Variables aléatoires

Un employé d'un centre d'appels doit joindre N personnes.

- Le premier jour, il appelle ces N personnes, chacune ayant une probabilité $p \in]0;1[$ de répondre, et ce de façon indépendante des autres personnes. On note alors X_1 le nombre de personnes ayant répondu ce premier jour.
- Le deuxième jour, l'employé rappelle tous ceux qui n'ont pas répondu à son premier appel, mais ne rappelle pas ceux qui ont déjà répondu le premier jour. On suppose que la probabilité p de répondre est inchangée. On note alors X_2 le nombre de personnes ayant répondu ce deuxième jour.
- L'employé, voulant suivre les directives du centre d'appels, est tenace : il rappelle le troisième jour tous ceux qui n'ont pas répondu lors des deux jours précédents et ainsi de suite...
- Il rappelle donc le jour n toutes les personnes n'ayant pas répondu les $(n-1)$ jours précédents. La probabilité de répondre est toujours de p pour chaque personne, et le fait de répondre est indépendant des autres. On note X_n le nombre de personnes ayant répondu le jour n.
- Enfin, on note Z_n le nombre total de personnes qui ont pu être jointes durant les n premiers jours.

On supposera que les variables aléatoires sont toutes définies sur un même espace probabilisé $(\Omega, \mathbb{P})$.

Partie A. *Les deux premiers jours*

A1. (a) Reconnaître la loi suivie par X_1 (justifier).

Correction : X_1 compte le nombre de succès dans N répétitions indépendantes (d'après l'énoncé) d'une même épreuve de Bernoulli (à savoir, la i-ième personne répond), donc $\boxed{X_1 \sim \mathscr{B}(N, p)}$.

(b) En déduire l'ensemble des valeurs $X_1(\Omega)$ ainsi que l'espérance et la variance de X_1.

Correction : On en déduit que $\boxed{X_1(\Omega) = [\![0;N]\!],\ \mathbb{E}(X_1) = Np \text{ et } \mathbb{V}(X_1) = Np(1-p).}$

A2. Soit $i \in [\![0;N]\!]$.

(a) Sachant $(X_1 = i)$, quelles sont les valeurs prises par X_2 ?

Correction : Sachant $(X_1 = i)$ (c'est-à-dire, i personnes ont répondu le premier jour), l'employé va donc rappeler les $N - i$ personnes restantes le deuxième jour.

Ainsi, $\boxed{\text{les valeurs prises par } X_2 \text{ sachant } (X_1 = i) \text{ sont données par } [\![0;N-i]\!].}$

(b) En déduire la loi conditionnelle de X_2 sachant $(X_1 = i)$ en donnant les valeurs $\mathbb{P}(X_2 = j \mid X_1 = i)$ pour tout $j \in [\![0;N]\!]$.

Correction : Sachant $(X_1 = i)$, la variable aléatoire X_2 compte le nombre de succès (les personnes qui répondent à l'appel) après $N - i$ répétitions indépendantes d'une même épreuve de Bernoulli (raisonnement similaire à la question (A 1a)).

Donc, $\boxed{\text{la loi conditionnelle de } X_2 \text{ sachant } (X_1 = i) \text{ est la loi binomiale de paramètres } N - i \text{ et } p.}$

Ceci signifie :

$$\boxed{\mathbb{P}(X_2 = j \mid X_1 = i) = \begin{cases} \dbinom{N-i}{j} p^j (1-p)^{N-i-j} & \text{si } j \in [\![0;N-i]\!] \\ 0 & \text{sinon.} \end{cases}}$$

ⓒ Conclure en donnant la loi du couple (X_1, X_2).

Correction : Comme $X_1(\Omega) = X_2(\Omega) = [\![0;N]\!]$, il s'agit de calculer les valeurs $\mathbb{P}(X_1 = i, X_2 = j)$ pour tout $(i,j) \in [\![0;N]\!]^2$. Grâce à la question précédente, on a :

$$\begin{aligned}\mathbb{P}(X_1 = i, X_2 = j) &= \mathbb{P}(X_1 = i)\,\mathbb{P}(X_2 = j \mid X_1 = i) \quad \text{(définition d'une probabilité conditionnelle)}\\ &= \begin{cases} \binom{N}{i} p^i (1-p)^{N-i} \binom{N-i}{j} p^j (1-p)^{N-i-j} & \text{si } j \in [\![0;N-i]\!],\\ 0 & \text{sinon.}\end{cases} \quad \text{(d'après (A1a) et (A2b))}\\ &\boxed{= \begin{cases} \binom{N}{i}\binom{N-i}{j} p^{i+j}(1-p)^{2N-2i-j} & \text{si } j \in [\![0;N-i]\!],\\ 0 & \text{sinon.}\end{cases}}\end{aligned}$$

Partie B. *Un cas particulier des deux premiers jours.*

On suppose dans cette partie que $N = 2$ et $p = \frac{1}{2}$. On donne le tableau décrivant la loi de (X_1, X_2) ci-dessous :

$\mathbb{P}(X_1 = i, X_2 = j)$		$j = 0$	$j = 1$	$j = 2$
i	0	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$
	1	λ	λ	0
	2	λ	0	0

B1. Que vaut λ ?

Correction : La somme des probabilités devant valoir 1, on a :

$$3\lambda + \frac{1}{16} + \frac{1}{8} + \frac{1}{16} = 1 \iff 3\lambda + \frac{1}{4} = 1 \iff 3\lambda = \frac{3}{4} \iff \boxed{\lambda = \frac{1}{4}.}$$

B2. Par un exemple de calcul de probabilité (qui ne renvoie pas la valeur 0), vérifier la cohérence de ce tableau avec le résultat obtenu en (ⓒ).

Correction : Par exemple, calculons $\mathbb{P}(X_1 = 0, X_2 = 0)$ à l'aide de (ⓒ) (en prenant $N = 2$ et $p = \frac{1}{2}$) :

$$\begin{aligned}\mathbb{P}(X_1 = 0, X_2 = 0) &= \tbinom{2}{0}\tbinom{2-0}{0}\left(\tfrac{1}{2}\right)^{0+0}\left(1-\tfrac{1}{2}\right)^{2\cdot 2-2\cdot 0-0}\\ &= 1\cdot 1\cdot 1\cdot\left(\tfrac{1}{2}\right)^4\\ &= \tfrac{1}{16},\end{aligned}$$

ce qui $\boxed{\text{coïncide bien avec la valeur indiquée dans le tableau.}}$

B3. Déterminer la loi de X_2 en justifiant soigneusement (*un calcul littéral* et *numérique est attendu*).

Correction : Comme $X_2(\Omega) = \{0,1,2\}$, il s'agit de calculer les valeurs $\mathbb{P}(X_2 = j)$ pour $j \in \{0,1,2\}$. Grâce à la formule des probabilités totales (en utilisant le système complet d'événements $((X_1 = i))_{1\leqslant i\leqslant 3}$) on a :

$$\begin{aligned}
\mathbb{P}(X_2 = 0) &= \mathbb{P}(X_1 = 0, X_2 = 0) + \mathbb{P}(X_1 = 1, X_2 = 0) + \mathbb{P}(X_1 = 2, X_2 = 0)\\
&= \frac{1}{16} + \lambda + \lambda = \frac{9}{16}.\\
\mathbb{P}(X_2 = 1) &= \mathbb{P}(X_1 = 0, X_2 = 1) + \mathbb{P}(X_1 = 1, X_2 = 1) + \mathbb{P}(X_1 = 2, X_2 = 1)\\
&= \frac{1}{8} + \lambda = \frac{3}{8}.\\
\mathbb{P}(X_2 = 2) &= \mathbb{P}(X_1 = 0, X_2 = 2) + \mathbb{P}(X_1 = 1, X_2 = 2) + \mathbb{P}(X_1 = 2, X_2 = 2)\\
&= \frac{1}{16}.
\end{aligned}$$

(Vis-à-vis du tableau, il s'agissait donc de sommer chacune des lignes). Résumons :

j	0	1	2
$\mathbb{P}(X_2 = j)$	$\frac{9}{16}$	$\frac{3}{8}$	$\frac{1}{16}$

B4. Les variables aléatoires X_1 et X_2 sont-elles indépendantes ?

Correction : En procédant comme précédemment, on en déduit la loi de X_1 (sommation des colonnes cette fois-ci). Résumons le tout :

$\mathbb{P}(X_1 = i, X_2 = j)$		$j = 0$	$j = 1$	$j = 2$	$\mathbb{P}(X_1 = i)$
i	0	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{4}$
	1	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$
	2	$\frac{1}{4}$	0	0	$\frac{1}{4}$
$\mathbb{P}(X_2 = j)$		$\frac{9}{16}$	$\frac{3}{8}$	$\frac{1}{16}$	

On constate que $\mathbb{P}(X_1 = 2, X_2 = 2) = 0 \neq \mathbb{P}(X_1 = 2)\mathbb{P}(X_2 = 2) = \frac{1}{4}\cdot\frac{1}{16}$, donc $\boxed{X_1 \text{ et } X_2 \text{ ne sont pas indépendantes.}}$

B5. Calculer $\mathbb{E}(X_2)$ et $\mathbb{V}(X_2)$.

Correction : D'après la loi de X_2, on a :

$$\boxed{\mathbb{E}(X_2)} = 0\cdot\mathbb{P}(X_2 = 0) + 1\cdot\mathbb{P}(X_2 = 1) + 2\cdot\mathbb{P}(X_2 = 2) = \frac{3}{8} + \frac{2}{16}\ \boxed{= \frac{1}{2}.}$$

Grâce au théorème de transfert :

$$\mathbb{E}(X_2^2) = 0^2\cdot\mathbb{P}(X_2 = 0) + 1^2\cdot\mathbb{P}(X_2 = 1) + 2^2\cdot\mathbb{P}(X_2 = 2) = \frac{3}{8} + \frac{4}{16} = \frac{5}{8}.$$

Finalement :

$$\boxed{\mathbb{V}(X_2)} = \mathbb{E}(X_2^2) - \mathbb{E}(X_2)^2 = \frac{5}{8} - \frac{1}{4}\ \boxed{= \frac{3}{8}.}$$

Partie C. *Les prémices de la loi de Z_n.*

On revient au cas général où $N \in \mathbb{N}^*$ et $p \in]0;1[$.

C1. Si $n \in \mathbb{N}^*$, exprimer Z_n en fonction des X_k.

Correction : Comme Z_n est le nombre total de personnes ayant répondu, on a $\boxed{Z_n = X_1 + \ldots + X_n = \sum_{k=1}^{n} X_k.}$

C2. Soit $k \geqslant 2$. Calculer $\mathbb{P}(X_k = 0 \mid X_1 = 0, \dots, X_{k-1} = 0)$.

Correction : L'événement $(X_1 = 0, \dots, X_{k-1} = 0)$ signifie que personne n'a répondu les $(k-1)$ premiers jours. Donc, la loi conditionnelle de X_k sachant $(X_1 = 0, \dots, X_{k-1} = 0)$ est la loi binomiale $\mathscr{B}(N, p)$ (raisonnement identique à (A 1a)). Ainsi,

$$\boxed{\mathbb{P}(X_k = 0 \mid X_1 = 0, \dots, X_{k-1} = 0)} = \binom{N}{0} p^0 (1-p)^{N-0} \boxed{= (1-p)^N.}$$

C3. En déduire que $\mathbb{P}(Z_n = 0) = (1-p)^{nN}$ pour tout $n \in \mathbb{N}^*$.

Correction : Rappelons qu'une somme d'entiers naturels est nulle, si et seulement si tous les entiers sont nuls. Par conséquent :

$$\begin{aligned}
\mathbb{P}(Z_n = 0) &= \mathbb{P}(X_1 + \dots + X_n = 0) \\
&= \mathbb{P}(X_1 = 0, \dots, X_n = 0) \\
&= \mathbb{P}(X_1 = 0)\, \mathbb{P}(X_2 = 0 \mid X_1 = 0) \dots \mathbb{P}(X_n = 0 \mid X_1 = 0, \dots, X_{n-1} = 0) \quad \text{(probabilités composées)} \\
&= \underbrace{(1-p)^N}_{\text{cf. (A 1a)}} \times \underbrace{(1-p)^N \times \dots \times (1-p)^N}_{\text{d'après (C2)}} \\
&\boxed{= (1-p)^{nN}} \qquad \text{(car il y a } n \text{ facteurs).}
\end{aligned}$$

C4. Que vaut $\lim\limits_{n \to +\infty} \mathbb{P}(Z_n = 0)$?

Correction : Vu que $(1-p)^N \in \left]-1;1\right[$ (car $0 < 1-p < 1$ puis stricte croissance de la fonction $x \mapsto x^N$ sur $]0;1[$), on en déduit que $((1-p)^N)^n = (1-p)^{nN} \underset{n \to +\infty}{\longrightarrow} 0$ c'est-à-dire $\boxed{\mathbb{P}(Z_n = 0) \underset{n \to +\infty}{\longrightarrow} 0.}$

Partie D. *Loi de* Z_n.

D1. À l'aide de la question (b), montrer que pour tout $(i, j) \in [\![0; N]\!]^2$,

$$\mathbb{P}(Z_2 = j \mid Z_1 = i) = \begin{cases} \dbinom{N-i}{j-i} p^{j-i} (1-p)^{N-j} & \text{si } j \geqslant i \\ 0 & \text{sinon.} \end{cases}$$

Correction : Par définition de Z_1 et Z_2 :

$$\begin{aligned}
\boxed{\mathbb{P}(Z_2 = j \mid Z_1 = i)} &= \mathbb{P}(X_1 + X_2 = j \mid X_1 = i) \\
&= \mathbb{P}(X_2 = j - i \mid X_1 = i) \\
&\underset{(b)}{=} \begin{cases} \dbinom{N-i}{j-i} p^{j-i} (1-p)^{N-j} & \text{si } 0 \leqslant j - i \leqslant N - i, \\ 0 & \text{sinon.} \end{cases} \\
&\boxed{= \begin{cases} \dbinom{N-i}{j-i} p^{j-i} (1-p)^{N-j} & \text{si } j \geqslant i, \\ 0 & \text{sinon.} \end{cases}}
\end{aligned}$$

D2. Si $j \in [\![0; N]\!]$, en déduire que :

$$\mathbb{P}(Z_2 = j) = p^j (1-p)^{2N-2j} \sum_{i=0}^{j} \binom{N-i}{j-i} \binom{N}{i} (1-p)^{j-i}.$$

Correction : En utilisant le système complet d'événements $((Z_1 = i))_{0\leqslant i\leqslant N}$, la formule des probabilités totales nous donne :

$$\begin{aligned}
\boxed{\mathbb{P}(Z_2 = j)} &= \sum_{i=0}^{N} \mathbb{P}(Z_1 = i)\,\mathbb{P}(Z_2 = j \mid Z_1 = i)\\
&= \sum_{i=0}^{N} \mathbb{P}(X_1 = i)\,\mathbb{P}(Z_2 = j \mid Z_1 = i)\\
&= \sum_{i=0}^{j} \binom{N}{i} p^i (1-p)^{N-i} \binom{N-i}{j-i} p^{j-i}(1-p)^{N-j} \qquad \text{(d'après (A 1a) et (D 1))}\\
&\boxed{= p^j(1-p)^{2N-2j} \sum_{i=0}^{j} \binom{N}{i}\binom{N-i}{j-i}(1-p)^{j-i}.}
\end{aligned}$$

D3. Simplifier $\dfrac{\binom{N-i}{j-i}\binom{N}{i}}{\binom{j}{i}}$.

Correction : Si $0 \leqslant i \leqslant j \leqslant N$, on a :

$$\frac{\binom{N-i}{j-i}\binom{N}{i}}{\binom{j}{i}} = \frac{(N-i)!}{(j-i)!(N-j)!}\,\frac{N!}{i!(N-i)!}\,\frac{i!(j-i)!}{j!} = \frac{N!}{(N-j)!j!}\,\boxed{=\binom{N}{j}.}$$

D4. En déduire que $Z_2 \sim \mathscr{B}(N, p_2)$ où $p_2 = p(2-p)$.

Correction : Grâce aux deux questions précédentes, on a donc, pour $j \in [\![0;N]\!]$:

$$\begin{aligned}
\mathbb{P}(Z_2 = j) &= p^j(1-p)^{2N-2j}\sum_{i=0}^{j}\binom{N}{j}\binom{j}{i}(1-p)^{j-i}\\
&= \binom{N}{j} p^j (1-p)^{2N-2j}\sum_{i=0}^{j}\binom{j}{i}(1-p)^{j-i} \qquad \text{(mise en facteur)}\\
&= \binom{N}{j} p^j (1-p)^{2N-2j}(1+(1-p))^j \qquad \text{(formule du binôme)}\\
&= \binom{N}{j} \left(p(2-p)\right)^j \left((1-p)^2\right)^{N-j}\\
&= \binom{N}{j} \left(p(2-p)\right)^j \left(1-p(2-p)\right)^{N-j} \qquad (\text{car } 1-p(2-p) = 1-2p+p^2 = (1-p)^2)\\
&= \binom{N}{j} p_2^j (1-p_2)^{N-j} \qquad (\text{où } p_2 = p(2-p)).
\end{aligned}$$

On reconnaît donc une $\boxed{\text{loi binomiale de paramètres } N \text{ et } p_2 = p(2-p).}$

On admet que, pour tout $n \in \mathbb{N}^*$, on a $Z_n \sim \mathscr{B}(N, p_n)$ où (p_n) est une suite de réels définie par la relation de récurrence suivante :

$$p_1 = p \quad \text{et} \quad \forall\, n \in \mathbb{N}^*, \quad p_{n+1} = (1-p)p_n + p.$$

D5. Déterminer une expression de p_n en fonction de p et de n.

Correction : Il s'agit d'une suite arithmético-géométrique. La solution générale de l'équation homogène (qui est $\forall\, n \in \mathbb{N}^*,\ p_{n+1} = (1-p)p_n$) est donnée par $\left(\lambda(1-p)^n\right)_{n\in\mathbb{N}^*}$ avec $\lambda \in \mathbb{R}$. Une solution particulière constante est 1 (car $1 = (1-p)\cdot 1 + p$). Toute solution de cette équation linéaire est somme d'une solution particulière et de la solution générale de l'équation homogène. Ainsi, il existe $\lambda \in \mathbb{R}$ tel que :

$$\forall\, n \in \mathbb{N}^*, \quad p_n = 1 + \lambda(1-p)^n.$$

Or $p_1 = p \iff p = 1 + \lambda(1-p) \iff \lambda = -1$. Finalement :

$$\boxed{\forall\ n \in \mathbb{N}^*, \quad p_n = 1 - (1-p)^n.}$$

D6. Retrouver le résultat de la question (C4.).

Correction : Comme $\mathrm{Z}_n \sim \mathscr{B}(\mathrm{N}, p_n)$, on a :

$$\boxed{\mathbb{P}(\mathrm{Z}_n = 0)} = \binom{\mathrm{N}}{0} p_n^0 (1-p_n)^{\mathrm{N}-0} = (1-p_n)^{\mathrm{N}} = \left((1-p)^n\right)^{\mathrm{N}} \boxed{= (1-p)^{n\mathrm{N}}.}$$