MATHÉMATIQUES II - ESSEC E 2007

Proposition de corrigé par David Meneu

Lycée Champollion - Grenoble, pour Major-Prépa

Partie I : Modélisation poissonienne

On considère une société d'assurance comptant N clients et garantissant à chacun d'entre eux un capital d'un montant de s euros en cas de décès. On suppose que le nombre de décès annuel suit une loi de Poisson de paramètre entier k. Le revenu annuel de la société fourni par la perception des primes d'assurance des N clients est au total de $ks(1+\lambda)$ eurons, où λ est un réel strictement positif représentant le taux de sécurité que la société s'accorde afin de faire face à un nombre de sinistres plus élevé que la moyenne. La société dispose également d'un fond de réserve R dans lequel elle peut puiser exceptionnellement. Un bilan financier de la société est effectué tous les 5 ans.

On note Y le nombre de décès enregistrés sur une période de 5 ans.

A. Résultats généraux :

I.A.1. Chaque décès coûte s euros à la société, donc la somme totale due par la société aux clients au moment du blian financier est $\boxed{sY}$.

I.A.2. Si sur une période de 5 ans, Y_1, Y_2, Y_3, Y_4, Y_5 sont les variables aléatoires égales aux nombres de décès enregistrés chaque année de la période, alors d'après l'énoncé ces variables aléatoires suivent toutes la loi de Poisson de paramètre k : alors $Y = Y_1 + Y_2 + Y_3 + Y_4 + Y_5$, et si ces 5 variables aléatoires sont mutuellement indépendantes, le rappel préliminaire fait par le sujet assure que Y suit la loi de Poisson de paramètre $5k$.

On supposera dorénavant que Y suit une loi de Poisson de paramètre $5k$.

I.A.3. D'après le cours sur la loi de Poisson : $E(Y) = V(Y) = 5k$.

I.A.4. Soit Z une variable aléatoire suivant la loi normale centrée, réduite : on sait qu'une densité de Z est la fonction $\varphi :\ x \mapsto \dfrac{1}{\sqrt{2\pi}}.e^{-x^2/2}$ qui est continue et strictement positive sur $\mathbb{R}$, ce qui garantit que sa fonction de répartition $\Phi :\ t \mapsto \displaystyle\int_{-\infty}^{t} \frac{\exp(-x^2/2)}{\sqrt{2\pi}}\mathrm{d}x$ est de classe $\mathcal{C}^1$ et strictement croissante sur $\mathbb{R}$. Elle réalise donc une bijection de $\mathbb{R}$ dans $]-\lim\limits_{x\to-\infty}\Phi(x); \lim\limits_{x\to+\infty}\Phi(x)[=]0;1[$.

À ce titre, il existe bien un unique réel t_0 tel que $\Phi(t_0) = 0.99$. Comme $0.99 = 0.5$, on a bien : $\Phi(t_0) > \Phi(0) \iff t_0 > 0$ par stricte croissance de Φ sur $\mathbb{R}$.

I.A.5. On fait aparaître ici un résultat lié au Théorème de la Limite Centrée, en écrivant :

$$P\big(Y - 5k > t_0\sqrt{5k}\big) = P\Big(\frac{Y-5k}{\sqrt{5k}} > t_0\Big) = P\Big(k.\frac{\frac{1}{k}Y - 5}{\sqrt{5k}} > t_0\Big) = P\Big(\sqrt{k}.\frac{\frac{1}{k}Y-5}{\sqrt{5}} > t_0\Big)$$

Or Y qui suit la loi de Poisson de paramètre $5k$, a même loi que la somme $S_k = \displaystyle\sum_{i=1}^{k} X_i$ de k variables aléatoires mutuellement indépendantes, de même loi de Poisson de paramètre 5, et ainsi en notant $\overline{S_k} = \dfrac{1}{k}\displaystyle\sum_{i=1}^{k} X_i$: $\quad P(Y - 5k > t_0\sqrt{5k}) = P\Big(\sqrt{k}.\dfrac{\overline{S_k} - 5}{\sqrt{5}} > t_0\Big)$ qui tend bien vers

$P(Z > t_0) = 0.01$ lorsque k tend vers $+\infty$, d'après le théorème de la limite centrée appliqué à la suite de variables aléatoires $(X_i)_{i\in\mathbb{N}^*}$, mutuellement indépendantes et de même loi, d'espérance 5 et d'écart-type $\sqrt{5}$.

Pour la fin de cette partie, on supposera k assez grand pour utiliser l'approximation

$$P\big(Y - 5k > t_0\sqrt{5k}\big) = 0.01 \qquad (A)$$

B. Exemples d'application :

Dans cette partie il s'agit d'exploiter l'approximation (A).

I.B.1. La société d'assurance peut faire face à toutes les indemnistions requises sur l'exercice de 5 ans si et seulement si ses fonds disponibles : rentrées $5sk(1+\lambda)$ auxquelles s'ajoute le fond de réserve R, excèdent les sommes dues à ses clients : sY. D'où l'inégalité voulue.

I.B.2. On cherche ici la valeur de R pour laquelle $P\big(5sk(1+\lambda) + R < sY\big) = 0.01$.

Or : $P\big(5sk(1+\lambda) + R < sY\big) = P\big(Y - 5k > 5k\lambda + \frac{1}{s}R\big)$, on cherche donc R tel que :

$$5k\lambda + \frac{1}{s}R = t_0\sqrt{5k} \iff R = s(t_0\sqrt{5k} - 5k\lambda)$$

I.B.3. On note $\mu = \frac{k}{N}$ le taux de mortalité dans l'ensemble des clients. Pour que la société puisse se dispenser d'un fond de réserve, il faut que :

$P\big(5sk(1+\lambda) < sY\big) = 0.01 \iff P\big(Y - 5k > 5k\lambda\big) = 0.01$. On cherche donc k tel que :

$$t_0\sqrt{5k} = 5k\lambda \iff \sqrt{5k} = \frac{t_0}{\lambda} \iff k = \frac{(t_0)^2}{5\lambda^2}$$

or $\mu = \frac{k}{N} \iff N = \frac{k}{\mu}$, donc il faut que $N = \dfrac{(t_0)^2}{5\lambda^2\mu}$ (au moins).

Partie II : Médianes

Soit X une variable aléatoire réelle. On définit l'ensemble

$$\mathcal{M}(X) = \Big\{m \in \mathbb{R} \big|\ P(X < m) \leqslant \frac{1}{2} \leqslant P(X \leqslant m)\Big\}.$$

Un élément de $\mathcal{M}(X)$ est appelé **médiane** de X.

II.1. Soit X une variable aléatoire réelle : sa fonction de répartition F est définie par

$$\forall x \in \mathbb{R}, \quad F(x) = P(X \leqslant x)$$

II.2. Soit X une variable aléatoire de Bernoulli de paramètre $\frac{1}{2}$: $\quad P(X=0) = \dfrac{1}{2} = P(X=1)$, donc :

- pour tout $m < 0$, $P(X < m) = 0$ et $P(X \geqslant m) = 1$
- pour $m = 0$: $P(X < 0) = 0$ et $P(X \geqslant 0) = 1$
- pour $m \in]0,1[$: $P(X < m) = P(X=0) = \frac{1}{2}$ et $P(X \geqslant m) = P(X=1) = \frac{1}{2}$.
- pour $m = 1$: $P(X < 1) = P(X=0) = \frac{1}{2}$ et $P(X \geqslant 1) = P(X=1) = \frac{1}{2}$.
- pour $m > 1$: $P(X < m) = 1$ et $P(X \geqslant m) = 0$

On en déduit que dans ce cas : $\quad \mathcal{M}(X) =]0;1]$.

II.3. Soit X une variable aléatoire de loi exponentielle de paramètre $\alpha > 0$, de fonction de répartition F_X.

Pour tout $m < 0$, $P(X < m) = 0$ donc $m \notin \mathcal{M}(X)$; pour tout réel $m \geqslant 0$, comme X est une variable à densité, $P(X < m) = P(X \leqslant m)$, donc : $m \in \mathcal{M}(X)$ si et seulement si $P(X \leqslant m)$ est à la fois supérieur et inférieur à $\frac{1}{2}$, soit : $m \in \mathcal{M}(X) \iff m \geqslant 0$ et $F_X(m) = \frac{1}{2}$

On résout donc, pour $m \geqslant 0$:

$F_X(m) = \frac{1}{2} \iff 1 - e^{-\alpha m} = \frac{1}{2} \iff e^{-\alpha m} = \frac{1}{2} \iff -\alpha m = -\ln(2) \iff m = \frac{\ln(2)}{\alpha}$, soit :

$$\mathcal{M}(X) = \left\{\frac{\ln(2)}{\alpha}\right\}$$

on revient au cas général où X est une variable aléatoire réelle.

II.4. Soient a et b deux éléments de $\mathcal{M}(X)$ avec $a \leqslant b$; pour tout $c \in [a, b]$, par croissance de la probabilité :

$$P(X < c) \leqslant P(X < b) \leqslant \frac{1}{2} \quad \text{et} \quad \frac{1}{2} \leqslant P(X \leqslant a) \leqslant P(X \leqslant c) \quad \text{donc} \quad P(X < c) \leqslant \frac{1}{2} \leqslant P(X \leqslant c)$$

et par conséquent, c appartient bien encore à $\mathcal{M}(X)$.

On a ainsi démontré que $\mathcal{M}(X)$ est un intervalle.

II.5. Supposons que f possède une densité f continue sur $\mathbb{R}$ telle que $f(x) > 0$ pour tout x réel. Dans ce cas :

- On sait que pour tout réel m, $P(X < m) = P(X \leqslant m)$, donc :
  $P(X < m) \leqslant \frac{1}{2} \leqslant P(X \leqslant m) \iff P(X \leqslant m) = \frac{1}{2}$, et donc dans ce cas :
  $$\mathcal{M}(X) = \{m \in \mathbb{R} |\ F_X(m) = \frac{1}{2}\}$$
- La fonction de répartition F_X est continue, strictement croissante sur $\mathbb{R}$, et réalise d'après le théorème éponyme, une bijection de $\mathbb{R}$ dans $]0, 1[$. Il existe donc un unique réel m_0 tel que $F_X(m_0) = \frac{1}{2}$, et ainsi : $\mathcal{M}(X) = \{m_0\}$ est réduit à un seul réel.

En particulier, lorsque X suit la loi normale centrée, réduite : la densité φ : $x \mapsto \frac{1}{\sqrt{2\pi}}.e^{-x^2/2}$ de X, est effectivement continue, strictement positive sur $\mathbb{R}$. On sait aussi que sa fonction de répartition Φ a pour valeur particulière : $\Phi(0) = \frac{1}{2}$, donc dans ce cas :

$$\mathcal{M}(X) = \{1/2\}$$

II.6. L'exemple de la loi exponentielle montre déjà que si X admet une espérance, celle-ci n'est pas forcément une médiane de X : dans ce cas, $\mathcal{M}(X) = \{\frac{\ln(2)}{\alpha}\}$, mais $E(X) = \frac{1}{\alpha} \neq \frac{\ln(2)}{\alpha}$.

Partie III : Médiane d'une variable poissonienne

A. Préliminaires d'analyse :

On étudie la suite $(u_n)_{n \in \mathbb{N}^*}$ définie par :

$$\forall n \in \mathbb{N}^*, \quad u_n = \exp(-n).\frac{n^n}{n!}$$

III.A.1. On considère ici la fonction $f: x \mapsto x - \dfrac{x^2}{4} - \ln(1+x)$, bien définie et dérivable sur $[0;1]$, avec :

$$\forall x \in [0;1], \quad f'(x) = 1 - \frac{x}{2} - \frac{1}{1+x} = \frac{2(1+x) - x(1+x) - 2}{2(1+x)} = \frac{x - x^2}{2(1+x)}$$

Pour tout $x \in [0;1]$: $\quad x^2 \leqslant x \iff x - x^2 \geqslant 0$, et $2(1+x) > 0$, donc f est croissante sur $[0;1]$. Comme $f(0) = 0 - 0 - \ln(1) = 0$ est alors le minimum de f sur $[0;1]$, la fonction f est positive sur cet intervalle :

$$\forall x \in [0;1], \quad f(x) \geqslant 0 \iff x - \frac{x^2}{4} \geqslant \ln(1+x)$$

III.A.2. Pour tout $n \in \mathbb{N}^*$:

$$\frac{u_{n+1}}{u_n} = \frac{\exp(-n-1).\dfrac{(n+1)^{n+1}}{(n+1)!}}{\exp(-n).\dfrac{n^n}{n!}} = \exp(-1).\frac{(n+1)^{n+1}}{(n+1)!}.\frac{n!}{n^n} = \exp(-1).\frac{(n+1)^n.\cancel{(n+1)}.\cancel{n!}}{\cancel{n!}.\cancel{(n+1)}.n^n}$$

$$= \exp(-1).\left(\frac{n+1}{n}\right)^n = \exp(-1).\left(1 + \frac{1}{n}\right)^n = \exp\left(n\ln\left(1 + \frac{1}{n}\right) - 1\right)$$

III.A.3. De ce qui précède, on déduit que pour tout $n \in \mathbb{N}^*$:

$$\ln(u_n) - \ln(u_{n+1}) = \ln\left(\frac{u_n}{u_{n+1}}\right) = -\ln\left(\frac{u_{n+1}}{u_n}\right) = 1 - n\ln\left(1 + \frac{1}{n}\right)$$

D'après l'inégalité obtenue en III.A.1., avec $x = \dfrac{1}{n} \geqslant 0$ pour tout $n \in \mathbb{N}^*$:

$$\ln\left(1 + \frac{1}{n}\right) \leqslant \frac{1}{n} - \frac{1}{4n^2} \iff n\ln\left(1 + \frac{1}{n}\right) \leqslant 1 - \frac{1}{4n} \iff \frac{1}{4n} \leqslant 1 - n\ln\left(1 + \frac{1}{n}\right)$$

Or la série $\displaystyle\sum_{n \geqslant 1} \frac{1}{4n}$ diverge puisque c'est, à un facteur constant près, la série harmonique.

Par comparaison de séries à termes positifs, la série $\displaystyle\sum_{n \geqslant 1} \ln(u_n) - \ln(u_{n+1})$ est donc elle-même divergente.

III.A.4. On peut même préciser, s'agissant d'une série à termes positifs divergente, que :

$$\lim_{n \to +\infty} \sum_{k=0}^{n-1} \ln(u_k) - \ln(u_{k+1}) = +\infty \iff \lim_{n \to +\infty} \ln(u_0) - \ln(u_n) = +\infty \iff \lim_{n \to +\infty} \ln(u_n) = -\infty$$

Et ainsi : $\quad \displaystyle\lim_{n \to +\infty} u_n = \lim_{n \to +\infty} e^{\ln(u_n)} = \lim_{X \to -\infty} e^X = 0.$

B. Probabilités :

Pour tout $n \in \mathbb{N}$, on note P_n la fonction définie par :

$$P_n(\lambda) = \exp(-\lambda).\sum_{k=0}^{n} \frac{\lambda^k}{k!} = \exp(-\lambda)\left(1 + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^n}{n!}\right)$$

pour tout réel λ.

III.B.1. Soit $n \in \mathbb{N}^*$: la fonction P_n est bien de classe $\mathcal{C}^2$ sur $\mathbb{R}$ comme produit de deux fonctions de classe $\mathcal{C}^2$ sur $\mathbb{R}$, et :

$$\forall \lambda \in \mathbb{R}, \quad P_n'(\lambda) = -\exp(-\lambda).\sum_{k=0}^{n}\frac{\lambda^k}{k!} + \exp(-\lambda).\sum_{k=\cancel{0}1}^{n}\frac{k\lambda^{k-1}}{k!} = \exp(-\lambda).\Big(-\sum_{k=0}^{n}\frac{\lambda^k}{k!} + \sum_{k=1}^{n}\frac{\lambda^{k-1}}{(k-1)!}\Big)$$

$$= \exp(-\lambda).\Big(\sum_{j=0}^{n-1}\frac{\lambda^j}{j!} - \sum_{k=0}^{n}\frac{\lambda^k}{k!}\Big) = -\exp(-\lambda).\frac{\lambda^n}{n!}$$

$$\forall \lambda \in \mathbb{R}, \quad P_n''(\lambda) = \exp(-\lambda).\frac{\lambda^n}{n!} - \exp(-\lambda).\frac{n\lambda^{n-1}}{n!} = \exp(-\lambda).\frac{\lambda^{n-1}}{n!}.(\lambda - n)$$

III.B.2. Pour tout $n \in \mathbb{N}^*$:

$$P_n(n-1)+P_n'(n-1) = \exp(-\lambda).\sum_{k=0}^{n}\frac{(n-1)^k}{k!} - \exp(-\lambda).\frac{(n-1)^n}{n!} = \exp(-\lambda).\sum_{k=0}^{n-1}\frac{(n-1)^k}{k!} = P_{n-1}(n-1)$$

III.B.3. Soit $Q : \mathbb{R} \to \mathbb{R}$ une application de classe $\mathcal{C}^2$ sur $\mathbb{R}$.

(i) Soit $n \in \mathbb{N}^*$: on réalise une intégration par parties dans $\displaystyle\int_{n-1}^{n}(n-t)Q''(t)\mathrm{d}t$, en posant :

$$u(t) = (n-t) \quad \longrightarrow \quad u'(t) = -1$$
$$v'(t) = Q''(t) \quad \longrightarrow \quad v(t) = Q'(t)$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$, donc par intégration par parties :

$$\int_{n-1}^{n}(n-t)Q''(t)\mathrm{d}t = \Big[(n-t).Q'(t)\Big]_{n-1}^{n} + \int_{n-1}^{n}Q'(t)\mathrm{d}t = -\big(n-(n-1)\big).Q'(n-1) + \Big[Q(t)\Big]_{n-1}^{n}$$

ce qui donne bien, pour tout $n \in \mathbb{N}^*$:

$$\int_{n-1}^{n}(n-t).Q''(t)\mathrm{d}t = -Q'(n-1)+Q(n)-Q(n-1) \iff Q(n) = Q(n-1)+Q'(n-1)+\int_{n-1}^{n}(n-t).Q''(t)\mathrm{d}t \quad (E)$$

(ii) En appliquant (E) à $Q = P_n$, on obtient, d'après la relation obtenue en III.B.2., et pour tout $n \in \mathbb{N}^*$:

$$P_n(n) = P_n(n-1) + P_n'(n-1) + \int_{n-1}^{n}(n-t).P_n''(t)\mathrm{d}t$$

$$= P_{n-1}(n-1) + \int_{n-1}^{n}(n-t).\exp(-t).\frac{t^{n-1}}{n!}(t-n)\mathrm{d}t$$

$$= P_{n-1}(n-1) - \int_{n-1}^{n}(n-t).\exp(-t).\frac{t^{n-1}}{n!}\mathrm{d}t$$

Pour tout entier $n \geqslant 2$, et pour tout réel $t \in [n-1;n]$: $\quad (n-t)^2 \geqslant 0$, $\exp(-t) > 0$ et $\dfrac{t^{n-1}}{n!} \geqslant 0$. La fonction dans l'intégrale précédente est donc continue, positive sur $[n-1;n]$: la propriété de positivité de l'intégrale assure alors que $\displaystyle\int_{n-1}^{n}(n-t)^2.\exp(-t).\frac{t^{n-1}}{n!}\mathrm{d}t \geqslant 0$, et donc :

$$\forall n \geqslant 2, \quad P_n(n) \leqslant P_{n-1}(n-1)$$

ce qui assure que la suite $\big(P_n(n)\big)_{n\in\mathbb{N}^*}$ est décroissante.

(iii) En appliquant (E) à $Q = P_{n-1}$, on obtient, pour tout $n \geqslant 2$:

$$\begin{aligned} P_{n-1}(n) &= P_{n-1}(n-1) + P'_{n-1}(n-1) + \int_{n-1}^{n} (n-t)P''_{n-1}(t)\mathrm{d}t \\ &= P_{n-2}(n-1) + \int_{n-1}^{n} (n-t).(t-n+1)\exp(-t).\frac{t^{n-2}}{(n-1)!}\mathrm{d}t \end{aligned}$$

Pour tout $n \geqslant 2$ et tout réel $t \in [n-1; n]$:
$(n-t) \geqslant 0$ et $n-1 \leqslant t \leqslant n \iff 0 \leqslant t-n+1 \leqslant 1$; la fonction intégrée est continue et positive sur $[n-1; n]$: par positivité de l'intégrale, on en déduit :

$$\forall n \geqslant 2, \quad \int_{n-1}^{n} (n-t).(t-n+1)\exp(-t).\frac{t^{n-2}}{(n-1)!}\mathrm{d}t \iff P_{n-1}(n) \geqslant P_{n-2}(n-1)$$

ce qui prouve bien que la suite $\big(P_{n-1}(n)\big)_{n\geqslant 1}$, est croissante.

III.B.4. Pour tout $n \in \mathbb{N}^*$:

$$P_n(n) - P_{n-1}(n) = \exp(-n).\Big(\sum_{k=0}^{n} \frac{n^k}{k!} - \sum_{k=0}^{n-1} \frac{n^k}{k!}\Big) = \exp(-n).\frac{n^n}{n!} = u_n$$

Ainsi :

- La suite $\big(P_n(n)\big)_{n\in\mathbb{N}^*}$ est décroissante, tandis que la suite $\big(P_{n-1}(n)\big)_{n\geqslant 1}$ est croissante.
- $\lim\limits_{n\to+\infty} P_n(n) - P_{n-1}(n) = \lim\limits_{n\to+\infty} u_n = 0$ comme on l'a vu à la fin de la partie A.

Les suites $\big(P_n(n)\big)_{n\geqslant 1}$ et $\big(P_{n-1}(n)\big)_{n\in\mathbb{N}^*}$ sont donc adjacentes : d'après le théorème associé, elles sont donc convergentes et de même limite.

On considère dorénavant Z une variable aléatoire réelle de loi de Poisson de paramètre $n \in \mathbb{N}^*$.

III.B.5. Pour tout $n \in \mathbb{N}^*$:

$$P\Big(\frac{Z-n}{\sqrt{n}} \leqslant 0\Big) = P(Z-n \leqslant 0) = P(Z \leqslant n) = \sum_{k=0}^{n} P(Z=k) = \sum_{k=0}^{n} \exp(-n).\frac{n^k}{k!} = P_n(n)$$

Tout comme on l'a fait dans la partie I.A., il suffit alors de remarquer que si $(Z_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires mutuellement indépendantes, de même loi de Poisson de paramètre 1 : la variable aléatoire $S_n = \sum\limits_{k=1}^{n} Z_k$ suit la même loi que Z, et on peut écrire :

$$\lim_{n\to+\infty} P\Big(\frac{Z-n}{\sqrt{n}} \leqslant 0\Big) = \lim_{n\to+\infty} P\Big(\frac{S_n-n}{\sqrt{n}} \leqslant 0\Big) = \Phi(0) = \frac{1}{2}$$

d'après le Théorème de la Limite Centrée, ce qui donne bien :

$$\lim_{n\to+\infty} P_n(n) = \frac{1}{2}$$

III.B.7. Les deux suites adjacentes de la question III.B.4. encadrent leur limite commune, dont on vient de voir qu'elle vaut $\frac{1}{2}$; comme $\big(P_n(n)\big)_{n\in\mathbb{N}^*}$ est la suite décroissante, alors :

$$\forall n \in \mathbb{N}^*, \quad P_{n-1}(n) \leqslant \frac{1}{2} \leqslant P_n(n)$$

III.B.8. La double inégalité précédente se réécrit :

$$\forall n \in \mathbb{N}^*, \quad P(Z \leqslant n-1) \leqslant \frac{1}{2} \leqslant P(Z \leqslant n) \iff P(Z < n) \leqslant \frac{1}{2} \leqslant P(Z \leqslant n)$$

ce qui prouve bien que n est une médiane de la variable aléatoire Z.

III.B.9. L'énoncé admet finalement que $\mathcal{M}(Z) = \{n\}$.

À la lumière de ce résultat : il y a donc une chance sur deux que le nombre de décès (annuel, ou sur une période de 5 ans) soit supérieur à la moyenne. Dans ces conditions, prendre un taux de sécurité λ nul apparaît extrêmement risqué !

Partie IV : Inégalité maximale de Lévy

Soit J un entier strictement supérieur à 1. On considère J variables aléatoires mutuellement indépendantes $X_1, \ldots, X_J$. On suppose que pour chaque entier $j \in [\![1, J]\!]$, la variable aléatoire X_j suit une loi de Poisson de paramètre entier non nul k et on définit $Y_j = X_j - k$. On pose enfin pour tout entier i tel que $1 \leqslant i \leqslant J$: $\quad S_i = \sum_{j=1}^{i} Y_j$.

IV.1. Soit i un entier compris entre 1 et J : $\quad S_J - S_i = \sum_{j=i+1}^{J} Y_i = \sum_{j=i+1}^{J} X_i - (J-i)k$, où $\sum_{j=i+1}^{J} X_i$ suit la loi de Poisson de paramètre $(J-i)k$, comme somme de $(J-i)$ variables de Poisson de paramètre k, mutuellement indépendantes. D'après III.B.8., $(J_i)k$ est donc une médiane (et même la médiane) de cette variable aléatoire, seul réel vérifiant :

$$P\Big(\sum_{j=i+1}^{J} X_j < (J-i)k\Big) \leqslant \frac{1}{2} \leqslant P\Big(\sum_{j=i+1}^{J} X_j \leqslant (J-i)k\Big) \iff P(S_J - S_i < 0) \leqslant \frac{1}{2} \leqslant P(S_J - S_i \leqslant 0)$$

ce qui prouve bien que 0 est une médiane de $S_J - S_i$.

Soit x un nombre réel positif. On considère les événements :

$$\Omega_0 = \Big\{ \max_{1 \leqslant j \leqslant J} S_j \leqslant x \Big\} \quad \text{et} \quad \Omega_1 = \{S_1 > x\}$$

puis pour tout entier i tel que $2 \leqslant i \leqslant J$, on note

$$\Omega_i = \Big\{ \max_{1 \leqslant j < i} S_j \leqslant x \Big\} \cap \{S_i > x\}$$

IV.2. Interprétons concrètement le sens des événements $\Omega_0, \Omega_1, \Omega_i$:

- Ω_0 est réalisé si et seulement si la plus grande valeur prise par une variable de l'échantillon $(S_1, \ldots, S_J)$ est inférieure ou égale à x : on sait que c'est équivalent au fait que **toutes** les variables de l'échantillon, prennent une valeur inférieure ou égale à x.

- L'événement Ω_1 est réalisé si et seulement si la variable $S_1 = Y_1 - k$ prend une valeur strictement supérieure à x : il est donc en particulier incompatible avec Ω_0.

- Pour tout entier i compris entre 2 et J : l'événement Ω_i est réalisé si et seulement si aucune des $i-1$ premières variables aléatoires S_i ne dépasse x, tandis que S_i est strictement supérieure à x : cela correspond donc au fait que dans l'échantillon $(S_1, \ldots, S_J)$, S_i est la *première* dans l'ordre, à dépasser x.

Il est alors clair, au vu de ces reformulations, que $(\Omega_0, \Omega_1, \Omega_2, \ldots, \Omega_J)$ est un système complet d'événements :

Soit aucune des variables $(S_i)_{1\leqslant i\leqslant J}$ ne dépasse x, auquel cas Ω_0 est réalisé, soit il existe un indice $i \in [\![1, J]\!]$ tel que S_i soit la première dans l'ordre à dépasser x, auquel cas l'événement correspondant Ω_i est réalisé, et lui seul.

Chaque issue réalise bien un et un seul de ces événements.

IV.3. Soit i un entier compris entre 1 et J : l'événement $\{S_J - S_i \geqslant 0\} \cap \Omega_i$ est réalisé si et seulement si S_i est la première variable de l'échantillon à dépasser la valeur x, **et** S_J prend une valeur supérieure à S_i, donc à x : cela implique bien que $\{S_J > x\}$ est réalisé, et Ω_i aussi !

$$\{S_J - S_i \geqslant 0\} \cap \Omega_i \subset \{S_J > x\} \cap \Omega_i$$

IV.4. L'inclusion précédente et la croissance de la probabilité donnent alors :

$$\forall i \in [\![1; J]\!], \quad P\big(\{S_J - S_i \geqslant 0\} \cap \Omega_i\big) \leqslant P\big(\{S_J > x\} \cap \Omega_i\big)$$

On peut alors passer à la somme dans cette inégalité, ce qui donne :

$$\sum_{i=1}^{J} P\big(\{S_J - S_i \geqslant 0\} \cap \Omega_i\big) \leqslant \sum_{i=1}^{J} P\big(\{S_J > x\} \cap \Omega_i\big) \iff P(S_J > x) \geqslant \sum_{i=1}^{J} P\big(\{S_J - S_i \geqslant 0\} \cap \Omega_i\big)$$

d'après la formule des probabilités totales, puisque $(\Omega_i)_{0\leqslant i\leqslant J}$ est un système complet d'événements !

IV.5. Soit i un entier compris entre 1 et J : l'événement $\{S_J - S_i \geqslant 0\}$ concerne la variable aléatoire $S_J - S_i = \displaystyle\sum_{j=i+1}^{J} X_i - (J-i)k$, tandis que l'événement Ω_i concerne les variables $S_1, \ldots, S_i$ qui ne dépendent que des variables $X_1, \ldots X_i$: comme les $(X_i)_{1\leqslant i\leqslant J}$ sont mutuellement indépendantes, le lemme des coalitions assure alors que les événements $\{S_J - S_i \geqslant 0\}$ et Ω_i sont **indépendants**.

I.V.6. De ce qu'on vient de voir, on déduit que :

$$\sum_{i=1}^{J} P\big(\{S_J - S_i \geqslant 0\} \cap \Omega_i\big) = \sum_{i=1}^{J} P(S_J - S_i \geqslant 0) \times P(\Omega_i) \geqslant \sum_{i=1}^{J} \frac{1}{2} P(\Omega_i) = \frac{1}{2}\big(1 - P(\Omega_0)\big)$$

puisque d'après IV.1., $P(S_J - S_i \geqslant 0) \geqslant \dfrac{1}{2}$ ($\dfrac{1}{2}$ est une médiane de $S_J - S_i$) et puisque les $(\Omega_i)_{0\leqslant i\leqslant J}$ forment un système complet d'événements.

D'après IV.4. et par transitivité de l'inégalité, on en déduit :

$$P(\overline{\Omega_0}) \leqslant 2P(S_J > x) \iff P\Big(\max_{1\leqslant j\leqslant J} S_j > x\Big) \leqslant 2P(S_J > x)$$

IV.7. D'après ce qui précède , avec $k = 10$: $\quad P\Big(\max_{1\leqslant j\leqslant 4} S_j > 15\Big) \leqslant 2P(S_4 > 15)$, où $S_4 = \displaystyle\sum_{i=1}^{4} X_i - 40$;

la variable aléatoire $\overline{X_4} = \displaystyle\sum_{i=1}^{4} X_i$ suit la loi de Poisson de paramètre 40, une valeur suffisamment grande pour qu'on puisse considérer que $\dfrac{\overline{X_4} - 40}{\sqrt{40}} = \dfrac{S_4}{\sqrt{40}}$ suit la loi lormale centrée, réduite, et écrire :

$$P(S_4 > 15) = P\Big(\frac{S_4}{\sqrt{40}} > \frac{15}{\sqrt{40}}\Big) \approx P\Big(T > \frac{15}{\sqrt{40}}\Big) \leqslant P(T > 2) \approx 2/5\% \quad \text{car } \frac{15}{\sqrt{40}} > \frac{15}{7} > 2$$

Donc : $\quad P\Big(\max_{1\leqslant j\leqslant 4} S_j > 15\Big) \leqslant 5\%$.

En admettant que $k = 10$, il y a donc moins de 5% de chances d'observer 15 décès au cours d'une des 4 premières années. La valeur $k = 10$ est donc très probablement sous-estimée.

★★★ FIN DU SUJET ★★★