

Partiel du module RFIDEC

C. Gonzales

Durée : 2 heures

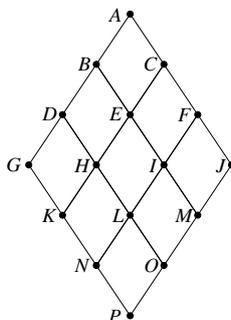
Seuls documents autorisés :

Tous documents autorisés, sauf la copie et les brouillons des voisins.

Calculatrices, ordinateurs et téléphones portables interdits.

Exercice 1 (6 pts)

Un robot doit se rendre du point A au point P en passant par les arêtes du graphe ci-dessous. Le robot est limité dans ses mouvements, aussi ne peut-il que descendre (par exemple, lorsqu'il est en E , il ne peut aller qu'en H ou en I , mais pas en B). Lorsqu'il est sur un nœud du graphe, il peut descendre soit sur l'arête de gauche, soit sur celle de droite. Son programme lui fait choisir 7 fois sur 10 l'arête de gauche et 3 fois sur 10 celle de droite.



Q 1.1 Calculez la probabilité que le robot passe en G pour aller vers P . Faites de même avec H , I , et enfin J .

Il est clair que le choix du robot d'aller à droite ou à gauche sur un embranchement ne dépend pas des embranchements précédents, les choix sont donc tous indépendants. De plus, il y a toujours une probabilité de 7/10ème d'aller à gauche et de 3/10ème d'aller à droite. Par conséquent, les choix représentent des épreuves de Bernoulli et la variable aléatoire $X =$ « nombre de fois où l'on a pris à droite pour atteindre G , H , I ou J » suit une loi binômiale de paramètres 3 et 3/10.

Pour aller en G , il faut aller 3 fois à gauche. Donc $P(G) = P(X = 0) = \frac{3!}{3! \times 0!} \times \left(\frac{3}{10}\right)^0 \left(\frac{7}{10}\right)^3 = 0,343$.

De même, $P(H) = P(X = 2) = 0,441$, $P(I) = P(X = 1) = 0,189$ et $P(J) = P(X = 0) = 0,027$.

On pouvait trouver ces probabilités différemment : il est évident que $P(B) = 0,7$ et $P(C) = 0,3$.

Pour arriver en D , il faut partir de B , donc $P(D) = P(D, B) = P(D|B)P(B) = 0,7 \times 0,7 = 0,49$.

De même, pour arriver en E , il faut soit venir de B , soit de C , et ces deux conditions sont en exclusion mutuelle. Par conséquent, $P(E) = P(E, B) + P(E, C) = P(E|B)P(B) + P(E|C)P(C) = 0,3 \times 0,7 + 0,7 \times 0,3 = 0,42$. Enfin, $P(F) = P(F|C)P(C) = 0,3 \times 0,3 = 0,09$. On peut maintenant appliquer le même procédé sur la ligne contenant les nœuds G , H , I et J :

$P(G) = P(G|D)P(D) = 0,7 \times 0,49 = 0,343$, $P(H) = P(H|D)P(D) + P(H|E)P(E) = 0,3 \times 0,49 + 0,7 \times 0,42 = 0,441$, $P(I) = P(I|E)P(E) + P(I|F)P(F) = 0,3 \times 0,42 + 0,7 \times 0,09 = 0,189$. Enfin, $P(J) = P(J|F)P(F) = 0,3 \times 0,09 = 0,027$.

Q 1.2 Soit X une variable aléatoire valant 0 si le robot est passé en G , 1 s'il est passé en H , 2 en I et 3 en J . Quelle est la loi de probabilité suivie par X ? Justifiez votre réponse.

Comme nous l'avons vu précédemment, les déplacements en chaque nœud constituent une épreuve de Bernoulli. En remarquant que, pour aller en G , on ne se déplace qu'à gauche, que pour aller en H , on se déplace exactement une fois vers la droite, que pour aller en I , on se déplace exactement 2 fois vers la droite, et que pour aller en J , on se déplace 3 fois vers la droite, on peut déduire que la variable X correspond au nombre de succès des épreuves de Bernoulli pour arriver jusqu'en G , H , I ou J . Pour arriver sur ces nœuds, on a fait trois déplacements, autrement dit on a effectué trois épreuves de Bernoulli, dont on a vu qu'elles étaient indépendantes les unes des autres et avaient toutes la même probabilité de succès : 0,3. Par conséquent, $X \sim \text{Bin}(3, 0.3)$.

Exercice 2 (4 pts)

Soit trois variables aléatoires X , Y , Z , ayant respectivement pour domaines $\{x_1, x_2\}$, $\{y_1, y_2\}$ et $\{z_1, z_2\}$. On a pu déterminer la probabilité jointe $P(X, Y, Z)$ de ces 3 variables :

	y_1		y_2	
	x_1	x_2	x_1	x_2
z_1	0,060	0,140	0,060	0,140
z_2	0,192	0,048	0,288	0,072

Montrez que X est indépendante de Y conditionnellement à Z .

X indépendant de Y conditionnellement à $Z \iff P(X|Y, Z) = P(X|Z)$. Or, d'après la table ci-dessus, $P(Y = y_1, Z = z_1) = 0,06 + 0,14 = 0,2$, $P(Y = y_2, Z = z_1) = 0,06 + 0,14 = 0,2$, $P(Y = y_1, Z = z_2) = 0,192 + 0,048 = 0,24$ et $P(Y = y_2, Z = z_2) = 0,288 + 0,072 = 0,36$. Étant donné que $P(X|Y, Z) = P(X, Y, Z)/P(Y, Z)$, $P(X|Y, Z)$ est égal à :

	y_1		y_2	
	x_1	x_2	x_1	x_2
z_1	0,060/0,2	0,140/0,2	0,060/0,2	0,140/0,2
z_2	0,192/0,24	0,048/0,24	0,288/0,36	0,072/0,36

Autrement dit, $P(X|Y, Z)$ est égal à :

	y_1		y_2	
	x_1	x_2	x_1	x_2
z_1	0,3	0,7	0,3	0,7
z_2	0,8	0,2	0,8	0,2

On remarque que le sous-tableau correspondant à $Y = y_1$, autrement dit $P(X|Z, Y = y_1)$, est identique à celui correspondant à $Y = y_2$, i.e., $P(X|Z, Y = y_1) = P(X|Z, Y = y_2)$.

Or $P(X|Z) = P(X|Z, Y = y_1)P(Y = y_1|Z) + P(X|Z, Y = y_2)P(Y = y_2|Z) = P(X|Z, Y = y_1) \times [P(Y = y_1|Z) + P(Y = y_2|Z)] = P(X|Z, Y = y_1)$ car, quelle que soit la valeur z_i de Z , $P(Y = y_1|Z = z_i) + P(Y = y_2|Z = z_i) = 1$. Par conséquent, $P(X|Z) = P(X|Z, Y)$ et on a donc indépendance de X et de Y conditionnellement à Z .

On peut aussi faire la même déduction en observant que la connaissance de la valeur de Y (qui permet de savoir dans quelle colonne du tableau ci-dessus on se trouve) n'apporte aucune information sur X puisque les sous-tableaux à $Y = y_1$ et à $Y = y_2$ sont identiques. Par conséquent, X est bien indépendant de Y conditionnellement à Z .

Exercice 3 (5 pts)

Vous voulez investir à la bourse. Afin d'optimiser vos profits, vous relevez pendant 16 jours le cours du CAC40. Au début de ces deux semaines, celui-ci vaut 5715 points. Dans l'échantillon de 16 jours, le CAC40 vaut en moyenne 5726,025 points, avec un écart-type de 6 points. Vous ne voulez investir que si le CAC40 est à la hausse.

Q 3.1 Sachant que la variable $X = \text{« valeur du CAC40 »}$ suit une loi normale de variance 36, effectuez un test d'hypothèse de niveau de confiance 99% pour savoir si le CAC40 a augmenté. Vous préciserez bien les hypothèses H_0 et H_1 .

Ici, on s'intéresse au fait que le CAC40 puisse augmenter, par conséquent, $H_0 = \text{« } \mu = 5715 \text{ »}$ et $H_1 = \text{« } \mu > 5715 \text{ »}$. Un test d'hypothèse que niveau de confiance 99% signifie que α , la probabilité de rejeter H_0 sachant que H_0 est vraie, est égal à 0,01. Autrement dit :

$$\begin{aligned} 0,01 &= P(\text{rejeter } H_0 | H_0 \text{ est vraie}) \\ &= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > c \mid \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1) \text{ et } \mu = 5715\right) \\ &= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > 2,35 \mid \mu = 5715\right) \\ &= P(\bar{X} > \mu + 2,35\sigma/\sqrt{n} \mid \mu = 5715) \\ &= P(\bar{X} > 5715 + 2,35 \times 6/4) = P(\bar{X} > 5718,525). \end{aligned}$$

La règle du test est donc : si $\bar{x} > 5718,525$, alors rejeter H_0 , sinon rejeter H_1 .

Q 3.2 D'après le test précédent, peut-on conclure que le CAC40 a augmenté ?

Étant donné que la moyenne de notre échantillon est de 5726,025, qui est supérieur à 5718,525, notre test d'hypothèse nous dit de rejeter H_0 et donc de considérer qu'il est effectivement plus probable que la bourse ait augmenté.

Q 3.3 Calculez la puissance du test pour $\mu = 5726,025$. Pour vous aider, vous pourrez supposer que si une variable $Y \sim \mathcal{N}(0, 1)$, alors :

$$P(Y > -1) = 0,8413 \quad P(Y > -2) = 0,9772 \quad P(Y > -3) = 0,9986 \quad P(Y > -4) \approx 1.$$

$$\begin{aligned} 1 - \beta(5726,025) &= P(\text{rejeter } H_0 | H_1 \text{ est vraie}) \\ &= P(\bar{X} > 5718,525 \mid \mu = 5726,025) \\ &= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{5718,525 - \mu}{\sigma/\sqrt{n}} \mid \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0; 1) \text{ et } \mu = 5726,025\right) \\ &= P\left(\frac{\bar{X} - 5726,025}{\sigma/\sqrt{n}} > \frac{5718,525 - 5726,025}{\sigma/\sqrt{n}} \mid \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0; 1)\right) \\ &= P\left(\frac{\bar{X} - 5726,025}{6/4} > \frac{-7,5}{6/4} \mid \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0; 1)\right) \\ &= P\left(\frac{\bar{X} - 5726,025}{6/4} > -5 \mid \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0; 1)\right) \approx 1. \end{aligned}$$

La puissance du test est donc approximativement de 1.

Exercice 4 (5 pts)

Tirer des nombres aléatoires sur ordinateur est, contrairement à ce que l'on pourrait croire, une tâche assez complexe à réaliser. Un nouvel algorithme a été développé à cet effet et nous nous demandons si les nombres qu'il fournit sont bien aléatoires. Il nous faudra une batterie de tests pour pouvoir décider si l'algorithme est bon. Parmi ceux-ci, il en existe un relativement simple : supposons que l'algorithme ne nous fournit que des nombres entre 0 et 9. Si ces nombres sont vraiment tirés aléatoirement, alors en en tirant suffisamment, on devrait se rapprocher d'une loi uniforme : probabilité 1/10ème de choisir n'importe quel nombre entre 0 et 9. On a donc exécuté l'algorithme 200 fois et on a obtenu les résultats suivants :

Valeur	effectif
0	14
1	24
2	18
3	20
4	23
5	13
6	23
7	18
8	24
9	23

Q 4.1 Calculez la moyenne de l'échantillon.

$$\bar{x} = \frac{1}{200} [0 \times 14 + 1 \times 24 + 2 \times 18 + 3 \times 20 + 4 \times 23 + 5 \times 13 + 6 \times 23 + 7 \times 18 + 8 \times 24 + 9 \times 23] = 4,7.$$

Q 4.2 À l'aide d'un test d'ajustement, testez si les nombres 0 à 9 sont distribués uniformément. Utilisez un niveau de test $\alpha = 0,05$ (niveau de confiance du test = 95%).

Bien évidemment notre test d'hypothèse sera : $H_0 =$ « l'algorithme tire uniformément les chiffres » v.s. $H_1 =$ « le tirage n'est pas uniforme ». Si l'on avait eu une loi uniforme, chaque chiffre aurait du apparaître exactement 20 fois. Si l'on appelle X_i la variable « effectif » recensé pour le i ème chiffre, la statistique d'ajustement est donc :

$$D^2 = \sum_{i=0}^9 \frac{(X_i - 20)^2}{20}.$$

On sait que cette statistique suit une loi χ_9^2 , c'est-à-dire une loi du χ^2 à 9 degrés de liberté.

$$\begin{aligned} \alpha = 0,05 &= P(\text{rejeter } H_0 | H_0 \text{ est vraie}) \\ &= P(D^2 > d_\alpha | D^2 \sim \chi_9^2). \end{aligned}$$

D'après la table du χ^2 , $d_\alpha = 16,9$. Autrement dit, si la valeur d^2 de D^2 observée sur l'échantillon est supérieure à 16,9, alors on conclura que l'algorithme ne tire pas les chiffres uniformément, sinon on considèrera le tirage uniforme. Calculons maintenant la valeur de D^2 observée :

$$d^2 = \frac{1}{20}[(14 - 20)^2 + (24 - 20)^2 + (18 - 20)^2 + (20 - 20)^2 + (23 - 20)^2 + (13 - 20)^2 + (23 - 20)^2 + (18 - 20)^2 + (24 - 20)^2 + (23 - 20)^2] = 7,6.$$

Donc on peut effectivement estimer que le tirage est uniforme.