Statistiques Inférentielles II
Correction du TD 1 : Rappels

Licence 3 MIASHS - IDS (2021-2022)

Jairo Cuglirari, Guillaume Metzler
Institut de Communication (ICOM)
Université de Lyon, Université Lumière Lyon 2
Laboratoire ERIC UR 3083, Lyon, France

jairo.cugliari@univ-lyon2.fr guillaume.metzler@univ-lyon2.fr

Exercice 1 : Fonction de masse

Soit X une variable une variable aléatoire discrète selon la fonction de masse de probabilité p définie comme

$$p(x) = \begin{cases} 1/2 & \text{si } x = -1, \\ 1/10 & \text{si } x \in \{2, 3, 4, 5, 10\}, \\ 0 & \text{sinon.} \end{cases}$$

1. Montrer que p est une fonction de masse de probabilité.

   On doit simplement vérifier que la somme des masses est bien égale à 1. Ce que l'on vérifie aisément ici :

   $$\mathbb{P}[X = -1]+\mathbb{P}[X = 2]+\mathbb{P}[X = 3]+\mathbb{P}[X = 4]+\mathbb{P}[X = 5]+\mathbb{P}[X = 10] = \frac{1}{2}+5\times\frac{1}{10} = 1.$$

2. Déterminer l'espérance de X.

   On utilise à nouveau ce que nous avons vu dans le précédent TD pour déterminer le moment d'ordre 1 de notre variable aléatoire.

   $$\mathbb{E}[X] = \sum_{k\in\{-1,2,3,4,5,10\}} k\,\mathbb{P}[X = k],$$

$$
\begin{aligned}
&= -1 \times \frac{1}{2} + \frac{1}{10}(2+3+4+5+10), \\
&= -\frac{1}{2} + \frac{24}{10}, \\
&= \frac{19}{10}.
\end{aligned}
$$

3. Déterminer la variance de X.

   On utilisera la *formule de König-Huygens* qui, si la variable aléatoire admet un moment d'ordre 2, affirme :

$$Var[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

   Nous avons déjà calculé $\mathbb{E}[X]$ précédemment, il nous faut encore calculer $\mathbb{E}[X^2]$,

$$
\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k \in \{-1,2,3,4,5,10\}} k^2 \, \mathbb{P}[X = k], \\
&= 1 \times \frac{1}{2} + \frac{1}{10}(2^2 + 3^2 + 4^2 + 5^2 + 10^2), \\
&= \frac{1}{2} + \frac{154}{10}, \\
&= \frac{159}{10}.
\end{aligned}
$$

   Ainsi la variance de notre variable aléatoire X est

$$Var[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{159}{10} - \frac{361}{100} = \frac{1590}{100} - \frac{361}{100} = \frac{1229}{100}.$$

4. Déterminer la fonction de répartition F_X de la variable aléatoire X.

   Dans le cas discret, la fonction de répartition garde la même définition que dans le cas continue, *i.e.*

$$F_X(t) = \mathbb{P}[X \leq t].$$

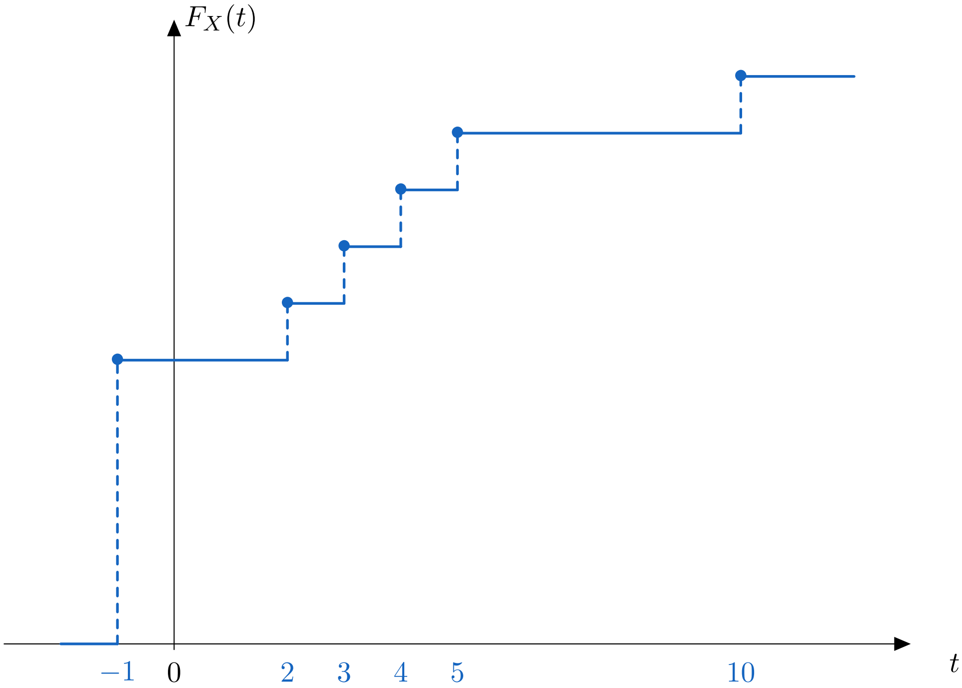
   On va synthétiser les valeurs de cette fonction dans un tableau

t	-1	2	3	4	5	10
$\mathbb{P}[X = t]$	1/2	1/10	1/10	1/10	1/10	1/10
$\mathbb{P}[X \leq t]$	1/2	6/10	7/10	8/10	9/10	1

   On pourrait également donner l'expression de cette fonction de répartition

$$F_X(t) = \sum_{k \in \mathbb{Z}} p(k) \mathbb{1}_{\{k \leq t\}}.$$

Une représentation graphique de cette fonction de répartition est donnée ci-dessous.



On y reconnaît une fonction en forme d'escalier.

5. Obtenir la fonction quantile Q, *i.e.* l'inverse de la fonction de répartition F.

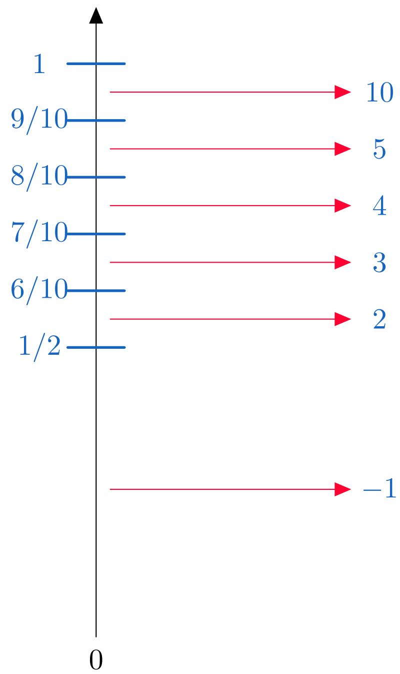
L'inverse de la fonction de répartition est définie pour tout $s \in [0,1]$ par la relation

$$Q_X(s) = F_X^{-1}(s) = \inf\{x : F(x) \geq s\}.$$

On va refaire un tableau analogue à notre précédent tableau pour donner les valeurs de cette fonction quantile.

s	$[0,1/2]$	$]1/2,6/10]$	$]6/10,7/10]$	$]7/10,8/10]$	$]8/10,9/10]$	$]9/10,1]$
$Q_X(s)$	-1	2	3	4	5	10

On peut représenter cette fonction inverse avec l'interprétation graphique suivante où Q_X représente les valeurs prises par la variable aléatoire et s une probabilité.



Ainsi, ce graphique remontre bien que l'on a une probabilité de $1/2$ de tirer la valeur -1 et de $1/10$ pour les autres valeurs.

Cette fonction quantile peut donc s'avérer très utile lorsque l'on cherchera à générer un échantillon issu d'une loi inconnue. On pourra le faire avec le processus suivant :

(a) tirer aléatoirement (selon une loi uniforme) un nombre dans $[0, 1]$

(b) calculer l'image de ce nombre (que l'on voit comme une probabilité) par la fonction quantile.

Exercice 2 : Fonction de densité

Soit X une variable aléatoire dont la densité f est donnée par $f_X(x) = kx^2$ pour tout $x \in [0, 2]$ et $k > 0$.

1. Rappeler les conditions pour qu'une fonction soit une fonction de densité.
   La fonction f_X doit vérifier deux choses :

   (a) elle doit positive pour tout x appartenant au support de la fonction et

   (b) l'intégrale de cette fonction sur sn support doit être égale à 1, *i.e.*

$$\int_I f_X(x)\, dx = 1.$$

2. Déterminer k pour que la fonction f_X soit une fonction de densité.

   Etant donnée la question précédente, on voit déjà que $f_X(x) \geq 0$ pour tout $x \in [0,2]$, pour que ce soit densité, il faut déterminer k telle que

$$\int_0^2 f_X(x)\, dx = 1 \iff \int_0^2 kx^2\, dx = 1 \iff \frac{k}{3}2^3 = 1.$$

   Donc f_X est une densité si $k = \frac{3}{8}$.

3. Déterminer l'espérance de X.

   On applique simplement la définition de l'espérance étudiée lors du précédent TD, ce qui nous donne :

$$\mathbb{E}[X] = \int_I x f_X(x)\, dx = \int_0^2 \frac{3}{8}x^3\, dx = \frac{3}{32} \times 2^4 = \frac{3}{2}.$$

4. Expliquer comment vous obtiendrez la variance de X.

   On utilisera la *formule de König-Huygens* qui, si la variable aléatoire admet un moment d'ordre 2 (ce qui est le cas ici car nous avons une densité continue sur un compact), affirme :

$$Var[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

   Nous avons déjà calculé $\mathbb{E}[X]$ précédemment, il nous faut encore calculer $\mathbb{E}[X^2]$, qui vaut

$$\mathbb{E}[X^2] = \int_I x^2 f_X(x)\, dx = \int_0^2 \frac{3}{8}x^4\, dx = \frac{3}{40} \times 2^5 = \frac{12}{5}.$$

   Ainsi la variance de notre variable aléatoire X est

$$Var[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{12}{5} - \frac{9}{4} = \frac{3}{20}.$$

5. Obtenir la borne M telle que $f_X(x) \leq M$ pour tout $x \in [0,2]$.

   La fonction f_X est continue sur un compact, elle est donc bornée et sa borne supérieure (qui est atteinte) est donnée par $f_X(2) = \frac{3}{2} = M$.

6. Déterminer F_X, la fonction de répartition de la variable aléatoire X.

   On se souvient de la définition de la fonction de répartition

$$F_X(t) = \mathbb{P}[X \leq t] = \int_0^t f_X(x)\, dx = \int_0^t \frac{3}{8}x^2\, dx = \frac{1}{8}t^3.$$

7. Déterminer la fonction F_X^{-1} (on la note parfois Q_X pour fonction quantile) inverse de la fonction de répartition de répartition F_X.

   L'inverse de la fonction de répartition est définie pour tout $s \in [0,1]$ par la relation

   $$Q_X(s) = F_X^{-1}(s) = \inf\{x : F(x) \geq s\}.$$

   Ainsi $Q_X(s)$ est le quantile d'ordre s de la loi de la variable aléatoire X. D'un point de vue pratique, on va calculer l'inverse de la fonction de répartition qui est une fonction bijective sur $[0,2]$.

   $$F(t) = s \iff \frac{1}{8}t^3 = s \iff t^3 = 8s \iff t = 2\sqrt[3]{s} \iff t = F^{-1}(s).$$

   Ainsi la fonction quantile Q_X est la fonction définie de $[0,1]$ dans $[0.2]$ par la relation $Q_X(s) = 2\sqrt[3]{s}$.