Devoir maison 13 - Probabilités

Première partie : Temps d'attente du n-ième succès

On effectue une succession d'expériences de Bernoulli indépendantes, avec une probabilité de succès égale à $p \in]0,1[$.
On note T_n le nombre d'expériences nécessaires pour obtenir le n-ième succès.

1. Identifier la loi de T_1.
On reconnait une expérience type de la loi géométrique $\mathscr{G}(p)$.

2. Donner la loi de T_2, puis celle de T_n pour n quelconque.
Pour $i \in \mathbb{N}^*$, on note S_i l'événement "avoir un succès au rang i" et $E_i = \overline{S_i}$.
On a : $T_2(\Omega) = [\![2, +\infty[\![$ (car il faut au moins deux expériences pour avoir deux succès !).
Pour $k \in [\![2, +\infty[\![$: $(T_2 = k) = \displaystyle\bigcup_{i=1}^{k-1} (E_1 \cap E_2 \cap \cdots \cap E_{i-1} \cap S_i \cap E_{i+1} \cap \cdots \cap E_{k-1} \cap S_k)$.
Cette union disjointe donne les possibilités de combinaisons de 2 succès et $k-2$ échecs, se terminant par un succès.
Les expériences étant indépendantes, tous ces événements (qui comptent 2 succès et $k-1$ échecs) ont la même probabilité : $p^2 q^{k-2}$. De plus, il y en a autant que de façons de placer le premier succès parmi les $k-1$ premières expériences, c'est-à-dire $\displaystyle\binom{k-1}{1} = k-1$.
Finalement, $\forall k \in [\![2, +\infty[\![, \quad \mathbb{P}(T_2 = k) = \displaystyle\binom{k-1}{1} p^2 q^{k-2}$.
Ce résultat se généralise ainsi :

$$\forall k \in [\![n, +\infty[\![, \quad \mathbb{P}(T_n = k) = \binom{k-1}{n-1} p^n q^{k-n}$$

3. Pour $n \in \mathbb{N}^*$ et $k \in T_n(\Omega)$, déterminer la loi de $T_{n+1} - T_n$ conditionnée par $(T_n = k)$.
En déduire la loi de $T_{n+1} - T_n$.
Soit $n \in \mathbb{N}^*$. La variable aléatoire $T_{n+1} - T_n$ correspond à l'intervalle de temps entre le n-ième et le $n+1$-ième succès.
Le bon sens voudrait que cette variable aléatoire suive une loi géométrique de paramètre p. C'est ce que nous allons démontrer...
Soient $k \in T_n(\Omega)$ et $j \in \mathbb{N}^*$.
$\mathbb{P}((T_n = k) \cap (T_{n+1} - T_n = j)) = \mathbb{P}((T_n = k) \cap (T_{n+1} = j + k))$
$= \mathbb{P}((T_n = k) \cap E_{k+1} \cap \cdots \cap E_{j+k-1} \cap S_{j+k})$.
La formule des probabilités composées donne :
$\mathbb{P}((T_n = k) \cap (T_{n+1} - T_n = j)) = \mathbb{P}(T_n = k) q^{j-1} p$.
On en déduit : $\mathbb{P}_{(T_n=k)}(T_{n+1} - T_n = j) = pq^{j-1}$.
La formule des probabilités totales avec le système complet d'événements $((T_n = k))_{k \in T_n(\Omega)}$ donne : $\mathbb{P}(T_{n+1} - T_n = j) = \displaystyle\sum_{k=n}^{+\infty} \mathbb{P}_{(T_n=k)}(T_{n+1} - T_n = j)\mathbb{P}(T_n = k)$ d'où :
$\mathbb{P}(T_{n+1} - T_n = j) = \displaystyle\sum_{k=n}^{+\infty} q^{j-1} p \mathbb{P}(T_n = k) = q^{j-1} p \left(\sum_{k=n}^{+\infty} \mathbb{P}(T_n = k) \right) = q^{j-1} p \times 1$.
On retrouve la loi géométrique attendue !

4. En utilisant la question précédente, calculer $\mathbb{E}(T_n)$.

En remarquant que $T_n = \sum_{k=2}^{n}(T_k - T_{k-1}) + T_1$, et sachant que les lois géométriques admettent une espérance finie, on obtient l'existence de l'espérance de T_n.
De plus, en utilisant la linéarité de l'espérance, on obtient :
$\mathbb{E}(T_n) = \sum_{k=2}^{n} (\mathbb{E}(T_k) - \mathbb{E}(T_{k-1})) + \mathbb{E}(T_1) = \frac{n}{p}$.

5. Pour $n \in \mathbb{N}^*$, démontrer que $T_{n+1} - T_n$ et T_n sont indépendants. En déduire $V(T_n)$.

Soient $n \in \mathbb{N}^*$, $k \in [\![n, +\infty[\![$, et $j \in \mathbb{N}^*$.
On a montré que $\mathbb{P}((T_n = k) \cap (T_{n+1} - T_n = j)) = q^{j-1}p\mathbb{P}(T_n = k) = \mathbb{P}(T_{n+1} - T_n = j)\mathbb{P}(T_n = k)$.
Donc les variables aléatoires T_n et $T_{n+1} - T_n$ sont indépendantes.
On a donc : $\mathrm{V}(T_{n+1}) = \mathrm{V}(T_{n+1} - T_n + T_n) = \mathrm{V}(T_{n+1} - T_n) + \mathrm{V}(T_n) = \frac{q}{p^2} + \mathrm{V}(T_n)$.
Par télescopage, on obtient : $\mathrm{V}(T_n) = n\frac{q}{p^2}$

6. En utilisant l'indépendance de T_{n+1} et T_n, calculer la fonction génératrice de T_n.

Sot $n \in \mathbb{N}^*$. On note g_n la fonction génératrice de T_n.
On sait que la fonction génératrice de $T_{n+1} - T_n$ est définie pour $t \in \left]-\frac{1}{q}, \frac{1}{q}\right[$ par $g(t) = \frac{pt}{1-qt}$.
De l'indépendance de T_n et $T_{n+1} - T_n$, on déduit que $g_{n+1} = g \times g_n$.
Comme $g_1 = g$, une récurrence immédiate donne :

$$\forall n \in \mathbb{N}^*, \quad \forall t \in \left]-\frac{1}{q}, \frac{1}{q}\right[, \quad g_n(t) = \left(\frac{pt}{1-qt}\right)^n$$

7. En déduire la formule du binôme négatif :

$$\forall n \in \mathbb{N} \quad \forall x \in]-1, 1[, \quad \frac{1}{(1-x)^{n+1}} = \sum_{k=n}^{+\infty} \binom{k}{n} x^{k-n}$$

La définition d'une fonction génératrice donne : $\forall t \in \left]-\frac{1}{q}, \frac{1}{q}\right[$:
$\sum_{k=n}^{+\infty} t^k \mathbb{P}(T_n = k) = \sum_{k=n}^{+\infty} \binom{k-1}{n-1} p^n q^{k-n} t^k = \left(\frac{pt}{1-qt}\right)^n$.
En utilisant ce résultat à l'ordre $n+1$ et en posant $x = qt$, on trouve pour $n \geq 2$ et $x \in]-1, 1[$:
$\frac{p^{n+1}x^{n+1}}{q^{n+1}(1-x)^{n+1}} = \sum_{k=n+1}^{+\infty} \binom{k-1}{n} \frac{p^{n+1}}{q^{n+1}} x^k$ puis, en divisant par $\frac{p^{n+1}}{q^{n+1}} x^{n+1}$:
$\frac{1}{(1-x)^{n+1}} = \sum_{k=n+1}^{+\infty} \binom{k-1}{n} x^{k-1-n}$, soit finalement :

$$\frac{1}{(1-x)^{n+1}} = \sum_{k=n}^{+\infty} \binom{k}{n} x^{k-n}$$

Il reste à considérer les cas $n = 0$ et $n = 1$:
Pour $n = 0$, on retrouve la série géométrique : $\sum_{k=0}^{+\infty} \binom{k}{0} x^{k-0} = \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}$;
Pour $n = 1$ on retrouve la série géométrique dérivée :
$\sum_{k=1}^{+\infty} \binom{k}{1} x^{k-1} = \sum_{k=1}^{+\infty} kx^{k-1} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{k=0}^{+\infty} x^k\right) = \frac{1}{(1-x)^2}$

Deuxième partie : Loi de Pascal

Dans cette partie, on sera amené à utiliser la formule du binôme négatif établie dans première partie. Soient $n \in \mathbb{N}^*$ et $p \in]0,1[$. On note $q = 1 - p$.

1. Montrer que la suite de réels

$$\forall k \in \mathbb{N}, \quad p_k = \binom{k+n-1}{n-1} p^n q^k$$

définit la loi de probabilité d'une variable aléatoire à valeurs dans $\mathbb{N}$.
On l'appelle *loi de Pascal de paramètres n et p*.
Les réels p_k étant positifs, il suffit de prouver que la série de terme général p_k converge, et que sa somme vaut 1.
On a, pour $N \in \mathbb{N}$: $\displaystyle\sum_{k=0}^{N} p_k = \sum_{k=0}^{N} \binom{k+n-1}{n-1} p^n q^k = p^n \sum_{k=0}^{N} \binom{k+n-1}{n-1} q^k = p^n \sum_{i=n-1}^{N+n-1} \binom{i}{n-1} q^{i-(n-1)}$.
La formule du binôme négatif assure la convergence, et donne : $\displaystyle\sum_{k=0}^{+\infty} p_k = p^n \times \frac{1}{(1-q)^n} = 1$.

2. Soit X une variable aléatoire suivant une telle loi. Déterminer la fonction génératrice de X.
Soit t dans le domaine de définition de G_X. Par définition :

$$G_X(t) = \sum_{k=0}^{+\infty} t^k \mathbb{P}(X=k) = \sum_{k=0}^{\infty} \binom{k+n-1}{n-1} p^n (qt)^k = \frac{p^n}{(1-qt)^n}$$

en utilisant la formule du binôme négatif.
En particulier, G_X est définie sur $\left]-\frac{1}{q}, \frac{1}{q}\right[$.

3. En déduire que X admet une espérance et une variance, et les calculer.
Comme $\frac{1}{q} > 1$, la fonction G_X est deux fois dérivable en 1. On en déduit que X admet une espérance et une variance, et on trouve :

$$\mathbb{E}(X) = G_X'(1) = \frac{nq}{p}; \quad \mathrm{V}(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2 = \frac{nq}{p^2}$$

4. Deux variables aléatoires indépendantes X et Y suivent deux lois de Pascal de paramètres respectifs (n,p) et (m,p). Déterminer la loi de $X+Y$.
Les variables aléatoires X et Y étant indépendantes, $G_{X+Y} = G_X \times G_Y$; on a donc :

$$\forall t \in \left]-\frac{1}{q}, \frac{1}{q}\right[, \quad G_{X+Y}(t) = \frac{p^n}{(1-qt)^n} \times \frac{p^m}{(1-qt)^m} = \frac{p^{n+m}}{(1-qt)^{n+m}}$$

$X+Y$ suit donc une loi de Pascal de paramètres $(n+m, p)$.