Correction de l'interrogation écrite N°1

06 mars 2012. Durée 2h

Question de cours

Soit X une variable aléatoire réelle. Pour tout $a > 0$, on a

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}[|X|]}{a}.$$

Exercice

1) L'espace probabilisé de ce modèlle est $\Omega := \{(a_i)_{i \leq n} \in \{1, 2, \ldots, n\}^n \,:\, a_i \neq a_j \,\forall i \neq j\}$, muni de l'équiprobabilité. Les boules rouges sont numérotées de 1 à k, et les boules vertes de $k+1$ à n. On cherche le cardinal des suites de n éléments distincts telles que $a_i \leq k$. On a k choix pour l'élément a_i, puis $(n-1)!$ choix pour les autres éléments. Or CardΩ=n! donc $\mathbb{P}(X_i = 1) = \frac{k(n-1)!}{n!} = \frac{k}{n}$. Puis $\mathbb{P}(X_i = 0) = 1 - \mathbb{P}(X_i = 1) = \frac{n-k}{n}$.

2) On a $\{T = i\} = \cap_{j=1}^{i-1}\{X_j = 0\} \cap \{X_i = 1\}$. On a donc k choix pour l'élément a_i, puis $A_{n-k}^{i-1} = \frac{(n-k)!}{(n-k-(i-1))!}$ pour les éléments $(a_1, \ldots, a_{i-1})$ (nombre d'arrangements de $i-1$ éléments pris parmi les $n-k$ boules vertes), puis $(n-i)!$ pour le reste des boules. On remarque que l'on doit avoir $n - k - i + 1 \geq 0$. Ainsi, si $i \leq n - k + 1$,

$$\mathbb{P}(T = i) = \frac{k(n-k)!(n-i)!}{(n-k-i+1)!n!}.$$

Si $i \geq n - k + 2$, on a nécessairement $\mathbb{P}(T = i) = 0$.

Problème

Partie I

1) On a

$$\sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} C_{n+k-1}^{n-1} p^n (1-p)^k = \sum_{k=0}^{\infty} \frac{(n+k-1)!}{(n-1)!k!} p^n (1-p)^k = \frac{p^n}{(n-1)!} \sum_{k=0}^{\infty} \frac{(n+k-1)!}{k!}(1-p)^k.$$

On utilise l'indication (1) avec $r = 1 - p$ ce qui donne

$$\sum_{k=0}^{\infty} \frac{(n+k-1)!}{k!}(1-p)^k = \frac{(n-1)!}{p^n}$$

et on conclut que $\sum_{k=0}^{\infty} p_k = 1$.

2) On a

$$G_X(s) = \sum_{k\geq 0} s^k \mathbb{P}(X = k) = \sum_{k\geq 0} s^k C_{n+k-1}^{n-1} p^n (1-p)^k = \frac{p^n}{(n-1)!} \sum_{k=0}^{\infty} \frac{(n+k-1)!}{k!} (s(1-p))^k.$$

On utilise cette fois (1) avec $r = s(1-p)$. On obtient que

$$G_X(s) = \left(\frac{p}{1 - s(1-p)} \right)^n.$$

On a $\mathbb{E}[X] = G'_X(1) = n\frac{1-p}{p}$. De même $\mathbb{E}[X(X-1)] = G''_X(1) = n(n+1)\left(\frac{1-p}{p}\right)^2$. Donc $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = n(n+1)\left(\frac{1-p}{p}\right)^2 + n\frac{1-p}{p} - n^2\left(\frac{1-p}{p}\right)^2 = n\frac{1-p}{p^2}$.

3) On a $G_X(s) = \left(\frac{p}{1-s(1-p)}\right)^n$ et $G_Y(s) = \left(\frac{p}{1-s(1-p)}\right)^m$. Comme X et Y sont indépendantes, $G_{X+Y}(s) = G_X(s)G_Y(s) = \left(\frac{p}{1-s(1-p)}\right)^{n+m}$. Donc $X + Y$ suit une loi binômiale négative de paramètre $n + m$.

<u>*Partie II*</u>

4) On a $\{T_1 = k\} = \cap_{i=1}^{k-1}\{\varepsilon_i = 0\} \cap \{\varepsilon_k = 1\}$. Comme les variables $(\varepsilon_i, i \geq 1)$ sont indépendantes, on a $\mathbb{P}(T_1 = k) = (1-p)^{k-1}p$ pour $k \geq 1$. Puisque $\mathbb{P}(T < \infty) = \sum_{k\geq 0} \mathbb{P}(T = k) = 1$, on a $\mathbb{P}(T < \infty) = 1$. On calcule

$$\mathbb{E}[T_1] = \sum_{k\geq 1} k(1-p)^{k-1}p = \frac{1}{p}.$$

De plus, $\mathbb{E}[T_1^2] = \sum_{k\geq 1} k^2(1-p)^{k-1}p = \frac{2-p}{p^2}$ d'où $\mathrm{Var}[T_1] = \frac{1-p}{p^2}$.

5) On cherche la probabilité $\mathbb{P}(T_1 = k, T_2 - T_1 = \ell)$. On a $\{T_1 = k\} \cap \{T_2 - T_1 = \ell\} = \cap_{i=1}^{k-1}\{\varepsilon_i = 0\} \cap \{\varepsilon_k = 1\} \cap_{j=k+1}^{k+\ell-1} \{\varepsilon_i = 0\} \cap \{\varepsilon_{k+\ell} = 1\}$. En utilisant l'indépendance des $(\varepsilon_i, i \geq 1)$, on obtient que

$$\mathbb{P}(T_1 = k, T_2 - T_1 = \ell) = (1-p)^{k-1}p(1-p)^{\ell-1}p$$

pour $k, \ell \geq 1$ et 0 si $k = 0$ ou $\ell = 0$. En particulier, T_1 et $T_2 - T_1$ sont indépendantes et $\mathbb{P}(T_2 - T_1 = \ell) = (1-p)^{\ell-1}p$ donc $T_2 - T_1$ a la même loi que T_1.

6) On écrit, pour $r \geq 2$,

$$\mathbb{P}(T_2 = r) = \sum_{k=1}^{r-1} \mathbb{P}(T_2 = r, T_1 = k) = \sum_{k=1}^{r-1} \mathbb{P}(T_2 - T_1 = r - k, T_1 = k) = \sum_{k=1}^{r-1} p^2(1-p)^{r-2} = (r-1)p^2(1-p)^{r-2}.$$

Ainsi, pour $k \geq 0$, $\mathbb{P}(T_2 - 2 = k) = (k+1)p^2(1-p)^k$. On remarque que c'est la loi d'une binômiale négative de paramètres 2 et p.

7) Plus généralement, calculons la probabilité $\mathbb{P}(T_n = k)$. Pour que $T_n = k$, il faut et il suffit qu'il y ait $n - 1$ "1" parmi les $k - 1$ premiers ε_i et que $\varepsilon_k = n$. On le réécrit $\{T_n = k\} = \{\sum_{i=1}^{k-1} \varepsilon_i = n - 1\} \cap \{\varepsilon_k = 1\}$. On sait que $\sum_{i=1}^{k-1} \varepsilon_i$ suit une loi $B(k-1, p)$. Donc par indépendance de $\sum_{i=1}^{k-1} \varepsilon_i$

et ε_k, on obtient que, pour $k \geq n$,

$$\mathbb{P}(T_n = k) = \mathbb{P}(\sum_{i=1}^{k-1} \varepsilon_i = n-1)\mathbb{P}(\varepsilon_k = 1) = C_{k-1}^{n-1} p^n (1-p)^{k-n}.$$

Il est alors facile de vérifier que $T_n - n$ suit une loi binômiale de paramètres n et p.

Partie III

8) Par linéarité de l'espérance, $\mathbb{E}[S_n] = n\mathbb{E}[X_1]$. On a $\mathbb{E}[X_1] = \sum_{k\geq 0} k\mathbb{P}(X_1 = k) = \sum_{k\geq 0} k(1-p)^k p = \frac{(1-p)}{p}$ donc $\mathbb{E}[S_n] = n\frac{1-p}{p}$. De plus, comme les $(X_i, i \geq 1)$ sont indépendantes, on a $\mathrm{Var}[S_n] = n\mathrm{Var}[X_1]$. On calcule $\mathbb{E}[X_1^2] = \sum_{k\geq 0} k^2(1-p)^k p = \frac{(1-p)(2-p)}{p^2}$ et $\mathrm{Var}[X_1] = \mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2 = \frac{1-p}{p^2}$. Finalement $\mathrm{Var}[S_n] = n\frac{1-p}{p^2}$.

9) C'est une conséquence de l'inégalité de Bienaymé-Tchebychev.

10) Comme les variables $(X_i, i \geq 1)$ sont indépendantes, on a $G_{S_n}(s) = G_{X_1}(s)^n$. On calcule $G_{X_1}(s)$. On a

$$G_{X_1}(s) = \sum_{k\geq 0}(1-p)^k p s^k = \frac{p}{1-s(1-p)}.$$

Donc $G_{S_n}(s) = \left(\frac{p}{1-s(1-p)}\right)^n$. On retrouve bien la fonction génératrice de la loi binômiale négative de paramètres n et p.

Question bonus. Calculons la fonction génératrice de S_N:

$$G_{S_N}(s) = \mathbb{E}\left[s^{S_N}\right] = \sum_{k=0}^{\infty} \mathbb{E}\left[s^{S_k} \mathbb{1}_{\{N=k\}}\right].$$

Comme N et les variables $(X_i, i \geq 1)$ sont indépendantes, on a $\mathbb{E}\left[s^{S_k}\mathbb{1}_{\{N=k\}}\right] = \mathbb{E}\left[s^{S_k}\right]\mathbb{P}(N = k)$. On a déjà calculé $\mathbb{E}\left[s^{S_k}\right] = \left(\frac{p}{1-s(1-p)}\right)^k$. De plus, $\mathbb{P}(N = k) = a^k(1-a)$. Donc

$$G_{S_N}(s) = (1-a)\sum_{k=0}^{\infty} a^k \left(\frac{p}{1-s(1-p)}\right)^k = \frac{1-a}{1-\frac{ap}{1-s(1-p)}} = \lambda\frac{1-\mu}{1-\mu s} + (1-\lambda)$$

avec $\lambda = a$ et $\mu = \frac{1-p}{1-ap}$. C'est la fonction génératrice de la variable YZ avec Y Bernoulli de paramètre λ et Z géométrique de paramètre μ.