Probabilités discrètes et statistiques descriptives

L2 Université Lyon 1 2019

Correction du Contrôle continu 2 du mardi 2 avril 2019

Première partie
(à rendre sur la première copie)

Questions de cours (10 minutes maxi, 4 points) : (cf cours)

Exercice 1. (3 points+ Bonus : 2 points)
On fixe un réel $p \in]0,1[$.

1. Énoncer le théorème de Fubini-Tonelli pour les familles sommables. (cf cours, c'est le cas positif avec une équivalence pour les sommabilités. Il fallait à la fois l'équivalence et la formule)
2. Montrer que la famille $(p^{i+j})_{(i,j)\in\mathbb{N}^2}$ est sommable et calculer sa somme :
$$S_1 = \sum_{(i,j)\in\mathbb{N}^2} p^{i+j}.$$
3. Calculer la somme
$$S_2 = \sum_{n=0}^{\infty} np^n.$$
4. **(Bonus : 2 points)** Montrer que la famille $\big((i+j)p^{i+j}\big)_{(i,j)\in\mathbb{N}^2}$ est sommable et calculer sa somme :
$$S_3 = \sum_{(i,j)\in\mathbb{N}^2} (i+j)p^{i+j}.$$

Solution. 1. Rappelons que pour $p \in]-1,1[$ et c quelconque, on a[1] : $\displaystyle\sum_{n\geq 0} cp^n = \frac{c}{1-p}$.

S'agissant de réels positifs, on peut appliquer le théorème de Fubini-Tonelli :
$$S_1 = \sum_{i\geq 0}\left(\sum_{j\geq 0} p^{i+j}\right) = \sum_{i\geq 0}\frac{p^i}{1-p} = \frac{1}{1-p}\sum_{i\geq 0} p^i = \frac{1}{(1-p)^2}.$$
Cela entraîne que la famille est sommable.

Autre méthode. On groupe ensemble les points (i,j) de $\mathbb{N}^2$ selon la valeur de $i+j$. Plus précisément, comme dans la figure 1, on note, pour $n \in \mathbb{N}$,
$$D_n = \big\{(i,j)\in\mathbb{N}^2,\ i+j=n\big\} = \big\{(0,n),(1,n-1),\ldots,(n-1,1),(n,0)\big\}.$$
On constate que D_n contient $n+1$ points pour tout n. On calcule alors, grâce au théorème de sommation par paquets et au théorème de dérivation des séries entières :
$$\begin{aligned} S_1 &= \sum_{n\in\mathbb{N}}\sum_{(i,j)\in D_n} p^{i+j} = \sum_{n\in\mathbb{N}}\sum_{(i,j)\in D_n} p^n = \sum_{n\in\mathbb{N}} p^n|D_n| \\ &= \sum_{n\in\mathbb{N}}(n+1)p^n = \sum_{n\in\mathbb{N}^*}\frac{\mathrm{d}}{\mathrm{d}p}p^{n+1} = \frac{\mathrm{d}}{\mathrm{d}p}\sum_{n\in\mathbb{N}^*}p^{n+1} = \frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{p}{-p+1}\right) \\ &= \frac{1}{(1-p)^2}. \end{aligned}$$

1. Pour une somme infinie : « premier terme divisé par (1 − raison) ».

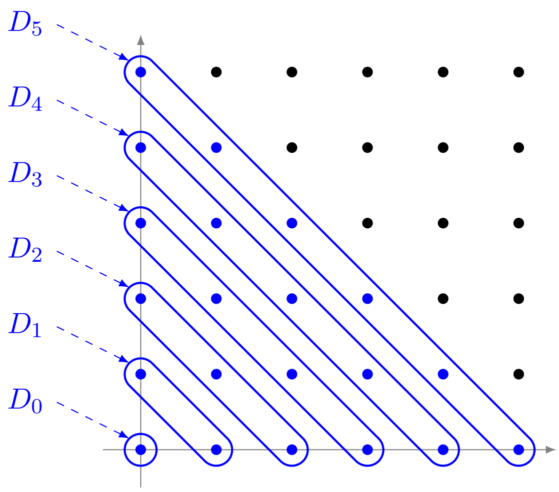

FIGURE 1 – Les diagonales D_n

2. On commence par calculer

$$S_2 = \sum_{n\geq 0} np^n = \sum_{n\geq 1}(n-1+1)p^n = \sum_{n\geq 1}(n-1)p^{n-1}\cdot p + \sum_{n\geq 1} p^n = pS_2 + \frac{p}{1-p},$$

d'où

$$S_2 = \frac{p}{(1-p)^2}.$$

Autre méthode, on utilise le calcul par sommation par paquet de la méthode 2 du 1.

$$S_1 = \sum_{n\in\mathbb{N}} (n+1)p^n$$

donc

$$S_2 = \sum_{n\geq 0} np^n = \sum_{n\geq 1} np^n = \sum_{m\geq 0}(m+1)p^{m+1} = pS_1.$$

3. Par suite :

$$S_3 = \sum_{i\geq 0}\sum_{j\geq 0}(ip^ip^j + p^ijp^j) = \sum_{i\geq 0}\frac{ip^i}{1-p} + \sum_{i\geq 0} p^iT = \frac{S_2}{1-p} + \frac{S_2}{1-p} = \frac{2p}{(1-p)^3}.$$

Autre méthode. Le calcul de S_3 est analogue mais il faut reconnaître (puis calculer) une dérivée seconde :

$$\begin{aligned} S_2 &= \sum_{n\geq 0}\sum_{(i,j)\in D_n}(i+j)p^{i+j} = \sum_{n\geq 0}\sum_{(i,j)\in D_n} np^n = \sum_{n\geq 0} np^n|D_n| \\ &= \sum_{n\geq 0} n(n+1)p^n = p\sum_{n\geq 0} n(n+1)p^{n-1} \\ &= p\sum_{n\geq 1}\frac{\mathrm{d}^2}{\mathrm{d}p^2}p^{n+1} = p\frac{\mathrm{d}^2}{\mathrm{d}p^2}\sum_{n\geq 1} p^{n+1} = p\frac{\mathrm{d}^2}{\mathrm{d}p^2}\left(\frac{p}{-p+1}\right) \\ &= p\frac{\mathrm{d}}{\mathrm{d}p}\frac{1}{(1-p)^2} = \frac{2p}{(1-p)^3}. \end{aligned}$$

Exercice 2. (3 points)

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Soit $p \in]0,1[$. On suppose qu'il existe une constante $c > 0$ telle que pour tout n,

$$P(X=n) = c\,p^n.$$

1. Calculer la valeur de c.

2. Calculer $P(X \geq 2 | X \geq 1)$.
3. Donner l'ensemble des valeurs prises par $Y = (-1)^X$ puis calculer la loi de Y.

Solution. 1. On doit avoir : $\sum_{n\geq 0} P(X = n) = 1$. On calcule donc :

$$\sum_{n\geq 0} P(X = n) = \sum_{n\geq 0} cp^n = \frac{c}{1-p}$$

d'où $c = 1 - p$.

2. On calcule brutalement :

$$P(X \geq 2) = \sum_{n\geq 2} P(X = n) = \sum_{n\geq 2} (1-p)p^n = \frac{(1-p)p^2}{1-p} = p^2.$$

On aurait pu plus habilement calculer :

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - (1-p) - (1-p)p = 1 - 1 + p - p + p^2 = p^2.$$

De même

$$P(X \geq 1) = 1 - P(X \leq 0) = 1 - (1-p) = p$$

donc en utilisant la définition d'une probabilité conditionnelle (et le fait que $X \geq 2$ implique $X \geq 1$) :

$$P(X \geq 2 | X \geq 1) = \frac{P(X \geq 2, X \geq 1)}{P(X \geq 1)} = \frac{P(X \geq 2)}{P(X \geq 1)} = \frac{p^2}{p} = p.$$

3. Comme X est à valeurs entières, $(-1)^X$ est à valeurs dans $\{-1, 1\}$. Plus précisément :

$$\begin{cases} Y = 1 \iff X \text{ est pair,} \\ Y = -1 \iff X \text{ est impair.} \end{cases}$$

Autrement dit, l'événement $Y = 1$ est la réunion disjointe des événements $X = 2k$ ($k \in \mathbb{N}$) et l'événement $Y = -1$ est la réunion disjointe des événements $X = 2k + 1$ ($k \in \mathbb{N}$).
On en déduit la loi de Y :

$$P(Y = 1) = \sum_{n \text{ pair}} P(X = n) = \sum_{k\geq 0} P(X = 2k) = \sum_{k\geq 0} (1-p)p^{2k} = \frac{1-p}{1-p^2} = \frac{1}{1+p} \ ;$$

$$P(Y = -1) = \sum_{n \text{ impair}} P(X = n) = \sum_{k\geq 0} P(X = 2k+1) = \sum_{k\geq 0} (1-p)p^{2k+1} = \frac{(1-p)p}{1-p^2} = \frac{p}{1+p}.$$

Ces deux calculs séparés donnent un moyen de contrôle : $P(Y = 1) + P(Y = -1) = 1$ (ouf !).

Exercice 3. (4 points) On considère la série statistique $(1, 3, 9, 19, 7, 3)$.

1. Calculer sa moyenne empirique et sa variance empirique.
2. Calculer sa médiane et ses quartiles.

Solution. 1. Moyenne empirique :

$$m_6 = \frac{1}{6}(1 + 3 + 9 + 19 + 7 + 3) = \frac{42}{6} = 7.$$

On calcule la variance empirique de deux façons pour faire bonne mesure :

$$\sigma_6^2 = \frac{1}{6}\sum_{i=1}^{6}(x_i - m_6)^2 = \frac{1}{6}\big((-6)^2 + (-4)^2 + 2^2 + 12^2 + 0^2 + (-4)^2\big) = \frac{216}{6} = 36,$$

ou bien

$$\sigma_6^2 = \frac{1}{6}\sum_{i=1}^{6} x_i^2 - m_6^2 = \frac{1}{6}(1^2 + 3^2 + 9^2 + 19^2 + 7^2 + 3^2) - 7^2 = \frac{510}{36} - 49 = 36.$$

2. Vu le petit effectif, il est commode de commencer par classer l'échantillon :

$$(y_1, y_2, y_3, y_4, y_5, y_6) = (1, 3, 3, 7, 9, 19).$$

La moitié de l'effectif est formée de trois valeurs. La troisième valeur est $y_3 = 3$, la quatrième est $y_4 = 7$, donc la médiane est

$$\frac{y_3 + y_4}{2} = \frac{3+7}{2} = 5.$$

Le quart de l'effectif représente $1,5$. Au moins un quart de l'effectif, c'est donc la deuxième valeur ($\lceil 1{,}5 \rceil = 2$), donc le premier quartile vaut

$$Q_1 = y_2 = 3.$$

Au moins trois quarts de l'effectif, c'est au moins 4,5 valeurs, c'est-à-dire 5 valeurs : le troisième quartile vaut

$$Q_3 = y_5 = 9.$$

On peut retrouver ces valeurs à partir de la représentation graphique de la fonction de répartition empirique (fig. 2).

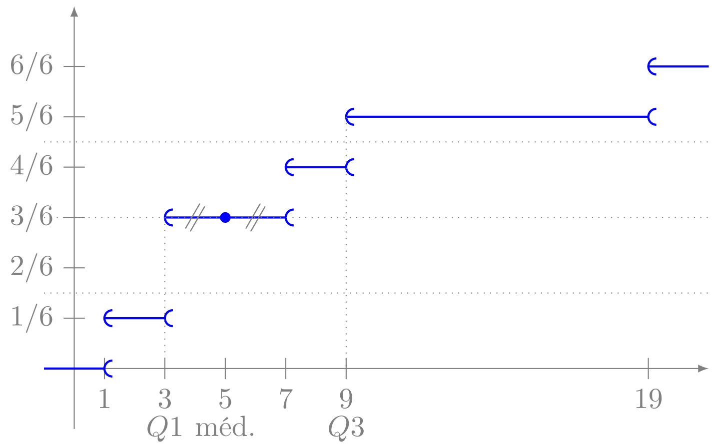

FIGURE 2 – Fonction de répartition empirique

Exercice 4. (4 points) Une compagnie d'assurance assure des voitures et répartit ses clients conducteurs en trois classes de risques R1, R2, R3 : les risques faibles, les risques moyens et les risques élevés. Les effectifs de ces trois classes représentent 10 % de la population totale pour la classe R1, 60 % pour R2 et 30 % pour R3. Les statistiques indiquent que les probabilités d'avoir un accrochage au cours de l'année pour une personne de l'une de ces trois classes sont respectivement de 0,05, de 0,15 et de 0,30.

1. Quelle est la probabilité qu'une personne choisie au hasard dans la population ait un accrochage dans l'année ?
2. Si M. Kantorovich, qui est assuré dans cette compagnie, a eu un accrochage cette année, quelle est la probabilité qu'il soit du type R1 ?

Solution. 1. On munit l'ensemble des clients de la probabilité uniforme. Pour $k \in \{1, 2, 3\}$, on note R_k la classe de risque Rk, c'est-à-dire l'ensemble (l'événement) des clients que la compagnie met dans la classe de risque Rk. On note A l'ensemble des clients qui vont avoir un accident dans l'année. Les hypothèses s'écrivent :

$$P(R_1) = 0{,}1\ ; \qquad P(R_2) = 0{,}6\ ; \qquad P(R_3) = 0{,}3\ ;$$
$$P(A \mid R_1) = 0{,}05\ ; \qquad P(A \mid R_2) = 0{,}15\ ; \qquad P(A \mid R_3) = 0{,}3.$$

D'après la formule des probabilités totales, on a :

$$P(A) = \sum_{k=1}^{3} P(R_k) \cdot P(A \mid R_k) = 0{,}1 \times 0{,}05 + 0{,}6 \times 0{,}15 + 0{,}3 \times 0{,}3 = 0{,}005 + 0{,}09 + 0{,}09 = 0{,}185.$$

2. La probabilité que M. Kantorovich soit du type R1 est $P(R_1 \mid A)$. Or on a :

$$P(R_1 \mid A) = \frac{P(R_1 \cap A)}{P(A)} = \frac{P(A \mid R_1)P(R_1)}{P(A)} = \frac{0{,}1 \times 0{,}05}{0{,}185} = \frac{0{,}005}{0{,}185} = \frac{1}{37} \simeq 0{,}027.$$

NB : Par contraste, on a :

$$P(R_2 \mid A) = P(R_3 \mid A) = \frac{0{,}09}{0{,}185} \simeq 0{,}4865.$$

Exercice 5. (2 points+ Bonus : 2 points)

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On suppose que X suit une de loi de Poisson de paramètre $\lambda > 0$, c'est-à-dire que pour tout $n \in \mathbb{N}$:

$$P(X = n) = \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}.$$

1. Calculer l'espérance $\mathbb{E}(X)$.
2. **(Bonus : 2 points)**Calculer $\mathbb{E}\Big(\dfrac{1}{1+X}\Big)$.

Solution. 1. On a :

$$\mathbb{E}(X) = \sum_{n\geq 0} n \cdot P(X = n) = \sum_{n\geq 1} n\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda} = \lambda\mathrm{e}^{-\lambda}\sum_{n\geq 1}\frac{\lambda^{n-1}}{(n-1)!} \overset{m=n-1}{=} \lambda\mathrm{e}^{-\lambda}\sum_{m\geq 0}\frac{\lambda^m}{m!} = \lambda\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = \lambda.$$

2. La fonction $f : x \mapsto 1/(1+x)$ est strictement décroissante sur $\mathbb{R}^+$ donc injective donc l'ensemble des valeurs prises par $1/(1+X)$ est $\big\{1/(1+n),\ n \in \mathbb{N}\big\}$ et, pour tout $n \in \mathbb{N}$,

$$P\Big(\frac{1}{1+X} = \frac{1}{1+n}\Big) = P(X = n).$$

Il vient alors :

$$\begin{aligned}\mathbb{E}\Big(\frac{1}{1+X}\Big) &= \sum_{n\geq 0}\frac{1}{1+n}P\Big(\frac{1}{1+X}\Big) = \sum_{n\geq 0}\frac{1}{n+1}\cdot\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda} \\ &= \frac{\mathrm{e}^{-\lambda}}{\lambda}\sum_{n\geq 0}\frac{\lambda^{n+1}}{(n+1)!} = \frac{\mathrm{e}^{-\lambda}}{\lambda}\sum_{m\geq 1}\frac{\lambda^m}{m!} = \frac{\mathrm{e}^{-\lambda}}{\lambda}\left(\sum_{m\geq 0}\frac{\lambda^m}{m!} - 1\right) \\ &= \frac{\mathrm{e}^{-\lambda}}{\lambda}(\mathrm{e}^{\lambda} - 1) = \frac{1 - \mathrm{e}^{-\lambda}}{\lambda}.\end{aligned}$$