Introduction aux processus stochastiques

Indépendances conditionnelles

Factorisation de distributions conjointes.

Soient 4 v.a. discrètes $\mathcal{W}$, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$. Leur d.d.p. conjointe est $P(\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z})$.

1. Décrire les nombreuses façons de factoriser cette distribution.
   (cf. Annexes à cette répétition, qui traitent le cas de 2 v.a. $\mathcal{X}$ et $\mathcal{Y}$)
2. Décomposer l'espérance $\mathbb{E}\{f(\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z})\}$ d'une fonction f des 4 v.a. en espérance d'espérances conditionnelles, utilisant la factorisation de la d.d.p. conjointe suivante :

$$P(\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}) = P(\mathcal{W}, \mathcal{Y})P(\mathcal{X}|\mathcal{W}, \mathcal{Y})P(\mathcal{Z}|\mathcal{W}, \mathcal{X}, \mathcal{Y}).$$

   De quelles variables les espérances conditionnelles utilisées sont-elles des fonctions ?

Utilisation d'indépendances conditionnelles dans la factorisation de d.d.p. conjointes.

Soient 4 v.a. discrètes $\mathcal{W}$, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$. Leur d.d.p. conjointe est $P(\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z})$.
La variable $\mathcal{W}$ est indépendante de $\mathcal{Y}$ connaissant $\mathcal{X}$, ce qui se note $\mathcal{W} \perp \mathcal{Y}|\mathcal{X}$, et signifie

$$P(\boxed{\mathcal{W} = W, \mathcal{Y} = Y}|\mathcal{X} = X) = P(\boxed{\mathcal{W} = W}|\mathcal{X} = X)P(\boxed{\mathcal{Y} = Y}|\mathcal{X} = X),$$

ce qu'on note $P(\mathcal{W}, \mathcal{Y}|\mathcal{X}) = P(\mathcal{W}|\mathcal{X})P(\mathcal{Y}|\mathcal{X})$.

1. Montrer que $\mathcal{W} \perp \mathcal{Y}|\mathcal{X}$ signifie aussi $P(\mathcal{Y}|\mathcal{W}, \mathcal{X}) = P(\mathcal{Y}|\mathcal{X})$ et $P(\mathcal{W}|\mathcal{X}, \mathcal{Y}) = P(\mathcal{W}|\mathcal{X})$.
2. Donner 2 exemples de factorisation de $P(\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z})$ qui exploitent $\mathcal{W} \perp \mathcal{Y}|\mathcal{X}$. Décomposer ensuite $\mathbb{E}\{f(\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z})\}$ de la même façon.

Chaînes de Markov et indépendances conditionnelles.

On dit que les v.a. $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$ forment une chaîne de Markov (d'ordre 1) lorsque leur d.d.p. conjointe se factorise selon

$$\begin{aligned} P(\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n) &= P(\mathcal{X}_1)P(\mathcal{X}_2|\mathcal{X}_1)P(\mathcal{X}_3|\mathcal{X}_2)\ldots P(\mathcal{X}_n|\mathcal{X}_{n-1}) \\ &= P(\mathcal{X}_1)\prod_{i=2}^{n} P(\mathcal{X}_i|\mathcal{X}_{i-1}). \end{aligned}$$

1. Montrer que si $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3$ forment une chaîne de Markov,
   alors $\mathcal{X}_3, \mathcal{X}_2, \mathcal{X}_1$ forment aussi une chaîne de Markov.
   Cette propriété justifie d'ailleurs la notation $\mathcal{X}_1 \leftrightarrow \mathcal{X}_2 \leftrightarrow \mathcal{X}_3$ lorsque $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3$ forment une chaîne de Markov.
2. Montrer que si $\mathcal{X}_1, \ldots, \mathcal{X}_n$ forment une chaîne de Markov,
   alors $\mathcal{X}_i, \mathcal{X}_j, \mathcal{X}_k$ $(i < j < k)$ forment une chaîne de Markov.

Réseaux bayesiens et indépendances conditionnelles.

Soient 4 v.a. discrètes $\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$.
Leur d.d.p. conjointe se factorise selon
$P(\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}) = P(\mathcal{W})P(\mathcal{X}|\mathcal{W})P(\mathcal{Y}|\mathcal{W})P(\mathcal{Z}|\mathcal{X}, \mathcal{Y})$.

1. Tracer le *réseau bayesien* correspondant à cette distribution de probabilité.
   C'est un graphe dirigé, dont les noeuds correspondent aux variables aléatoires et portent le nom correspondant. Toute probabilité de la forme $P(\mathcal{X}_0|\mathcal{X}_1, \ldots, \mathcal{X}_n)$ induit n arcs dirigés de $\mathcal{X}_i$ à $\mathcal{X}_0$, $1 \leq i \leq n$.
   Une fois tous les arcs dirigés placés, il n'y a pas de cycle orienté. On place généralement les noeuds sans arc incident en haut ou à gauche du graphe.
2. A-t-on nécessairement
   (a) $\mathcal{X} \perp \mathcal{Y}$?
   (b) $\mathcal{X} \perp \mathcal{Y}|\mathcal{W}$?
   (c) $\mathcal{X} \perp \mathcal{Y}|\mathcal{W}, \mathcal{Z}$?

   On justifiera les réponses en recourant aux notions de chemins bloqués et de d-séparation. (cf explications orales en répétition).

Réseaux bayesiens et indépendances conditionnelles (encore).

Soient 7 v.a. discrètes $\mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$.
Leur d.d.p. conjointe se factorise selon
$P(\mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}) = P(\mathcal{X})P(\mathcal{Z})P(\mathcal{Y}|\mathcal{X}, \mathcal{Z})P(\mathcal{W}|\mathcal{Z})P(\mathcal{U}|\mathcal{Y})P(\mathcal{T}|\mathcal{Y}, \mathcal{W})P(\mathcal{V}|\mathcal{T})$.

1. Tracer le *réseau bayesien* correspondant à cette distribution de probabilité.
2. A-t-on nécessairement
   (a) $\mathcal{W} \perp \mathcal{Y}|\mathcal{Z}$?
   (b) $\mathcal{U} \perp \{\mathcal{W}, \mathcal{Z}\}|\mathcal{Y}$?
   (c) $\mathcal{U} \perp \mathcal{W}|\{\mathcal{T}, \mathcal{Z}\}$?
   (d) $\mathcal{Y} \perp \mathcal{W}|\mathcal{V}$?

Annexes

1. Formules de Bayes.
   Soient 2 variables aléatoires (v.a.) discrètes $\mathcal{X} = \{X_1, X_2, \ldots, X_m\}$ et $\mathcal{Y} = \{Y_1, \ldots, Y_n\}$.
   Leur distribution de probabilité (d.d.p.) conjointe est encodée dans une table dont les éléments sont $P(\mathcal{X} = X_i, \mathcal{Y} = Y_j)$. La d.d.p. conjointe est notée $P(\mathcal{X}, \mathcal{Y})$.
   Il est possible de *factoriser* une probabilité conjointe $P(\mathcal{X} = X_i, \mathcal{Y} = Y_j)$ de 2 façons :

   $$\begin{aligned} P(\mathcal{X} = X_i, \mathcal{Y} = Y_j) &= P(\mathcal{X} = X_i)P(\mathcal{Y} = Y_j|\mathcal{X} = X_i) \qquad \forall i, j, \\ P(\mathcal{X} = X_i, \mathcal{Y} = Y_j) &= P(\mathcal{Y} = Y_j)P(\mathcal{X} = X_i|\mathcal{Y} = Y_j) \qquad \forall i, j, \end{aligned}$$

   de sorte que la d.d.p. conjointe se factorise en

   $$P(\mathcal{X}, \mathcal{Y}) = P(\mathcal{X})P(\mathcal{Y}|\mathcal{X}) \qquad P(\mathcal{X}, \mathcal{Y}) = P(\mathcal{Y})P(\mathcal{X}|\mathcal{Y}).$$

   De là, les formules de Bayes pour exprimer une probabilité conditionnelle :

   $$\begin{aligned} P(\mathcal{Y} = Y_j|\mathcal{X} = X_i) &= \frac{P(\mathcal{X} = X_i, \mathcal{Y} = Y_j)}{P(\mathcal{X} = X_i)} \quad \text{(peut être écrit si } P(\mathcal{X} = X_i) > 0) \\ P(\mathcal{X} = X_i|\mathcal{Y} = Y_j) &= \frac{P(\mathcal{X} = X_i, \mathcal{Y} = Y_j)}{P(\mathcal{Y} = Y_j)} \quad \text{(peut être écrit si } P(\mathcal{Y} = Y_j) > 0) \end{aligned}$$

   Souvent, les indices i, j sont omis. Ainsi, on écrit

   $$P(\mathcal{Y} = Y|\mathcal{X} = X) = \frac{P(\mathcal{X} = X, \mathcal{Y} = Y)}{P(\mathcal{X} = X)}$$

   en supposant que X renvoie à une des valeurs parmi $\{X_1, \ldots, X_m\}$, et Y à une des valeurs parmi $\{Y_1, \ldots, Y_n\}$.
   Comme on peut décomposer $P(\mathcal{X} = X)$ en

   $$P(\mathcal{X} = X) = \sum_{j=1}^{n} P(\mathcal{X} = X, \mathcal{Y} = Y_j) = \sum_{Y} P(\mathcal{X} = X, \mathcal{Y} = Y),$$

   à partir des formules de Bayes on peut calculer les probabilités conditionnelles en utilisant uniquement les probabilités conjointes :

   $$\begin{aligned} P(\mathcal{Y} = Y|\mathcal{X} = X) &= \frac{P(\mathcal{X} = X, \mathcal{Y} = Y)}{\sum_Y P(\mathcal{X} = X, \mathcal{Y} = Y)} \\ P(\mathcal{X} = X|\mathcal{Y} = Y) &= \frac{P(\mathcal{X} = X, \mathcal{Y} = Y)}{\sum_X P(\mathcal{X} = X, \mathcal{Y} = Y))} \end{aligned}$$

2. Distribution de probabilité conditionnelle.
   La d.d.p. de $\mathcal{Y}$ conditionnellement à $\mathcal{X} = X$ renvoie à la table des n probabilités

   $$P(\mathcal{Y} = Y_j|\mathcal{X} = X),$$

   et est notée $P(\mathcal{Y}|X)$. La somme des éléments de la table vaut 1.
   La d.d.p. de $\mathcal{Y}$ conditionnellement à $\mathcal{X}$ renvoie à m tables de n probabilités. Chaque table est relative à une des valeurs possibles pour X. Le tout est désigné par $P(\mathcal{Y}|\mathcal{X})$.