Examen du module MGDE

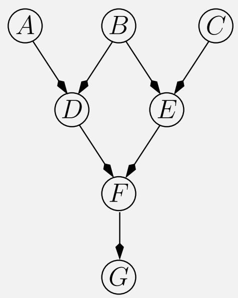
Durée : 3 heures

Exercice 1

La loi jointe des 7 v.a. (A, B, C, D, E, F, G) vérifie les relations d'indépendance suivantes :

$$B \perp\!\!\!\perp A \;; \quad C \perp\!\!\!\perp (A, B) \;; \quad D \perp\!\!\!\perp C \mid (A, B) \;; \quad E \perp\!\!\!\perp (A, D) \mid (B, C) \;;$$
$$F \perp\!\!\!\perp (A, B, C) \mid (D, E) \;; \quad G \perp\!\!\!\perp (A, B, C, D, E) \mid F.$$

a) Construire le graphe d'indépendance correspondant lorsque les variables sont rangées dans l'ordre A, B, C, D, E, F, G.



b) Qu'est-ce que le critère de d-séparation permet d'affirmer concernant les propriétés suivantes :

$$A \perp\!\!\!\perp C \mid (D, E) \;; \quad A \perp\!\!\!\perp C \mid (B, D, E) \;; \quad A \perp\!\!\!\perp C \mid G \;; \quad A \perp\!\!\!\perp C \mid (B, D, E, F).$$

D'après le critère de d-séparation :

$NON[A \mid d \mid C/(D,E)]$, car la chaîne $ADBEC$ est active ; il en résulte, d'après Pearl & Verma, que $NON[A \perp\!\!\!\perp C \mid (D,E)]$ dans presque toutes les lois jointes décomposables selon le graphe.

$A \mid d \mid C/(B,D,E)$ car $ADBEC$ est bloquée en B et $ADFEC$ est bloquée en F ; d'où, $A \perp\!\!\!\perp C \mid (B,D,E)$.

$NON[A \mid d \mid C/G]$, car la chaîne $ADFEC$ est active (G est descendant de F) ; il en résulte que $NON[A \perp\!\!\!\perp C \mid G]$ dans presque toutes les lois jointes décomposables selon le graphe.

$A \mid d \mid C/(B,D,E,F)$ car $ADBEC$ est bloquée en B et $ADFEC$ est bloquée en D et en E ; d'où, $A \perp\!\!\!\perp C \mid (B,D,E,F)$.

c) À l'aide des propriétés générales de la relation d'indépendance conditionnelle, démontrer la validité de l'assertion :

$$D \perp\!\!\!\perp (C,E) \mid (A,B).$$

D'après les propriétés générales de la relation d'indépendance conditionnelle :

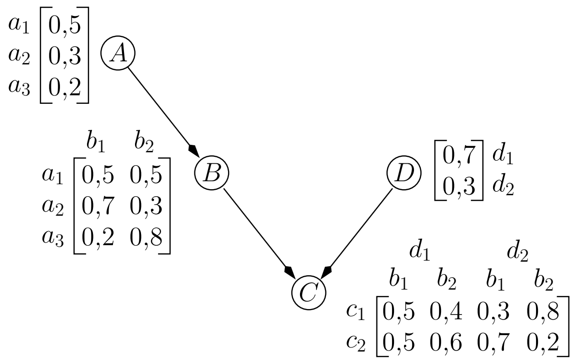
$E \perp\!\!\!\perp (A, D) \mid (B, C) \Longrightarrow_{\text{UNFAIBLE}} E \perp\!\!\!\perp D \mid (A, B, C)$

$E \perp\!\!\!\perp D \mid (A, B, C) \Longrightarrow_{\text{SYM}} D \perp\!\!\!\perp E \mid (A, B, C)$

$$\left.\begin{array}{c} D \perp\!\!\!\perp E \mid (A, B, C) \\ D \perp\!\!\!\perp C \mid (A, B) \end{array}\right\} \Longrightarrow_{\text{CONT}} D \perp\!\!\!\perp (C, E) \mid (A, B).$$

Exercice 2

On considère le réseau bayésien suivant :



1) En utilisant la méthode de Pearl, calculez les probabilités *a priori* de tous les nœuds du réseau.

2) Soit e_A l'information « A ne peut plus prendre la valeur a_1 ». Propagez cette information dans le réseau et calculez les probabilités *a posteriori* de chacun des nœuds.
3) Aurait-on pu prévoir la valeur de $P(D|e_A)$ sans faire de calcul ? Si oui, pourquoi ?

Exercice 3

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

1°) Le traitement est considéré comme efficace dans la population étudiée lorsque :

$$p(g/t) \stackrel{\text{DEF}}{=} P(G = g/T = t) > p(g/\bar{t}) \stackrel{\text{DEF}}{=} P(G = g/T = \bar{t}).$$

L'est-il ?

$$[p(t,g)] = \begin{array}{c|cc} & g & \bar{g} \\ \hline t & 0,244 & 0,269 \\ \bar{t} & 0,284 & 0,203 \end{array} \qquad \begin{array}{rc} & \begin{array}{cc} g & \bar{g} \end{array} \\ [p(g/t)] = & [0,476 \;\; 0,524] \\ [p(g/\bar{t})] = & [0,583 \;\; 0,417] \end{array}$$

Le traitement paraît inefficace dans la population totale puisque

$$p(g/t) = 48/101 = 0,476 < p(g/\bar{t}) = 56/96 = 0,583.$$

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 :

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

Les probabilités sont encore estimées par les fréquences relatives observées. On note $p(g/t, a) \stackrel{\text{DEF}}{=} P(G = g/T = t, A = a)$, 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.

A partir des données par catégories d'âge ($A = a$, si le malade est un adulte, $A = \bar{a}$ si c'est un enfant) on construit les tableaux suivants :

$$A = a :$$

$$[p(t, g/a)] = \begin{array}{c} \\ t \\ \bar{t} \end{array} \begin{array}{c} \begin{array}{cc} g & \bar{g} \end{array} \\ \begin{bmatrix} 0,291 & 0,469 \\ 0,083 & 0,156 \end{bmatrix} \end{array} \qquad \begin{array}{l} \\ [p(g/t, a)] = \\ [p(g/\bar{t}, a)] = \end{array} \begin{array}{c} \begin{array}{cc} g & \bar{g} \end{array} \\ \begin{matrix} [0,383 & 0,617] \\ [0,347 & 0,653] \end{matrix} \end{array}$$

$$A = \bar{a} :$$

$$[p(t,g/\bar a)] = \begin{array}{c c} & \begin{array}{cc} g & \bar g \end{array} \\ \begin{array}{c} t \\ \bar t \end{array} & \begin{bmatrix} 0,198 & 0,079 \\ 0,475 & 0,247 \end{bmatrix} \end{array} \qquad \begin{array}{l} \\ [p(g/t,\bar a)] = \\ [p(g/\bar t,\bar a)] = \end{array} \begin{array}{c} \begin{array}{cc} g & \bar g \end{array} \\ [0,715 \;\; 0,285] \\ [0,658 \;\; 0,342] \end{array}$$

Le traitement apparaît comme efficace dans chacune des deux sous-populations :

$$p(g/t,a) > p(g/\bar t,a) \text{ et } p(g/t,\bar a) > p(g/\bar t,\bar a).$$

C'est un exemple du paradoxe de SIMPSON. Cette apparente contradiction entre les conclusions tirées de l'échantillon et celles tirées des deux sous-échantillons s'explique par le fait que :
$p(g/t) = p(a/t)p(g/t,a) + p(\bar a/t)p(g/t,\bar a)$ alors que :
$p(g/\bar t) = p(a/\bar t)p(g/\bar t,a) + p(\bar a/\bar t)p(g/\bar t,\bar a)$,
donc que l'on fait des moyennes pondérées avec des poids différents, ce qui n'implique pas de conservation des inégalités.

3°a) Quel est le graphe d'indépendance Γ obtenu pour l'ordre d'énumération ATG des variables ?

3°b) 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 :

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

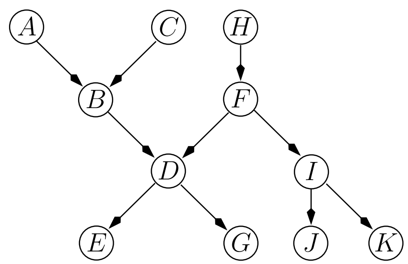
Donner l'expression de la *probabilité conditionnelle par intervention* $p(g \parallel t) \stackrel{\text{DEF}}{=} q(g/t)$ en fonction de $p(g/t, a)$ et $p(g/t, \bar{a})$; calculer de même $p(g \parallel \bar{t})$, correspondant à l'intervention imposant $T = \bar{t}$. Pourquoi le paradoxe a-t-il disparu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

La probabilité conditionnelle par intervention $T = t$ est :
$p(g \parallel t) =_{\text{DEF}} q(g/t) = q(a/t)q(g/t, a) + q(\bar{a}/t)q(g/t, \bar{a})$;
or, puisque, dans Γ', $A \perp\!\!\!\perp T$,
$q(a/t) = q(a) = \sum_{t,g} q(a, t, g) = p(a) \sum_{t,g} p(g/t, a') = p(a)$
et $q(g/t, a) = \dfrac{q(a, t, g)}{q(a, t)} = \dfrac{q(a, t, g)}{q(a)q(t)} = p(g/t, a)$;
de même, $q(\bar{a}/t) = p(\bar{a})$ et $q(g/t, \bar{a}) = p(g/t, \bar{a})$;
d'où : $p(g \parallel t) = p(a)p(g/t, a) + p(\bar{a})p(g/t, \bar{a})$; d'où $p(g \parallel t) = 0,553$.
On montrerait de même que :

$p(g \parallel \bar{t}) = p(a)p(g/\bar{t}, a) + p(\bar{a})p(g/\bar{t}, \bar{a})$; $p(g \parallel \bar{t}) = 0,506$.
Cette fois-ci, les poids sont les mêmes et l'on vérifie bien que de l'efficacité du traitement sur les adultes comme sur les enfants il découle bien l'efficacité du traitement en général ; son intensité apparaît dans la comparaison :
$p(g \parallel t) = 0,553 > p(g \parallel \bar{t}) = 0,506$.

Exercice 4

Soit le réseau bayésien ci-dessous :



1) Moralisez ce réseau.
2) Triangulez le graphe moral en utilisant la séquence d'élimination suivante : $A, K, I, J, G, H, C, D, B, F, E$. Vous indiquerez pour chaque nœud éliminé le graphe non orienté obtenu après élimination.
3) Dessinez un arbre de jonction correspondant à cette séquence d'élimination et indiquez à côté des cliques les probabilités conditionnelles que vous stockerez dans ces cliques.
4) Indiquez les contenus des messages transitant dans les deux sens des arêtes sur chaque séparateur pour le calcul des probabilités *a priori* par l'algorithme de Shafer-Shenoy.
5) Montrer que, quel que soit l'arbre de jonction et quelles que soient deux cliques voisines C_i et C_j d'intersection S_{ij}, les variables de $C_i \backslash S_{ij}$ sont indépendantes de $C_j \backslash S_{ij}$ conditionnellement à S_{ij}.
   S_{ij} sépare l'arbre de jonction en deux sous-arbres T_1 et T_2. Montrez que les variables de $T_1 \backslash S_{ij}$ sont indépendantes de $T_2 \backslash S_{ij}$ conditionnellement à S_{ij}.
6) En utilisant la *d*-séparation, montrez que la propagation d'une information e_A concernant A dans l'arbre de jonction obtenu en 4) ne nécessite le calcul que de deux nouveaux messages.