Examen de septembre du module MGDE

C. Gonzales / J.-Y. Jaffray

Durée : 3 heures

Seuls documents autorisés :
Notes et transparents de cours

Exercice 1

Les probabilités élémentaires de la loi jointe de 3 v.a. X,Y et Z, de domaines $\mathcal{X} = \{x_1, x_2\}$, $\mathcal{Y} = \{y_1, y_2, y_3\}$ et $\mathcal{Z} = \{z_1, z_2\}$ sont donnés par les tableaux ci-dessous :

$$| p(x,y,z_1) |= \begin{array}{c} \\ x_1 \\ \\ x_2 \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} \frac{3}{34} & \frac{1}{34} & \frac{3}{34} \\ \\ \frac{3}{34} & \frac{1}{34} & \frac{5}{68} \end{bmatrix} \end{array}$$

$$| p(x,y,z_2) |= \begin{array}{c} \\ x_1 \\ \\ x_2 \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} \frac{3}{17} & \frac{1}{17} & \frac{3}{17} \\ \\ \frac{3}{34} & \frac{1}{34} & \frac{5}{68} \end{bmatrix} \end{array}$$

Q 1.1 Construire les tableaux correspondants des couples (X,Y), (X,Z) et (Y,Z).
A-t-on $X \perp\!\!\!\perp Y$? $X \perp\!\!\!\perp Z$? $Y \perp\!\!\!\perp Z$?

Q 1.2 A-t-on $X \perp\!\!\!\perp Y \mid Z$? A-t-on $X \perp\!\!\!\perp Z \mid Y$? A-t-on $Y \perp\!\!\!\perp Z \mid X$?

Exercice 2

On s'intéresse à l'effet d'un certain traitement sur la rapidité de guérison d'une maladie virale. Les données hospitalières, portant sur 197 patients, sont résumées par le tableau suivant :

$$\begin{array}{cc} & \begin{array}{cc} g & \bar{g} \end{array} \\ \begin{array}{c} t \\ \bar{t} \end{array} & \begin{bmatrix} 48 & 53 \\ 56 & 40 \end{bmatrix} \end{array}$$

$$\text{où } \begin{cases} T = t, \text{ si le malade reçoit le traitement, } T = \bar{t} \text{ sinon} \\ G = g, \text{ si le malade guérit en moins de 8 jours, } G = \bar{g} \text{ sinon.} \end{cases}$$

On estime les probabilités $p(t,g) =_{\text{DEF}} P(T = t, G = g)$, etc. par les fréquences relatives observées données ci-dessus.

Q 2.1 Le traitement est considéré comme efficace dans la population étudiée lorsque
$p(g/t) =_{\text{DEF}} P(G = g/T = t) > p(g/\bar{t}) =_{\text{DEF}} P(G = g/T = \bar{t})$. L'est-il ?

Q 2.2 On dispose en fait de données plus précises, permettant de répartir les patients en deux catégories d'âge : $A = a$, si le malade est un adulte, $A = \bar{a}$ si c'est un enfant. Les tableaux correspondants sont les suivants :

$$\text{Lorsque } A = a : \quad \begin{array}{c|cc} & g & \bar{g} \\ \hline t & 28 & 45 \\ \bar{t} & 8 & 15 \end{array} \qquad \text{Lorsque } A = \bar{a} : \quad \begin{array}{c|cc} & g & \bar{g} \\ \hline t & 20 & 8 \\ \bar{t} & 48 & 25 \end{array}$$

Les probabilités sont encore estimées par les fréquences relatives observées. On note $p(g/t,a) =_{\text{DEF}} P(G = g/T = t, A = a)$, etc. Le traitement apparaît-il comme efficace dans chacune des sous-populations, c.-à-d. a-t-on :

$$p(g/t,a) > p(g/\bar{t},a) \; ? \qquad p(g/t,\bar{a}) > p(g/\bar{t},\bar{a}) \; ?$$

Que constate-t-on ? Par quelle relation $p(g/t)$ est-elle reliée à $p(g/t,a)$ et $p(g/t,\bar{a})$? Par quelle relation $p(g/\bar{t})$ est-elle reliée à $p(g/\bar{t},a)$ et $p(g/\bar{t},\bar{a})$? Commentez.

Q 2.3 Quel est le graphe d'indépendance Γ obtenu pour l'ordre d'énumération ATG des variables ?

Q 2.4 On suppose qu' *intervenir sur T en imposant $T = t$* a pour effet :
i) de supprimer dans Γ le lien entre A et T ;
ii) de mettre pour loi de probabilité décomposable selon le nouveau graphe Γ' la loi $q(.)$ donnée par :

$$\begin{cases} q(a',t,g') = p(a')p(g'/t,a') \\ q(a',\bar{t},g') = 0 \end{cases} \quad \text{pour } a' \in \{a,\bar{a}\} \text{ et } g' \in \{g,\bar{g}\}.$$

Donner l'expression de la *probabilité conditionnelle par intervention* $p(g \parallel t) =_{\text{DEF}} q(g/t)$ en fonction de $p(g/t,a)$ et $p(g/t,\bar{a})$; calculer de même $p(g \parallel \bar{t})$,correspondant à l'intervention imposant $T = \bar{t}$. Pourquoi le paradoxe a-t-il disparu ?

Exercice 3

On considère le réseau bayésien ci-dessous dans lequel les modalités des variables sont indiquées à côté de celles-ci.

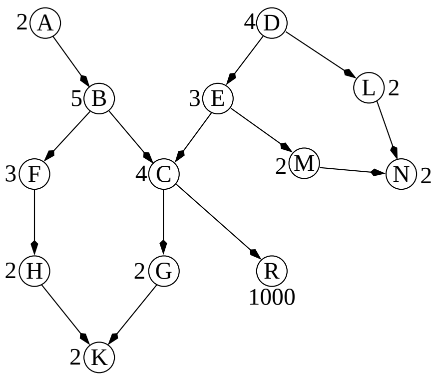

Fig. 1 – Un réseau bayésien.

1) À quelle décomposition de la loi jointe correspond ce graphe ?
2) Dessinez le graphe obtenu après moralisation et triangulation selon l'algorithme de kjærulff (vous indiquerez le graphe et les cliques obtenus après chaque élimination de variable).
3) Dessinez un arbre de jonction correspondant. Vous indiquerez dans quels noeuds vous placerez les différentes probabilités conditionnelles du réseau de la figure 1.
4) En effectuant dans cet arbre de jonction une phase de collecte de racine la clique BCE, quels calculs meneriez-vous ?