L2 MASS Université Lyon 1, Probabilités-statistique 2
Année Universitaire 2020 – 2021

Correction TD 1 - Probabilités

Exercice 1 :
On va noter E l'ensemble des numéros des 40000 étudiant.e.s inscrit.e.s, et on pose $\Omega = E^{100}$. L'univers Ω correspond donc à l'ensemble des listes de 100 numéros d'étudiant.e.s inscrit.e.s. Au passage, cela signifie qu'il peut y avoir 2 fois le même numéro, donc que l'on cherche à modéliser un tirage *avec remise.*
Il reste à définir $\mathbb{P}$. Comme on veut que chaque tirage soit équiprobable, et qu'il y a 40000^{100} tirages possibles (avec remise), il suffit de définir :

$$\forall A \in \mathcal{A}, \mathbb{P}(A) = \frac{Card(A)}{40000^{100}}.$$

On note ensuite, pour un numéro d'étudiant i, b_i la variable qui vaut 1 si l'étudiant.e est fumeur.euse et 0 sinon.
On peut alors définir

$$\forall \omega \in \Omega, X(\omega) = \Big(b_{\omega_1}, \cdots, b_{\omega_{100}}\Big).$$

On a donc $X_i = b_{\omega_i}$.
La variable aléatoire X_i prend ses valeurs dans $\{0,1\}$, il s'agit donc d'une variable aléatoire de Bernoulli, et il n'y a plus qu'à calculer son paramètre $p = \mathbb{P}(X_i = 1)$:
On a

$$\mathbb{P}(X_i = 1) = \mathbb{P}\Big(\{\omega \in \Omega, X_i(\omega) = 1\}\Big) = \mathbb{P}\Big(\{\omega \in \Omega, b_{\omega_i} = 1\}\Big) = \frac{\text{Nombre de fumeurs}}{40000}.$$

Exercice 2 :
On va calculer les espérances pour les lois usuelles :

- Loi de Bernoulli : $X \sim \mathcal{B}(p), p \in [0,1]$

$$\mathbb{E}[X] = (1-p).0 + p.1 = p$$

- Loi Binomiale :$X \sim \mathcal{B}(N,p), N \in \mathbb{N}^*, p \in [0,1]$

  Le plus simple est d'utiliser le fait que X a même loi qu'une somme de N variables aléatoires $\mathcal{B}(p)$ indépendantes, donc même espérance, puis d'utiliser la linéarité de l'espérance. On a donc, si $Y_1, \cdots Y_N$ désigne N variables aléatoires $\mathcal{B}(p)$ indépendantes :

$$\mathbb{E}[X] = \mathbb{E}\Big[\sum_{i=1}^{N} Y_i\Big] = \sum_{i=1}^{N} \mathbb{E}\Big[Y_i\Big] = \sum_{i=1}^{N} p = pN.$$

  L'autre façon consiste à essayer de faire le calcul direct en utilisant la formule du Binôme de Newton :
  On a

$$\mathbb{E}[X] = \sum_{k=0}^{N} k \binom{N}{k} p^k (1-p)^{N-k},$$

  Pour calculer cette somme, il suffit d'écrire

$$(x+y)^N = \sum_{k=0}^{N} \binom{N}{k} x^k y^{N-k},$$

  et de dériver partiellement par rapport à x :

$$N(x+y)^{N-1} = \sum_{k=0}^{N} k \binom{N}{k} x^{k-1} y^{N-k},$$

On reconnaît alors la formule que l'on cherche à calculer, ce qui donne

$$\begin{aligned}
\mathbb{E}[X] &= p\sum_{k=0}^{N} k\binom{N}{k}p^{k-1}(1-p)^{N-k} \\
&= pN(p+(1-p))^{N-1} \\
&= Np
\end{aligned}$$

- Loi de Poisson $X \sim \mathcal{P}(\lambda), \lambda > 0$

  Pour cette question, il faut voir que l'on va traiter avec des sommes infinies, donc c'est mieux d'avoir en tête le cours de séries entières. Ici on va sommer seulement des termes positifs, donc on peut donner un sens à la somme, qu'elle soit finie ou infinie.

  On peut écrire
  $$\mathbb{E}[X] = \sum_{k\in\mathbb{N}} k\frac{\lambda^k}{k!}e^{-\lambda},$$

  Pour calculer la valeur de cette série, on peut remarquer qu'elle ressemble à la série entière $e^x = \sum_{k\in\mathbb{N}} \frac{x^k}{k!}$ qui a un rayon de convergence infini. On peut dériver terme à terme à l'intérieur (strictement) du disque de convergence, c'est à dire partout ici, cela donne $e^x = \sum_{k\in\mathbb{N}} k\frac{x^{k-1}}{k!}$, et onn reconnaît la somme cherchée. On obtient alors

  $$\begin{aligned}
  E[X] &= \lambda e^{-\lambda}\sum_{k\in\mathbb{N}} k\frac{\lambda^{k-1}}{k!} \\
  &= \lambda e^{-\lambda}e^{\lambda} \\
  &= \lambda
  \end{aligned}$$

  Il est aussi possible de simplifier le $\frac{k}{k!} = \frac{1}{(k-1)!}$, puis de faire un changement d'indices pour obtenir le même résusltat.

- Loi Uniforme discrète $X \sim \mathcal{U}(\llbracket 1, N \rrbracket), N \in \mathbb{N}^*$. Le calcul peut ici se faire directement :
  $$\mathbb{E}[X] = \sum_{k=1}^{N} k\frac{1}{N} = \frac{1}{N}\frac{N(N+1)}{2} = \frac{N+1}{2}.$$

  On peut vérifier que pour un dé équilibré à 6 faces par exemple, $N = 6$, et on retrouve $\mathbb{E}[X] = \frac{7}{2} = 3.5$.

- Loi Géométrique $X \sim \mathcal{G}(p), p \in]0, 1]$, *on exclue ici $p = 0$, car dans ce cas X est infinie avec probabilité 1 : le premier succès n'arrive jamais.*

  Ici encore, il faudra être prudent, car on va manipuler des séries. Mais il s'agit encore de termes positif, donc on peut écrire
  $$\mathbb{E}[X] = \sum_{k\in\mathbb{N}} kp(1-p)^{k-1},$$

  Cette série peut se calculer en dérivant terme à terme la série $\frac{1}{1-x} = \sum_{k\in\mathbb{N}} x^k$ qui a un rayon de convergence de 1, donc on peut dériver terme à terme pour $|x| < 1$, et on obtient
  $$\forall x \in \mathbb{R}, |x| < 1, \frac{1}{(1-x)^2} = \sum_{k\in\mathbb{N}} kx^{k-1}.$$

  Il ne reste plus qu'à appliquer la formule:

  $$\begin{aligned}
  \mathbb{E}[X] &= p\sum_{k\in\mathbb{N}^*} k(1-p)^{k-1} \\
  &= p\frac{1}{\Big(1-(1-p)\Big)^2} \\
  &= \frac{p}{p^2} \\
  &= \frac{1}{p}.
  \end{aligned}$$

Ce résultat est un peu intuitif, puisqu'il signifie par exemple que si l'on a $1/100$ de gagner à un jeu, alors il faudra y jouer *en moyenne* 100 fois avant de gagner la première partie.

- Loi Exponentielle $X \sim \mathcal{E}(\lambda), \lambda > 0$, On va devoir manipuler des intégrales généralisées (entre 0 et $+\infty$). On rappelle que cette intégrale est définie pour g positive comme $\int_0^{+\infty} g(x)\mathrm{d}x = \lim_{M\to+\infty} \int_0^M g(x)\mathrm{d}x$. Dans les calculs qui suivent, toutes ces limites sont sous-entendues.

$$
\begin{aligned}
\mathbb{E}[X] &= \int_0^{+\infty} x\lambda e^{-\lambda x}\mathrm{d}x \\
&\overset{IPP}{=} \left[x\lambda \frac{e^{-\lambda x}}{-\lambda}\right]_0^{+\infty} - \int_0^{+\infty} \lambda \frac{e^{-\lambda x}}{-\lambda}\mathrm{d}x \\
&= 0 + \int_0^{+\infty} e^{-\lambda x}\mathrm{d}x \\
&= \left[\frac{e^{-\lambda x}}{-\lambda}\right]_0^{+\infty} \\
&= 0 - \frac{e^{-\lambda 0}}{-\lambda} \\
&= \frac{1}{\lambda}
\end{aligned}
$$

- Loi Normale $X \sim \mathcal{N}(m, \sigma^2), m \in \mathbb{R}, \sigma > 0$, attention, on va ici intégrer une fonction non nécessairement positive. Pour pouvoir donner un sens à $\mathbb{E}[X]$, il faut tout d'abord vérifier que $\mathbb{E}[|X|] < +\infty$, i.e. que l'intégrale $\int_{-\infty}^{+\infty} |x| \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}\mathrm{d}x$ est convergente. Cela peut se faire par exemple en utilisant $|x|e^{-\frac{(x-m)^2}{2\sigma^2}} = o_{\pm\infty}(\frac{1}{x^2})$.
  On a alors
$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}\mathrm{d}x.$$
  On peut remarquer ici qu'apparaît presque la dérivée $\left(e^{-\frac{(x-m)^2}{2\sigma^2}}\right)' = \frac{x-m}{\sigma^2} e^{-\frac{(x-m)^2}{2\sigma^2}}$. Donc pour pouvoir primitiver directement, on va artificiellement la faire apparaître :
$$
\begin{aligned}
\mathbb{E}[X] &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} (x-m)e^{-\frac{(x-m)^2}{2\sigma^2}}\mathrm{d}x + \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} me^{-\frac{(x-m)^2}{2\sigma^2}}\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} \sigma^2\Big(-e^{-\frac{(x-m)^2}{2\sigma^2}}\Big)'\mathrm{d}x + m,
\end{aligned}
$$
  où l'on a utilisé pour le terme de droite, le fait que $x \mapsto \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$ est une densité, donc d'intégrale vaut 1.
  On obtient alors
$$
\begin{aligned}
\mathbb{E}[X] &= \frac{\sigma}{\sqrt{2\pi}}\left[-e^{-\frac{(x-m)^2}{2\sigma^2}}\right]_{-\infty}^{+\infty} + m \\
&= m.
\end{aligned}
$$
  Une façon alternative de calculer le terme de gauche est de faire un changement de variables $y = x - m$ afin d'obtenir $\int_{-\infty}^{+\infty} ye^{-\frac{y^2}{2\sigma^2}}\mathrm{d}y$, puis d'utiliser le fait que l'intégrale d'une fonction impaire sur un intervalle symétrique par rapport à 0 est nulle.

- Loi Uniforme continue $X \sim \mathcal{U}([a,b]), a < b$:

  On peut faire le calcul directement :

$$\begin{aligned}
\mathbb{E}[X] &= \int_a^b \frac{x}{b-a}\mathrm{d}x \\
&= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b \\
&= \frac{b^2-a^2}{2(b-a)} \\
&= \frac{(b-a)(b+a)}{2(b-a)} \\
&= \frac{a+b}{2},
\end{aligned}$$

  ce qui est plutôt intuitif : si on tire uniformément entre deux nombre a et b, alors l'espérance du résultat correspond à la moyenne de ces nombres.

Exercice 3 :

1. - La variable X est à valeurs dans $\mathbb{N}$, mais à cause du terme $\mathbb{1}_{1\leq k\leq 6}$, on a $X(\Omega) = \{1, 2, \cdots, 6\}$.
   - Il faut que $\mathbb{P}(X \in \mathbb{R}) = 1$, donc $\sum_{k=1}^{6} \mathbb{P}(X = k) = 1$. On peut utiliser la formule donnée pour calculer C :

$$\begin{aligned}
\sum_{k=1}^{6} \frac{C}{k} = 1 \Leftrightarrow C &= \frac{1}{\sum_{k=1}^{6} \frac{1}{k}} \\
\Leftrightarrow C &= \frac{1}{\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{6}} \\
\Leftrightarrow C &= \frac{20}{49} \approx 0.408
\end{aligned}$$

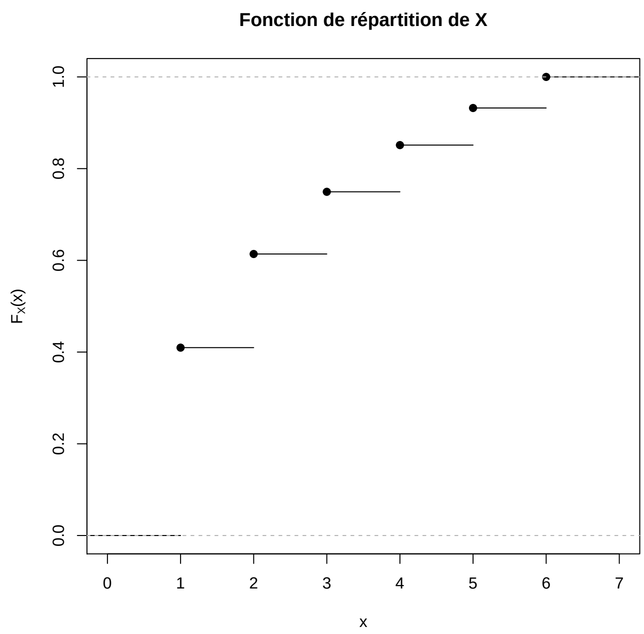
   - On peut alors calculer

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^{6} k\frac{C}{k} \\
&= \sum_{k=1}^{6} C \\
&= 6C \\
&= \frac{120}{49} \\
&\approx 2.449
\end{aligned}$$

   - On peut calculer

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k=1}^{6} k^2\frac{C}{k} \\
&= \sum_{k=1}^{6} Ck \\
&= 21C \\
&= \frac{420}{49} \\
&\approx 8.571
\end{aligned}$$

- Pour représenter F_X, il faut remarquer que, comme X est une variable discrète, F_X est constante par morceaux avec des "sauts" en $k \in [\![1,6]\!]$ de hauteur $\mathbb{P}(X = k)$. Par exemple, le premier "saut" en $k = 1$ a pour hauteur $\mathbb{P}(X = 1) = \frac{C}{1} \approx 0.408$. On obtient alors



- La variable aléatoire X est une variable aléatoire continue, d'ensemble image $X(\Omega) = [1, 6]$ (là où la densité est strictement positive).
- Pour calculer C, il faut utiliser le fait que $\int_{x\in\mathbb{R}} f_X(x)\mathrm{d}x = 1$, ce qui donne

$$\begin{aligned} C &= \frac{1}{\int_{x\in\mathbb{R}} \frac{1}{x} \mathbb{1}_{x\in[1,6]}\mathrm{d}x} \\ &= \frac{1}{\int_{x=1}^{6} \frac{\mathrm{d}x}{x}} \\ &= \frac{1}{\Big[\log(x)\Big]_1^6} \\ &= \frac{1}{\ln(6) - \ln(1)} \\ &= \frac{1}{\ln(6)} \approx 0.558 \end{aligned}$$

- On peut directement calculer

$$\begin{aligned} \mathbb{E}[X] &= C \int_1^6 x \frac{1}{x} \mathrm{d}x \\ &= C \int_1^6 \mathrm{d}x \\ &= 5C \\ &= \frac{5}{\ln(6)} \approx 2.79 \end{aligned}$$

- On peut directement calculer

$$
\begin{aligned}
\mathbb{E}[X^2] &= C\int_1^6 x^2 \frac{1}{x}\mathrm{d}x \\
&= C\int_1^6 x\mathrm{d}x \\
&= C\Big[\frac{x^2}{2}\Big]_1^6 \\
&= C\Big(\frac{6^2}{2} - \frac{1}{2}\Big) \\
&= \frac{35}{2\ln(6)} = 9.77
\end{aligned}
$$

- La fonction de répartition de X est donnée par

$$
F_X(t) = \begin{cases} 0 \text{ si } t \leq 1 \\ C\int_1^t \frac{1}{x}\mathrm{d}x = C\log(t) = \frac{\log(t)}{\log(6)} \text{ si } 1 \leq t \leq 6 \\ 1 \text{ sinon} \end{cases}
$$

  On peut alors la représenter :