Soit X une variable aléatoire suivant la loi binomiale $\mathscr{B}(n\,;\,p)$.
On pose $Y = n - X$ et $Z = X - Y$.

1. Quelle est la loi suivie par Y ?
2. Quelle est la loi suivie par Z ?
3. Calculer l'espérance et la variance de Z puis en déduire $\text{cov}(X, Y)$.

Analyse

Un exercice où la loi binomiale est prétexte à divers calculs comme autant de variations autour des opérateurs espérance et variance.

Résolution

Question 1.

Dans un schéma de Bernoulli consistant en la répétition de *n* expériences de Bernoulli indépendantes et de même paramètre *p*, la variable aléatoire X peut correspondre au nombre de SUCCES. Dès lors, la variable aléatoire Y comptabilise simplement le nombre d'ECHEC. Elle suit donc elle-même une loi binomiale de paramètre *n* et $1-p$.

Plus explicitement (mais la justification précédente seule suffit !), lorsque X prend ses valeurs dans $[\![0\,;\,n]\!]$, il en va de même pour Y.
On a alors :

$$\begin{aligned}
p(Y=k) &= p(n-X=k) \\
&= p(X=n-k) \\
&= \binom{n}{n-k} p^{n-k}(1-p)^{n-(n-k)} \\
&= \binom{n}{n-(n-k)}(1-p)^k\, p^{n-k} \\
&= \binom{n}{k}(1-p)^k\, p^{n-k} \\
&= \binom{n}{k}(1-p)^k\left[1-(1-p)\right]^{n-k}
\end{aligned}$$

La variable aléatoire Y suit bien la loi binomiale de paramètre n et $1-p$.

La variable aléatoire Y suit la loi binomiale de paramètre n et $1-p$.

Question 2.

Si X prend la valeur k dans $[\![0\,;n]\!]$ alors Y prend la valeur $n-k$ dans $[\![0\,;n]\!]$ également. On en déduit ainsi que $\mathrm{Z}=\mathrm{X}-\mathrm{Y}$ prend la valeur $k-(n-k)=2k-n$ dans l'ensemble :

$$\{-n\,;-n+2\,;-n+4\,;\ldots;n-2\,;n\}$$

Réciproquement, tout entier de la forme $2k-n$ se décompose de façon unique sous la forme $k-(n-k)$.

On a ainsi : $\mathrm{Z}(\Omega)=\{-n+2k\,/\,k\in[\![0\,;n]\!]\}=\{-n\,;-n+2\,;-n+4\,;\ldots;n-2\,;n\}$ et :

$$p(\mathrm{Z}=2k-n)=p(\mathrm{X}=k)=p(\mathrm{Y}=n-k)=\binom{n}{k}p^k(1-p)^{n-k}$$

La loi de probabilité de la variable aléatoire Z est définie par :

- $\mathrm{Z}(\Omega)=\{-n+2k\,/\,k\in[\![0\,;n]\!]\}=\{-n\,;-n+2\,;-n+4\,;\ldots;n-2\,;n\}$.
- $\forall k\in[\![0\,;n]\!],\ p(\mathrm{Z}=2k-n)=\binom{n}{k}p^k(1-p)^{n-k}$.

ATTENTION ! Z ne suit pas une loi binomiale !

Question 3.

La linéarité de l'espérance nous permet d'écrire : $\mathrm{E}(\mathrm{Z})=\mathrm{E}(\mathrm{X}-\mathrm{Y})=\mathrm{E}(\mathrm{X})-\mathrm{E}(\mathrm{Y})$.

Comme X suit la loi binomiale de paramètres n et p, on a immédiatement : $\mathrm{E}(\mathrm{X})=np$ et comme Y suit la loi binomiale de paramètre n et $1-p$, on a aussi : $\mathrm{E}(\mathrm{Y})=n(1-p)$.

Il vient donc : $\mathrm{E}(\mathrm{Z})=\mathrm{E}(\mathrm{X})-\mathrm{E}(\mathrm{Y})=np-n(1-p)=n(2p-1)$.

On pouvait également écrire :

$$\mathrm{E}(\mathrm{Z})=\mathrm{E}(\mathrm{X}-\mathrm{Y})=\mathrm{E}(\mathrm{X}-(n-\mathrm{X}))=\mathrm{E}(2\mathrm{X}-n)=2\mathrm{E}(\mathrm{X})-n=2\times np-n=n(2p-1)$$

$$\boxed{\mathrm{E}(\mathrm{Z})=n(2p-1)}$$

On a, en utilisant les propriétés fondamentales de la variance :

$$\mathrm{V}(\mathrm{Z})=\mathrm{V}(2\mathrm{X}-n)=\mathrm{V}(2\mathrm{X})=4\mathrm{V}(\mathrm{X}).$$

Comme X suit la loi binomiale de paramètres *n* et *p*, on a : $\mathrm{V}(\mathrm{X})=np(1-p)$.
D'où :

$$\mathrm{V}(\mathrm{Z})=4\mathrm{V}(\mathrm{X})=4np(1-p)$$

$$\boxed{\mathrm{V}(\mathrm{Z})=4np(1-p)}$$

On a enfin :

$$\mathrm{V}(\mathrm{Z})=\mathrm{V}(\mathrm{X}-\mathrm{Y})=\mathrm{V}(\mathrm{X})+\mathrm{V}(\mathrm{Y})-2\operatorname{cov}(\mathrm{X},\mathrm{Y})$$

On a : $\mathrm{V}(\mathrm{Y})=\mathrm{V}(n-\mathrm{X})=\mathrm{V}(-\mathrm{X})=\mathrm{V}(\mathrm{X})$.
D'où, en tenant compte de $\mathrm{V}(\mathrm{Z})=4np(1-p)$:

$$\begin{aligned}
&\mathrm{V}(\mathrm{Z})=\mathrm{V}(\mathrm{X}-\mathrm{Y})=\mathrm{V}(\mathrm{X})+\mathrm{V}(\mathrm{Y})-2\operatorname{cov}(\mathrm{X},\mathrm{Y})\\
\Leftrightarrow\ &4np(1-p)=np(1-p)+np(1-p)-2\operatorname{cov}(\mathrm{X},\mathrm{Y})\\
\Leftrightarrow\ &2np(1-p)=-2\operatorname{cov}(\mathrm{X},\mathrm{Y})\\
\Leftrightarrow\ &\operatorname{cov}(\mathrm{X},\mathrm{Y})=-np(1-p)
\end{aligned}$$

$$\boxed{\operatorname{cov}(\mathrm{X},\mathrm{Y})=-np(1-p)}$$

Les variables X et Y sont négativement corrélées. Ce résultat est évident dès lors que l'on se « penche » un peu sur la relation $\mathrm{Y}=n-\mathrm{X}$!

On pouvait d'ailleurs mener directement le calcul de la covariance de X et Y en procédant comme suit :

$$\begin{aligned}
\operatorname{cov}(\mathrm{X},\mathrm{Y})&=\mathrm{E}(\mathrm{XY})-\mathrm{E}(\mathrm{X})\mathrm{E}(\mathrm{Y})\\
&=\mathrm{E}\left[\mathrm{X}(n-\mathrm{X})\right]-\mathrm{E}(\mathrm{X})\mathrm{E}(n-\mathrm{X})\\
&=\mathrm{E}\left(n\mathrm{X}-\mathrm{X}^2\right)-\mathrm{E}(\mathrm{X})\left[n-\mathrm{E}(\mathrm{X})\right]\\
&=\mathrm{E}(n\mathrm{X})-\mathrm{E}\left(\mathrm{X}^2\right)-n\mathrm{E}(\mathrm{X})+\left[\mathrm{E}(\mathrm{X})\right]^2\\
&=\cancel{n\mathrm{E}(\mathrm{X})}-\mathrm{E}\left(\mathrm{X}^2\right)\cancel{-n\mathrm{E}(\mathrm{X})}+\left[\mathrm{E}(\mathrm{X})\right]^2\\
&=-\mathrm{E}\left(\mathrm{X}^2\right)+\left[\mathrm{E}(\mathrm{X})\right]^2\\
&=-\left\{\mathrm{E}\left(\mathrm{X}^2\right)-\left[\mathrm{E}(\mathrm{X})\right]^2\right\}\\
&=-\mathrm{V}(\mathrm{X})
\end{aligned}$$

Mais on pouvait également tirer parti encore plus efficacement de la relation $\mathrm{Y} = n - \mathrm{X}$ en utilisant la linéarité de la covariance par rapport à chacune des composantes :

$$\begin{aligned}
\operatorname{cov}(\mathrm{X}, \mathrm{Y}) &= \operatorname{cov}(\mathrm{X}, n - \mathrm{X}) \\
&= \operatorname{cov}(\mathrm{X}, n) - \operatorname{cov}(\mathrm{X}, \mathrm{X}) \\
&= 0 - \mathrm{V}(\mathrm{X}) \\
&= -\mathrm{V}(\mathrm{X})
\end{aligned}$$