DS N°5

Le but de ce problème est la modélisation du passage de bus à un arrêt.

On notera $\mathbb{N}$ l'ensemble des entiers naturels, et $\mathbb{N}^*$ l'ensemble des entiers naturels non nuls. p désigne dans tout le problème un réel tel que $0 < p < 1$, et l'on pose $q = 1 - p$.
On dit qu'une variable aléatoire X suit la loi géométrique sur $\mathbb{N}$ de paramètre p lorsque son univers – image est égal à $\mathbb{N}$, et que l'on a pour tout $n \in \mathbb{N}$: $\mathbb{P}(X = n) = p\, q^n$.

On pourra alors noter : $X \hookrightarrow \mathcal{G}_{\mathbb{N}}(p)$.

On prendra garde à ne pas confondre la loi $\mathcal{G}_{\mathbb{N}}(p)$ avec la loi géométrique usuelle $\mathcal{G}(p)$ (d'univers – image égal à $\mathbb{N}^*$).

A. Préliminaires

Les résultats de cette partie seront utiles dans toute la suite. Ils pourront être admis si nécessaire.

A.1.a. Démontrer la formule de Pascal :

$$\forall (a, b) \in \mathbb{N}^2,\ a \leq b,\ \binom{b+1}{a+1} = \binom{b}{a+1} + \binom{b}{a}.$$

Soit x un réel tel que $0 < x < 1$. Pour tout entier $n \geq 0$, on considère la série $\sum_{k \geq 0} \binom{k+n}{n} x^k$.

b. Montrer que la série $\sum_{k \geq 0} \binom{k+n}{n} x^k$ converge.

On note alors $S(n, x) = \sum_{k=0}^{+\infty} \binom{k+n}{n} x^k$.

A.2. a. Calculer $S(0, x)$.

b. Montrer que pour tout entier naturel n :

$$(1 - x)\, S(n + 1, x) = S(n, x).$$

c. En déduire que, pour tout entier naturel n, $S(n,x) = \dfrac{1}{(1-x)^{n+1}}$.

En déduire la valeur de $\displaystyle\sum_{k=0}^{+\infty} \binom{k+n-1}{n-1} x^k$, pour tout $n \in \mathbb{N}^*$, et de $\displaystyle\sum_{n=1}^{+\infty} \binom{k+n-1}{n-1} x^n$, pour tout $k \in \mathbb{N}$.

Les résultats de cette question sont essentiels dans tout le problème.

A.3. Soit un entier naturel n**.** Dans cette question, on cherche à établir la formule suivante :

$$\forall (m,\ell) \in \mathbb{N}^2 \;,\; \mathcal{H}(m,\ell) : \sum_{k=0}^{\ell} \binom{n+k}{n}\binom{m+\ell-k}{m} = \binom{m+n+\ell+1}{m+n+1}.$$

a. Établir $\mathcal{H}(m,0)$, pour tout $m \in \mathbb{N}$.

b. Montrer par récurrence que, pour tout $\ell \in \mathbb{N}$, $\mathcal{H}(0,\ell)$ est vérifiée :

$$\sum_{k=0}^{\ell} \binom{n+k}{n} = \binom{n+\ell+1}{n+1}.$$

c. Soit $\ell \in \mathbb{N}$. On suppose que $\mathcal{H}(m,\ell)$ est vérifiée pour tout $m \in \mathbb{N}$.

On fixe alors un entier $m \in \mathbb{N}$, et l'on suppose également $\mathcal{H}(m,\ell+1)$ vérifiée.

Montrer qu'alors, $\mathcal{H}(m+1,\ell+1)$ est elle aussi vérifiée.

On pourra remarquer que :

$$\forall k \in \llbracket 0, \ell+1 \rrbracket \;,\; \binom{m+\ell-k+2}{m+1} = \binom{m+\ell-k+1}{m+1} + \binom{m+\ell-k+1}{m}.$$

d. Conclure.

B. Lois binomiales négatives

B.1. Soit U une variable aléatoire réelle, et soit n un entier naturel strictement positif (on rappelle que $p \in]0,1[$, et que $q = 1-p$). On dit que U **suit la loi binomiale négative de paramètres** n **et** p si et seulement si U est une variable aléatoire discrète à valeurs dans $\mathbb{N}$ dont la loi de probabilité est définie par :

$$\forall k \in \mathbb{N},\; \mathbb{P}(U=k) = \binom{n+k-1}{n-1} p^n q^k.$$

On note alors $U \hookrightarrow \mathcal{B}^-(n,p)$.

a. Vérifier que l'on a bien défini ainsi une loi de probabilité.

Soit alors U une variable aléatoire de loi $\mathcal{B}^-(n, p)$.

b. Montrer que $\mathbb{E}(U+n)$ et $\mathbb{E}((U+n)(U+n+1))$ existent, et calculer ces deux espérances.

c. En déduire que U possède une espérance et une variance, et déterminer et $\mathbb{E}(U)$ et $\mathbb{V}(U)$.

B.2. Soient n et m deux entiers naturels non nuls.

On considère deux variables aléatoires réelles X_n et X_m indépendantes, de lois respectives $\mathcal{B}^-(n, p)$ et $\mathcal{B}^-(m, p)$.

a. Montrer que $X_n + X_m \hookrightarrow \mathcal{B}^-(n+m, p)$. *Le résultat de la question* **A.3.** *pourra être utile.*

b. On se donne n variables aléatoires mutuellement indépendantes $U_1, \ldots, U_n$ de même loi géométrique sur $\mathbb{N}$ de paramètre p.

i – Vérifier que la loi géométrique sur $\mathbb{N}$ de paramètre p est une loi binomiale négative dont on précisera les paramètres.

ii – En déduire que $\sum_{j=1}^{n} U_j$ est une variable aléatoire de loi $\mathcal{B}^-(n, p)$.

Le résultat de cette question sera largement utilisé dans la suite du problème.

iii – Retrouver alors les résultats de la question **B.1.c.** .

B.3. On se donne maintenant une suite $(U_j)_{j \in \mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes, toutes de loi géométrique sur $\mathbb{N}$ de paramètre p.

On considère également une variable aléatoire N, indépendante des variables aléatoires U_j, et suivant une loi géométrique sur $\mathbb{N}^*$ de paramètre α, avec $0 < \alpha < 1$. On pose $\beta = 1 - \alpha$.

On considère la variable aléatoire $V = \sum_{j=1}^{N} U_j$: V est donc une somme d'un nombre *aléatoire* de variables aléatoires.

a. Soient $k \in \mathbb{N}$ et $n \in \mathbb{N}^*$. Déterminer la probabilité conditionnelle $\mathbb{P}(V = k \mid N = n)$.

b. Pour $k \in \mathbb{N}$, exprimer $\mathbb{P}(V = k)$ sous forme de somme.

c. Calculer cette somme, et vérifier que V suit la loi géométrique sur $\mathbb{N}$ de paramètre $\dfrac{\alpha p}{q + \alpha p}$.

C. Modélisation du passage des bus

On s'intéresse aux instants de passages successifs des bus à un arrêt donné. Dans cette modélisation, le passage d'un bus à un arrêt est considéré comme instantané (le bus arrive et repart au même instant).

Le service commence à l'instant $T_0 = 0$. Le premier bus du matin passe à l'instant $T_1 \in \mathbb{N}$. On pose $U_1 = T_1 - T_0$, qui représente donc le temps entre l'ouverture du service et le passage du premier bus de la journée. Le temps écoulé entre les passages du premier et du second bus de la journée est modélisé par une variable aléatoire U_2 à valeurs dans $\mathbb{N}$. T_2 désigne l'instant auquel ce second bus arrive ; on a donc $U_2 = T_2 - T_1$. Le bus suivant passe ensuite à l'instant $T_3 \in \mathbb{N}$ au bout d'un temps U_3, et ainsi de suite.

Pour $n \in \mathbb{N}^*$, T_n désigne l'instant où le n – ième bus arrive à l'arrêt, et U_{n+1} le temps écoulé entre les passage du n – ième et du $(n+1)$ – ième bus de la journée.

On a donc, pour tout $n \in \mathbb{N}$, $U_{n+1} = T_{n+1} - T_n$.

On suppose que les variables aléatoires U_n, $n \geq 1$, sont mutuellement indépendantes. On suppose de plus qu'elles suivent la même loi géométrique sur $\mathbb{N}$ de paramètre p.

On notera que ceci implique que plusieurs bus peuvent arriver au même instant ; dans ce cas, on considère qu'ils arrivent l'un derrière l'autre, parler du premier, du deuxième,…, du n – ième bus a donc toujours un sens.

On rappelle que l'on a posé $T_0 = 0$. On a donc pour tout $n \in \mathbb{N}^*$, $T_n = \sum_{j=1}^{n} U_j$.

On définit enfin la fonction de comptage N de la façon suivante : pour tout $m \in \mathbb{N}$, N_m est le nombre de bus qui sont passés à l'arrêt dans l'intervalle de temps $0, m$.

C.1. On cherche tout d'abord à se faire une idée des propriétés élémentaires du modèle.

Soit $m \in \mathbb{N}$. Montrer que

i – $N_m = 0$ si et seulement si $m < T_1$.

ii – Pour tout $n \in \mathbb{N}^*$, $N_m = n$ si et seulement si $T_n \leq m < T_{n+1}$.

C.2. Soit $m \in \mathbb{N}$ fixé et soit $n \in \mathbb{N}^*$. On admettra que N_m est une variable aléatoire réelle discrète, et on s'intéresse ici à sa loi.

a. Donner la loi de T_n.

b. Montrer que $N_m \geq n$ si et seulement si $T_n \leq m$. En déduire une expression de $\mathbb{P}\left(N_m \geq n\right)$ utilisant une somme.

c. En déduire que N_m suit la loi binomiale négative de paramètres $m+1$ et q.

C.3. On suppose que les bus de la ligne passant à l'arrêt considéré peuvent avoir deux terminus A et B différents. Pour $m \in \mathbb{N}$, on note A_m (respectivement B_m) le nombre de bus allant au terminus A (respectivement au terminus B) qui sont passés dans l'intervalle de temps $\llbracket 0, m \rrbracket$. N_m désigne comme ci – dessus le nombre total de bus, tous terminus confondus. On a donc $N_m = A_m + B_m$.

Lorsqu'un bus se présente à l'arrêt, on suppose qu'il a pour terminus A avec une probabilité $\alpha \in \left]0, 1\right[$, et B avec la probabilité $\beta = 1 - \alpha$, indépendamment des autres bus.

a. Soit $n \in \mathbb{N}$. Déterminer la probabilité conditionnelle $\mathbb{P}\left(A_m = k \mid N_m = n\right)$ pour $m \in \mathbb{N}$ et k entier, $0 \leq k \leq n$.

b. En déduire que $A_m \hookrightarrow \mathcal{B}^-\left(m+1, \dfrac{q}{\alpha p + q}\right)$. Quelle est la loi de B_m ?

c. Montrer que $\mathbb{P}\left(A_m = k\right) = \dbinom{k+m}{k}\left(\dfrac{q}{q + \alpha p}\right)^{m+1}\left(\dfrac{\alpha p}{q + \alpha p}\right)^{k}$.

Préciser $\mathbb{P}\left(B_m = 0\right)$, et déterminer $\mathbb{P}\left(A_m = k, B_m = 0\right)$.

Les variables aléatoires A_m et B_m sont – elles indépendantes ? *On pourra faire tendre k vers $+\infty$.*

D. Absence de mémoire

Dans cette partie, on fixe un instant $a \in \mathbb{N}^*$. Pour des raisons d'étude statistique, un employé de la compagnie de bus se poste chaque jour à l'instant a à l'arrêt étudié et note les heures de passage des bus à partir de cet instant (on convient que l'employé voit tous les bus passant à l'instant a, s'il y en a). On appelle U'_1 le temps que cet employé attend avant de voir passer un premier bus, puis U'_2, U'_3, etc. les intervalles de temps entre chacun des bus suivants. On pose $T'_1 = U'_1$, et plus généralement, pour tout $n \in \mathbb{N}^*$, $T'_n = \sum_{j=1}^{n} U'_j$. Enfin, on définit une nouvelle fonction de comptage M_m, de sorte que, pour tout $m \in \mathbb{N}$, M_m représente le nombre de bus que l'employé a vu passer à l'arrêt au bout d'un temps m à partir de son arrivée à l'instant a, c'est – à – dire le nombre de bus s'étant présentés dans l'intervalle $\llbracket a, a+m \rrbracket$.

Les variables aléatoires N_m, T_n, U_n intervenant dans la suite de cette partie sont définies dans la partie **C** .

D.1. On cherche à déterminer la loi de T'_1.

a. Soit $m \in \mathbb{N}$.

i – Pour $(u, v) \in \mathbb{N}^2$ fixés, $u \leq v$, que représente la quantité $N_v - N_u$?

En déduire que $M_m = N_{a+m} - N_{a-1}$.

ii – Montrer que l'événement $(T'_1 = m)$ est la réunion des événements

$\left[\left(T_n < a\right) \cap \left(T_{n+1} = a + m\right)\right]$ lorsque n parcourt $\mathbb{N}$.

iii – Soit $n \in \mathbb{N}$. Justifier que T_n et U_{n+1} sont indépendantes ; déterminer la loi du couple $\left(T_n\,,\ U_{n+1}\right)$. *On traitera séparément le cas* $n = 0$.

b. Soit $m \in \mathbb{N}$. Montrer que $P\left(T'_1 = m\right) = p\,q^{a+m} + \sum_{n=1}^{+\infty} \sum_{k=0}^{a-1} \binom{k+n-1}{n-1} p^n q^k\, p\, q^{a+m-k}$.

c. En déduire que T'_1 suit la loi géométrique sur $\mathbb{N}$ de paramètre p.

d. On admettra que ce résultat entraîne que pour tout n, T'_n a la même loi que T_n, et que M_m a la même loi que N_m pour $m \in \mathbb{N}$. Cette dernière propriété est appelée **absence de mémoire**. Pourquoi ?

D.2. Paradoxe du bus

On s'intéresse ici à la quantité Δ égale au temps écoulé entre le passage du dernier bus arrivant à l'arrêt dans l'intervalle $0, a$ et l'instant a d'arrivée de l'employé, si au moins un bus est passé dans cet intervalle, égale à a sinon. On admettra que Δ est une variable aléatoire réelle discrète.

a. Pour tout $m \in \mathbb{N}$, que représente la quantité T_{N_m} ? Montrer que $\Delta = a - T_{N_a}$.

b. A quoi correspond l'événement $(\Delta = a)$? Calculer $\mathbb{P}(\Delta = a)$.

c. Pour entier j tel que $0 \leq j < a$, montrer que

$$\mathbb{P}(\Delta = j) = \sum_{n=1}^{+\infty} \mathbb{P}\left(\left(T_n = a - j\right) \cap \left(U_{n+1} > j\right)\right),$$

et calculer cette somme.

d. Montrer que, pour tout $n \in \mathbb{N}^*$, $\mathbb{E}(\Delta) + \mathbb{E}(U'_1) > \mathbb{E}\left(U_n\right)$.

En quoi ce résultat est – il paradoxal ? (*Si vous ne le trouvez pas paradoxal, expliquez aussi pourquoi*).

Ce résultat est couramment appelé paradoxe du bus.

E. Nombre d'usagers utilisant un bus donné

Dans cette partie, on considère un entier $k \in \mathbb{N}$. On suppose que le nombre de personnes présentes à l'arrêt à l'instant k est modélisé par une variable aléatoire Z suivant une loi de Poisson de paramètre $\lambda > 0$.
Si au moins un bus passe à l'arrêt à l'instant k, chacune des personnes présentes choisit au hasard l'un des bus passant à cet instant ; les choix des usagers sont indépendants.

On note R_k la variable aléatoire représentant le nombre de bus passant à l'arrêt à l'instant k.

E.1. Dans cette question, et dans cette question seulement on considère un entier $r \geq 1$, et l'on suppose que le nombre de bus passant à l'arrêt à l'instant k est égal à r (autrement dit, on suppose l'évènement $\left(R_k = r\right)$ réalisé).

On note alors W_r la variable aléatoire représentant le nombre de personnes montant dans le premier bus.

a. Pour $n \in \mathbb{N}$ et pour $w \in \mathbb{N}$ tel que $0 \leq w \leq n$, déterminer $\mathbb{P}\left(W_r = w \mid Z = n\right)$.

b. Pour $w \in \mathbb{N}$, justifier l'égalité : $\mathbb{P}\left(W_r = w\right) = \sum_{n=w}^{+\infty} \mathbb{P}\left(W_r = w \mid Z = n\right)\mathbb{P}\left(Z = n\right)$.

c. En déduire que $W_r \hookrightarrow \mathcal{P}\left(\dfrac{\lambda}{r}\right)$.

Dans tout ce qui suit, on ne suppose plus le nombre de bus passant à l'instant k égal à r ; on rappelle que ce nombre de bus est modélisé par la variable aléatoire R_k.

E.2. En utilisant les questions **C.2.c.** et **D.1.d.**, montrer que R_k suit la loi $\mathcal{G}_{\mathbb{N}}(q)$.

E.3. ***Formule de l'espérance totale***

Soit (X, Y) un couple de variables aléatoires discrètes, toutes deux à valeurs dans $\mathbb{N}$. On suppose que Y admet une espérance, et que pour tout $r \in \mathbb{N}$, $\mathbb{P}(X = r) \neq 0$. Soit $r \in \mathbb{N}$.

a. Montrer qu'il existe une variable aléatoire Y_r à valeurs dans $\mathbb{N}$ telle que :

$$\forall\, n \in \mathbb{N},\ \mathbb{P}\left(Y_r = n\right) = \mathbb{P}\left(Y = n \mid X = r\right).$$

b. Montrer que la variable aléatoire Y_r possède une espérance.

On note $\mathbb{E}_{(X=r)}(Y)$ l'espérance de Y_r.

c. Montrer que $\mathbb{E}(Y) = \sum_{r=0}^{+\infty} \mathbb{E}_{(X=r)}(Y) . \mathbb{P}(X = r)$.

d. Montrer que réciproquement, si pour tout $r \in \mathbb{N}$, Y_r possède une espérance, et si la série

$\sum \mathbb{E}_{(X=r)}(Y).\mathbb{P}(X=r)$ converge, alors Y possède une espérance, et

$\mathbb{E}(Y) = \sum_{r=0}^{+\infty} \mathbb{E}_{(X=r)}(Y).\mathbb{P}(X=r)$ *cette formule est appelée formule de l'espérance totale.*

On note W la variable aléatoire représentant le nombre de personnes montant dans le premier bus passant à l'instant k, si au moins un bus passe à cet instant, et égale à 0 si aucun bus ne passe.

E.4.a. Justifier l'égalité $\mathbb{E}(W) = \sum_{r=1}^{+\infty} \frac{\lambda}{r}.p^r q$.

b. Soit $K \in \mathbb{N}^*$. Vérifier que $\sum_{r=1}^{K} \frac{p^r}{r} = \int_0^p \frac{\mathrm{d}t}{1-t} - \int_0^p \frac{t^K \,\mathrm{d}t}{1-t}$.

Montrer que $\lim_{K \to +\infty} \int_0^p \frac{t^K \,\mathrm{d}t}{1-t} = 0$.

En déduire que $\mathbb{E}(W) = -\lambda q \ln q$.

c. Montrer le nombre moyen de passagers montant dans le premier bus arrivant à l'instant k est toujours inférieur à $\frac{\lambda}{\mathrm{e}}$, et est égal à $\frac{\lambda}{\mathrm{e}}$ lorsque $p = 1 - \frac{1}{\mathrm{e}}$.

Fin