Correction des Travaux dirigés de Probabilités et Statistiques - Lois discrètes usuelles et fonctions génératrices

Julian Tugaut*

*Si vous trouvez des erreurs de français ou de mathématiques ou bien si vous avez des questions et/ou des suggestions, envoyez-moi un mail à julian.tugaut@univ-st-etienne.fr

Table des matières

Exercice 1

Énoncé

X désigne une variable aléatoire réelle.

1. Soit $\mathcal{L}(X) = \mathcal{B}(20, 0.3)$. Calculer $\mathbb{P}(3 < X < 12)$ et $\mathbb{P}(X \leq 6)$.

2. Soit $\mathcal{L}(X) = \mathcal{B}(30, 0.3)$. Calculer $\mathbb{P}(4 \leq X < 8)$ et $\mathbb{P}(X < 5)$.

3. Soit $\mathcal{L}(X) = \mathcal{B}(30, 0.8)$. Calculer $\mathbb{P}(12 < X < 16)$ et $\mathbb{P}(X = 16)$.

Correction

Correction du 1)

Ici, $\mathcal{L}(X) = \mathcal{B}(20; 0.3)$. Il y a deux manières de calculer :

$$\begin{array}{rcl} \mathbb{P}(3 < X < 12) & = & \mathbb{P}(4 \leq X \leq 11) \\ & = & \mathbb{P}(X = 4) + \mathbb{P}(X = 5) + \mathbb{P}(X = 6) + \mathbb{P}(X = 7) \\ & + & \mathbb{P}(X = 8) + \mathbb{P}(X = 9) + \mathbb{P}(X = 10) + \mathbb{P}(X = 11) \\ & = & 0.1304 + 0.1789 + 0.1916 + 0.1643 \\ & + & 0.1144 + 0.0653 + 0.0309 + 0.0120 = 0.8878\,. \end{array}$$

Ou alors, on utilise la seconde colonne de la table :

$$\begin{array}{rcl} \mathbb{P}(3 < X < 12) & = & \mathbb{P}(3 < X \leq 11) = \mathbb{P}\Big(\{X \leq 11\} \bigcap \overline{\{X \leq 3\}}\Big) \\ & = & \mathbb{P}(X \leq 11) - \mathbb{P}(X \leq 3) = 0.9949 - 0.1071 = 0.8878\,. \end{array}$$

Puis $\mathbb{P}(X \leq 6) = 0.6080$.

Correction du 2)

On procède de même avec $\mathcal{L}(X) = \mathcal{B}(30; 0.3)$:

$$\begin{array}{rcl} \mathbb{P}(4 \leq X < 8) & = & \mathbb{P}(3 < X \leq 7) = \mathbb{P}\Big(\{X \leq 7\} \bigcap \overline{\{X \leq 3\}}\Big) \\ & = & \mathbb{P}(X \leq 7) - \mathbb{P}(X \leq 3) = 0.2814 - 0.0093 = 0.2721 \\ \text{et} \quad \mathbb{P}(X < 5) & = & \mathbb{P}(X \leq 4) = 0.0302\,. \end{array}$$

Correction du 3)

Ici, $\mathcal{L}(X) = \mathcal{B}(30; 0.8)$. Comme $0.8 > 0.5$, on ne peut utiliser directement les tables statistiques. Aussi, on pose $Y := 30 - X$ et l'on en déduit $\mathcal{L}(Y) = \mathbb{B}(30; 0.2)$ ainsi que $X = 30 - Y$:

$$\begin{array}{rcl} \mathbb{P}(X = 16) & = & \mathbb{P}(30 - Y = 16) = \mathbb{P}(Y = 14) = 0.0007 \\ \text{et} \quad \mathbb{P}(12 < X < 16) & = & \mathbb{P}(12 < 30 - Y < 16) \\ & = & \mathbb{P}(14 < Y < 18) = \mathbb{P}(14 < Y \leq 17) \\ & = & \mathbb{P}\Big(\{Y \leq 17\} \bigcap \overline{\{Y \leq 14\}}\Big) = \mathbb{P}(Y \leq 17) - \mathbb{P}(Y \leq 14) \\ & = & 1 - 0.9998 = 0.0002\,. \end{array}$$

Exercice 2

Énoncé

Soient deux variables aléatoires réelles X_1 et X_2, indépendantes. On suppose qu'elles suivent respectivement les lois $\mathcal{B}(10, 0.4)$ et $\mathcal{B}(20, 0.4)$.

Calculer $\mathbb{P}(X_1 + X_2 < 8)$.

Correction

Première méthode

Comme les variables aléatoires X_1 et X_2 sont indépendantes, $Y := X_1 + X_2$ suit la loi $\mathbb{B}(10 + 20; 0.4) = \mathbb{B}(30; 0.4)$. D'où $\mathbb{P}(X_1 + X_2 < 8) = \mathbb{P}(Y < 8) = \mathbb{P}(Y \leq 7) = 0.0435$.

Seconde méthode

On pourrait aussi obtenir le résultat sans utiliser le résultat du cours :

$$\begin{aligned}
\mathbb{P}(X_1 + X_2 < 8) &= \mathbb{P}(X_1 + X_2 \leq 7) \\
&= \sum_{k=0}^{7} \mathbb{P}(X_1 + X_2 = k) \\
&= \sum_{k=0}^{7} \sum_{j=0}^{k} \mathbb{P}(X_1 = j, X_1 + X_2 = k) \\
&= \sum_{k=0}^{7} \sum_{j=0}^{k} \mathbb{P}(X_1 = j, X_2 = k - j) \\
&= \sum_{k=0}^{7} \sum_{j=0}^{k} \mathbb{P}(X_1 = j)\, \mathbb{P}(X_2 = k - j) \ .
\end{aligned}$$

Néanmoins, cette méthode nécessite d'effectuer beaucoup de calculs. On choisira donc la première méthode.

Exercice 3

Énoncé

X désigne une variable aléatoire réelle.
1. Soit $\mathcal{L}(X) = \mathcal{P}(0.4)$. Calculer $\mathbb{P}(11 < X < 42)$.
2. Soit $\mathcal{L}(X) = \mathcal{P}(0.9)$. Calculer $\mathbb{P}(2 \leq X \leq 4)$ et $\mathbb{P}(X \geq 4)$.

Correction

Correction du 1)

Ici, $\mathcal{L}(X) = \mathcal{P}(0.4)$. On utilise à nouveau la seconde colonne du tableau :

$$\begin{aligned} \mathbb{P}(11 < X < 42) &= \mathbb{P}(11 < X \leq 41) = \mathbb{P}\left(\{X \leq 41\} \bigcap \overline{\{X \leq 11\}}\right) \\ &= \mathbb{P}(X \leq 41) - \mathbb{P}(X \leq 11) = 1 - 1 = 0\,. \end{aligned}$$

Correction du 2)

Ici, $\mathcal{L}(X) = \mathcal{P}(0.9)$. On procède de la même façon :

$$\begin{aligned} \mathbb{P}(2 \leq X \leq 4) &= \mathbb{P}(1 < X \leq 4) = \mathbb{P}\left(\{X \leq 4\} \bigcap \overline{\{X \leq 1\}}\right) \\ &= \mathbb{P}(X \leq 4) - \mathbb{P}(X \leq 1) = 0.9977 - 0.7725 = 0.2252\,. \end{aligned}$$

Puis, $\mathbb{P}(X \geq 4) = 1 - \mathbb{P}(X \leq 3) = 1 - 0.9866 = 0.0134$.

Exercice 4

Énoncé

Soient deux variables aléatoires réelles X_1 et X_2, indépendantes. On suppose qu'elles suivent respectivement les lois $\mathcal{P}(0.4)$ et $\mathcal{P}(0.6)$.

Calculer $\mathbb{P}(X_1 + X_2 < 4)$.

Correction

Première méthode

Comme les variables aléatoires X_1 et X_2 sont indépendantes, $Y := X_1 + X_2$ suit la loi $\mathcal{P}(0.4 + 0.6) = \mathcal{P}(1)$. D'où $\mathbb{P}(X_1 + X_2 < 4) = \mathbb{P}(Y < 4) = \mathbb{P}(Y \leq 3) = 0.9810$.

Seconde méthode

On pourrait aussi obtenir le résultat sans utiliser le résultat du cours :

$$\begin{aligned}
\mathbb{P}(X_1 + X_2 < 4) &= \mathbb{P}(X_1 + X_2 \leq 3) \\
&= \sum_{k=0}^{3} \mathbb{P}(X_1 + X_2 = k) \\
&= \sum_{k=0}^{3} \sum_{j=0}^{k} \mathbb{P}(X_1 = j, X_1 + X_2 = k) \\
&= \sum_{k=0}^{3} \sum_{j=0}^{k} \mathbb{P}(X_1 = j, X_2 = k - j) \\
&= \sum_{k=0}^{3} \sum_{j=0}^{k} \mathbb{P}(X_1 = j)\, \mathbb{P}(X_2 = k - j)\ .
\end{aligned}$$

Néanmoins, cette méthode nécessite d'effectuer beaucoup de calculs. On choisira donc la première méthode.

Exercice 5

Énoncé

Soient X et Y deux variables aléatoires réelles telles que X suit la loi $\mathcal{B}(n,p)$ et Y suit la loi $\mathcal{P}(\lambda)$.

1. *Calculer l'espérance et la variance de X en utilisant sa fonction génératrice.*

2. *Calculer l'espérance et la variance de Y en utilisant sa fonction génératrice.*

Rappels

On rappelle que l'on a

$$G_X(s) = (ps + 1 - p)^n$$

ainsi que

$$G_Y(s) = \exp\left[\lambda(s-1)\right] .$$

Remarque

On pourrait calculer l'espérance ainsi que la variance de X et de Y sans passer par les fonctions génératrices.

Correction

Correction du 1)

On sait que l'on a

$$G'_X(1) = \mathbb{E}[X] .$$

Donc, $\mathbb{E}[X] = G'_X(1)$. Or le calcul nous donne

$$G'_X(s) = np\,(ps + 1 - p)^{n-1} .$$

D'où $G'_X(1) = np$. Il vient $\mathbb{E}[X] = np$.
De même, on a

$$G''_X(1) = \mathbb{E}[X^2] - \mathbb{E}[X] .$$

Donc, $\mathbb{E}[X^2] = G''_X(1) + G'_X(1)$. Or le calcul nous donne

$$G''_X(s) = n(n-1)p^2\,(ps + 1 - p)^{n-2} .$$

D'où $G''_X(1) = n(n-1)p^2$. Il vient $\mathbb{E}[X^2] = n(n-1)p^2 + np = np((n-1)p+1)$. Puis, il s'ensuit

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = np((n-1)p+1) - n^2p^2 = np(1-p) .$$

Correction du 2)

On procède de même avec la loi de Poisson de paramètre λ.

On sait que l'on a

$$G'_Y(1) = \mathbb{E}[Y]\,.$$

Donc, $\mathbb{E}[Y] = G'_Y(1)$. Or le calcul nous donne

$$G'_Y(s) = \lambda \exp\left[\lambda\,(s-1)\right]\,.$$

D'où $G'_Y(1) = \lambda$. Il vient $\mathbb{E}[X] = \lambda$.

De même, on a

$$G''_Y(1) = \mathbb{E}[Y^2] - \mathbb{E}[Y]\,.$$

Donc, $\mathbb{E}[Y^2] = G''_Y(1) + G'_Y(1)$. Or le calcul nous donne

$$G''_Y(s) = \lambda^2 \exp\left[\lambda\,(s-1)\right]\,.$$

D'où $G''_Y(1) = \lambda^2$. Il vient $\mathbb{E}[Y^2] = \lambda^2 + \lambda$. Puis, il s'ensuit

$$\mathrm{Var}[Y] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda\,.$$

Exercice 6

Énoncé

1. Soit U suivant une loi uniforme discrète sur $[\![1;n]\!]$ avec $n \geq 1$ c'est-à-dire que l'on a $\mathbb{P}_U = \frac{1}{n}\sum_{k=1}^{n} \delta_k$. Calculer la fonction génératrice de U.
2. On lance deux dés (à six faces) de façon indépendante et on note S le résultat de la somme des deux résultats obtenus. Calculer la fonction génératrice de S.

Correction

Correction du 1)

Par définition, $G_U(t) := \mathbb{E}\left(t^U\right) = \sum_{k=1}^{n} t^k \mathbb{P}(U = k)$. Or, ici, $\mathbb{P}(U = k) = \frac{1}{n}$ pour tout $k \in [\![1;n]\!]$. Il s'ensuit

$$G_U(t) = \frac{1}{n}\sum_{k=1}^{n} t^k = \frac{t}{n}\frac{1-t^n}{1-t} \, ,$$

où la dernière égalité n'est valable que si $t \neq 1$.

Correction du 2)

Notons U_1 et U_2 les résultats du premier dé et du deuxième dé respectivement. Alors U_1 et U_2 sont indépendantes et de même loi à savoir la loi uniforme discrète sur $[\![1;6]\!]$. Il vient $G_{U_1}(t) = G_{U_2}(t) = \frac{t}{6}\frac{1-t^6}{1-t}$, pour $t \neq 1$.

Puis, comme $S = U_1 + U_2$, on a par indépendance de U_1 et U_2 : $G_S(t) = G_{U_1}(t)G_{U_2}(t) = \frac{t^2}{36}\left(\frac{1-t^6}{1-t}\right)^2$, pour $t \neq 1$.

Remarque 1

Pour la seconde question, on aurait également pu calculer la loi de probabilité de S :

$$\begin{aligned}
\mathbb{P}_S &= \mathbb{P}_{U_1} * \mathbb{P}_{U_2} \\
&= \frac{1}{36}\left(\delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5 + \delta_6\right) * \left(\delta_1 + \delta_2 + \delta_3 + \delta_4 + \delta_5 + \delta_6\right) \\
&= \frac{1}{36}\left(\delta_2 + 2\delta_3 + 3\delta_4 + 4\delta_5 + 5\delta_6 + 6\delta_7 + 5\delta_8 + 4\delta_9 + 3\delta_{10} + 2\delta_{11} + \delta_{12}\right) \, .
\end{aligned}$$

Puis, en appliquant la définition de G_S :

$$G_S(t) = \frac{1}{36}\left(t^2 + 2t^3 + 3t^4 + 4t^5 + 5t^6 + 6t^7 + 5t^8 + 4t^9 + 3t^{10} + 2t^{11} + t^{12}\right) \, .$$

Ce résultat est cohérent avec le précédent car $\frac{1-t^6}{1-t} = 1 + t + t^2 + t^3 + t^4 + t^5$, pour $t \neq 1$. Ainsi, pour tout t différent de 1 :

$$
\begin{aligned}
G_S(t) &= \frac{t^2}{36}\left(\frac{1-t^6}{1-t}\right)^2 \\
&= \frac{t^2}{36}\left(1+t+t^2+t^3+t^4+t^5\right)^2 \\
&= \frac{t^2}{36}\left(1+2t+3t^2+4t^3+5t^4+6t^5+5t^6+4t^7+3t^8+2t^9+t^{10}\right) \\
&= \frac{1}{36}\left(t^2+2t^3+3t^4+4t^5+5t^6+6t^7+5t^8+4t^9+3t^{10}+2t^{11}+t^{12}\right) \\
&= G_S(t)\,.
\end{aligned}
$$

Remarque 2

On pourrait calculer l'espérance et la variance de S à partir de sa fonction génératrice. En effet, G_S est bien définie en 1 par continuité. Il en est de même pour G'_S et G''_S.

Exercice 7 (*)

Énoncé

On modélise la désintégration radioactive d'un élément instable par un PPP (Processus Ponctuel de Poisson). Pour tout intervalle de temps $\mathcal{I}$, $N_{\mathcal{I}}$ désigne le nombre de désintégrations durant l'intervalle $\mathcal{I}$. Ce processus admet les trois propriétés fondamentales suivantes :

1. Soient $\mathcal{I}_1$ et $\mathcal{I}_2$ deux intervalles disjoints. Alors, les évènements "il y a k_1 désintégrations dans I_1" et "il y a k_2 désintégrations dans I_2" sont indépendants pour tous k_1, k_2.
2. La probabilité de désintégration au cours d'un intervalle de temps infiniment petit $[t; t+h]$ est petite d'ordre 1 : $\lambda(t)h + o(h)$.
3. La probabilité de plus d'une désintégration au cours d'un intervalle de temps petit $[t; t+h]$ est petite d'ordre strictement supérieur à 1 : $o(h)$.

Le processus peut se déterminer de la manière suivante. Soit $P_n(t_1, t_2) := \mathbb{P}\left(N_{[t_1,t_2]} = n\right)$ la probabilité de n désintégrations dans l'intervalle $[t_1; t_2]$. Alors, on a

$$\mathbb{P}\left(N_{[t,t+h]} = 1\right) = P_1(t, t+h) = \lambda(t)h + o(h)\,,$$

ainsi que

$$\sum_{n=2}^{\infty} \mathbb{P}\left(N_{[t,t+h]} = n\right) = \sum_{n=2}^{\infty} P_n(t, t+h) = o(h)\,.$$

Ici, λ est une fonction positive. Soit t_0 l'instant initial. On pose $P_n(t) := P_n(t_0, t)$ pour tout $t \geq t_0$.

1. *En calculant $P_0(t+h)$ avec h petit, trouver une relation entre P_0' et P_0.*

2. *En calculant $P_m(t+h)$ avec h petit, trouver une relation entre P_m', P_m, P_{m-1} et λ.*

3. *Résoudre cette équation différentielle pour $m = 0$, $m = 1$ et $m = 2$. Puis, généraliser la solution. On peut procéder à un raisonnement par récurrence.*

On suppose maintenant que la fonction λ est constante : $\lambda(t) = \lambda$ et $t_0 := 0$.

4. *Montrer que l'on a alors : $P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$.*

5. *Donner l'espérance mathématique et la variance de cette loi.*

On suppose maintenant $\lambda := 10^5$.

6. *Quelle est la probabilité qu'il y ait exactement 10^6 désintégrations durant un laps de temps de dix secondes ?*

7. *Trouver un majorant simple de la probabilité que le nombre de désintégrations en dix secondes soit d'au moins 2.10^6.*

Correction

Correction du 1)

On calcule $P_0(t+h) = \mathbb{P}\left(N_{[t_0,t+h]} = 0\right) = \mathbb{P}\left(N_{[t_0,t]} = 0, N_{[t,t+h]} = 0\right)$:

$$\mathbb{P}\left(N_{[t_0,t+h]} = 0\right) = \mathbb{P}\left(N_{[t_0,t]} = 0\right)\mathbb{P}\left(N_{[t,t+h]} = 0\right)\,,$$

d'après la première propriété (l'indépendance des évènements). Ainsi :

$$P_0(t+h) - P_0(t) = P_0(t)\left(P_0(t,t+h) - 1\right) = -P_0(t)\sum_{j=1}^{\infty} P_j(t,t+h) = -P_0(t)\lambda(t)h + o(h)\,.$$

On en déduit

$$P_0'(t) = -\lambda(t)P_0(t)\,.$$

Correction du 2)

On procède de même. Il y a m désintégrations dans l'intervalle $[t_0; t+h]$ s'il y a m désintégrations dans l'intervalle $[t_0; t]$ et 0 dans l'intervalle $[t; t+h]$ ou (exclusif) $m-1$ dans l'intervalle $[t_0; t]$ et 1 dans l'intervalle $[t; t+h]$. On néglige le cas où il y a deux désintégrations dans l'intervalle $[t; t+h]$ d'après la propriété trois :

$$P_m(t+h) = P_m(t)P_0(t,t+h) + P_{m-1}(t)P_1(t,t+h) + o(h)\,.$$

On en déduit ainsi

$$\begin{aligned}\frac{P_m(t+h) - P_m(t)}{h} &= P_m(t)\frac{P_0(t,t+h) - 1}{h} + P_{m-1}(t)\frac{P_1(t,t+h)}{h} + o(1)\\ &= \lambda(t)\left(P_{m-1}(t) - P_m(t)\right) + o(1)\,.\end{aligned}$$

Conséquemment, on a

$$P_m'(t) = \lambda(t)\left(P_{m-1}(t) - P_m(t)\right)\,.$$

Correction du 3)

Pour $m = 0$, On est amené à résoudre

$$P_0'(t) = -\lambda(t)P_0(t)\,,$$

ce qui donne $P_0(t) = Ce^{-\Lambda(t)}$ où Λ est la primitive de λ qui s'annule en t_0. Or, $P_0(t_0) = 1$ d'où $P_0(t) = e^{-\Lambda(t)}$. L'équation en $m = 1$ est alors

$$P_1'(t) = -\lambda(t)P_1(t) + \lambda(t)P_0(t)\,.$$

La solution homogène est $e^{-\Lambda(t)}$. Une solution particulière est immédiatement $\Lambda(t)e^{-\Lambda(t)}$. Puis, comme $P_1(t_0) = 0$, il vient

$$P_1(t) = \Lambda(t)e^{-\Lambda(t)}\,.$$

Puis de même, pour $m = 2$, on doit résoudre :

$$P_2'(t) = -\lambda(t)P_2(t) + \lambda(t)\Lambda(t)e^{-\Lambda(t)}\,.$$

On trouve $P_2(t) = \frac{\Lambda(t)^2}{2}e^{-\Lambda(t)}$. On peut prouver par récurrence que l'on a $P_n(t) = \frac{\Lambda(t)^n}{n!}e^{-\Lambda(t)}$.

Correction du 4)

Ici, $\lambda(t) = \lambda$ est une constante donc Λ, sa primitive nulle en $t_0 = 0$, vaut $\Lambda(t) = \lambda t$. On trouve donc

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t},$$

la loi de Poisson de paramètre λt.

Correction du 5)

On calcule d'abord le moment d'ordre un. Soit N_t la variable aléatoire correspondant au nombre de désintégrations en un temps t. On calcule son espérance comme suit.

$$\mathbb{E}[N_t] = \sum_{n=0}^{\infty} nP_n(t) = \sum_{n=1}^{\infty} nP_n(t) = \sum_{n=1}^{\infty} \frac{(\lambda t)^n}{(n-1)!} e^{-\lambda t} = \lambda t \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \lambda t\,.$$

De la même manière, on trouve

$$\mathbb{E}[N_t(N_t - 1)] = \sum_{n=0}^{\infty} n(n-1)P_n(t) = \sum_{n=2}^{\infty} \frac{(\lambda t)^n}{(n-2)!} e^{-\lambda t} = (\lambda t)^2 \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} = (\lambda t)^2\,.$$

On en déduit $\mathbb{E}[N_t^2] = (\lambda t)^2 + \lambda t$ d'où $\sigma^2[N_t] = \lambda t = \mathbb{E}[N_t]$. On aurait également pu dire directement $\mathbb{E}[N_t] = \mathrm{Var}[N_t] = \lambda t$ puisque N_t suit la loi de Poisson de paramètre λt.

Correction du 6)

On cherche ici à calculer $P_{10^6}(10)$ avec $\lambda := 10^5$. Cette probabilité est égale à

$$P_{10^6}(10) = \frac{\left(10^6\right)^{10^6}}{10^6!} e^{-10^6}\,.$$

On utilise la formule de Stirling : $n! \approx n^n e^{-n}\sqrt{2\pi n}$ et il vient $P_{10^6}(10) \approx \frac{10^{-3}}{\sqrt{2\pi}} \approx 3.99 \times 10^{-4}$.

Correction du 7)

On applique l'inégalité de Markov :

$$\mathbb{P}\left(N_{10} \geq 2.10^6\right) \leq \frac{\mathbb{E}[N_{10}]}{2.10^6} = \frac{10\lambda}{2.10^6} = \frac{1}{2}\,.$$

On peut trouver une majoration plus précise avec l'inégalité de Bienaymé-Tchebychev :

$$\mathbb{P}\left(N_{10} \geq 2.10^6\right) \leq \mathbb{P}\left(|N_{10} - \mathbb{E}[N_{10}]| \geq 10^6\right) \leq \frac{\sigma^2[N_{10}]}{\left(10^6\right)^2} = \frac{\lambda 10}{10^{12}} = 10^{-6}\,.$$

On pourrait obtenir une majoration encore meilleure en utilisant les grandes déviations (que l'on ne voit pas dans ce cours) :

$$\mathbb{P}\left(N_{10} \geq 2.10^6\right) \leq 2\exp\left(-(2\log(2) - 1).10^6\right)\,.$$

Exercice 8 (*)

Énoncé

Une poule pond N œufs, et N est une distribution de Poisson de paramètre λ. Chaque œuf éclôt avec probabilité p et les éclosions sont des évènements indépendants. Soit K le nombre de poussins. Calculer la fonction génératrice de K et en déduire que K suit une loi de Poisson de paramètre λp.

Correction

On a $G_X(s) = ps + (1-p)$ où X vaut 1 si l'œuf éclôt et 0 sinon et $G_N(s) = e^{\lambda(s-1)}$. On va montrer que $G_K(s) = G_N(G_X(s)) = e^{\lambda(ps+(1-p)-1)} = e^{\lambda p(s-1)}$ ce qui impliquera immédiatement que K suit une loi de Poisson de paramètre λp.

Puisque K prend ses valeurs dans $\mathbb{N}$, on peut écrire pour tout $s \in [-1;1]$:

$$G_K(s) = \sum_{k=0}^{\infty} s^k \mathbb{P}(K=k) \ ,$$

ce qui peut aussi s'écrire

$$G_K(s) = \sum_{k\in\mathbb{N}} s^k \sum_{n\in\mathbb{N}} \mathbb{P}(K=k, N=n) \ .$$

Mais,

$$\sum_{k\in\mathbb{N}} |s|^k \sum_{n\in\mathbb{N}} \mathbb{P}(K=k, N=n) \leq \sum_{k\in\mathbb{N}} \sum_{n\in\mathbb{N}} \mathbb{P}(K=k, N=n) = 1 < +\infty \, .$$

La famille $\left\{s^k \mathbb{P}(K=k, N=n)\right\}_{(k,n)\in\mathbb{N}\times\mathbb{N}}$ est donc sommable, et d'après la propriété de Fubini, on a :

$$G_K(s) = \sum_{n\in\mathbb{N}} \sum_{k\in\mathbb{N}} s^k \mathbb{P}(K=k, N=n) \ .$$

Mais, on a l'inclusion $\{N=0\} \subset \{K=0\}$, et, par conséquent :

$$G_K(s) = \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*} \sum_{k\in\mathbb{N}} s^k \mathbb{P}(K=k, N=n) \ .$$

Si, pour $n \geq 1$, on note $K_n = \sum_{j=1}^n X_j$, on a $\{K=k\} \cap \{N=n\} = \{K_n = k\} \cap \{N=n\}$ et donc :

$$G_K(s) = \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*} \sum_{k\in\mathbb{N}} s^k \mathbb{P}(K_n=k, N=n) \ .$$

En tenant compte de l'indépendance de K_n et N, il vient :

$$\begin{aligned}
G_K(s) &= \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*} \left[\sum_{k\in\mathbb{N}} s^k \mathbb{P}(K_n=k)\right] \mathbb{P}(N=n) \\
&= \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*} G_{K_n}(s) \mathbb{P}(N=n) \\
&= \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*} (G_X(s))^n \, \mathbb{P}(N=n) \\
&= G_N[G_X(s)] \ .
\end{aligned}$$

Remarque

Il était possible de trouver la loi de probabilité de S sans passer par la fonction génératrice : en utilisant à nouveau le système complet d'évènements dont on s'est servi plus haut.

Exercice 9 (*)

Énoncé

Le nombre N de clients entrant pendant une journée dans un grand magasin est une variable aléatoire de loi de Poisson de paramètre $\lambda > 0$. Les probabilités respectives pour chaque client d'acheter zéro, un ou deux articles de marque A sont $\frac{1}{6}$, $\frac{1}{2}$ et $\frac{1}{3}$. Le nombre d'articles en question achetés pendant une journée est une variable aléatoire S. On étudie la loi de S grâce à sa fonction génératrice.

On modélise le problème de façon plus précise de la manière suivante. Soit une variable aléatoire N de loi de Poisson de paramètre $\lambda > 0$ et une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ indépendantes, de même loi (X_n représente le nombre d'articles achetés par le n-ième client) définie par $\mathbb{P}X_1^{-1} = \frac{1}{6}\delta_0 + \frac{1}{2}\delta_1 + \frac{1}{3}\delta_2$. On suppose de plus que N et les X_n, $n \in \mathbb{N}^*$, constituent une famille de variables aléatoires indépendantes. On définit enfin la variable aléatoire S par

$$S := \mathbb{1}_{\{N\geq 1\}} \sum_{j=1}^{N} X_j \, ,$$

où on fait la convention d'écriture $\sum_{j=1}^{0} = 0$.

1. *Calculer la fonction génératrice G_S de la variable aléatoire S, en tout $s \in [-1;1]$.*

2. *En déduire la probabilité $\mathbb{P}(S = 3)$ et la calculer numériquement dans le cas $\lambda = 6$.*

3. *Justifier l'existence de la moyenne $\mathbb{E}[S]$ et de la variance σ_S^2. Les calculer numériquement dans le cas $\lambda = 6$.*

Correction

Correction du 1)

Puisque S prend ses valeurs dans $\mathbb{N}$, on peut écrire pour tout $s \in [-1;1]$:

$$G_S(s) = \sum_{k=0}^{\infty} s^k \mathbb{P}\left(S = k\right) \, ,$$

ce qui peut aussi s'écrire

$$G_S(s) = \sum_{k\in\mathbb{N}} s^k \sum_{n\in\mathbb{N}} \mathbb{P}\left(S = k, N = n\right) \, .$$

Mais,

$$\sum_{k\in\mathbb{N}} |s|^k \sum_{n\in\mathbb{N}} \mathbb{P}\left(S = k, N = n\right) \leq \sum_{k\in\mathbb{N}} \sum_{n\in\mathbb{N}} \mathbb{P}\left(S = k, N = n\right) = 1 < +\infty \, .$$

La famille $\left\{s^k \mathbb{P}\left(S = k, N = n\right)\right\}_{(k,n)\in\mathbb{N}\times\mathbb{N}}$ est donc sommable, et d'après la propriété de Fubini, on a :

$$G_S(s) = \sum_{n\in\mathbb{N}} \sum_{k\in\mathbb{N}} s^k \mathbb{P}\left(S = k, N = n\right) \, .$$

Mais, on a l'inclusion $\{N=0\} \subset \{S=0\}$, et, par conséquent :

$$G_S(s) = \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*}\sum_{k\in\mathbb{N}} s^k \mathbb{P}\left(S=k, N=n\right) \, .$$

Si, pour $n \geq 1$, on note $S_n = \sum_{j=1}^n X_j$, on a $\{S=k\}\cap\{N=n\} = \{S_n=k\}\cap\{N=n\}$ et donc :

$$G_S(s) = \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*}\sum_{k\in\mathbb{N}} s^k \mathbb{P}\left(S_n=k, N=n\right) \, .$$

En tenant compte de l'indépendance de S_n et N, il vient :

$$\begin{aligned} G_S(s) &= \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*}\left[\sum_{k\in\mathbb{N}} s^k \mathbb{P}\left(S_n=k\right)\right]\mathbb{P}\left(N=n\right) \\ &= \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*} G_{S_n}(s)\mathbb{P}\left(N=n\right) \\ &= \mathbb{P}(N=0) + \sum_{n\in\mathbb{N}^*} \left(G_{X_1}(s)\right)^n \mathbb{P}\left(N=n\right) \\ &= G_N\left[G_{X_1}(s)\right] \, . \end{aligned}$$

On rappelle que $G_N(s) = e^{\lambda(s-1)}$. Par ailleurs, on a :

$$G_{X_1}(s) = \mathbb{E}\left[s^{X_1}\right] = \frac{1}{6} + \frac{1}{2}s + \frac{1}{3}s^2 \, .$$

Conséquemment, on a

$$G_S(s) = \exp\left\{\lambda\left(-\frac{5}{6} + \frac{1}{2}s + \frac{1}{3}s^2\right)\right\} \, .$$

Correction du 2)

De l'unicité du développement en série entière dans $]-1;1[$ de la fonction G_S résulte que l'on a :

$$\mathbb{P}(S=3) = \frac{G_S^{(3)}(0)}{3!} \, .$$

Or, $G_S(s) = e^{f(s)}$ avec $f(s) = \lambda\left(-\frac{5}{6} + \frac{1}{2}s + \frac{1}{3}s^2\right)$. D'où

$$\begin{aligned} & G_S'(s) = f'(s)e^{f(s)} \, , \\ \text{puis} \quad & G_S''(s) = \left(f''(s) + (f'(s))^2\right)e^{f(s)} \\ \text{et} \quad & G_S^{(3)}(s) = \left(f^{(3)}(s) + 3f'(s)f''(s) + (f'(s))^3\right)e^{f(s)} \, . \end{aligned}$$

Or, $f'(s) = \lambda(\frac{1}{2} + \frac{2}{3}s)$, $f''(s) = \frac{2}{3}\lambda$ puis $f^{(3)}(s) = 0$. Il s'ensuit : $f'(0) = \frac{1}{2}\lambda$, $f''(0) = \frac{2}{3}\lambda$ et $f^{(3)}(0) = 0$. Conséquemment :

$$f^{(3)}(0) + 3f'(0)f''(0) + (f'(0))^3 = \lambda^2 + \frac{1}{8}\lambda^3 \, ,$$

ce qui nous donne :

$$\mathbb{P}(S=3) = \frac{\lambda^2}{6}\left[1+\frac{\lambda}{8}\right]\exp\left[-\frac{5}{6}\lambda\right] .$$

Dans le cas où $\lambda = 6$, on obtient $\mathbb{P}(S=3) = \frac{21}{2}e^{-5} \approx 0.067$.

Correction du 3)

On a $0 \leq S \leq 2N$. Or, la variable aléatoire N admet des moments de tout ordre donc S aussi. La formule de Wald peut être appliquée dans ce contexte puisque l'on a $G_S(s) = G_N\left(G_{X_1}(s)\right)$. On obtient alors

$$\mathbb{E}[S] = \mathbb{E}[N] \times \mathbb{E}[X_1] = \frac{7}{6}\lambda .$$

Dans le cas $\lambda = 6$, on obtient $\mathbb{E}[S] = 7$.
De même, on a $\sigma_S^2 = \frac{11}{6}\lambda$. Dans le cas $\lambda = 6$, on obtient $\sigma_S^2 = 11$.

Exercice 10 (*)

Énoncé

On étudie la propagation d'un nom de famille à travers les générations. On note Z_n le nombre d'hommes de la génération n portant ce nom. On suppose $Z_0 = 1$. On suppose que le $i^{\text{ème}}$ homme donne naissance à X_i garçons portant le même nom. On suppose que les variables aléatoires X_i sont indépendantes et de même loi. On note G la fonction génératrice de X_1. On note H_n celle de Z_n. On suppose par ailleurs qu'il n'y a ni immigration ni mortalité infantile. Le but de l'exercice est de déterminer la probabilité que le nom s'éteigne, c'est-à-dire la probabilité $p_n = \mathbb{P}(Z_n = 0)$.

1. Donner une relation de récurrence reliant H_n à H_{n-1}.
2. Exprimer p_{n+1} en fonction de p_n.
3. Montrer que l'on a $\lim\limits_{n\to+\infty} p_n = x_0$ où $x_0 := \inf\{x \in [0;1] \ : \ G(x) = x\}$.
4. Montrer que si $\mathbb{E}(X_1) < 1$, alors $p_n \longrightarrow 1$ quand n tend vers l'infini.

Correction

Correction du 1)

Les hommes de la génération n étant issus de la génération $n-1$, on a

$$Z_n = X_1 + \cdots + X_{Z_{n-1}}\,.$$

En d'autres termes, Z_n est une somme aléatoire de Z_{n-1} variables aléatoires de même loi. On a ainsi pour tout $t \in [0;1[$:

$$G_{Z_n}(t) = G_{Z_{n-1}}(G(t))\ ,$$

où G est la fonction génératrice de X_i pour tout $i \in \mathbb{N}$. Or, $Z_1 = X_1$ vu qu'il y a exactement un homme à la génération 0. Il s'ensuit

$$G_{Z_1}(t) = G(t)\,.$$

Puis, en procédant à une récurrence, on obtient $H_n = H_{n-1} \circ G$ d'où

$$H_n(t) = \Bigg(\underbrace{G \circ \cdots \circ G}_{n \text{ fois}}\Bigg)(t)\,.$$

Correction du 2)

On a $\mathbb{P}(Z_n = 0) = H_n(0)$ par propriété sur la fonction génératrice. On en déduit $p_{n+1} = H_{n+1}(0) = \Bigg(\underbrace{G \circ \cdots \circ G}_{n+1 \text{ fois}}\Bigg)(0) = G\left(\Bigg(\underbrace{G \circ \cdots \circ G}_{n \text{ fois}}\Bigg)(0)\right) = G(H_n(0)) = G(p_n)$.

Par conséquent, la suite $(p_n)_n$ est définie par $p_0 = G(0) \in \mathbb{R}_+$ et $p_{n+1} = G(p_n)$.

Correction du 3)

Commençons par montrer que la suite $(p_n)_n$ est convergente. En tant que probabilité d'un évènement, on sait : $p_n \leq 1$. La suite est donc majorée. Or, comme la fonction génératrice est toujours croissante et comme $p_0 \geq 0$, on a $p_1 = G(p_0) \geq G(0) = p_0$. De manière générale, en procédant à une récurrence, on peut montrer $p_{n+1} \geq p_n$. La suite est donc croissante. Comme elle est majorée, elle est donc convergente. Notons l la limite.

Comme la fonction G est continue sur $[0;1]$, par passage à la limite, on obtient $G(l) = l$.

Il reste maintenant à montrer que toute solution de l'équation $G(x) = x$ majore la suite $(p_n)_n$. Soit ainsi $\lambda \in [0;1]$ tel que $G(\lambda) = \lambda$. Alors $\lambda = G(\lambda) \geq G(0) = p_0$. Puis, $\lambda = G(\lambda) \geq G(p_0) = p_1$. On peut ainsi montrer par récurrence $\lambda \geq p_n$ pour tout $n \in \mathbb{N}$. On en déduit par passage à la limite : $\lambda \geq l$. Par conséquent, $l = x_0 := \inf \{x \in [0;1] \ : \ G(x) = x\}$, ce qui achève de prouver la convergence de la suite de terme général p_n vers x_0.

Correction du 4)

D'après les propriétés sur les fonctions génératrices, $\mathbb{E}[X] = G'(1)$. Or, $\mathbb{E}[X] < 1$. Posons $\psi(t) := G(t) - t$. Alors, $\psi'(1) = G'(1) - 1 < 0$. Or, $\psi''(t) = G''(t) = \mathbb{E}[X(X-1)t^{X-2}]$. Comme $X \in \mathbb{N}$, $X(X-1) \geq 0$ et donc $\psi''(x) \geq 0$. Il s'ensuit $\psi'(t) < 0$ pour tout $t \in [0;1]$ d'où la fonction ψ est décroissante. Or, $\psi(1) = G(1) - 1 = 0$. Conséquemment, $\{x \in [0;1] \ : \ G(x) = x\} = \{1\}$ donc $x_0 = 1$ et il vient

$$\lim_{n \to +\infty} p_n = 1 \,.$$