Statistiques et Probabilités, L2 MASS, Partiel
28/04/2014, Durée : deux heures
Le barème est indicatif. Documents, téléphones portables et calculatrices ne sont pas autorisés.

Exercice 1. Loi normale *(7,5 points)*

Soit $X_1, \dots, X_n$ une suite des variables aléatoires i.i.d. de loi normale de moyenne θ et de variance $\theta(1-\theta)$, où θ est un paramètre inconnu appartenant à l'intervalle $]0,1[$.

On rappelle que la densité de la loi $\mathcal{N}(\mu, \sigma^2)$ est $\frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)$. Si X suit la loi $\mathcal{N}(\mu, \sigma^2)$, alors $\mathbb{E}(X^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$.

On rappelle aussi que pour une variable aléatoire X suivant la loi gaussienne de moyenne 0 et de variance 1 on a

$$\mathbb{P}(X \leq -2.32) = 0.01; \mathbb{P}(X \leq -1.96) = 0.025; \mathbb{P}(X \leq -1.64) = 0.05; \mathbb{P}(X \leq -1.28) = 0.10.$$

1. Écrire la vraisemblance d'une réalisation $(x_1, \dots, x_n)$ de $(X_1, \dots, X_n)$ en fonction de θ et montrer que la détermination du maximum de vraisemblance se ramène à la résolution d'une équation du 3ème degré (que l'on ne cherchera pas à résoudre).

2. On considère les statistiques suivantes

$$T_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad U_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

Montrer que T_n et U_n sont des estimateurs sans biais de θ et ils convergent vers θ en probabilité lorsque n tend vers l'infini. (Indication : vous pouvez utiliser l'inégalité de B.T. ou la loi des grands nombres pour montrer la convergence en probabilité.)

3. Calculer les risques quadratiques moyens de T_n et U_n et en déduire lequel des deux estimateurs est meilleur.

4. Application : On observe une réalisation $(x_1, \dots, x_{100})$ de 100-échantillon et on trouve que $\frac{1}{100}\sum_{i=1}^{100} x_i^2 = 0.5$. En faisant une approximation gaussienne de la loi de U_n, donner un intervalle de confiance de niveau 90% pour θ.

5. On veut tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$ à l'aide de l'échantillon $X_1, \dots, X_n$. Donner la région de rejet du test pour le niveau du test $\alpha = 5\%$ en utilisant, au choix, l'une quelconque des deux statistiques précédentes.

Exercice 2. Chaîne de Markov *(5 points)*

On considère la suite $X_0, \dots, X_n, \dots$ telle que $X_0 = 2$ et qui vérifie l'équation

$$X_{n+1} = int\left(\frac{X_n}{2} + e_n\right)$$

où $int(\cdot)$ est la fonction partie entière ($int(x)$ est l'entier inférieur à x le plus proche de x) et $(e_n)_{n\in\mathbb{N}}$ est une suite des variables aléatoires i.i.d. issues de loi de Bernoulli de paramètre 1/3.

1. Montrer que $X_0, X_1, X_2, \dots$ est une chaîne de Markov.

2. Quels sont les états possibles pour $(X_n)_{n\in\mathbb{N}}$? Lesquels sont récurrents ? lesquels ne le sont pas ?

3. Calculer la matrice de transition de la chaîne de Markov. Justifier votre réponse.

Tournez la page s.v.p. –>

Exercice 3. Couple d'entiers *(7,5 points)*

Soit (X,Y) un couple de variables aléatoires à valeurs dans $\{0,1\} \times \mathbb{N}^*$, dont la loi jointe est donnée par

$$\mathbb{P}(X=0, Y=k) = \frac{3^k - 1}{6^k}, \quad \mathbb{P}(X=1, Y=k) = \frac{1}{6^k}, \quad \text{pour tous} \quad k \in \mathbb{N}^*.$$

(Indication : $\sum_{i=n}^{\infty} x^i = x^n/(1-x)$ pour tous $n \in \mathbb{N}$, si $|x| < 1$. $\mathbb{N}^* = \mathbb{N}\backslash\{0\}$.)

1. Donner les lois marginales de X et Y. Les variables X et Y sont elles indépendantes ?
2. Calculer l'espérance et la variance de X.
3. Calculer $\mathbb{P}(X = Y)$.
4. Quelle est la loi de la variable $Z = X + Y$?
5. Quelle est la loi de la variable $W = Y - X$?

Tournez la page s.v.p. –>