

Préparation à l'agrégation externe de Sciences Sociales

Probabilités 2

2020-2021

Exercice 1 (2015)

1. Soit $X \sim U[0, 1]$,
— Fonction de densité :

$$f_X(x) = \begin{cases} 1 & \text{si } x \in [0, 1] \\ 0 & \text{sinon} \end{cases}$$

- Fonction de répartition :

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

- $E(X) = 1/2$
— $V(X) = 1/12$

2. La variable aléatoire U représente le temps d'attente du voyageur avant de faire le premier trajet.
3. L'événement $(U > x)$ peut s'écrire comme :

$$\begin{aligned} (U > x) &= (\min\{X, Y\} > x) \\ &= (X > x) \cap (Y > x). \end{aligned}$$

Sa probabilité est :

- Si $x < 0$, $P(U > x) = 1$
— Si $0 \leq x \leq 1$,

$$\begin{aligned} P(U > x) &= P((X > x) \cap (Y > x)) \\ &= P(X > x)P(Y > x) \\ &= (1 - P(X \leq x))(1 - P(Y \leq x)) \\ &= (1 - x)^2 \\ &= 1 - 2x + x^2. \end{aligned}$$

- Si $x > 1$, $P(U > x) = 0$

La fonction de répartition peut s'écrire comme $F(x) = P(U \leq x) = 1 - P(U > x)$. Donc,

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 2x - x^2 & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

4. Il faut montrer que $F(x) = \int_{-\infty}^x f(t)dt$ où

$$f(t) = \begin{cases} 2 - 2t & \text{si } t \in [0, 1] \\ 0 & \text{sinon} \end{cases}$$

— Si $x < 0$,

$$\begin{aligned}\int_{-\infty}^x f(t)dt &= \int_{-\infty}^x 0dt \\ &= 0 \\ &= F(x)\end{aligned}$$

— Si $0 \leq x \leq 1$,

$$\begin{aligned}\int_{-\infty}^x f(t)dt &= \int_0^x (2 - 2t)dt \\ &= 2x - x^2 \\ &= F(x)\end{aligned}$$

— Si $x > 1$,

$$\begin{aligned}\int_{-\infty}^x f(t)dt &= \int_0^1 (2 - 2t)dt \\ &= 1 \\ &= F(x)\end{aligned}$$

5. L'espérance et la variance de U sont :

$$\begin{aligned}E(U) &= \int_{\mathbb{R}} xf(x)dx \\ &= \int_0^1 x(2 - 2x)dx \\ &= 1/3.\end{aligned}$$

$$\begin{aligned}V(U) &= E(U^2) - E(U)^2 \\ &= \int_{\mathbb{R}} x^2 f(x)dx - 1/9 \\ &= \int_0^1 x^2(2 - 2x)dx - 1/9 \\ &= 1/6 - 1/9 \\ &= 1/18.\end{aligned}$$

6. On peut écrire la durée totale (en minutes) comme :

$$\begin{aligned}DT &= 10 \times U + T_1 + C + 10 \times Z + T_2 \\ &= 125 + 10 \times U + 10 \times Z,\end{aligned}$$

où :

- U représente le temps d'attente avant le premier trajet (en dizaines de minutes),
- T_1 représente la durée du premier trajet qui est égale à 60 minutes,
- C représente le temps de changement qui est égale à 5 minutes,
- Z représente le temps d'attente avant le deuxième trajet (en dizaines de minutes),
- T_2 représente la durée du deuxième trajet qui est égale à 60 minutes.

Donc,

$$\begin{aligned}E(DT) &= E(125 + 10 \times U + 10 \times Z) \\ &= 125 + 10 \times E(U) + 10 \times E(Z) \\ &= 125 + 3,3 + 5 \\ &= 133,3.\end{aligned}$$

Exercice 2 (2008)

1. On peut définir pour $i = 1, \dots, 100$ les variables

$$Z_i = \begin{cases} 1 & \text{si l'internaute } i\text{-ème a passé au moins une commande dans l'année 2006} \\ 0 & \text{sinon} \end{cases}$$

Chaque Z_i suit une loi Bernoulli avec paramètre $p = 0,3$.

Donc on peut écrire

$$X = \sum_{i=1}^{100} Z_i$$

où les variables Z_1, \dots, Z_{100} suivent une loi $Ber(0,3)$ et peuvent être considérés indépendants¹.
Pourtant :

$$X \sim Bin(100; 0,3),$$

$$E(X) = 0,3 \times 100 = 30,$$

$$V(X) = 100 \times 0,3 \times (1 - 0,3) = 21.$$

2. Par une loi normale $N(30, \sqrt{21})$.
3. Soit $Z \sim N(0, 1)$,

$$\begin{aligned} P(X > 30) &= P(X - 30 > 0) \\ &= P\left(\frac{X - 30}{\sqrt{21}} > 0\right) \\ &= P(Z > 0) \\ &= 1/2 \end{aligned}$$

$$\begin{aligned} P(|X - E(X)| \leq 2) &= P(-2 \leq X - 30 \leq 2) \\ &= P\left(\frac{-2}{\sqrt{21}} \leq \frac{X - 30}{\sqrt{21}} \leq \frac{2}{\sqrt{21}}\right) \\ &= P\left(\frac{-2}{\sqrt{21}} \leq Z \leq \frac{2}{\sqrt{21}}\right) \\ &= \Phi(0.44) - \Phi(-0.44) \\ &= \Phi(0.44) - [1 - \Phi(0.44)] \\ &= 2 \times \Phi(0.44) - 1 \\ &= 2 \times 0,67 - 1 \\ &= 0,34 \end{aligned}$$

$$\begin{aligned} P(X = 80) &= P\left(\frac{X - 30}{\sqrt{21}} = \frac{50}{\sqrt{21}}\right) \\ &= P(Z = 10.9) \\ &= 0 \end{aligned}$$

1. Comme le tirage est fait avec remise, rien empêche la possibilité de qu'on prend deux fois le même internaute, donc les variables Z_1, \dots, Z_{100} ne sont pas forcément indépendants. Mais, comme le nombre d'internautes est beaucoup plus grand que 100, la probabilité de choisir deux fois le même internaute est très très faible et, pourtant, on peut considérer Z_1, \dots, Z_{100} comme variables indépendants.