Université de Paris-Saclay
Faculté des sciences d'Orsay
L3 E3A

3 heures
2023-2024
Probabilités

Partiel

On rappelle les égalités suivantes :

$$\sum_{k=0}^{+\infty} \frac{x^k}{k!} = e^x, \qquad \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x} \text{ si } |x| < 1, \qquad \sum_{k=1}^{+\infty} k x^{k-1} = \frac{1}{(1-x)^2} \text{ si } |x| < 1,$$

$$\sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6}, \qquad \sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6}, \qquad \sum_{k=1}^{+\infty} \frac{x^{-k}}{k} = -\ln\left(1 - \frac{1}{x}\right) \text{ si } |x| > 1.$$

Exercice 1 (10 minutes)

Les 275 passagers d'un vol long-courrier s'apprêtent à embarquer dans un avion possédant 55 sièges en classe confort et 220 sièges en classe économique. Les voyageurs partent soit pour un séjour court, soit pour un séjour long. Sachant que l'on voyage en classe économique, 35 % des passagers partent pour un séjour long alors que, sachant que l'on voyage en classe confort, 70 % des passagers ont opté pour un séjour long. On choisit au hasard un passager du vol.

1. Calculer la probabilité pour que le passager choisit voyage en classe économique.
2. Sachant que le passager voyage en classe confort, calculer la probabilité pour qu'il effectue un séjour court.
3. Calculer la probabilité que le passager choisi parte en classe économique et pour un séjour long.
4. Sachant que le passager part pour un long séjour, quelle est la probabilité que ce passager voyage en classe économique ?

Exercice 2 (10 minutes)

On suppose qu'à la naissance, la probabilité qu'un nouveau-né soit un garçon est égale à 1/2. On suppose que tous les couples ont des enfants jusqu'à obtenir un garçon. On souhaite évaluer la proportion de garçons dans une génération de cette population. On note X la variable aléatoire correspondant au nombre d'enfants d'un couple et $R = \frac{1}{X}$ la variable aléatoire correspondant à la proportion de garçons d'un couple. On supposera $R \in L^1$.

1. Donner la distribution de probabilité de X et donner son ou ses paramètres.
2. Calculer, en moyenne, la proportion de garçon. Qu'en pensez-vous ?

Exercice 3 (30 minutes)

Pour améliorer la sûreté de fonctionnement d'un serveur informatique, on envisage d'introduire de la redondance, c'est-à-dire d'avoir plusieurs exemplaires des composants importants. On peut réaliser les opérations suivantes :

Cas 1 : Soit on utilise trois alimentations de 300 Watts chacune : le serveur peut continuer à fonctionner avec une alimentation en panne car il consomme au maximum 500 Watts.

Cas 2 : Soit on place les quatre disques durs en configuration [1] dite « RAID 5 » et le serveur peut continuer à fonctionner avec, au plus, un disque dur en panne.

On suppose que la probabilité de panne d'une seule alimentation est p et que celle d'une panne d'un seul disque dur est q. On suppose en outre que tous les composants sont indépendants.

1. Une configuration RAID (Redundant Array of Independent Disks) permet d'utiliser plusieurs disques durs pour augmenter : la sécurité en configuration RAID 1 (évite la perte des données en cas de crash d'un disque dur), les performances en configuration RAID 0 (en écrivant une partie des données sur un disque dur et les autres sur d'autres disques durs), ou les deux (sécurité et performances) en configuration RAID 5 qui est un mélange de RAID 0 et RAID 1. Le coût au mégaoctet des disques durs ayant été divisé par 1 300 000 en 29 ans, aujourd'hui le RAID est choisi pour d'autres raisons que le coût de l'espace de stockage. Plus d'information ici https ://fr.wikipedia.org/wiki/RAID_(informatique)

1. Soit un serveur avec alimentations redondantes (cas 1) : on pose X la variable aléatoire du nombre total d'alimentation en panne. Soit les variables aléatoires A_i avec $i = 1, 2$ ou 3 telles que $A_i = 1$ si l'alimentation i est en panne (et $A_i = 0$ sinon). En reliant la variable aléatoire X aux variables aléatoires A_i calculer la probabilité de panne du serveur, notée α, en supposant qu'aucun autre composant que les alimentations ne peut tomber en panne.

2. Soit un serveur avec disques durs en configuration RAID 5 (cas 2). On pose Y la variable aléatoire représentant le nombre de disque dur en panne. Calculer la probabilité de panne du serveur, notée β, en supposant qu'aucun autre composant que les disques dur ne peut tomber en panne.

3. Si $p = q$, en étudiant le signe de la fonction $\alpha - \beta$, quelle solution de redondance est la plus intéressante ?

Exercice 4 (40 minutes)

Dans cet exercice [2], on cherche quelle est la probabilité, pour que, parmi N personnes présentes dans une assemblées, deux au moins aient la même date d'anniversaire (**c'est-à-dire ayant le même jour et même mois de naissance quelque soit l'année**). (Note : on fera l'hypothèse simplificatrice qu'une année comporte toujours 365 jours et que les dates d'anniversaires sont équiprobables). On introduit la variable aléatoire X_i pour $i = \{1, \ldots, N\}$ qui correspond au numéro du jours de naissance de la personne i (par exemple une personne née le 12 juillet est née le 193ème jours de l'année).

1. Donner la distribution de probabilité de X_i.

2. En justifiant soigneusement, pour $i \neq i'$, donner la probabilité de l'évènement $X_i = X_{i'}$? En déduire la probabilité de l'évènement $X_i \neq X_{i'}$.

3. Ecrire la relation reliant $P\left(X_1 \neq X_N \cap X_2 \neq X_N \cap \cdots \cap X_{N-1} \neq X_N \middle| X_1 \neq X_2 \neq \cdots \neq X_{N-1}\right)$ à $P\left(X_1 \neq X_2 \neq \cdots \neq X_{N-1}\right)$ et $P\left(X_1 \neq X_2 \neq \cdots \neq X_N\right)$

4. On suppose $N = 3$. Donner, par dénombrement, la probabilité $P\left(X_1 \neq X_3 \cap X_2 \neq X_3 \middle| X_1 \neq X_2\right)$. En écrivant l'événements $X_1 \neq X_2 \neq X_3$ à l'aide des événements $X_1 \neq X_2$, $X_2 \neq X_3$ et $X_1 \neq X_3$, calculer la probabilité $P\left(X_1 \neq X_2 \neq X_3\right)$. Les événements $X_1 \neq X_2$, $X_2 \neq X_3$ et $X_1 \neq X_3$ sont ils 2-à-2 indépendants ? Sont-ils conjointement indépendants ? En déduire la probabilité que deux personnes au moins aient la même date d'anniversaire.

5. De la même manière, étudier le cas $N = 4$. En déduire une formule pour le cas général, c'est-à-dire quelque soit N.

Exercice 5 (40 minutes)

Tous les jours, après son travail, Rémi va boire un verre au bar puis rentre chez lui. Un jour sur deux, en probabilité, il est ivre lorsqu'il sort du bar. Un jour sur dix, en probabilité, la police effectue un contrôle d'alcoolémie au près de lui lorsqu'il rentre chez lui. On suppose que ces deux évènements (ivresse et contrôle de la part de la police) sont indépendants, et que leur survenue un jour donné ne dépend pas de ce qui se passe les autres jours. Si le test d'alcoolémie se révèle positif, Rémi perd un point sur son permis. On supposera que le test ne se trompe jamais. On note X_i la variable aléatoire qui correspond au nombre de points perdus le jour i.

2. Le paradoxe des anniversaires résulte de l'estimation probabiliste du nombre de personnes que l'on doit réunir pour avoir au moins une chance sur deux que deux personnes de ce groupe aient leur anniversaire le même jour. Ce nombre est 23, ce qui ne correspond pas à l'intuition commune. À partir d'un groupe de 57 personnes, la probabilité est supérieure à 99 %. Il s'agit d'un paradoxe non pas dans le sens de contradiction logique, mais dans le sens où c'est une vérité mathématique qui contredit l'intuition : la plupart des gens estiment que cette probabilité est très inférieure à 50 %. Cette étude est due à Richard von Mises. Pour trouver d'autres paradoxes de ce genre, on peut calculer la probabilité en remplaçant 365 par n'importe quel nombre K, et pour n'importe quelle taille de groupe N. Il existe une application concrète de ce paradoxe. Il existe des méthodes (comme le ≪ MD5 checksum ≫ ou le ≪ SHA 256≫) qui permettent d'associer un identifiant numérique à un fichier. Si vous voulez que ce nombre identifie de manière quasi-unique ce fichier parmi tous vos fichiers, vous devez vous assurer que la probabilité que deux fichiers différents aient le même MD5 (ou autre) soit suffisamment faible. Cela fixe la taille K de l'ensemble des identifiants. Dès que N devient grand, il faut aller chercher des K astronomiques pour avoir des probas de conflit disons de 1% ! Pour le MD5, il y 2ˆ128 combinaisons ! (source : https ://scienceetonnante.com/2012/05/28/le-paradoxe-des-anniversaires/)

1. Question préliminaire : soit $x \in]-1, 1[$ et $r \in \mathbb{N}^*$. Calculer pour $r = 1, 2, 3$ la dérivée r-ième $\frac{d^r}{dx^r}\left(\frac{1}{1-x}\right)$ et en déduire que

$$\sum_{n=r}^{+\infty} n\,(n-1)\,(n-2)\cdots(n-r+1)\,x^{n-r} = \frac{r!}{(1-x)^{r+1}}$$

.

2. Donner la distribution des X_i.

3. Pour tout $n \geq 1$, on note $S_n = \sum_{i=1}^{n} X_i$. Que représente S_n ? Donner sa distribution de probabilité, son espérance, sa variance.

4. En tant que jeune conducteur, Rémi ne dispose que de 6 points sur son permis. On note T le nombre de jours où Rémi a gardé son permis (donc au minimum $T = 6$). Quelles sont les conditions sur S_{n-1} et X_n pour que $T = n$? Quelle est la distribution de T ?

5. Est-ce que $T \in L^1$? Si oui, alors calculer $\mathbb{E}\,(T)$.

Exercice 6 (10 minutes)

Une population comporte, en moyenne, une personne mesurant plus de $1m90$ sur 80 personnes. Sur 100 personnes, en utilisant la loi de Poisson, calculer la probabilité qu'il y ait au moins deux personnes mesurant plus de $1.90m$.

Exercice 7 (40 minutes)

Soit X une variable aléatoire discrète telle que sa fonction de répartition est donnée par $F_X\,(n) = P\,(X \leq n) = \sum_{k=1}^{n} \frac{\alpha}{k(k+1)(k+2)}$ avec $n \in \mathbb{N}^*$ et α une constante telle que $\underset{n\to+\infty}{lim} F_X\,(n) = 1$.

1. Donner la distribution de probabilité de X.

2. Décomposer en éléments simples la fraction rationnelle définie par $\frac{1}{x(x+1)(x+2)}$ et, en justifiant soigneusement, montrer que $\sum_{k=1}^{n} \frac{1}{k(k+1)(k+2)} = \lambda - \frac{1}{2(n+1)} + \frac{1}{2(n+2)}$ où λ est une constante positive que l'on précisera.

3. En déduire la valeur de α.

Un critère de convergence des séries est qu'une série $\sum_{k=1}^{+\infty} u_k$ converge si et seulement si $\forall \epsilon > 0$, il existe un rang à partir duquel $|S_{2n} - S_n| < \epsilon$ où $S_n = \sum_{k=1}^{n} u_k$. Sinon la série diverge.

4. Soit la série, dite série harmonique, $\sum_{k=1}^{+\infty} \frac{1}{k}$. Montrer que $\forall i \in \{1, 2, \ldots, n\}$, $\frac{1}{n+i} \geq \frac{1}{2n}$ et en déduire que l'on a toujours $|S_{2n} - S_n| > \frac{1}{2}$. Conclure quant à la convergence ou la divergence de la série harmonique.

5. Sachant que $\sum\limits_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$, est-ce que $X \in L^1$? Si oui, calculer $\mathbb{E}\,(X)$.

6. D'après la question 4, est-ce que $X \in L^2$? Si oui, calculer $Var\,(X)$.