Probabilités
Université Clermont Auvergne
IUT Clermont Auvergne
2A BUT Info, 2024-2025

A. El Ouni
A. Khaldi
C. Samir
A. Wohrer

TD4 – Lois de probabilités discrètes usuelles

Exercice 1.

On place un hamster dans une cage. Il se trouve face à 5 portillons dont un seul lui permet de sortir de la cage. A chaque essai infructueux, on le replace à l'endroit initial.

1. En supposant que le hamster ne soit pas doué d'apprentissage et qu'il choisisse donc de façon équiprobable entre les 5 solutions à chaque nouvel essai, déterminer la probabilité des évènements :
   a) le hamster sort au premier essai,
   b) le hamster sort au troisième essai,
   c) le hamster sort au septième essai.

Solution: Dans ce cas, il s'agit d'une **loi géométrique**, $X \sim Geo(1/5)$. (On répète la même expérience binaire jusqu'à obtenir le premier succès.) Ainsi :

$$P(X = 1) = \frac{1}{5}$$
$$P(X = 2) = \frac{4}{5}.\frac{1}{5}$$
$$P(X = 3) = \left(\frac{4}{5}\right)^2.\frac{1}{5}$$
$$P(X = 7) = \left(\frac{4}{5}\right)^6.\frac{1}{5}$$

avec toujours la même structure : on doit commencer par $k - 1$ échecs (d'où le $\left(\frac{4}{5}\right)^{k-1}$) et observer le premier succès au k-ème essai (d'où le $\frac{1}{5}$ final).

2. Le hamster mémorise maintenant les essais infructueux et choisit de façon équiprobable entre les portillons qu'il n'a pas encore essayés. On désigne par X la variable aléatoire égale au nombre d'essais effectués.
   a) Quelles valeurs peut prendre X ? Déterminer sa loi de probabilité.
   b) Déterminer l'espérance mathématique $E(X)$ et interpréter le résultat.
   c) Déterminer la variance $V(X)$.

Solution: Dans ce cas, il s'agit d'une **loi uniforme** $X \sim U(5)$. La façon la plus simple de le comprendre est la suivante : le hamster va essayer les 5 portes, dans un ordre aléatoire (et sans répétition, car il ne visite jamais deux fois la même porte). La "bonne" porte peut se trouver en n'importe quelle position (de 1 à 5) dans cet ordre de visite. Donc le hamster a autant de chances de sortir en 1,2,3,4 ou 5 coups : c'est bien une loi uniforme.

$$\text{Pour } k = 1 \dots 5, \quad P(X = k) = \frac{1}{5}.$$

L'espérance de cette loi est

$$E[X] = \frac{1 + 2 + 3 + 4 + 5}{5} = 3.$$

Sa variance est

$$V[X] = E[X^2] - (E[X])^2 = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2}{5} - 3^2 \quad = 11 - 9 = 2.$$

Exercice 2.
On suppose que la probabilité pour qu'une personne soit allergique à un médicament donné est égale à 10^{-3}. On s'intéresse à un échantillon de 1000 personnes. On appelle X la variable aléatoire dont la valeur est le nombre de personnes allergiques dans l'échantillon.

1. Déterminer, en la justifiant, la loi de probabilité de X, puis les valeurs $E(X)$ et $V(X)$.

Solution: On répète $n = 1000$ fois une expérience binaire ayant une probabilité $p = 10^{-3}$ de succès ('la personne est allergique'), et on compte le nombre de succès. Il s'agit donc d'une **loi binomiale** : $X \sim Bin(1000, 0.001)$.
L'espérance d'UNE expérience est $p = 10^{-3}$. Ici, on somme les résultat de $n = 1000$ expériences. L'espérance totale est donc $E[X] = np = 1000 \times 0.001 = 1$. (En moyenne, on détecte exactement 1 personne allergique sur notre échantillon.)
La variance d'UNE expérience est $p(1-p)$. Ici, on somme les résultat de $n = 1000$ expériences. De plus, les expériences sont *indépendantes*, donc leurs variances s'ajoutent tout simplement. La variance totale est donc $V[X] = np(1-p) = 1000 \times 0.001 \times 0.999 = 0.999$.
Ces deux formules (espérance et variance d'une loi binomiale) font partie de votre cours, elles peuvent aussi être utilisées sans justification.

2. En utilisant une approximation que l'on justifiera, calculer les probabilités des évènements suivants :
   a) Il y a exactement deux personnes allergiques dans l'échantillon.
   b) Il y a au moins deux personnes allergiques dans l'échantillon.
   c) Calculer et interpréter les valeurs de $E(X)$ et $V(X)$.

Solution: On est ici dans le cas particulier où
- le réservoir de personnes est très grand ($n = 1000$),
- la probabilité individuelle de succès est très faible ($p = 0.001$),
- l'espérance résultante est 'équilibrée' ($\lambda = np = 1$).

Dans ce cas (propriété du cours), la loi binomiale $Bin(n, p)$ est très bien approchée par une loi de Poisson $Pois(\lambda)$. Donc, ici $X \sim Pois(1)$ est une bonne approximation. Avec cette approximation :

a) $P(X=2) = e^{-2}\frac{1}{2!} \simeq 6.8\%$

b) $P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - e^{-2}\left(\frac{1}{0!} + \frac{1}{1!}\right) \simeq 73\%$.
(On rappelle que $0! = 1$.)

c) Dans la loi de Poisson, $E(X) = \lambda = 1$ et $V(X) = \lambda = 1$. On constate que ces valeurs sont très proches des valeurs exactes (pour la loi binomiale, question 1). Ça semble justifier que l'approximation de Poisson est valide.

Remarque : Expliquons intuitivement pourquoi la loi de Poisson est une bonne approximation de la loi binomiale dans ce régime.

Dans la loi binomiale, on a un « réservoir » d'objets (ici $n = 1000$) qui peuvent apparaître au hasard avec probabilité p. En moyenne, il apparaît donc un nombre d'objets $E[X] = np$ (ici, $E[X] = 1$).

Dans la loi de Poisson, on sait juste que des objets apparaissent au hasard, et qu'en moyenne il apparaît λ objets. La seule différence avec la loi binomiale est qu'ici, le réservoir d'objets est « infini ». Naturellement, lorsque la loi binomiale possède un très grand « réservoir » d'objets potentiels (ici, $n = 1000$) dont très peu vont apparaître (en moyenne, $\lambda = 1$), on peut sans problème considérer le réservoir comme infini – ce qui ramène à la loi de Poisson $Pois(1)$.

Exercice 3.
On s'intéresse à un bureau de poste, un mardi aux alentours de 11 heures du matin. À cette heure-ci, les arrivées des clients au guichet sont totalement aléatoires, avec en moyenne un client toutes les deux minutes. On s'intéresse à la variable aléatoire X : nombre de clients entrés dans l'agence entre 10h et 10h10.

1. Parmi les trois lois suivantes : *Poisson / uniforme / binômiale* ,
   laquelle faut-il utiliser pour modéliser la variable X ? Justifiez votre réponse.
2. Que vaut $\mathrm{E}(X)$? Déduisez-en la valeur du paramètre à utiliser dans la loi de X.
3. Calculez la probabilité qu'il rentre strictement moins de trois clients dans l'agence entre 10h et 10h10.

Exercice 4.
Pour chacune des variables aléatoires suivantes, dites quelle loi pourrait être utilisée pour la modéliser : *uniforme , géométrique , binômiale , Poisson , aucune des 4* ?

1. Le nombre d'insectes tapés pendant un trajet en voiture.

Solution: Poisson

2. Le nombre de minutes avant de taper le premier insecte pendant un trajet en voiture.

Solution: Géométrique. À chaque minute il y a une certaine chance p de taper un insecte, et la même expérience se répète de manière indépendante chaque minute, jusqu'au premier “succès”=insecte tapé.

3. La valeur du “numéro chance” au tirage du loto de samedi prochain.

Solution: Uniforme (entre 1 et 10)

4. Le plus grand nombre qui sortira dans le tirage du loto de samedi prochain.

Solution: Aucune des 4. $Y = \max(X_1, \ldots, X_5)$ où $\{X_1, \ldots, X_5\}$ est un tirage sans remise de 5 numéros entre 1 et 49. Le max de plusieurs variables ne suit généralement pas une loi usuelle (mais on peut quand même trouver des formules).

5. Le nombre de jours de couvaison d'un œuf de poule.

Solution: Aucune des 4. La durée de couvaison est relativement fixe, entre 19 et 21 jours, suivant les races et la taille de l'œuf.

6. Le nombre d'années avant la désintégration d'un atome de carbone 14.

Solution: Géométrique.

7. Le nombre de malades du Covid parmi 100 personnes choisies au hasard.

Solution: Binômiale. Il y a 100 personnes qui ont chacune une probabilité p d'être malades, *indépendamment les unes des autres*, et on compte le nombre total de malades.

8. Le nombre de malades du Covid parmi une classe de maternelle de 23 élèves.

Solution: Aucune des 4. Ce serait une binômiale si les 23 variables “X_i :=l'élève i est malade” étaient *indépendantes*. Mais en pratique ces variables ne sont pas indépendantes, car le présence d'un élève malade tend à augmenter les chances que d'autres le soient aussi. Et donc, la somme des X_i n'obéit pas à une loi binômiale.

9. Le nombre de malades du Covid détectés par un laboratoire d'analyses, une matinée au hasard.

Solution: Poisson. Les malades sont comme des “moucherons” qui arrivent au hasard dans le laboratoire, avec une certaine intensité moyenne (par exemple : 1 malade toutes les 20 min en moyenne). Il peut y avoir des matinées plus ou moins chargées juste du fait du hasard.

10. Dans un centre d'appels de démarchage téléphonique :
    (a) Le nombre d'appels à effectuer avant de trouver une personne qui reste en ligne.

    Solution: Géométrique.

    (b) Le nombre d'appels à effectuer avant de vendre 5 produits.

Solution: Aucune des 4. Si on note $X :=$ nombre d'appels à effectuer avant de vendre *un* produit, alors X peut sans doute se modéliser par une loi géométrique. Mais ici, il faut vendre 5 produits, donc la variable prend la forme $Y = X_1 + X_2 + X_3 + X_4 + X_5$ où chaque X_i suit une loi géométrique. Or la somme de plusieurs variables géométriques ne suit pas une loi géométrique. (Elle suit une distribution plus générale appelée *loi binômiale négative.*)

Exercice 5 (mathématiques de la loi de Poisson)**.**

1. Soit X une variable aléatoire distribuée suivant une loi de Poisson de paramètre λ. Rappelez la formule pour les nombres $p_k = P(X = k)$.

Solution:

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda}$$

pour $k = 0, 1, 2, \ldots$ (nombre entiers)

2. On admettra la formule suivante pour la fonction exponentielle :

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Déduisez-en que la loi de Poisson est bien *normalisée*, c'est-à-dire que la somme des probabilités p_k fait bien 1.

Solution:

$$\begin{aligned} p_0 + p_1 + p_2 + \ldots &= \sum_{k=0}^{\infty} p_k \\ &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} e^{\lambda} \\ &= 1 \end{aligned}$$

3. Rappelez la définition générale de $\mathrm{E}(X)$, et prouvez que pour une loi de Poisson, $\mathrm{E}(X) = \lambda$.

Solution:

$$\begin{aligned} \mathrm{E}(X) &= 0.p_0 + 1.p_1 + 2.p_2 + \ldots \\ &= \sum_{k=0}^{\infty} k.p_k \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{\infty} k.\frac{\lambda^k}{k!}e^{-\lambda} \\
&= e^{-\lambda}\sum_{k=1}^{\infty} k.\frac{\lambda^k}{k!} && \text{(car } k=0 \text{ ne contribue pas)} \\
&= e^{-\lambda}\lambda\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} && \text{(simplifie le } k/k! \text{ et sort un } \lambda) \\
&= e^{-\lambda}\lambda\left(1+\lambda+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\dots\right) \\
&= e^{-\lambda}\lambda e^{\lambda} \\
&= \lambda
\end{aligned}
$$

En conclusion, $\mathrm{E}(X) = \lambda$ comme annoncé dans le cours.

4. Prouvez que pour une loi de Poisson, $\mathrm{E}(X^2) = \lambda^2 + \lambda$. (*Indice :* utilisez le fait que $k^2 = k(k-1) + k$.) Déduisez-en la valeur de $\mathrm{Var}(X)$.

Solution:

$$
\begin{aligned}
\mathrm{E}(X^2) &= 0^2.p_0 + 1^2.p_1 + 2^2.p_2 + \dots \\
&= \sum_{k=0}^{\infty} k^2.p_k \\
&= \sum_{k=0}^{\infty} \left(k(k-1)+k\right).p_k && \text{(indice)} \\
&= \sum_{k=2}^{\infty} k(k-1)p_k \quad + \sum_{k=0}^{\infty} kp_k && (k=0 \text{ et } k=1 \text{ ne contribuent pas}) \\
&= e^{-\lambda}\sum_{k=2}^{\infty} k(k-1)\frac{\lambda^k}{k!} \quad + \lambda && \Big(\sum_{k=0}^{\infty} kp_k = \lambda, \text{ cf. question précédente}\Big) \\
&= e^{-\lambda}\lambda^2\sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} \quad + \lambda && \text{(simplifie le } k(k-1)/k! \text{ et sort un } \lambda^2) \\
&= e^{-\lambda}\lambda^2 e^{\lambda} \quad + \lambda \\
&= \lambda^2 + \lambda
\end{aligned}
$$

On en déduit que

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

En conclusion, $\mathrm{Var}(X) = \lambda$ comme annoncé dans le cours.

5. Soient deux variables aléatoires *indépendantes* $X \sim \mathrm{Pois}(\lambda)$ et $Y \sim \mathrm{Pois}(\theta)$. Soit la variable aléatoire $Z = X + Y$.

   (a) Montrez que Z suit encore une loi de Poisson, dont vous préciserez le paramètre.

Solution: Soit un entier k fixé. Voici toutes les configurations de X et Y qui mènent à une somme de $Z = k$:

- $X = 0$ et $Y = k$
- $X = 1$ et $Y = k - 1$
- $\ldots$
- $X = k$ et $Y = 0$

On peut donc écrire

$$P(Z = k) = P(X = 0 \text{ et } Y = k) + P(X = 1 \text{ et } Y = k-1) + \cdots + P(X = k \text{ et } Y = 0)$$

puis on développe ce calcul :

$$\begin{aligned}
P(Z = k) &= \sum_{i=0}^{k} P(X = i \text{ et } Y = k - i) \\
&= \sum_{i=0}^{k} P(X = i)P(Y = k - i) && \text{(car } X \text{ et } Y \text{ sont indépendantes)} \\
&= \sum_{i=0}^{k} e^{-\lambda}\frac{\lambda^i}{i!} e^{-\theta}\frac{\theta^{k-i}}{(k-i)!} \\
&= e^{-\lambda-\theta}\sum_{i=0}^{k} \frac{1}{k!}\binom{k}{i}\lambda^i\theta^{k-i} && \text{car } \binom{k}{i} = k!/i!(k-i)! \\
&= e^{-(\lambda+\theta)}\frac{(\lambda+\theta)^k}{k!} && \text{(formule du binôme : } (x+y)^k = \sum_{i=0}^{k}\binom{k}{i}x^i y^{k-i}\text{)}
\end{aligned}$$

Donc, Z suit une loi de Poisson de paramètre $\lambda + \theta$.

(b) Fournissez une interprétation intuitive de ce résultat.

Solution: Pensons à la loi de Poisson comme la loi du “nombre de moucherons sur un pare-brise”. Si X représente le nombre de moucherons collecté sur un premier trajet, et Y le nombre de moucherons collecté sur un second trajet, alors Z représente le nombre de moucherons collecté sur le trajet total. Les évènements (taper un moucheron) étant tous indépendants et survenant au hasard pendant le trajet total, il est clair que Z doit toujours suivre une loi de Poisson.

De plus son paramètre correspond à son espérance (nombre moyen de moucherons attendu), qui est forcément la somme des espérances sur les 2 parties du trajet total, donc $\lambda + \theta$.