Tutorat MAP311 : Feuille d'exercice 2

Nicolas Kielbasiewicz*

15 juin 2007

3 Probabilités conditionnelles

Exercice 3.1 *On suppose que l'on a autant de chance d'avoir une fille ou un garçon à la naissance. Votre voisin de palier vous dit qu'il a deux enfants.*

1. *Quelle est la probabilité qu'il ait au moins un garçon ?*
2. *Quelle est la probabilité qu'il ait un garçon, sachant que l'ainée est une fille ?*
3. *Quelle est la probabilité qu'il ait un garçon, sachant qu'il a au moins une fille ?*
4. *Vous téléphonez à votre voisin. Une fille décroche le téléphone. Vous savez que dans une famille avec un garçon et une fille, la fille décroche avec une probabilité p. Quelle est la probabilité que votre voisin ait un garçon ?*
5. *Vous sonnez à la porte de votre voisin. Une fille ouvre la porte. Sachant que l'ainé(e) ouvre la porte avec probabilité p, et ce indépendamment de la répartition de la famille, quelle est la probabilité que votre voisin ait un garçon ?*

Correction de l'exercice 3.1 *On suppose que l'on a autant de chance d'avoir une fille ou un garçon à la naissance. Votre voisin de palier vous dit qu'il a deux enfants.*
On décrit donc une réalisation $\omega = (\omega_1, \omega_2)$ où ω_1 est le sexe de l'ainé(e) (F ou G), et ω_2 le sexe du cadet. L'espace d'état ou univers est donc :

$$\Omega = \{(G,G); (G,F); (F,G); (F,F)\}$$

Les naissances étant équiprobables, on choisit la probabilité uniforme sur Ω, qui sera utilisée dans presque toutes les questions de l'exercice.

1. *Quelle est la probabilité qu'il ait au moins un garçon ?*
   Il y a deux raisonnements possibles
   – On raisonne par dénombrement sur l'univers Ω
   – On raisonne sur l'événement contraire qui est «avoir deux filles », dont la probabilité est 1/4.
   On obtient donc dans les deux cas que la probabilité d'avoir au moins un garçon est $\frac{3}{4}$.
2. *Quelle est la probabilité qu'il ait un garçon, sachant que l'ainée est une fille ?*
   Cette fois-ci, on peut aussi bien raisonner par dénombrement, en prenant en compte l'information supplémentaire, ou utiliser la définition des probabilités conditionnelles, ce que nous ferons ici.

$$\mathbb{P}(\omega_2 = G|\omega_1 = F) = \frac{\mathbb{P}(\omega_2 = G, \omega_1 = F)}{\mathbb{P}(\omega_1 = F)} = \frac{1/4}{1/2} = \frac{1}{2}$$

*Unité de Mathématiques Appliquées, École Nationale Supérieure de Techniques Avancées

3. *Quelle est la probabilité qu'il ait un garçon, sachant qu'il a au moins une fille ?*
   Toujours les deux mêmes raisonnements possibles. En utilisant la définition des probabilités conditionnelles, et le résultat de la première question, on a :

$$\mathbb{P}(\exists G|\exists F) = \frac{\mathbb{P}(\exists G, \exists F)}{\mathbb{P}(\exists F)} = \frac{1/2}{3/4} = \frac{2}{3}$$

4. *Vous téléphonez à votre voisin. Une fille décroche le téléphone. Vous savez que dans une famille avec un garçon et une fille, la fille décroche avec une probabilité p. Quelle est la probabilité que votre voisin ait un garçon ?*
   Combien avons-nous d'informations ?

   «Une fille décroche le téléphone» : Le voisin a donc une fille, et elle décroche le téléphone

   «... probabilité p» : Si les deux enfants sont de sexes différents, alors la fille décroche avec une probabilité p

   Qui décroche ? : La façon dont est rédigé l'énoncé nous amène à supposer que c'est nécessairement l'un des enfants qui décroche.

   Cette fois-ci, on ne peut plus raisonner par dénombrement, car on a maintenant des événements d'une autre nature que le sexe des enfants. On applique donc la formule des probabilités conditionnelles :

$$\mathbb{P}(\exists G|\exists F, F\text{décroche}) = \frac{\mathbb{P}(\exists G, \exists F, F\text{décroche})}{\mathbb{P}(\exists F, F\text{décroche})}$$

   Calculons le numérateur. On va utiliser la définition des probabilités conditionnelles pour utiliser la probabilité p introduite :

$$\mathbb{P}(\exists G, \exists F, F\text{décroche}) = \mathbb{P}(F\text{décroche}|\exists G, \exists F)\mathbb{P}(\exists G, \exists F) = \frac{p}{2}$$

   Calculons maintenant le dénominateur. On sent qu'on va devoir de nouveau faire apparaître la probabilité p, mais on ne connait pas la probabilité que la fille décroche sachant qu'il existe au moins une fille. Il nous faut donc décomposer l'événement $\exists F$ en deux événements disjoints, à savoir «$\exists! F$»et «$\omega = (F, F)$», d'où, par indépendance :

$$\mathbb{P}(\exists F, F\text{décroche}) = \mathbb{P}(\omega = (F, F))\mathbb{P}(F\text{décroche}|\omega = (F, F)) + \mathbb{P}(\exists G, \exists F|\ F\text{décroche}) = \frac{1}{4} + \frac{p}{2}$$

   On obtient en définitive que :

$$\mathbb{P}(\exists G|\exists F, F\text{décroche}) = \frac{2p}{1 + 2p}$$

5. *Vous sonnez à la porte de votre voisin. Une fille ouvre la porte. Sachant que l'ainé(e) ouvre la porte avec probabilité p, et ce indépendamment de la répartition de la famille, quelle est la probabilité que votre voisin ait un garçon ?*
   Quelles sont les informations à prendre en compte ?

   «Une fille ouvre la porte » : Le voisin a donc une fille, et elle ouvre la porte.

   «l'ainé(e) ouvre la porte avec probabilité p » : Cette information est à prendre telle quelle.

   Qui ouvre la porte ? La formulation de la question nous amène à considérer que seuls les deux enfants ouvrent la porte.

   On va donc utiliser la définition des probabilités conditionnelles :

$$\mathbb{P}(\exists G|\exists F, F\text{ouvre la porte}) = \frac{\mathbb{P}(\exists G, \exists F, F\text{ouvre la porte})}{\mathbb{P}(\exists F, F\text{ouvre la porte})}$$

Calculons le numérateur. Les informations disponibles concernent non pas le sexe, mais l'âge de l'enfant. On va donc utiliser une formule de décomposition sur l'ainé/cadet, afin d'utiliser la probabilité p. On obtient donc, par indépendance :

$$\begin{aligned} \mathbb{P}(\exists G, \exists F, F\text{ouvre la porte}) &= \mathbb{P}(\omega_1\text{ouvre la porte})\mathbb{P}(\omega = (F,G)) + \mathbb{P}(\omega_2\text{ouvre la porte})\mathbb{P}(\omega = (G,F)) \\ &= \frac{p}{4} + \frac{1-p}{4} = \frac{1}{4} \end{aligned}$$

Calculons maintenant le numérateur. On obtient de la même manière que :

$$\begin{aligned} \mathbb{P}(\exists F, F\text{ouvre la porte}) &= \mathbb{P}(\omega_1\text{ouvre la porte})\mathbb{P}(\omega_1 = F) + \mathbb{P}(\omega_2\text{ouvre la porte})\mathbb{P}(\omega_2 = F) \\ &= \frac{p}{2} + \frac{1-p}{2} = \frac{1}{2} \end{aligned}$$

On obtient en définitive que :

$$\mathbb{P}(\exists G | \exists F, F\text{ouvre la porte}) = \frac{1}{2}$$

6 Lois conditionnelles

Exercice 6.1 *Soit $(X_1, \ldots, X_n)$ une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre $p \in]0,1[$. On note $S_n = \sum_{i=1}^{n} X_i$.*

1. *Calculer la loi de $(X_1, \ldots, X_n)$ conditionnellement à S_n.*
2. *Calculer la loi de X_i conditionnellement à S_n ($n > i$).*
3. *Les variables X_1 et X_2 sont-elles indépendantes conditionnellement à S_n ?*

Correction de l'exercice 6.1 *Soit $(X_1, \ldots, X_n)$ une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre $p \in]0,1[$. On note $S_n = \sum_{i=1}^{n} X_i$.*

1. *Calculer la loi de $(X_1, \ldots, X_n)$ conditionnellement à S_n.*
   Avant de commencer, petite précision sur les lois conditionnelles. Quand on calcule la probabilité d'un événement A sachant un autre événement B, on peut considérer que cet événement A est une réalisation possible x d'une variable aléatoire X. Par extension, on va pouvir déterminer la loi de X sachant l'événement B. Maintenant, on peut aussi considérer que l'événement B est une réalisation possible s d'une variable aléatoire S. En calculant la loi de X sachant l'événement$S = s$ pour toute valeur de s, on va ainsi calculer la loi de X conditionnellement à S. Il semble donc naturel que cette loi conditionnelle soit une expression dépendant de S. En pratique, on va calculer $\mathbb{P}(X = x | S = s)$ pour toute valeur de s possible, pour en déduire la relation cherchée.

   Forts de cette précision, revenons maintenant à la question posée, et calculons

   $$\mathbb{P}(X_1 = k_1, \ldots, X_n = k_n | S_n = k)$$

   On peut constater immédiatement que par définition de S_n, si $k \neq \sum_{i=1}^{n} k_i$, alors cette probabilité est nulle (probabilité conditionnelle entre deux événements impossibles). Dans le cas d'égalité, on va calculer cette probabilité en utilisant la définition des probabilités conditionnelles :

   $$\mathbb{P}(X_1 = k_1, \ldots, X_n = k_n | S_n = k) = \frac{\mathbb{P}(X_1 = k_1, \ldots, X_n = k_n, S_n = k)}{\mathbb{P}(S_n = k)}$$

Les variables X_i sont des variables aléatoires indépendantes de loi de Bernoulli de paramètre p, alors, par définition, S_n suit une loi binômiale de paramètres n et p. En ce qui concerne le numérateur, les variables X_i étant indépendantes, on va donc obtenir le produit des $\mathbb{P}(X_i = k_i)$, c'est à dire soit p, soit $1 - p$. Or, puisque $S_n = k$, on sait combien de k_i valent 1. D'où :

$$\mathbb{P}(X_1 = k_1, \dots, X_n = k_n | S_n = k) = \frac{p^k(1-p)^{n-k}}{\binom{n}{k}p^k(1-p)^{n-k}} = \frac{1}{\binom{n}{k}}$$

On contient donc une probabilité indépendante des k_i, mais seulement dépendante de k, et ce quelle que soit la valeur de k. Il s'agit donc d'une loi uniforme. Suite à la remarque préliminaire, on notera donc que la loi conditionnelle de$(X_1, \dots, X_n)$ sachant S_n, est définie par :

$$\mathbb{P}(X_1 = k_1, \dots, X_N = k_n | S_n) = \frac{1}{\binom{n}{S_n}} \mathbb{I}_{S_n = \sum k_i}$$

2. *Calculer la loi de X_i conditionnellement à S_n $(n > i)$.*
   On serait tenté de reprendre le raisonnement précédent, mais il y a cette fois-ci un raisonnement beaucoup plus fin. En effet, puisque $n > i$, la variable X_i peut prendre les valeurs 0 ou 1, que l'on conditionne par rapport à S_n ou non. La loi de X_i conditionnellement à S_n est donc une loi de Bernoulli de paramètre $\tilde{p}$. Il nous suffit donc de calculer $\mathbb{P}(X_i = 1 | S_n = k)$, et ce par un dénombrement après avoir appliqué la définition des probabilités variationnelles. En effet, avoir $X_i = 1$ et $S_n = k$ revient à dénombrer toutes les possibilités d'avoir $k-1$ sur les $X_{j \neq i}$, en supposant $k \neq 0$. On en déduit donc que :

$$\mathbb{P}(X_i = 1 | S_n = k) = \frac{\mathbb{P}(X_i = 1, S_n = k)}{\mathbb{P}(S_n = k)} = \frac{p\binom{n-1}{k-1}p^{k-1}(1-p)^{(n-1)-(k-1)}}{\binom{n}{k}p^k(1-p)^{n-k}} = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{k}{n}$$

   Cette relation reste vraie pour $k = 0$ (les événements «$X_i = 1$» et «$S_n = 0$» sont contraires). On obtient donc en définitive que $\tilde{p} = \frac{S_n}{n}$, ce qui termine de caractériser la loi de X_i conditionnellement à S_n.

3. *Les variables X_1 et X_2 sont-elles indépendantes conditionnellement à S_n ?*
   Par un raisonnement similaire à la question précédente, on montre que :

$$\mathbb{P}(X_1 = 1, X_2 = 1 | S_n = k) = \frac{k(k-1)}{n(n-1)} \neq \frac{k^2}{n^2}$$

   D'où :

$$\mathbb{P}(X_1 = 1, X_2 = 1 | S_n = k) \neq \mathbb{P}(X_1 = 1 | S_n = k)\mathbb{P}(X_2 = 1 | S_n = k)$$

   On en déduit donc que les variables X_1 et X_2 ne sont pas indépendantes conditionnellement à S_n, alors qu'elles sont indépendantes.

8 Variables aléatoires continues

Exercice 8.3 *Soit X et Y deux variables aléatoires indépendantes de lois respectives $\Gamma(\lambda, a)$ et $\Gamma(\lambda, b)$. Rappelons la définition de la* loi gamma $\Gamma(\lambda, a)$ de paramètres $\lambda > 0$ et $a > 0$ *par sa densité :*

$$f(x) = \frac{1}{\Gamma(a)} \lambda^a x^{a-1} e^{-\lambda x} \mathbb{I}_{x>0}$$

1. *Calculer la loi du couple $(X + Y, \frac{X}{X+Y})$.*
2. *Montrer que $X + Y$ et $\frac{X}{X+Y}$ sont indépendantes et identifier leurs lois.*

Correction de l'exercice 8.3 *Soit X et Y deux variables aléatoires indépendantes de lois respectives $\Gamma(\lambda, a)$ et $\Gamma(\lambda, b)$. Rappelons la définition de la* loi gamma $\Gamma(\lambda, a)$ de paramètres $\lambda > 0$ et $a > 0$ *par sa densité :*

$$f(x) = \frac{1}{\Gamma(a)}\lambda^a x^{a-1}e^{-\lambda x}\mathbb{I}_{x>0}$$

1. *Calculer la loi du couple $(X+Y, \dfrac{X}{X+Y})$.*

   Soit $S = X+Y$ et $T = \dfrac{X}{X+Y}$. D'après le cours, on va utiliser l'ésperance pour caractériser la loi de (S,T).

$$\begin{aligned}
\mathbb{E}[h(S,T)] &= \mathbb{E}[h(X+Y, \frac{X}{X+Y}] \\
&= \mathbb{E}[g(X,Y)] \\
&= \int_{\mathcal{D}} g(x,y)f_{(X,Y)}(x,y)dxdy \\
&= \int_{\mathcal{D}} h(x+y, \frac{x}{x+y})f_X(x)f_Y(y)dxdy \quad \text{par indépendance de } X \text{ et } Y \\
&= \int_{\mathcal{D}} h(s,t)\frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}e^{-\lambda(x+y)}x^{a-1}y^{b-1}dxdy
\end{aligned}$$

   On va donc maintenant effectuer le changement de variable $(x,y) \longmapsto (s,t)$. Le Jacobien de cette transformation est $\dfrac{1}{x+y}$. Et puisque $x = st$ et $y = s(1-t)$, on obtient :

$$\mathbb{E}[h(S,T)] = \int_{\mathcal{D}} h(s,t)\frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}e^{-\lambda s}s^{a+b-1}t^{a-1}(1-t)^{b-1}dsdt$$

   Ce qui permet d'obtenir la loi demandée, avec $s > 0$ et $0 < t < 1$.

2. *Montrer que $X+Y$ et $\dfrac{X}{X+Y}$ sont indépendantes et identifier leurs lois.*

   La densité obtenue à la question précédente est le produit d'une fonction de s et d'une fonction de t. CEla permet d'affirmer que les variables S et T sont indépendantes. On obtient les densiés de S et T par la formule des lois marginales, et on reconnait pour S la densité de loi Gamma de paramètres λ et $a+b$. Pour T, on obtient la densité :

$$f_T(t) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}t^{a-1}(1-t)^{b-1}\mathbb{I}_{]0,1[}$$

   Pour information, il s'agit de la densité de la loi béta de paramètres a et b.