FINAL - SQ20

La présentation, la lisibilité et la qualité de la rédaction entreront pour une part importante dans l'appréciation des copies. **L'utilisation de la calculatrice est conseillée.**

Exercice 1 : intervalle de confiance (5 points)

Pour estimer la précision d'un thermomètre, on réalise $n = 100$ mesures indépendantes de la température d'un liquide maintenu à une température constante de 20 degrés Celsius.
Chaque mesure est modélisée par une variable aléatoire X_i de loi normale $\mathcal{N}(\mu\,,\,\sigma)$ de **moyenne connue** $\mu = 20$. Les observations $x_1,\ x_2,\ \ldots, x_n$ ont conduit à la valeur $\displaystyle\sum_{i=1}^{n} x_i^2 = 40\,011$.

On se propose de construire un intervalle de confiance observé pour l'écart-type σ (qui représente la précision du thermomètre) au niveau de confiance de 99%.

1. Quel estimateur usuel choisit-on pour la variance σ^2 ?
2. Préciser la variable aléatoire de confiance Z utilisée pour obtenir un intervalle de confiance de la variance σ^2. Indiquer la loi de probabilité de Z.
3. Déterminer les réels a et b tels que $P(a \leqslant Z \leqslant b) = 0,99$ en répartissant le risque 1% de façon symétrique.
4. En déduire un intervalle de confiance aléatoire à 99% de la **variance** σ^2.
5. Calculer un intervalle de confiance observé pour l'**écart-type** σ au niveau 99%.

Exercice 2 : les 2 questions sont indépendantes (5 points)

1. Soit $n \in \mathbb{N}$. n individus ($n \geqslant 3$) jettent simultanément chacun une pièce de monnaie équilibrée. Une personne gagne si elle obtient le contraire de toutes les autres.
   (a) Exprimer en fonction de n, la probabilité p_n qu'une partie comporte un gagnant.
   (b) On effectue une succession de parties indépendantes les unes des autres jusqu'à l'obtention d'un gagnant. On note G_n la variable aléatoire égale au nombre de parties nécessaires à l'obtention d'un gagnant. Reconnaître la loi de G_n, donner son espérance.
2. Soit $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}^{+*}$. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles telles que
$$X_n \hookrightarrow \mathcal{N}\left(\mu\,,\,\frac{\sigma}{\sqrt{n}}\right)$$
   (a) On désigne par F_n la fonction de répartition de X_n. Pour tout réel x, exprimer $F_n(x)$ à l'aide de Φ, fonction de répartition d'une variable aléatoire de loi normale centrée réduite.
   (b) Déterminer $\displaystyle\lim_{n\to+\infty} F_n(x)$ en distinguant les cas $x > \mu$, $x < \mu$ et $x = \mu$.
   (c) Étudier la convergence en loi de la suite $(X_n)_{n\in\mathbb{N}^*}$.

Exercice 3 : loi de Pareto (10 points)

Soit a un réel strictement supérieur à 1. On considère la fonction f définie sur $\mathbb{R}$ par :

$$f(t) = \begin{cases} \dfrac{a}{t^{a+1}} & \text{si } t \geqslant 1 \\ \\ 0 & \text{sinon} \end{cases}$$

1. Soit B un réel supérieur à 1. On pose $I(B) = \displaystyle\int_1^B \frac{a}{t^{a+1}}\,\mathrm{d}t$.
   (a) Calculer $I(B)$ et donner $\displaystyle\lim_{B\to+\infty} I(B)$.
   (b) En déduire que f peut être considérée comme une densité de probabilité.

   Dans la suite de l'exercice, on désigne par X une variable aléatoire à densité admettant f pour densité.

2. (a) Calculer pour tout réel $B \geqslant 1$, l'intégrale $\displaystyle\int_1^B \frac{a}{t^a}\,\mathrm{d}t$.
   (b) En déduire que la variable aléatoire X admet une espérance $E(X)$ dont on donnera la valeur en fonction de a.
3. On note F la fonction de répartition de X.
   Calculer $F(x)$ pour tout réel x en distinguant les cas $x < 1$ et $x \geqslant 1$.
4. On pose $Y = \ln(X)$.
   On admet que Y est une variable aléatoire à densité et on note G sa fonction de répartition.
   (a) Montrer que pour tout réel x, $G(x) = F(\mathrm{e}^x)$
   (b) En déduire que pour tout réel positif x, $G(x) = 1 - \mathrm{e}^{-a\,x}$
   Donner de même $G(x)$ pour tout réel $x < 0$.
   (c) Reconnaître la loi de probabilité de Y. En déduire son espérance et sa variance sans calculs.
5. La paramètre a de X est inconnu. On cherche à l'estimer.
   Pour tout entier $n \geqslant 2$, on se donne n variables aléatoires $X_1, X_2, \dots, X_n$ mutuellement indépendantes suivant toutes la même loi que X. On pose enfin $T_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \ln(X_i)$.

   Montrer que T_n est un estimateur sans biais de $\dfrac{1}{a}$

   En déduire que T_n est un estimateur convergent de $\dfrac{1}{a}$.
6. On suppose que $n = 100$ et que les valeurs de l'échantillon observé ont été saisies dans une variable Scilab : `X=[X(1),X(2),...,X(100)]` matrice ligne à 100 colonnes.
   La commande `mean(log(X))` fournit le résultat 0,33.
   Proposer une estimation ponctuelle de a.