombre de balles de hockey défectueuses dans un paquet de 25 balles.
19. Le nombre total de buts marqués dans 25 matchs de foot, si, lors d'un seul match, il y a 2 buts en moyenne avec un écart-type de 1 but.
20. Parmi 25 journalistes, 10 sont brésiliens. Si 5 parmi les 25 sont choisis au hasard pour interviewer Usain Bolt, le nombre de journalistes brésiliens sélectionnés.

Les distributions :

A. Binomiale$(25, 0.2)$
B. Bernoulli(0.2)
C. Normale$(50, 25)$
D. Binomiale négative$(25, 0.2)$
E. Uniforme$(5, 25)$
F. Géométrique(0.2)
G. Exponentielle(25)
H. Hypergéométrique$(25, 10, 5)$
I. Normale$(2, 1.2)$
J. Poisson(0.2)

Résolution

16. F
17. G
18. A
19. C
20. H