Examen de 1ère session de MAPSI

Durée : 2 heures

Seuls documents autorisés : Calculatrices et votre antisèche recto-verso
– Barème indicatif –

Exercice 1 – Attente au distributeur de billets (4pt)

En supposant que le traitement d'une demande d'un client au distributeur est réalisé en temps constant (égal à 1 unité de temps) et que la distribution des arrivées des clients au distributeur suit une loi de Poisson de paramètre μ (unités de temps), la distribution de probabilité des durées X pendant lesquelles le distributeur est occupé suit une distribution de Borel de paramètre μ définie par :

$$P(X = n \text{ unités de temps}) = \frac{e^{-\mu n}(\mu n)^{n-1}}{n!}$$

où le paramètre μ appartient à l'intervalle $[0, 1]$. Sur une période de quelques jours, on a observé les durées d'occupation suivantes :

X	2	3	1	5	2	7

Q 1.1 (2pt) Déterminez par maximum de vraisemblance la valeur du paramètre μ.

Q 1.2 (2pt) Un expert de la banque affirme qu'*a priori* la distribution des paramètres μ dans la région où est situé le distributeur suit une loi Beta de paramètres $\alpha = 7$ et $\beta = 21$. On rappelle que la loi Beta est définie de la manière suivante :

$$\text{Beta}(x; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}.$$

En utilisant cet apriori, calculez par *maximum a posteriori* la valeur du paramètre μ.

Exercice 2 – Boursicotage (3pt)

Votre banquier vous propose un placement. Avant de l'accepter, vous avez observé le rendement de ce dernier sur chaque semaine pendant environ un an (exactement 50 semaines). Pour 1000 € investis, vous avez relevé les gains suivants :

gain	5 €	10 €	20 €	50 €
nombre de semaines	10	15	20	5

Autrement dit, pendant 10 semaines, un placement de 1000 € aurait rapporté chaque semaine 5 €, pendant 15 semaines il aurait rapporté chaque semaine 10 €, etc.

Q 2.1 (1pt) Calculez la moyenne et la variance de cet échantillon.

Q 2.2 (2pt) Le banquier vous annonce que des études menées par la banque montrent que le gain moyen de ce placement est de 20 € par semaine, avec un écart-type de 10 €. Selon un test d'hypothèse de niveau de confiance 95%, pouvez-vous confirmer que ce gain par semaine n'a pas baissé ?

Exercice 3 – Indépendance (3pt)

Un site de vente sur internet a recensé le volume de ventes d'un produit en fonction de l'âge de l'acheteur. Soit X la variable « âge de l'acheteur », dont le domaine est {jeune,vieux}, et Y la variable

« volume de ventes », dont le domaine est {petit,moyen,élevé}. Sur les 100 dernières ventes, le site a enregistré la répartition suivante :

$X \backslash Y$	petit	moyen	élevé
jeune	7	22	11
vieux	13	28	19

À un niveau de confiance de 95%, peut-on considérer que les variables X et Y sont indépendantes ? Vous justifierez votre réponse.

Exercice 4 – Mixtures et processeurs (6pt)

Sur un ordinateur muni d'un processeur à deux cœurs, on a noté les temps (en millisecondes) d'inaction des cœurs. On sait que, pour chaque cœur, la distribution de probabilité de ces temps est une distribution exponentielle (on rappelle que la distribution exponentielle de paramètre λ a pour fonction de densité $f(x;\lambda) = \lambda e^{-\lambda x}$). La distribution des temps d'inaction (variable X) sur le processeur est donc une mixture de distributions exponentielles :

$$p(x) = \pi_1 \lambda_1 e^{-\lambda_1 x} + (1 - \pi_1)\lambda_2 e^{-\lambda_2 x}.$$

Un logiciel a relevé les temps d'inaction suivants, sans indiquer à quels cœurs ils correspondent :

3	5	2

Q 4.1 (1pt) Soit $\Theta = \{\pi_1; \lambda_1; \lambda_2\}$. On introduit la variable de classe Y spécifiant à quelle loi exponentielle on appartient ($Y = 1$ si le temps correspond au 1er cœur, $Y = 2$ sinon). Cette variable est non observée. Soit x_i une valeur observée dans le tableau ci-dessus. Donnez l'expression mathématique de $p(x_i|Y = y, \Theta)$.

Q 4.2 (1pt) Indiquez l'expression de la fonction $Q_i(y)$ de l'algorithme EM en fonction de la $i$ème observation x_i et des paramètres de Θ.

Q 4.3 (1pt) Donnez l'expression mathématique de la fonction que l'on maximise dans l'étape M de l'algorithme EM en fonction des paramètres $\Theta = \{\pi_1; \lambda_1; \lambda_2\}$, des fonctions Q_i et des observations x_i.

Q 4.4 (3pt) En partant de $\Theta^0 = \{\pi_1 = 0,4; \lambda_1 = 2; \lambda_2 = 3\}$, appliquez une étape de l'algorithme EM.

Exercice 5 – Indépendance en probabilité (4pt)

Soit 5 variables aléatoires A, B, C, D, E dont la distribution de probabilité jointe se décompose selon l'expression suivante :

$$P(A, B, C, D, E) = P(A) \times P(B|A) \times P(C) \times P(D|B, C) \times P(E|D).$$

Q 5.1 (1pt) Exprimez la distribution $P(B, C)$ en fonction de $P(A, B, C, D, E)$.

Q 5.2 (0,5pt) Rappelez sous quelle condition sur $P(B, C)$ les deux variables B et C sont indépendantes.

Q 5.3 (2,5pt) Déterminez si B et C sont des variables aléatoires indépendantes. Vous justifierez mathématiquement votre réponse.