I.S.F.A. 2002-2003

Concours d'entrée

DEUXIEME EPREUVE DE MATHEMATIQUES

Durée: 4 heures
Calculatrices autorisées

OPTION B

Introduction

La première partie de l'épreuve est consacrée à la caractérisation des lois vérifiant la relation dite de Panjer. La seconde partie présente l'algorithme de Panjer qui permet de calculer certaines sommes aléatoires de variables aléatoires discrètes. La dernière partie est consacrée à un modèle très simple qui permet de comprendre pourquoi une tarification a posteriori en assurance automobile (c'est-à-dire prenant en compte le nombre d'accidents déclarés par un assuré) pénalise fortement les assurés qui ont eu un accident la première année, mais ne favorise que faiblement ceux qui n'en ont pas eu.

Même si certaines questions sont motivées par des problèmes actuariels, aucune connaissance dans le domaine de l'assurance n'est requise. Les trois parties sont indépendantes.

1 Etude et caractérisation des lois vérifiant la relation de récurrence de Panjer

On représente le nombre d'accidents à prendre en charge par une compagnie d'assurances sur une période donnée par une variable aléatoire discrète N, à valeurs dans $\mathbb{N}$, et dont la loi est décrite par les $p_k = P(N = k)$ pour $k \in \mathbb{N}$. Parmi ces distributions on s'intéressera ici à celles vérifiant la relation de récurrence de Panjer:

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

Le but de cette partie est de caractériser les lois vérifiant (1).
Rappelons que la loi de Poisson de paramètre λ est décrite par

$$\forall k \in \mathbb{N}, \quad p_k = e^{-\lambda} \frac{\lambda^k}{k!},$$

et que la loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in]0,1[$ est décrite par

$$\forall 0 \leq k \leq n, \quad p_k = \binom{n}{k} p^k (1-p)^{n-k}$$

et $p_k = 0$ pour $k > n$.
Rappelons le résultat suivant sur les séries entières: pour $c \in \mathbb{N}$ et $|z| < 1$,

$$\sum_{k=0}^{\infty} c(c+1)\dots(c+k-1)\frac{z^k}{k!} = \frac{1}{(1-z)^c}$$

1. Montrer que si N vérifie (1) et que $a = 0$, alors N suit une loi de Poisson dont on précisera le paramètre.
2. On définit la distribution binomiale négative de paramètres $\alpha > 0$ et $p \in]0,1[$ par

   $$\forall k \in \mathbb{N}, \quad p_k = \binom{\alpha + k - 1}{k} (1-p)^\alpha p^k$$

   où pour $x \in \mathbb{R}$ et $k \in \mathbb{N}$,

   $$\binom{x}{k} = \frac{x(x-1)\dots(x-k+1)}{k!}.$$

   Calculer sa moyenne et sa variance en fonction de α et de p.

3. Quelle distribution obtient-on lorsque $\alpha = 1$?
4. Lorsque N vérifie la relation de récurrence (1) avec $a \neq 0$, montrer que pour tout $k \in \mathbb{N}$,
$$p_k = p_0 \frac{a^k}{k!} \prod_{i=0}^{k-1} (\Delta + i)$$
pour un certain Δ à préciser.
5. Montrer que pour tout $k \in \mathbb{N}$,
$$p_k = \binom{\Delta + k - 1}{k} (1-p)^{\Delta} a^k (1-a)^{\Delta}$$
6. En déduire selon le signe de a la loi de N.
7. Conclure que les lois vérifiant la relation (1) sont exactement les distributions de Poisson, binomiales et binomiales négatives.
8. Parmi ces 3 types de distributions, lequel choisiriez-vous pour modéliser la loi de N si une étude statistique vous montrait que la variance estimée de N est beaucoup plus grande que la moyenne estimée de N?

2 Algorithme de Panjer

On s'intéresse maintenant à la variable aléatoire
$$X = \sum_{k=1}^{N} U_k$$
où N et les U_k, $k \in \mathbb{N}$ sont des variables aléatoires à valeurs entières. De plus les U_k sont indépendantes, identiquement distribuées, de loi commune décrite par les $q_k = P(U_i = k)$, et indépendantes de N. La somme est nulle par convention si $N = 0$. Définissons de plus l'espérance d'une variable aléatoire Y à valeurs dans $\mathbb{N}$ conditionnellement à un évènement A par:
$$E(Y \mid A) = \sum_{k \in \mathbb{N}} k P(Y = k \mid A)$$

1. Montrer que pour $j \in \mathbb{N}$ et $n \in \mathbb{N}^*$,
$$E\left(U_1 \mid \sum_{i=1}^{n} U_i = j\right) = \frac{j}{n}$$

2. Définissons pour $j \in \mathbb{N}$ et $n \in \mathbb{N}^*$ les probabilités

$$q_j^{*n} = P(U_1 + \cdots + U_n = j).$$

Montrer que pour j, $k \in \mathbb{N}$ et $n \in \mathbb{N}^*$,

$$P\left(U_1 = k \mid \sum_{i=1}^{n} U_i = j\right) = \frac{q_k q_{j-k}^{*(n-1)}}{q_j^{*n}}$$

3. Supposons dans la suite de cette partie que N vérifie la relation (1) de la partie 1.
Soit $r_k = P(X = k)$ pour $k \in \mathbb{N}$. Exprimer r_0 en fonction des q_k et des p_k.
4. A l'aide des questions précédentes, montrer que pour $j \in \mathbb{N}^*$,

$$r_j = \sum_{n=1}^{\infty} \left(a + bE\left(\frac{U_1}{j} \mid \sum_{i=1}^{n} U_i = j\right)\right) p_{n-1} q_j^{*n}.$$

5. En utilisant les questions précédentes, démontrer la formule récursive appelée "algorithme de Panjer":

$$\forall j \in \mathbb{N}^*, \quad r_j = C \sum_{k=1}^{j} \left(a + \frac{bk}{j}\right) q_k r_{j-k}$$

où C est une constante à préciser.

3 Modèle bons/mauvais conducteurs

Soit X_1 la variable aléatoire décrivant le temps en années jusqu'au premier accident d'un conducteur, X_2 le temps entre le premier et le deuxième accident et pour $i \geq 3$ soit X_i le temps entre le $i^{\text{ième}}$ accident et le $(i+1)^{\text{ième}}$ accident. On suppose que les $(X_i)_{i \in \mathbb{N}}$ sont indépendants et identiquement distribués selon une loi exponentielle de paramètre λ. On s'intéresse à la variable aléatoire N décrivant le nombre N d'accidents pendant une année.

1. On rappelle qu'une loi Gamma de paramètres $n \in \mathbb{N}^*$ et $\lambda > 0$ est donnée par sa densité:

$$\forall t \in \mathbb{R}^+, \quad f_{n,\lambda}(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}.$$

Montrer par récurrence sur n que pour $n \in \mathbb{N}^*$, $Y_n = X_1 + \cdots + X_n$ suit une loi Gamma dont on précisera les paramètres.

2. Calculer $P(N = 0)$.
3. Calculer $P(N = n)$ pour $n \in \mathbb{N}^*$. On pourra montrer que l'évènement $\{N \geq n\}$ s'écrit
$$\{X_1 + \cdots + X_n \leq 1\}.$$
4. En déduire que N suit une loi de Poisson de paramètre à préciser.
5. On distingue maintenant les bons et les mauvais conducteurs. Soit B l'évènement "être un bon conducteur" et M l'évènement "être un mauvais conducteur". Supposons qu'il y a autant de bons que de mauvais conducteurs, si bien que $P(B) = P(M) = \frac{1}{2}$. Pour les bons conducteurs, $\lambda = \lambda_B = 1/10$ et pour les mauvais conducteurs, $\lambda = \lambda_M = 1/5$. De façon plus rigoureuse, sachant que B (resp. M) est réalisé, la loi de N est décrite par (2) avec les X_i i.i.d., de loi exponentielle de paramètre λ_B (resp. λ_M).
Quel est le nombre moyen d'accidents en une année d'un bon conducteur? D'un mauvais conducteur? D'un conducteur pris au hasard?

6. De plus, à partir de maintenant, on suppose que le coût d'un accident n'est pas aléatoire et est toujours égal à $c > 0$.
Calculer la prime pure qu'un assureur pourrait faire payer à un nouvel assuré pour un an, c'est-à-dire $\pi = cE(N)$, où N est le nombre d'accidents subis par l'assuré en 1 an.
7. Sachant que l'assuré n'a pas eu d'accident la première année, calculer la nouvelle probabilité qu'il soit un bon conducteur: $P(B \mid N = 0)$. (Donner une formule littérale puis faire l'application numérique dans toutes les dernières questions, y compris celle-ci.)
8. Calculer $P(B \mid N = k)$ pour $k = 1,2,3,4$.
9. Comparez ces probabilités à $P(B)$. Qu'observez-vous? Comment pouvez-vous l'expliquer ?