SY01 A09 Corrigé de l'Examen médian

Exercice 1

Une famille compte 4 membres, qui peuvent être nés n'importe quel jour de l'année avec la même probabilité, et de manière indépendante les uns des autres.

(1) Quelle est la probabilité qu'au moins deux membres de la famille soient nés le même jour?

Correction (1,5 pt). L'espace fondamental est

$$\Omega = \{\text{quadruplets de jours d'anniversaires}\} = \{\text{4-listes d'éléments pris parmi 365}\}.$$

Son cardinal est donc 365^4. Soit A, l'évènement "Au moins deux membres sont nés le même jours". On va passer par l'événement complémentaire de A, qui est $\bar{A} =$ "les membres sont nés 4 jours différents". Donc, $\bar{A}$ est l'ensemble des sous-familles ordonnées de 4 éléments distincts pris parmi les 365 (la première date est l'anniversaire de l'individu 1, la 2ème, celui de l'individu 2, etc...), i.e. l'ensemble des arrangements de 4 éléments parmi 365. On a donc $\text{Card}\bar{A} = A_{365}^4$. Finalement, en vertu de l'hypothèse d'équiprobabilité,

$$\mathbf{P}[A] = 1 - \mathbf{P}\left[\bar{A}\right] = 1 - \frac{\text{Card}\bar{A}}{\text{Card}\Omega} = 1 - \frac{A_{365}^4}{365^4}.$$

□

(2) Je décide de sonner chez eux le 7 novembre, avec un cadeau dans les mains (on ne sait jamais!). Quelle est la probabilité qu'au moins un des membres de la famille fête effectivement son anniversaire ce jour-là?

Correction (1,5 pt). Sur le même espace de probabilité, on considère maintenant l'événement $B =$ "au moins un membre est né le 7/11". Là-encore, passons par son évènement complémentaire, i.e. $\bar{B} =$ "tous les membres sont nés un autre jour que le 7/11". Alors, $\bar{B}$ est l'ensemble des 4-listes d'éléments (donc, non forcément distincts) pris parmi les 364 autres que le 7/11. On a donc $\text{Card}\bar{B} = (364)^4$, et donc

$$\mathbf{P}[B] = 1 - \mathbf{P}\left[\bar{B}\right] = 1 - \frac{\text{Card}\bar{B}}{\text{Card}\Omega} = 1 - \left(\frac{364}{365}\right)^4.$$

□

Exercice 2

On répète, de manière indépendante, la même expérience aléatoire, pour laquelle une issue possible est donnée par l'évènement A, qui est de probabilité p (où $0 < p < 1$). On s'intéresse à la v.a. suivante: X compte la *deuxième* fois que l'évènement A se réalise.

(1) Quel est l'espace d'état E_X de X?

Correction (0,5 pt). Comme X compte le n° du la deuxième fois que A se réalise, X vaut au moins 2, et peut prendre après cela n'importe quelle valeur entière. Donc $E_X = \{2, 3, ...\} = \mathbb{N}^* \setminus \{1\}$. □

(2) Soit pour tout entier $i \geq 1$, B_i l'évènement "A se réalise la i-ème fois". Exprimer, pour toute valeur possible $k \in E_X$, l'évènement $\{X = k\}$ en fonction des B_i, $1 \leq i \leq k$.

Correction (1,5 pt). L'événement $\{X = k\}$ revient à dire que A s'est réalisé la k-ième fois (donc on est sur B_k), et qu'il s'est réalisé une fois, et seulement une fois auparavant. Donc, on est aussi sur l'un des $B_i, i = 1, ..., k-1$, et sur tous les $\bar{B}_j$, pour $j \neq i$. En d'autre terme, on a

$$\begin{aligned}\{X = k\} &= B_k \cap \left\{ \bigcup_{i=1}^{k-1} \left(B_i \cap \left(\bigcap_{j=1; j\neq i}^{k-1} \bar{B}_j \right) \right) \right\} \\ &= \bigcup_{i=1}^{k-1} \left(B_k \cap B_i \cap \left(\bigcap_{j=1; j\neq i}^{k-1} \bar{B}_j \right) \right)\end{aligned}$$

□

(3) En déduire la loi de X.

Correction (1,5 pt). On a donc, pour tout $k \geq 2$,

$$\mathbf{P}[X = k] = \sum_{i=1}^{k-1} \mathbf{P}\left[B_k \cap B_i \cap \left(\bigcap_{j=1; j\neq i}^{k-1} \bar{B}_j \right) \right] = \sum_{i=1}^{k-1} \mathbf{P}[B_k]\, \mathbf{P}[B_i]\, \Pi_{j=1; j\neq i}^{k-1} \mathbf{P}\left[\bar{B}_j\right] = \sum_{i=1}^{k-1} p^2 (1-p)^{k-2},$$

puisque $\mathbf{P}[B_\ell] = \mathbf{P}[A] = p$ et $\mathbf{P}\left[\bar{B}_\ell\right] = 1 - \mathbf{P}[A] = 1 - p$. Donc, comme les termes de la précédente somme ne dépendent pas de i(et il y en a $k-1$), on a donc

$$\mathbf{P}[X = k] = (k-1)p^2(1-p)^{k-2}.$$

□

Exercice 3 : CHANGER DE COPIE

On dispose d'une urne contenant N boules indistinguables au toucher, parmi lesquelles Np boules blanches et le reste sont des boules noires, où le réel $0 < p < 1$ est tel que Np est un entier, bien sûr. On prélève successivement et *sans remise* n boules de l'urne, où n est tel que $n < Np$ et $n < N(1-p)$. On note pour tout entier i, $1 \leq i \leq n$, X_i la variable aléatoire valant 1 si l'on prélève une blanche au i-ème coup et 0 sinon. Pour tout entier j, $1 \leq j \leq n$, on note

$$Y_j = \sum_{i=1}^{j} X_i.$$

(1) On s'intéresse d'abord aux v.a. X_1 et X_2.

(a) Donnner la loi de X_1;

Correction (0,5 pt). X_1 peut prendre les valeurs 0 et 1, elle suit donc une loi de Bernouilli. Pour en connaître le paramètre, on calcule

$$\mathbf{P}[X_1 = 1] = \mathbf{P}[\text{On tire une blanche la première fois}] = \frac{\text{Nombre de blanches}}{\text{Nombre total de boules}} = \frac{Np}{N} = p,$$

puisqu'il y a équiprobabilité. En conclusion, on a $X_1 \sim Ber(p)$. □

(b) Calculer $\mathbf{P}[X_2 = 1 \mid X_1 = i]$ pour $i \in \{0, 1\}$;

Correction (1 pt). On commence par calculer $\mathbf{P}[X_2 = 1 \mid X_1 = 0]$. Sur l'événement $\{X_1 = 0\}$, l'espace de probabilité restreint correspondant au 2ème tirage est l'ensemble des boules après le tirage d'une boule noire au premier tirage (puisque $X_1 = 0$). Donc,

$$\mathbf{P}[X_2 = 1 \mid X_1 = 0] = \frac{\text{Nombre de blanches après un tirage de noir}}{\text{Nombre total de boules après un tirage}} = \frac{Np}{N-1}.$$

De même,

$$\mathbf{P}[X_2 = 1 \mid X_1 = 1] = \frac{\text{Nombre de blanches après un tirage de blanche}}{\text{Nombre total de boules après un tirage}} = \frac{Np-1}{N-1}.$$

□

(c) En déduire que X_2 suit la loi $Ber(p)$ (on pourra distinguer suivant les différentes valeurs de X_1).

Correction (1 pt). X_2 prend les valeurs 0 et 1, donc elle suit une loi de Bernouilli. Par ailleurs,

$$\begin{aligned}\mathbf{P}[X_2=1] &= \mathbf{P}[\{X_2=1\}\cap\{X_1=1\}]+\mathbf{P}[\{X_2=1\}\cap\{X_1=0\}]\\ &= \mathbf{P}[X_2=1 \mid X_1=1]\,\mathbf{P}[X_1=1]+\mathbf{P}[X_2=1 \mid X_1=0]\,\mathbf{P}[X_1=0]\\ &= \frac{Np-1}{N-1}.p+\frac{Np}{N-1}.(1-p)\\ &= p,\end{aligned}$$

ce qui montre que X_2 suit également une loi $Ber(p)$. □

(2) Soit $j\in\{1,...,n\}$ fixé. Que représente la v.a. Y_j? Comment appelle-t-on communément sa loi? Donner explicitement cette loi en fonction de N, j et p.

Correction (1 pt). Pour tout j, Y_j compte le nombre de "1" jusqu'au $j$ème tirage, c'est à dire, le nombre de blanches tirées parmi les j boules prélevées au total. Par définition, Y_j suit donc une loi Hypergéométrique $\mathcal{H}(N,j,p)$ (N est la population totale, j est le nombre de boules prélevées et p est la proportion de blanches au départ). Pour tout $k\le j$ (en particulier, $k\le Np$ et $k\le N(1-p)$), on a donc

$$\mathbf{P}[Y_j=k]=\frac{C_{Np}^{k}C_{N(1-p)}^{j-k}}{C_N^j}.$$

□

(3) Soient P et Q, deux entiers strictement positifs et ℓ, un entier tel que $\ell\le P$ et $\ell\le Q$. Calculer la somme

$$\sum_{k=0}^{\ell} C_P^k C_Q^{\ell-k}$$

en fonction de ℓ, P et Q (on pourra pour cela s'inspirer de la question précédente).

Correction (1 pt). Considérons Z, une v.a. de loi hypergéométrique $\mathcal{H}(P+Q,\ell,\frac{P}{P+Q})$. Concrètement, Z compte le nombre d'individus de classe A prélevés après ℓ tirages sans remise dans une propulation de P individus de classe A et de Q autres individus. On a comme dans la question précédente, pour tout $k\le\ell$, $\mathbf{P}[Z=k]=\frac{C_P^kC_Q^{\ell-k}}{C_{P+Q}^k}$. Par ailleurs, comme on a affaire à une loi de probabilité, on a $\sum_{k=1}^{\ell}\mathbf{P}[Z=k]=1$, ce qui revient à dire que

$$\sum_{k=1}^{\ell}\frac{C_P^kC_Q^{\ell-k}}{C_{P+Q}^k}=1\Longleftrightarrow\sum_{k=1}^{\ell}C_P^kC_Q^{\ell-k}=C_{P+Q}^k.$$

□

(4) Soit $j\in\{1,...,n-1\}$ fixé.

(a) Déterminer $\mathbf{P}[X_{j+1}=1 \mid Y_j=k]$ pour toute valeur possible k de Y_j.

Correction (0,5 pt). Le raisonnement est le même qu'en 1.b). Sur l'évènement $\{Y_j=k\}$, on procède au $j+1$-ème coup à un tirage d'une boule parmi un ensemble de $N-j$ boules, composé de $Np-k$ blanches et $N(1-p)-(j-k)$ noires, puisque l'on a prélevé jusque-là k blanches (notez que l'on a clairement $j\ge k$!). Donc, comme en 1.b),

$$\mathbf{P}[X_{j+1}=1 \mid Y_j=k]=\frac{Np-k}{N-j}.$$

□

(b) En déduire, à l'aide de 3., la loi de X_{j+1} (on pourra s'inspirer du raisonnement fait en 1.(c)).

Correction (2 pts). La v.a. X_{j+1} suit forcément une loi de Bernouilli puisqu'elle vaut 0 ou 1. Comme en 1.c), on va partitionner l'événement $\{X_{j+1} = 1\}$ suivant les différentes valeurs de Y_j, dont on connaît la loi (voir (2)). En effet, la probabilité de tirer une blanche la $j+1$-ème fois dépend des j tirages précédents! On a

$$\mathbf{P}[X_{j+1} = 1] = \sum_{k=0}^{j} \mathbf{P}[X_{j+1} = 1 \mid Y_j = k]\, \mathbf{P}[Y_j = k] = \sum_{k=0}^{j} \frac{Np-k}{N-j} \frac{C_{Np}^{k} C_{N(1-p)}^{j-k}}{C_N^j}.$$

Il faut faire un peu d'arithmétique sur cette formule. On a donc

$$\begin{aligned}
\mathbf{P}[X_{j+1} = 1] &= \frac{1}{(N-j)C_N^j}.\sum_{k=0}^{j}(Np-k).\frac{Np!}{(Np-k)!k!}.C_{N(1-p)}^{j-k} \\
&= \frac{1}{(N-j)C_N^j}.\sum_{k=0}^{j}\frac{Np!}{(Np-k-1)!k!}.C_{N(1-p)}^{j-k} \\
&= \frac{Np}{(N-j)C_N^j}.\sum_{k=0}^{j}\frac{(Np-1)!}{(Np-k-1)!k!}.C_{N(1-p)}^{j-k} \\
&= \frac{Np}{(N-j)C_N^j}.\sum_{k=0}^{j} C_{Np-1}^{k}.C_{N(1-p)}^{j-k} \\
&= \frac{Np.C_{Np-1+N(1-p)}^{j}}{(N-j)C_N^j},
\end{aligned}$$

où l'on a appliqué le résultat de la partie 3) pour $\ell = j$, $P = Np-1$ et $Q = N(1-p)$. Finalement, on a donc

$$\mathbf{P}[X_{j+1} = 1] = p\frac{NC_{N-1}^{j}}{(N-j)C_N^j} = p\frac{\frac{N(N-1)!}{j!(N-1-j)!}}{\frac{N!(N-j)}{(N-j)!j!}} = p\frac{\frac{N!}{j!(N-1-j)!}}{\frac{N!}{(N-j-1)!j!}} = p.$$

Ouf! Donc, on a bien $X \sim Ber(p)$. □

(5) Déduire de la partie précédente que tous les X_i, $1 \leq i \leq n$, suivent la loi $Ber(p)$.

Correction (0,5 pt). On sait que $Y_j \sim \mathcal{H}(N, j, p)$ pour tout $j \geq 1$, et que $Y_j \sim \mathcal{H}(N, j, p) \Rightarrow X_{j+1} \sim Ber(p)$. Donc, par une récurrence immédiate, on a bien que $X_j \sim Ber(p)$ pour tout $j \geq 1$. □

(6) En déduire $\mathbf{E}[Y_j]$ pour tout $j \in \{1, ..., n\}$.

Correction (0,5 pt). Du coup, $\mathbf{E}[X_i] = p$ pour tout $i \geq 1$ et donc

$$\mathbf{E}[Y_j] = \mathbf{E}\left[\sum_{i=1}^{j} X_i\right] = \sum_{i=1}^{j} \mathbf{E}[X_i] = \sum_{i=1}^{j} p = jp,$$

puisque l'on somme j fois le terme constant p. On retrouve bien l'espérance de la loi Hypergéométrique. □

(7) Pourquoi Y_n ne suit-elle pas la loi $\mathcal{B}(n, p)$?

Correction (0,5 pt). Y_n suit la loi $\mathcal{H}(N, n, p)$ et consiste donc en la somme de n v.a. de loi $Ber(p)$. C'est la même chose pour la loi $\mathcal{B}(n, p)$, mais la différence ici est que les v.a. X_i ne sont pas *indépendantes*, puisque les tirages sont sans remise. □

Exercice 4 : CHANGER DE COPIE

Soit $\lambda > 0$. On suppose que X suit la loi exponentielle discrète: X est à valeurs dans $\mathbb{N}$ et pour tout $k \in \mathbb{N}$,

$$\mathbf{P}[X = k] = \left(1 - e^{-\frac{1}{\lambda}}\right) e^{-\frac{k}{\lambda}}.$$

Soit Y, une v.a. indépendante de X, et de même loi. On note finalement Z, la v.a. $X + Y$.

(1) Vérifier que la loi exponentielle discrète est bien une loi de probabilité.

Correction. On a clairement $\mathbf{P}[X = k] \in [0, 1]$, pour tout $k \in \mathbb{N}$. Par ailleurs,

$$\sum_{k \in \mathbb{N}} \mathbf{P}[X = k] = \left(1 - e^{-\frac{1}{\lambda}}\right) . \lim_{N \to \infty} \sum_{k=1}^{N} \left(e^{-\frac{1}{\lambda}}\right)^k = \left(1 - e^{-\frac{1}{\lambda}}\right) . \lim_{N \to \infty} \frac{1 - e^{-\frac{k}{\lambda}}}{1 - e^{-\frac{1}{\lambda}}} = 1,$$

ce qui prouve bien que la loi exponentielle discrète est une loi de probabilité. □

(2) Donner la valeur de la fonction de répartition de X en tout point $x \in \mathbb{R}$ (on pourra introduire la partie entière de x).

Correction. Comme X ne prend que des valeurs entières, pour tout $x \in \mathbb{R}+$, on a

$$F_X(x) = \mathbf{P}[X \leq x] = \sum_{k=1}^{[x]} \mathbf{P}[X = k],$$

où $[x]$ désigne la partie entière de x. En effet, être inférieur ou égal à x est équivalent à être inférieur ou égal au plus grand entier inférieur ou égal à x. Donc,

$$F_X(x) = \left(1 - e^{-\frac{1}{\lambda}}\right) . \sum_{k=1}^{[x]} \left(e^{-\frac{1}{\lambda}}\right)^k = \left(1 - e^{-\frac{1}{\lambda}}\right) . \frac{1 - e^{-\frac{[x]}{\lambda}}}{1 - e^{-\frac{1}{\lambda}}} = 1 - e^{-\frac{[x]}{\lambda}}.$$

Par ailleurs, il est clair que $F_X(x) = 0$ pour tout $x < 0$ puisque X ne prend que des valeurs positives. En conclusion, on a pour tout $x \in \mathbb{R}$, $F_X(x) = \left(1 - e^{-\frac{[x]}{\lambda}}\right) . \mathbf{1}_{\mathbb{R}+}(x)$. □

(3) Calculer la fonction génératrice de X, et en déduire l'espérance et la variance de X.

Correction. On a, pour tout $u \in [-1, 1]$,

$$g_X(u) = \mathbf{E}\left[u^X\right] = \sum_{k \in \mathbb{N}} \left(1 - e^{-\frac{1}{\lambda}}\right) . e^{-k/\lambda} u^k = \left(1 - e^{-\frac{1}{\lambda}}\right) . \sum_{k \in \mathbb{N}} \left(ue^{-\frac{1}{\lambda}}\right)^k = \frac{1 - e^{-\frac{1}{\lambda}}}{1 - ue^{-\frac{1}{\lambda}}}.$$

Donc, on a pour tout $u \in [-1, 1]$

$$g'_X(u) = \frac{\left(1 - e^{-\frac{1}{\lambda}}\right) e^{-\frac{1}{\lambda}}}{\left(1 - ue^{-\frac{1}{\lambda}}\right)^2} \text{ et } g''_X(u) = \frac{2\left(1 - e^{-\frac{1}{\lambda}}\right) e^{-\frac{2}{\lambda}}}{\left(1 - ue^{-\frac{1}{\lambda}}\right)^3}.$$

En particulier,

$$\mathbf{E}[X] = g'_X(1) = \frac{e^{-\frac{1}{\lambda}}}{1 - e^{-\frac{1}{\lambda}}}$$

et

$$Var[X] = g''_X(1) + g'_X(1) - (g'_X(1))^2 = \frac{e^{-\frac{2}{\lambda}}}{\left(1 - e^{-\frac{1}{\lambda}}\right)^2} + \frac{e^{-\frac{1}{\lambda}}}{1 - e^{-\frac{1}{\lambda}}} = \frac{e^{-\frac{1}{\lambda}}}{\left(1 - e^{-\frac{1}{\lambda}}\right)^2},$$

après simplification. □

(4) Déduire de (3), l'expression de la fonction génératrice de Z. Z suit-elle aussi une loi exponentielle discrète?

Correction. Comme X et Y sont indépendants, On a pour tout u, $g_Z(u) = g_X(u)g_Y(u)$. En outre, comme Y suit la même loi que X, elle a la même fonction génératrice, et donc, pour tout $u \in [-1, 1]$,

$$g_Z(u) = (g_X(u))^2 = \left(\frac{1 - e^{-\frac{1}{\lambda}}}{1 - ue^{-\frac{1}{\lambda}}}\right)^2. \tag{1}$$

Si Z avait suivi la loi exponentielle discrète de paramètre $\alpha > 0$, d'après (3), on aurait pour tout u

$$g_Z(u) = \frac{1 - e^{-\frac{1}{\alpha}}}{1 - ue^{-\frac{1}{\alpha}}}.$$

Or, la forme de (1) ne peut pas se ramener à cette dernière, pour aucune valeur de α. Donc, Z ne suit pas la loi exponentielle discrète. $\square$

(5) Déterminer la loi de Z en utilisant la convolution. Vérifiez le résultat de la question précédente.

Correction. Comme X et Y sont à valeurs dans $\mathbb{N}$, c'est aussi le cas de $Z = X + Y$. On a pour tout $n \in \mathbb{N}$, en vertu de l'indépendance de X et Y,

$$\begin{aligned}\mathbf{P}[Z = n] = \sum_{k=0}^{n} \mathbf{P}[X = k]\,\mathbf{P}[Y = n - k] &= \left(1 - e^{-\frac{1}{\lambda}}\right)^2 \sum_{k=0}^{n} e^{-\frac{k}{\lambda}} e^{-\frac{n-k}{\lambda}} \\ &= \left(1 - e^{-\frac{1}{\lambda}}\right)^2 \sum_{k=0}^{n} e^{-\frac{n}{\lambda}} = \left(1 - e^{-\frac{1}{\lambda}}\right)^2 (n+1)e^{-\frac{n}{\lambda}},\end{aligned}$$

et l'on retrouve clairement le fait que Z ne suit pas la loi exponentielle discrète, car cette dernière probabilité ne s'écrit en aucun cas $\left(1 - e^{-\frac{1}{\alpha}}\right) e^{-\frac{n}{\alpha}}$ pour une quelconque valeur de $\alpha > 0$. $\square$