Contrôle continu 2
Sujet A
Durée : 2 heures

Seules la calculatrice, une page de notes manuscrites et les tables des lois sont autorisées. Les exercices sont indépendants et peuvent être traités dans l'ordre de votre choix.

Exercice 1. La taille d'un nouveau-né suit une variable aléatoire normale de moyenne inconnue θ et d'écart-type égal à 5 cm. On suppose que les tailles des nouveau-nés sont indépendantes. Au cours de ce mois, il y a eu 64 naissances dans un hôpital. Un infirmier a mesuré ces nouveau-nés : la moyenne de leur taille est de 49.56 cm.

(1) Donner un estimateur de θ. En déduire une estimation du paramètre θ.

(2) Soit α dans $]0,1[$. Construire un intervalle de fluctuation pour θ au niveau de confiance $1-\alpha$.

(3) En déduire un intervalle de confiance au niveau 95% pour la taille moyenne d'un nouveau-né dans la population desservie par cet hôpital.

(4) Quel est le degré de confiance d'un intervalle de fluctuation de longueur 1 cm centré en la moyenne empirique ?

(5) Quelle est la taille minimale de l'échantillon pour que l'intervalle de longueur 1 cm centré en la moyenne empirique soit un intervalle de fluctuation de niveau de confiance 98% ?

Exercice 2. Soient $\gamma \in]0,1[$ et $(X_1, \dots, X_n)$ des variables aléatoires indépendantes identiquement distribuées à valeur dans $\{-1,0,1\}$ telles que $\mathbb{P}(X_1 = -1) = \gamma(1-\gamma)$ et $\mathbb{P}(X_1 = 0) = \gamma^2$.

(1) Calculer $\mathbb{E}[X_1]$ et $\mathbb{E}[X_1^2]$ en fonction de γ.

(2) Inverser sur $]0,1[$ les fonctions $g_1 : x \mapsto (1-x)^2$ et $g_2 : x \mapsto 1-x^2$.

(3) En déduire deux estimateurs de γ par la méthode des moments.

On note $e = (x_1, \dots, x_n)$ un échantillon d'observations associé à la modélisation $(X_1, \dots, X_n)$. Pour k dans $\{-1,0,1\}$, on note n_k le nombre d'observations qui sont égales à k. Plus formellement, on a $n_k = \sum_{i=1}^n \mathbf{1}_{x_i=k}$.

(4) Exprimer la vraisemblance de l'échantillon $(x_1, \dots, x_n)$ en fonction de γ et de n_{-1}, n_0 et n_1.

(5) En déduire l'estimateur par maximum de vraisemblance de γ.

Exercice 3. On note p la proportion d'individus touchés par la grippe lors de l'hiver 2015 parmis la population française (qui est donc de grande taille). En interrogeant 100 personnes choisies au hasard dans cette population, un enquêteur relève qu'il y en a 30 qui ont été touchées par la grippe. Pour tout i entre 1 et $n = 100$, on note $X_i = 1$ si la i-ème personne interrogée a été touchée par la grippe et $X_i = 0$ sinon. On suppose que les variables aléatoires X_i sont indépendantes.

(1) Quelle est la loi de X_1 ? Le justifier.

(2) Donner un estimateur de p. En déduire une estimation du paramètre p.

Dans la suite, on supposera que la taille de l'échantillon est assez grande pour justifier une approximation gaussienne.

(3) En le justifiant, dites par quelle variable aléatoire gaussienne on peut approximer la variable aléatoire

$$\sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - p}{\sqrt{p(1-p)}} \quad ?$$

Notons F la fonction de répartition de la loi $\mathcal{N}(0,1)$ et u_α tel que $F(u_\alpha) = 1 - \alpha/2$.

(4) En déduire une valeur approchée de

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - u_\alpha\sqrt{\frac{p(1-p)}{n}} \leq p \leq \frac{1}{n}\sum_{i=1}^{n} X_i + u_\alpha\sqrt{\frac{p(1-p)}{n}}\right).$$

(5) La réponse à la question précédente permet-elle de donner un intervalle de fluctuation pour p ? Si oui, le donner. Si non, expliquer quel est le problème et comment le contourner (sans faire de calculs).

Exercice 4. Soit $(X_1, \ldots, X_n)$ un échantillon de loi mère $\mathcal{N}(\mu, \sigma)$ où μ et σ sont inconnus. Notons

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad V_n^* = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 \quad \text{et} \quad V_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

(1) Quelles sont les lois respectives de $\frac{nV_n^*}{\sigma^2}$ et $\frac{(n-1)V_n}{\sigma^2}$? Le justifier pour $\frac{nV_n^*}{\sigma^2}$.

(2) Soit α dans $]0,1[$. Construire un intervalle de fluctuation pour l'écart-type σ au niveau de confiance $1-\alpha$.

(3) Sur un échantillon de taille $n = 4$, on observe : une fois la valeur 0.6, deux fois la valeur -1 et une fois la valeur 1. En déduire un intervalle de confiance pour σ de niveau 90%.