UE MA6031, 2011-2012
Devoir surveillé, Mardi 13 Mars 2012, 8h00-9h30

Texte (en italiques) et corrigé (en roman)

Exercice 1.

Une urne contient R boules rouges et B boules blanches indiscernables au toucher. L'épreuve aléatoire (dans les questions **1, 2, 3***) consiste à répéter indéfiniment l'opération consistant à tirer une boule de l'urne, puis à la remettre dans l'urne (après avoir noté sa couleur). Il s'agit donc de tirages successifs avec remise, donc indépendants.*

1. *Que vaut la probabilité de tirer exactement k boules rouges lors des N premiers tirages (k désignant un entier tel que $0 \leq k \leq N$) ? [on en donnera une expression littérale] ?*

Il s'agit de tirages indépendants et avec remise hors d'une urne. Le modèle probabiliste est donc celui de la section 1.2.3 du cours. Le paramètre p vaut ici $p = R/(B+R)$ (proportion de boules rouges présentes dans l'urne). Le nombre de boules rouges tirées S_N suit une loi binomiale $\mathscr{B}(N,p)$. On a donc :

$$P(\{S_N = k\}) = \binom{N}{k}\Big(\frac{R}{R+B}\Big)^k\Big(\frac{B}{R+B}\Big)^{N-k} \qquad \forall\, k \in \{0,...,N\}.$$

2. *On suppose (seulement dans cette question) $R = B$. On note S_N la VAR représentant le nombre de boules rouges tirées lors des N premiers tirages. Quelle loi de probabilité suit la variable S_N ? Par quelle intégrale peut on approcher la probabilité de l'évènement*

$$A_N = \Big\{\omega\,;\, \frac{N-\sqrt{N}}{2} \leq S_N(\omega) < \frac{N+\sqrt{N}}{2}\Big\}$$

lorsque N devient grand ? [on précisera à partir de quelle valeur de N cette approximation est licite].

Ici $p = 1/2$, et la variable S_N suit donc une loi binomiale $\mathscr{B}(N,1/2)$. D'après le théorème de Berry-Esseen (formule (1.8) du cours), on peut, si F_{S_N} désigne la fonction de répartition de S_N, approcher les valeurs $F_{S_N}((N+\sqrt{N})/2)$ et $F_{S_N}((N-\sqrt{N})/2)$ respectivement par :

$$F_{S_N}((N+\sqrt{N})/2) = P\big(\{S_N \leq (N+\sqrt{N})/2\}\big) \simeq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{1} e^{-u^2/2}, du$$

$$F_{S_N}((N-\sqrt{N})/2) = P\big(\{S_N \leq (N-\sqrt{N})/2\}\big) \simeq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-1} e^{-u^2/2}\, du$$

(on applique la formule (1.8) en prenant ici comme indiqué $p = 1/2$, ce qui implique que $\sqrt{Np(1-p)} = \sqrt{N}/2$). L'erreur absolue dans chacune de ces approximations est contrôlée (*cf.* l'inégalité (1.6) du cours) en :

$$|\text{erreur}| \leq \frac{.7655}{\sqrt{N}}$$

et on la considère (suivant le cours) ces approximations licites dès que $N > 30$ (on a alors en effet aussi $N/2 = Np = N(1-p) > 5$). L'approximation de $P(A_N)$ est donc :

$$P(A_N) \simeq F_{S_N}((N+\sqrt{N})/2) - F_{S_N}((N-\sqrt{N})/2) \simeq \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-u^2/2}\, du.$$

3. *On reprend R et B quelconques. Quelle loi suit la VAR T représentant le numéro du tirage où l'on tire une boule rouge pour la première fois ? [on donnera la valeur de la probabilité de l'évènement $\{\omega\,;\, T(\omega) = k\}$ pour $k \in \mathbb{N}^*$] Calculer en fonction de R et B l'espérance de la VAR T. Que vaut la probabilité de l'évènement $\{\omega\,;\, T(\omega)$ est pair$\}$?*

Si $R = 0$, la VAR T est identiquement égale à $+\infty$ (on ne tire jamais de boule rouge), et son esp̂erance vaut $+\infty$; on suppose donc toujours ensuite $R > 0$. La VAR T suit une loi géométrique de paramètre $p = R/(R+B)$. On a en effet, si $p = R/(B+R) > 0$,

$$P(\{T = k\}) = p(1-p)^{k-1} \qquad \forall\, k \in \mathbb{N}^*$$

(attention ici à la petite différence avec le cours, où la loi géométrique est considérée sur $\mathbb{N}$ et non, comme ici, sur $\mathbb{N}^*$, il faut savoir s'adapter). L'espérance de cette VAR T se calcule (par exemple) comme la valeur de la dérivée en $s = 1$ de la fonction caractéristique (*cf.* (2.7) dans le cours, et surtout l'exemple 2.6) ; cette fonction caractéristique est dans ce cas la fonction :

$$s \longmapsto E[s^T] = p \sum_{k=1}^{\infty} (1-p)^{k-1} s^k = ps \sum_{k=0}^{\infty} (1-p)^k = \frac{ps}{1-(1-p)s}.$$

On a donc $E[T] = 1 + p(1-p)/p^2 = 1/p = (R+B)/R$. La probabilité que T prenne une valeur paire vaut :

$$\begin{aligned} P\big(\{T \text{ pair}\}\big) &= \sum_{k=1}^{\infty} P\big(\{T = 2k\}\big) = p \sum_{k=1}^{\infty} (1-p)^{2k-1} \\ &= p(1-p) \sum_{k=0}^{\infty} (1-p)^{2k} = \frac{p(1-p)}{1-(1-p)^2} = \frac{1-p}{2-p} \\ &= \frac{B}{2B+R}. \end{aligned}$$

4. *On considère à nouveau une urne renfermant R boules rouges et B boules blanches indiscernables au toucher. On dispose d'autre part (hors de l'urne) d'un stock infini de boules des deux couleurs. On effectue une suite de tirages successifs suivant la règle suivante : si l'on tire une boule d'une certaine couleur (rouge ou blanche) au k-ième tirage, on observe sa couleur, puis on la remet dans l'urne, mais en y ajoutant en même temps $m \in \mathbb{N}^*$ boules de la couleur de la boule tirée (prises dans notre stock), ce avant de passer au tirage suivant. On note, pour $k \in \mathbb{N}^*$, X_k la VAR valant 1 si la boule tirée au k-ième tirage est rouge, 0 si elle est blanche.*

a) *Calculer les probabilités des évènements $\{\omega\,;\, X_k(\omega) = 1\}$ pour $k = 1, 2$, puis celle de l'évènement $\{\omega\,;\, X_1(\omega) = X_2(\omega) = 1\}$. Les VAR X_1 et X_2 sont-elles indépendantes ?*

On a $P(\{X_1 = 1\}) = R/(B + R)$. Pour calculer $P(\{X_2 = 1\}$, on utilise la formule des probabilités totales (Proposition 1.2, formule de Bayes (1.21) du cours). On a

$$\begin{aligned} P(\{X_2 = 1\}) &= P\big(\{X_2 = 1\} \mid \{X_1 = 1\}\big)\, P\big(\{X_1 = 1\}\big) \\ &\quad + P\big(\{X_2 = 1\} \mid \{X_1 = 0\}\big)\, P\big(\{X_1 = 0\}\big) \\ &= \frac{R+m}{R+B+m}\,\frac{R}{R+B} + \frac{R}{R+B+m}\,\frac{B}{R+B} \\ &= \frac{R(R+m+B)}{(R+B)(R+B+m)} = \frac{R}{R+B}. \end{aligned}$$

On a aussi :

$$\begin{aligned} P\big(\{X_1 = X_2 = 1\}\big) &= P\big(\{X_2 = 1\} \mid \{X_1 = 1\}\big)\, P\big(\{X_1 = 1\}\big) \\ &= \frac{(R+m)R}{(R+B)(R+B+m)}. \end{aligned}$$

Si $m \neq 0$, on a

$$\frac{(R+m)R}{(R+B)(R+B+m)} \neq \Big(\frac{R}{R+B}\Big)^2 = P\big(\{X_1 = 0\}\big)\, P\big(\{X_2 = 0\}\big),$$

ce qui prouve que les VAR X_1 et X_2 ne sont pas indépendantes (la condition (2.14) de la Définition 2.5 est en effet ici en défaut). Si $m = 0$ (tirages avec remise), on sait par contre que X_1 etX_2 sont indépendantes.

b) *On renote S_N la VAR $S_N := X_1 + \cdots + X_N$. Que représente S_N ? Calculer les probabilités conditionnelles $P(\{\omega\,;\, X_{N+1}(\omega) = 1\} \mid \{\omega\,;\, S_N(\omega) = k\})$ pour $k = 0, ..., N$. En déduire la formule*

$$P(\{\omega\,;\, X_{N+1}(\omega) = 1\}) = \frac{R + mE(S_N)}{R + B + Nm}.$$

La variable S_N représente le nombre de boules rouges tirées (puis remises, chaque fois avec m autres) au terme des N tirages. Pour tout $k = 0, ..., N$, on a :

$$P\big\{X_{N+1} = 1\} \mid \{S_N = k\}\big) = \frac{R + km}{R + B + mN} \ ;$$

en effet, au terme de N tirages qui ont vu sortir exactement k boules rouges, la composition de l'urne est de $R + km$ boules rouges et de $B + m(N - k)$ boules blanches. La formule des probabilités totales (formule de Bayes (1.21), Proposition 1.2 du cours) donne :

$$\begin{aligned} P\big(\{X_{N+1} = 1\}\big) &= \sum_{k=0}^{N} P\big(\{X_{N+1} = 1\} \mid \{S_N = k\}\big)\, P\big(\{S_N = k\}\big) \\ &= R \sum_{k=0}^{N} \frac{P\big(\{S_N = k\}\big)}{R + B + mN} + m \frac{\sum_{k=0}^{N} k P\big(\{S_N = k\}\big)}{R + B + mN} \\ &= \frac{R}{R + B + mN} + \frac{mE[S_N]}{R + B + mN} \end{aligned}$$

puisque l'ensemble $\{0, ..., N\}$ représente l'ensemble de toutes les valeurs possibles que peut prendre la VAR S_N.

c) *Montrer que toutes les variables X_k, $k \in \mathbb{N}^*$, ont même loi, et que cette loi ne dépend pas de m. Sont-elles mutuellement indépendantes ? deux à deux indépendantes ? En déduire la valeur de l'espérance $E(S_N)$.*

On montre que toutes les variables X_k, $k \in \mathbb{N}^*$, sont des VAR de Bernoulli de même paramètre, en montrant, par récurrence sur l'entier $N \in \mathbb{N}^*$, que, pour tout $N \in \mathbb{N}^*$, les VAR $X_1, ..., X_N$ sont toutes des VAR de Bernoulli de paramètre $R/(B + R)$. C'est le cas de X_1 (même de X_2, *cf.* le résultat de la question **a)**). Supposons le résultat vrai jusqu'au rang N. On a alors

$$E[S_N] = E[X_1] + E[X_2] + \cdots + E[X_N] = N \frac{R}{B + R}$$

d'après la linéarité de l'opération de prise d'espérance (*cf.* la sous-section 2.1.2 du cours). On a donc, en utilisant le résultat établi au **b)**,

$$\begin{aligned} P\big(\{X_{N+1} = 1\}\big) &= \frac{R}{R + B + mN} + \frac{mN}{R + B + mN} \frac{R}{R + B} \\ &= \frac{R}{R + B + mN}\Big(1 + \frac{mN}{R + B}\Big) = \frac{R}{R + B}, \end{aligned}$$

d'o`ù l'on conclut que les VAR X_1,...,X_{N+1}, ont bien même loi. Le résultat voulu est donc bien démontré par récurrence. Comme X_1 et X_2 ne sont pas indépendantes (*cf.* la question **a)**), les VAR X_k, $k \in \mathbb{N}^*$ ne sont pas deux-à-deux indépendantes. Elles ne sont *a fortiori* pas mutuellement indépendantes, puisque l'on sait (sous-section 2.1.5 du cours) que l'indépendance mutuelle entre VAR implique leur indépendance deux-à-deux. L'espérance de S_N vaut, on l'a vu, $E[S_N] = NE[X_1] = NR/(B+R)$.

d) *On reprend la même règle, mais cette fois avec $m = -1$ (ce qui revient à effectuer des tirages successifs sans remise), en conservant les mêmes notations. Quelle loi suit la VAR S_N lorsque $N \leq R+B$? [on précisera l'univers des possibles Ω et la valeur de $P(\{\omega\,;\, S_N(\omega) = k\})$ si $k \in \Omega$]. Justifier la formule $E(S_N) = NR/(R+B)$ en vous inspirant de la démarche adoptée aux questions* **a), b)**.

La VAR S_N suit dans ce cas une loi hypergéométrique. $\mathscr{H}(N, R, B)$ (d'après la sous-section 1.2.4 du cours). L'univers des possibles est $\Omega = \{\max(0, N - B), ..., \min(N, R)\}$, et l'on a, pour tout k dans cet ensemble :

$$P\big(\{S_N = k\}\big) = \frac{\binom{R}{k} \times \binom{B}{N-k}}{\binom{R+B}{N}}.$$

Les calculs effectués aux questions **a)** et **b)** précédentes peuvent être repris avec cette fois $m = -1$. La seule différence est que l'on n'utilise comme valeurs de k dans la question **b)** que les valeurs comprises entre $\max(0, N-B)$ et $\min(N, R)$ (correspondant à l'univers des possibles de la VAR S_N). Ce sont uniquement ces valeurs de k que l'on fait entrer en jeu dans la formule des probabilités totales utilisée aux questions **b)** et **c)** pour calculer $P\big(\{X_{N+1} = 1\}\big)$ lorsque les $P\big(\{X_k = 1\}\big)$ sont supposés tous égaux à $R/(R+B)$ lorsque $k = 1, ..., N$. On a donc encore $E[S_N] = NE[X_1] = NR/(R+B)$.

Exercice 2.

Mr X. rentre chez lui le soir avec un trousseau de $k \geq 2$ clefs dont seule une ouvre sa porte. Lorsqu'il est à jeun, il essaye une des clefs au hasard, puis, si elle ne fonctionne pas, la met de côté et essaye (toujours au hasard) une des clefs restantes ; il répète cette opération après chaque échec. Lorsqu'il est ivre, il essaye une clef au hasard, puis, si elle ne fonctionne pas, recommence avec cette fois le trousseau complet ; il répète cette opération après chaque échec. On note A l'évènement « Mr. X. est à jeun » et B l'évènement « Mr. X. est ivre ».

1. *Calculer, sachant que Mr. X. est à jeun, la probabilité qu'il parvienne à ouvrir la porte au terme exactement de n tentatives ($n \in \mathbb{N}^*$). On désigne par X_A la variable aléatoire représentant le nombre d'essais nécessaires à Mr. X. pour parvenir à ouvrir sa porte quand il est à jeun. Quelle loi suit la VAR X_A ? Que vaut l'espérance de X_A ?*

On a :
$$P\big(\{X_A = 1\}\big) = \frac{1}{k}$$
(car une seule clef du trouseau ouvre la porte et que le choix d'une clef est fait au hasard, ce qui signifie que les k choix possibles sont équiprobables).
On a aussi :
$$P(\big(\{X_A = 2\}\big) = \Big(1 - \frac{1}{k}\Big) \times \frac{1}{k-1} = \frac{1}{k}$$
puisque le trousseau ne comprend plus que $k-1$ clefs au second essai, le premier s'étant révélé infructeux : l'évènement $\{X_A = 2\}$ est intersection des deux évènements E_1 : ≪ *le premier essai se solde par un échec* ≫ et S_2 : ≪ *le second essai se solde par un succès* ≫. On utilise ici :
$$P\big(\{X_A = 2\}\big) = P(E_1 \cap \{X_A = 2\}\big) = P(E_1 \cap S_2) = P(E_1) \times P(S_2 \mid E_1).$$
On constate en itérant ce calcul que, pour $n \leq k$:
$$P\big(\{X_A = n\}\big) = \frac{k-1}{k} \times \frac{k-2}{k-1} \times \cdots \times \frac{k-(n-1)}{k-(n-2)} \times \frac{1}{k-(n-1)} = \frac{1}{k}.$$
Si $n > k$, on trouve $P\big(\{X_A = n\}\big) = 0$ (au terme de $k-1$ essais infructueux, il ne reste plus qu'une clef dans le trousseau, qui est forcément celle qui ouvre).
La VAR X_A suit donc la loi uniforme sur $\{1, ..., k\}$. Son espérance vaut :
$$E[X_A] = \frac{1}{k}(1 + 2 + \cdots + k) = \frac{k+1}{2}.$$
2. *Calculer, sachant que Mr. X. est ivre, la probabilité qu'il parvienne à ouvrir la porte au terme exactement de n tentatives ($n \in \mathbb{N}^*$). On désigne par X_B la variable aléatoire représentant le nombre d'essais nécessaires à Mr. X. pour parvenir à ouvrir sa porte quand il est ivre. Quelle loi suit la VAR X_B ? Que vaut l'espérance de X_B ?*
On a, pour tout $k \in \mathbb{N}^*$,
$$P\big(\{X_B = n\}\big) = \frac{1}{k}\Big(1 - \frac{1}{k}\Big)^{n-1}.$$
La VAR X_B suit donc une loi géométrique de paramètre $p = 1/k$ (*cf.* l'exercice I, question **3**), toujours considérée sur $\mathbb{N}^*$ et non (comme c'est le cas dans le cours) sur $\mathbb{N}$. L'espérance de X_B vaut donc $1/p = k$.
3. *On suppose $P(A) = P(B) = 1/2$. Calculer la probabilité π_n que Mr. X soit ivre, sachant qu'il est parvenu à ouvrir sa porte au terme d'exactement n essais ($n \in \mathbb{N}^*$) [on pensera à distinguer les cas où $n \leq k$ et $n > k$].*

Soit U_n l'évènement : ≪ *il a fallu a Mr. X. exactement* n *essais pour ouvrir la porte* ≫. Supposons tout d'abord $n \leq k$. On sait que $P(U_n \mid A) = 1/k$ et que $P(U_n \mid B) = (1/k)(1-1/k)^{n-1}$. On cherche à calculer $\pi_n = P(B \mid U_n)$. On utilise pour cela la formule des causes (Proposition 1.3) qui stipule que :

$$P(B \mid U_n) = \frac{P(U_n \mid B)\, P(B)}{P(U_n \mid A)P(A) + P(U_n \mid B)P(B)}.$$

Comme ici $P(A) = P(B) = 1/2$, on a donc :

$$\begin{aligned}
\pi_n &= P(B \mid U_n) = \frac{P(U_n \mid B)}{P(U_n \mid A) + P(U_n \mid B)} = \frac{\frac{1}{k}\left(1-\frac{1}{k}\right)^{n-1}}{\frac{1}{k} + \frac{1}{k}\left(1-\frac{1}{k}\right)^{n-1}} \\
&= \frac{k^{n-1}}{k^{n-1} + (k-1)^{n-1}}.
\end{aligned}$$

Si $n > k$, on a $P(U_n \mid A) = 0$ et, par conséquent $\pi_n = 1$.