UM - M1 MIND/Biostat 4 avril 2018

Examen de rattrapage - Théorie de l'information et de la décision
Partie 2, bayésienne (10 pts)

On rappelle que si la variable aléatoire θ suit une loi Beta de paramètre (a, b), elle admet comme densité

$$\pi(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}\mathbb{I}_{\theta\in[0,1]}$$

avec $a > 0$ et $b > 0$.

Exercice 1 (2 pts)

L'hôpital d'une petite ville du nord-ouest de l'Argentine compte parmi ses malades 4% qui sont d'origine basque, 58% d'origine espagnole, 32% d'origine indienne et 6% d'origine italienne. Sachant que 3% des indiens ont un sang de rhésus négatif, ainsi que 87% des basques et 22% des populations d'origine italienne et espagnole, quelle est la probabilité pour qu'une éprouvette de sang de rhésus négatif provienne d'un malade d'origine basque ?

Exercice 2 (4 pts)

Les aéroports se doivent de respecter certaines normes concernant les bruits émis par des avions au décollage et à l'atterrissage. Ainsi pour les zones habitées proches d'un aéroport, la limite tolérée se situe à environ 80 décibels. Les habitants d'un des villages proches d'un aéroport assurent que le bruit atteint la valeur limite de 84 décibels en moyenne pour un certain type d'avions. L'aéroport affirme qu'il n'est que de 70 décibels. Des experts sont convoqués pour trancher entre les deux parties en présence. Ils admettent que l'intensité du bruit causé par un avion de ce type suit une loi Gaussienne de moyenne θ et de variance 49. Ils enregistrent l'intensité du bruit provoqué par le passage de ces avions sur un échantillon de taille $n = 100$. Ils trouvent que $\bar{x}_{100} = \frac{1}{100}\sum_{i=1}^{100} x_i = 77$.

L'espace paramétrique est réduit à deux valeurs $\Theta \in \{70, 84\}$. Pour une loi a priori uniforme sur Θ et la fonction de perte ci-dessous (fonction de perte qui reflette le fait que le décideur souhaite avoir quelques certitudes quant à la mise en place couteuse d'un dispositif anti-bruit)

$$L(\theta, d) = \begin{cases} 0 & \text{si} \quad \theta = d \\ 1 & \text{si} \quad \theta = 84 \quad \text{et} \quad d = 70 \\ 2 & \text{si} \quad \theta = 70 \quad \text{et} \quad d = 84 \end{cases}$$

donner la réponse bayésienne.

Exercice 3 (4 pts)

Un laboratoire pharmaceutique a mis au point un nouveau médicament contre la migraine. Pour évaluer l'efficacité de ce traitement, une étude a été réalisée sur 1000 patients : 500 ont reçu le nouveau médicament et 500 le médicament *gold standard* de référence. Après deux mois de tests, les patients ont indiqué s'ils ont ressenti ou pas une amélioration de leur état migraineux, les résultats sont stockés dans la table ci-dessous.

	Amélioration	Pas amélioration
Nouveau médicament	295	205
Médicament de référence	305	195

Le laboratoire veut tester l'hypothèse selon laquelle le nouveau médicament a une efficacité similaire au médicament de référence. Si c'était le cas, il pourrait progressivement se substituer à lui dans la mesure où il génère beaucoup moins d'effets secondaires.

Nous nous plaçons dans le paradigme bayésien. Nous appelons θ_1, respectivement θ_2, la probabilité que le nouveau médicament, respectivement le médicament de référence, provoque une amélioration. Nous supposons des lois a priori uniformes pour les paramètres θ_1 et θ_2.

1 (2 pts) Donner les lois a posteriori de θ_1 et θ_2.

2 (2 pts) Donner des intervalles de crédibilité (à risques symétriques) de niveau 95% pour les paramètres θ_1 et θ_2 et indiquer comment nous pouvons utiliser ces deux intervalles pour répondre au problème posé.

1

Examen de rattrapage
HMMA 101
04/04/2018
Correction

Exercice 1

R = rhésus du patient

O = origine du patient

$$\mathbb{P}(O = \text{"Basque"} \mid R = \text{"Négatif"})$$

$$= \frac{\mathbb{P}(O = \text{"Basque"} \wedge R = \text{"Négatif"})}{\mathbb{P}(R = \text{"Négatif"})}$$

$$= \frac{0.87 \times 0.04}{0.87 \times 0.04 + 0.03 \times 0.32 + 0.22(0.58 + 0.06)}$$

$$\# \; 0.19$$

2

Exercice 2

$$\mathbb{P}(\theta = 70 | \underline{x}) \propto \mathbb{P}(\theta = 70) f(\underline{x} | \theta = 70)$$

$$\Leftrightarrow \mathbb{P}(\theta = 70 | \underline{x}) \propto \left(\frac{1}{2}\right) \prod_{i=1}^{100} e^{-\frac{1}{2\times 49}(x_i - 70)^2}$$

$$\Leftrightarrow \mathbb{P}(\theta = 70 | \underline{x}) \propto e^{-\frac{50}{49}(\bar{x} - 70)^2}$$

De même, $\mathbb{P}(\theta = 84 | \underline{x}) \propto e^{-\frac{50}{49}(\bar{x} - 84)^2}$

Comme $\bar{x} = 77$,

$$\mathbb{P}(\theta = 70 | \underline{x}) = \mathbb{P}(\theta = 84 | \underline{x}) = \frac{1}{2}$$

$$\mathbb{E}^{\pi}[L(\theta, d) | \underline{x}] =$$

$$\frac{1}{2} \mathbb{1}\{d = 70\} + \mathbb{1}\{d = 84\}$$

$$\Rightarrow \hat{d} = 70$$

Exercice 3

1) x = # individus parmi les 500 ayant reçu le nouveau médicament ressentant une amélioration

③

$x|\theta_1 \sim \mathcal{B}(500, \theta_1)$

y = # individus parmi les 500 ayant reçu le medicament de référence ressentant une amélioration

$y|\theta_2 \sim \mathcal{B}(500, \theta_2)$

$$\pi(\theta_1|x) \propto \pi(\theta_1) f(x|\theta_1)$$

$$\pi(\theta_1|x) \propto \theta_1^{x}(1-\theta_1)^{n-x} \mathbb{1}_{[0,1]}(\theta_1)$$

$$\Rightarrow \theta_1|x \sim Beta(296, 206)$$

De même $\theta_2|y \sim Beta(306, 196)$

2] $$I^{95\%}(\theta_1) = \left[F^{-1}_{Beta(296,206)}(0.025), F^{-1}_{Beta(296,206)}(0.975)\right]$$

$$\# \left[0.55, 0.63\right]$$

$$I^{95\%}(\theta_2) = \left[F^{-1}_{Beta(306,196)}(0.025), F^{-1}_{Beta(306,196)}(0.975)\right]$$

$$\# \left[0.57, 0.65\right]$$

Si $I^{95\%}(\theta_1) \cap I^{95\%}(\theta_2) \neq \emptyset$

alors le nouveau médicament

est considéré comme similaire au

précédent.