Contrôle

La calculatrice et les documents sont autorisés.
Toutes vos réponses doivent être soigneusement rédigées.
Le barème est donné à titre indicatif

Exercice 1 (3 points)

Un joueur lance un dé. S'il obtient un 6 il gagne 8 euros. Sinon, il perd 2 euros.

1. Calculer l'espérance de gain du joueur.

   On propose au joueur d'acheter pour n euros un deuxième dé avant de jouer. Il peut dans ce cas lancer les deux dés. Les gains sont les mêmes : 8 euros s'il obtient au moins un 6, et -2 euros sinon.

2. Pour quelles valeurs de n cet achat rend-il le jeu plus avantageux ?

Exercice 2 : loi de Pareto (10 points)

Soit $k > 1$. La loi de Pareto de paramètre k, notée $\mathcal{P}ar(k)$, est donnée par la densité de probabilité f nulle sur $]-\infty, 1[$ et définie sur $[1, +\infty]$ par

$$f(x) = \frac{k}{x^{k+1}}.$$

Soit $X \sim \mathcal{P}ar(k)$.

1. (a) Vérifier que f est bien une densité de probabilité.
   (b) Calculer l'espérance de X.
   (c) Soit $a > 1$. Calculer $\mathbf{P}(X > a)$.
   (d) Déterminer la médiane de X, *i.e.* la valeur m telle que $\mathbf{P}(X < m) = \mathbf{P}(X > m)$.
   (e) Soit $b > 1$. Calculer $\mathbf{P}(X > a + b | X > a)$.
   Montrer que cette probabilité croît avec a.
   (f) Si X représente une durée de vie, comment peut-on interpréter ce résultat ?

2. (a) Posons $Y = \ln(X)$. Montrer que Y suit la loi exponentielle de paramètre k.
   (b) Soit $n \geq 30$. Écrire le théorème central limite pour n tirages Y_j indépendants de la variable Y.
   (c) On dispose d'un échantillon de 50 valeurs de la variable X tirées indépendamment. On souhaite estimer la valeur du paramètre k. On sait simplement que le produit des 50 valeurs vaut approximativement 500000.

   Donner une estimation de k à l'aide de la variable Y.
   (d) Donner un intervalle de confiance à 95% pour k.

Exercice 3 : Valeur moyenne d'un signal bruité (7 points)

On reçoit un signal f entre les instants $t = 0$ et $t = 1$ et on souhaite calculer sa valeur moyenne donnée par $\int_0^1 f(t)dt$. Le résultat ne sera pas exact pour deux raisons. Le signal a été échantilloné, l'intégrale sera donc approchée par la moyenne des valeurs observées et de plus le signal émis a été perturbé par la présence d'un bruit. Le but de cet exercice est de mesurer l'impact qu'a ce bruit sur l'exactitude du résultat obtenu.

On note f le signal émis et ε le bruit. Ces fonctions sont définies sur $[0, 1]$. Le signal reçu est ainsi donné par $g = f + \varepsilon$. Le signal a été échantilloné avec un pas de $\frac{1}{n}$ avec $n \in \mathbf{N}^*$. On on a donc reçu les valeurs $g(\frac{k}{n})$ pour $k = 1 \ldots n$. La valeur moyenne de f sera donc estimée par $\frac{1}{n}\sum_{k=1}^{n} g(\frac{k}{n})$. En l'absence de bruit, on sait que l'erreur commise pour la fonction f considérée aurait été donnée par $|\int_0^1 f(t)dt - \frac{1}{n}\sum_{k=1}^{n} f(\frac{k}{n})| < \frac{10}{n}$.

Le bruit ε est modélisé à l'aide de la loi normale. On suppose que les variables $X_k = \varepsilon(\frac{k}{n})$ pour $k = 1 \ldots n$ sont des variables indépendantes de même loi $\mathcal{N}(0, 1)$.

On prendra $n = 100$ comme fréquence d'échantillonage (hormis dans les questions 3 et 4).

1. Appliquer le théorème central limite aux variables X_k.
2. Déterminer la probabilité p que $|\frac{1}{n}\sum_{k=1}^{n} g(\frac{k}{n}) - \frac{1}{n}\sum_{k=1}^{n} f(\frac{k}{n})| < 0,1$.
3. Déterminer une valeur de n, c'est-à-dire une fréquence d'échantillonage, telle que cette probabilité p soit de l'ordre de 0,99.
4. Que peut-on alors dire de la précision de la valeur moyenne de f obtenue ?
5. Si on ne peut pas augmenter la fréquence d'échantillonage $n = 100$, on peut essayer de diminuer le bruit. Déterminer une valeur de l'écart-type σ telle que si le bruit est décrit par une loi normale $\mathcal{N}(0, \sigma^2)$, alors la probabilité p sera de l'ordre de $0, 99$.