

DS2 (version B)

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.

*Les candidat·es sont invité·es à **encadrer** dans la mesure du possible les résultats de leurs calculs.*

*Aucun document n'est autorisé. **L'utilisation de toute calculatrice et de tout matériel électronique est interdite.** Seule l'utilisation d'une règle graduée est autorisée.*

Si au cours de l'épreuve, un candidat ou une candidate repère ce qui lui semble être une erreur d'énoncé, il la signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre.

On suppose, et c'est valable pour toute l'épreuve, que les librairies suivantes sont importées sous leurs alias habituels :

- `import numpy as np`
- `import numpy.random as rd`

Toutes les variables aléatoires de ce problème sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Introduction

On s'intéresse dans ce problème à la détermination de lois de probabilité composées qui interviennent en particulier dans la gestion du risque en assurance et en théorie de la ruine.

On étudie le modèle suivant :

- × le nombre de sinistres à prendre en charge par une compagnie d'assurances sur une période donnée est une variable aléatoire N à valeurs dans \mathbb{N} ;
- × les coûts des sinistres successifs sont modélisés par une suite de variables aléatoires $(U_k)_{k \in \mathbb{N}^*}$. On suppose que les variables U_k sont à valeurs dans \mathbb{N} , indépendantes et identiquement distribuées, et sont indépendantes de N ;
- × On pose, pour tout $n \in \mathbb{N}^*$, $X_n = \sum_{k=1}^n U_k$ et X_0 est la variable certaine de valeur 0 ;
- × la charge sinistrale totale pour la compagnie d'assurance sur une période est donnée par la variable aléatoire X définie par :

$$X = \sum_{k=1}^N U_k$$

et l'on précise que $X = X_0 = 0$ si N prend la valeur 0. **On dit que X suit une loi composée.**

- × pour tout entier naturel j , on pose $p_j = \mathbb{P}([N = j])$, $q_j = \mathbb{P}([U_1 = j])$ et $r_j = \mathbb{P}([X = j])$.

Partie I – Des exemples

Dans cette partie I, on suppose que les variables U_k suivent la loi de Bernoulli de paramètre p , où p est un réel de l'intervalle $]0, 1[$.

1. On suppose que l'on dispose d'une fonction **Python** `simulN()` (resp. `simulU()`) qui simule la variable aléatoire N (resp. la variable aléatoire U_1). Écrire une fonction **Python** `simulX()` qui simule la variable aléatoire X .

2. Pour n dans \mathbb{N}^* , quelle est la loi de X_n ?

3. Pour tout entier naturel j , établir : $r_j = \sum_{n=0}^{+\infty} \mathbb{P}([X_n = j]) p_n$.

Commentaire

Cette question sera un classique une fois le chapitre « couple de v.a.r. discrètes » fait. C'est normal d'avoir raté cette rédaction à ce moment de l'année.

4. Dans cette question 4., on suppose que N suit la loi binomiale de paramètres m , entier naturel, et π , réel dans $]0, 1[$. Soit j un entier naturel.

a) Justifier que $r_j = 0$ si $j > m$.

b) Établir que pour tout $j \in \llbracket 0, m \rrbracket$: $r_j = \sum_{n=j}^m \binom{n}{j} p^j (1-p)^{n-j} \binom{m}{n} \pi^n (1-\pi)^{m-n}$.

Commentaire

Il faut prendre l'habitude de se demander quels sont les univers images des variables aléatoires en présence lors d'un calcul de somme. Cela permet de simplifier la somme en enlevant les termes nuls.

- c) Vérifier que pour tous entiers j, n, m tels que $0 \leq j \leq n \leq m$: $\binom{n}{j} \binom{m}{n} = \binom{m}{j} \binom{m-j}{n-j}$.
- d) En déduire, pour tout $j \in \llbracket 0, m \rrbracket$: $r_j = \binom{m}{j} (p\pi)^j \sum_{\ell=0}^{m-j} \binom{m-j}{\ell} ((1-p)\pi)^\ell (1-\pi)^{m-j-\ell}$.
- e) Montrer finalement que X suit une loi binomiale et préciser ses paramètres en fonction de m, p et π .

5. On suppose dans cette question 5. que N suit la loi de Poisson de paramètre λ , réel strictement positif.

a) Montrer que pour tout entier naturel j , on a :

$$r_j = e^{-\lambda} \frac{(\lambda p)^j}{j!} \sum_{n=j}^{+\infty} \frac{1}{(n-j)!} (\lambda(1-p))^{n-j}$$

b) En déduire que X suit une loi de Poisson, et préciser son paramètre en fonction de p et λ .

Partie II – La loi binomiale négative

On généralise la définition des coefficients binomiaux aux nombres réels en posant, pour tout y réel et tout entier $k \in \mathbb{N}^*$: $\binom{y}{k} = \frac{1}{k!} \prod_{i=0}^{k-1} (y-i)$, et $\binom{y}{0} = 1$.

6. Écrire une fonction **Python** `coeffBin(y, k)` qui, prenant en arguments un réel y et un entier naturel non nul k , renvoie $\binom{y}{k}$.

7. La formule du binôme négatif.

Soit c un réel strictement positif, et x un réel de $[0, 1[$.

Pour tout entier naturel n , on pose $I_n = \int_0^x \frac{(x-t)^n}{(1-t)^{c+n+1}} dt$.

On admet le résultat suivant :

$$\forall n \in \mathbb{N}, \quad \frac{1}{(1-x)^c} = \sum_{k=0}^n \binom{c+k-1}{k} x^k + c \binom{c+n}{n} I_n$$

a) Vérifier que pour tout $t \in [0, x]$: $0 \leq \frac{x-t}{1-t} \leq x$.

En déduire l'encadrement : $0 \leq I_n \leq \frac{x^{n+1}}{(1-x)^{c+1}}$.

b) (i) Montrer, pour tout n dans \mathbb{N}^* : $\binom{c+n}{n} = \prod_{k=1}^n \left(1 + \frac{c}{k}\right)$.

Commentaire

Vous avez été nombreux à écrire $c!$ alors que cette notation n'a de sens que si c est un entier. Il fallait passer plus de temps à comprendre l'énoncé avant de se lancer dans les calculs. L'énoncé introduit justement une extension de la définition du coefficient binomial en utilisant le symbole de produit et pas le symbole de factorielle pour pouvoir manipuler des réels non entiers.

Ce genre de subtilité (sur des notations nouvellement introduites) fait partie de la difficulté des énoncés TOP3 et vous devez vous y habituer.

(ii) Montrer que pour tout réel t positif, $\ln(1+t) \leq t$.

Commentaire

L'énoncé n'était pas bien fait sur cette question car cette inégalité est en fait valable pour tout $t > -1$ et la question suivante se fait beaucoup plus simplement si on utilise cette inégalité sur $] -1, 0[$.

(iii) Établir, pour tout entier naturel $k \geq 2$: $\frac{1}{k} \leq \ln(k) - \ln(k-1)$.

En déduire : $\forall n \in \mathbb{N}^*, \sum_{k=1}^n \frac{1}{k} \leq 1 + \ln(n)$.

(iv) Montrer que pour tout n dans \mathbb{N}^* : $\ln \left(\binom{c+n}{n} \right) \leq c(1 + \ln(n))$.

En déduire : $\lim_{n \rightarrow +\infty} \binom{c+n}{n} x^{n+1} = 0$.

c) En conclure que la série $\sum_{k \geq 0} \binom{c+k-1}{k} x^k$ est convergente, et établir la formule du binôme négatif :

$$\sum_{k=0}^{+\infty} \binom{c+k-1}{k} x^k = \frac{1}{(1-x)^c}$$

8. Soit p un réel de $]0, 1[$ et r un réel strictement positif. Montrer que la suite de nombres $(p_k)_{k \in \mathbb{N}}$ définie par $p_k = \binom{r+k-1}{k} (1-p)^k p^r$ définit une loi de probabilité d'une variable aléatoire à valeurs dans \mathbb{N} . On l'appelle *loi binomiale négative* de paramètres r et p .

9. Si Y est une variable aléatoire suivant la loi binomiale négative de paramètres 1 et p , reconnaître la loi de $Y + 1$.

10. *Espérance et variance.*

Soit Z une variable aléatoire suivant la loi binomiale négative de paramètres r réel strictement positif et $p \in]0, 1[$.

a) Montrer que pour tout entier $k \geq 1$: $k \binom{r+k-1}{k} = r \binom{r+k-1}{k-1}$.

b) Montrer que Z admet une espérance et que l'on a : $\mathbb{E}(Z) = \frac{r(1-p)}{p}$.

c) Montrer que Z admet une variance et que l'on a : $\mathbb{V}(Z) = \frac{r(1-p)}{p^2}$.

On pourra commencer par calculer l'espérance de $Z(Z-1)$.

Partie III – Les lois de Panjer

On reprend les notations du début du problème : la variable aléatoire N à valeurs dans \mathbb{N} a sa loi donnée par $p_k = \mathbb{P}([N = k])$ pour $k \in \mathbb{N}$.

On suppose dans toute la suite du sujet que la loi de N vérifie la relation de Panjer : il existe deux réels a et b , avec $a < 1$ et $a + b > 0$, tels que :

$$\forall k \in \mathbb{N}^*, \quad p_k = \left(a + \frac{b}{k} \right) p_{k-1}$$

On dira alors que N suit la loi $\mathcal{P}(a, b)$.

11. *Détermination des lois de Panjer.*

- a) Montrer que pour tout entier k strictement positif, on a : $p_k = p_0 \prod_{i=1}^k \left(a + \frac{b}{i} \right)$.

Commentaire

Encore beaucoup d'erreurs de logique lors de l'initialisation, consistant à écrire comme point de départ la conclusion.

Pour éviter cet écueil, il faut soit utiliser la rédaction « D'une part... D'autre part... », soit partir du terme de gauche puis arriver au terme de droite au terme d'un calcul, soit partir du terme de droite puis arriver au terme de gauche au terme d'un calcul.

- b) Dans cette question, on suppose que $a = 0$.
Montrer que N suit une loi de Poisson de paramètre b .

Commentaire

Vous avez été nombreux à utiliser la formule de la question **11.a)** sans vous soucier de son domaine de validité. En particulier, elle n'est pas valable pour $k = 0$ a priori. Il fallait donc traiter ce cas à part dans vos calculs. Par chance, ici la formule restait valable pour $k = 0$, mais ce n'est pas toujours le cas !

- c) Dans cette question, on suppose que $a < 0$.
- (i) Montrer qu'il existe un unique entier naturel r , tel que : $\forall k > r, p_k = 0$ et $\forall k \leq r, p_k \neq 0$.
On pourra raisonner par l'absurde, et supposer les p_k tous strictement positifs.
- (ii) Montrer : $b = -a(r + 1)$.
- (iii) Établir que pour tout $k \in \llbracket 0, r \rrbracket, p_k = (-a)^k \binom{r}{k} p_0$.

En déduire que $p_0 = \frac{1}{(1-a)^r}$.

Commentaire

Comme dans beaucoup de questions de cette partie, il s'agissait d'une application de la méthode du SCE pour trouver la valeur du paramètre d'une loi discrète. Ici le paramètre s'appelle p_0 tandis que dans le TD ce paramètre s'appelait α . Le fait de ne pas reconnaître que c'est le même exercice doit vous pousser à passer plus de temps à réfléchir aux méthodes vues en classe pour prendre du recul et être capable de les remobiliser dans d'autres contextes.

- (iv) En conclure que N suit une loi binomiale et préciser les paramètres en fonction de a et b .
- d) Dans cette question, on suppose que $a > 0$.

- (i) Montrer que pour tout entier naturel k , on a : $p_k = \binom{\frac{b}{a} + k}{k} a^k p_0$.

Commentaire

Vous avez été plusieurs à faire appel à la question **7.b)(i)** mais vous n'avez pas fait attention au domaine de validité de cette question. Ce type d'erreurs a été très fréquent dans ce DS. Sur un sujet TOP3, c'est un des aspects qui feront la différence entre une copie correcte et une très bonne copie.

- (ii) En déduire que N suit une loi binomiale négative et préciser ses paramètres en fonction de a et b .

12. Montrer que, dans tous les cas, N admet une espérance et une variance, et qu'elles sont données par : $\mathbb{E}(N) = \frac{a+b}{1-a}$ et $\mathbb{V}(N) = \frac{a+b}{(1-a)^2}$.

Partie IV – L'algorithme de Panjer

- On reprend les notations de l'introduction du sujet et de la partie III.
- Si A est un évènement et Y une variable aléatoire, on note, si elle existe, $\mathbb{E}_A(Y)$ l'espérance de la loi conditionnelle de Y sachant A .

13. Pour tout $n \in \mathbb{N}$, exprimer $\mathbb{P}([X_n = 0])$ en fonction de q_0 puis établir que $r_0 = \sum_{n=0}^{+\infty} q_0^n p_n$.

14. Soit $j \in \mathbb{N}^*$.

a) Soit $n \in \mathbb{N}^*$, que vaut $\mathbb{E}_{[X_n=j]}(X_n)$? En déduire : $\mathbb{E}_{[X_n=j]}(U_1) = \frac{j}{n}$.

b) Établir : $r_j = \sum_{n=1}^{+\infty} \mathbb{P}([X_n = j]) \left(a + \frac{b}{n}\right) p_{n-1} = \sum_{n=1}^{+\infty} \mathbb{E}_{[X_n=j]} \left(a + \frac{b}{j} U_1\right) \mathbb{P}([X_n = j]) p_{n-1}$.

c) Montrer que pour tout $n \in \mathbb{N}^*$:

$$\mathbb{E}_{[X_n=j]} \left(a + \frac{b}{j} U_1\right) \mathbb{P}([X_n = j]) = \sum_{i=0}^j \left(a + \frac{b i}{j}\right) \mathbb{P}([U_1 = i]) \mathbb{P}([X_{n-1} = j - i])$$

d) En conclure : $r_j = \sum_{i=0}^j \left(a + \frac{b i}{j}\right) q_i r_{j-i}$, puis :

$$r_j = \frac{1}{1 - a q_0} \left(\sum_{i=1}^j \left(a + \frac{b i}{j}\right) q_i r_{j-i} \right)$$

Cette formule permet de calculer récursivement les nombre r_j et ainsi de déterminer la loi de X .

15. Des exemples d'application.

a) Dans cette question, les variables U_i suivent la loi de Bernoulli de paramètre p .

(i) Montrer que pour tout $j \in \mathbb{N}^*$, $r_j = \frac{p}{1 - a + a p} \left(a + \frac{b}{j}\right) r_{j-1}$.

En déduire que X suit une loi de Panjer.

(ii) Retrouver les résultats des questions 4. et 5. de la partie I.

b) Dans cette question, on suppose que $a = 0$, rappelons que cela entraîne que N suit la loi de Poisson de paramètre b .

Soit p un réel de l'intervalle $]0; 1[$.

(i) Montrer qu'il existe un unique réel α tel que la famille de nombre $(q_i)_{i \in \mathbb{N}^*}$ définie par $q_i = \alpha \frac{p^i}{i}$ définisse une loi de probabilité d'une variable aléatoire à valeurs dans \mathbb{N}^* (loi logarithme discrète). On pose $q_0 = 0$.

On suppose que les variables U_k suivent cette loi de probabilité.

(ii) Montrer que pour tout entier $j \geq 1$, on a : $r_j = \frac{b \alpha}{j} \sum_{i=1}^j p^i r_{j-i}$.

(iii) En utilisant un changement d'indice, établir pour tout $j \geq 2$: $r_j = \left(p + \frac{p(b\alpha - 1)}{j}\right) r_{j-1}$, puis montrer que cette égalité est encore vérifiée pour $j = 1$.

(iv) Conclure que X suit une loi binomiale négative et préciser ses paramètres en fonction de b , α et p .