Série 12 - inférence bayésienne

<u>Thèmes:</u> probabilité conditionnelle, théorème de Bayes, distribution a priori et a posteriori, décision bayésienne, distribution gaussienne multidimensionnelle

Exercice 1

Des bonbons sont emballés dans des boîtes différentes au regard du spécialiste, mais dont la marque n'est pas reconnaissable par le profane. Initialement, il y a 40 boîtes en tout, 20 boîtes de la marque A et 20 boîtes de la marque B. Parmi les 20 boîtes de la marque A, 14 boîtes contiennent des bonbons. Parmi les 20 boîtes de la marque B, 8 boîtes contiennent des bonbons. Les 40 boîtes sont fermées et mélangées de telle sorte qu'il soit impossible de reconnaître la marque des boîtes pour le profane. Ce dernier tire au hasard une boîte parmi les 40 et, par chance, il y a des bonbons dans la boîte. Quelle est la probabilité que cela soit une boîte de la marque A ?

Exercice 2

Calcul de probabilités conditionnelles. Soit la densité de probabilité bidimensionnelle qui est constante sur un triangle et nulle sinon d'expression

$$f(x,y) = \begin{cases} 2 & x + y \leq 1, x \geq 0, y > 0 \\ 0 & \text{sinon} \end{cases}$$

On demande d'évaluer

- $P(X < 0.5)$
- $P(X < Y)$
- $P(x < 0.5 \,|Y < 0.5)$

Exercice 3

Décision bayésienne. Soit les observations de réalisations de n variables aléatoires X_1, ..., X_n avec $x = (X_1, \ldots, X_n)$ sous deux hypothèses

$$
\begin{aligned}
H_0 : f(x|H_0) &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\frac{x_i^2}{\sigma^2}\right\} \\
H_1 : f(x|H_1) &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^2 \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}\right\}
\end{aligned}
$$

Déterminer le seuil de décision afin de minimiser le coût attendu avec les données

- Probabilité des hypothèses: $\pi(H_0) = 0.75$, $\pi(H_1) = 1 - \pi(H_0) = 0.25$
- Coûts : $C_{00} = C_{11} = 0$, $C_{01} = 1$ et $C_{10} = 2$

Exercice 4

Distribution gaussienne multidimensionnelle. Soit une variable aléatoire vectorielle de dimension 2 qui obéit à une distribution gaussienne multidimenssionnelle. On aimerait connaître la covariance a posteriori connaissant celle a priori en ayant observé le vecteur x. Voici les données du problème :

$$
A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \qquad C_0^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \qquad C^{-1} = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}
$$

avec C_0 la matrice de covariance a priori, C celle de la vraisemblance de l'échantillon, et A la matrice qui donne l'espérance de x en fonction de $\theta = (\theta_1\, \theta_2)^T$ inconnu, i.e.

$$
\mathbb{E}(x) = A\,\theta
$$

On demande :

- de calculer la covariance a posteriori donnée par la formule

$$
C = (C_0^{-1} + A^T C^{-1} A)^{-1}
$$

- de tracer la courbe donnant une distance de Mahalanobis constante de $r = 1$ pour la distribution a priori et pour celle a posteriori. Il n'est pas nécessaire de calculer le centre de la courbe. Vous supposerez donc que les centres sont situés dans un "repère relatif".