FINAL - SQ20

La présentation, la lisibilité et la qualité de la rédaction entreront pour une part importante dans l'appréciation des copies. **Le formulaire distribué en cours est le seul document autorisé. L'utilisation de la calculatrice est conseillée.**
On changera de copie à chaque exercice.

Exercice 1 : indépendance de variables aléatoires (5 points)

Soient $X_1\,,\,X_2\,,\,\ldots\,,\,X_n$ des variables aléatoires définies sur un même espace probabilisé $(\Omega\,,\,\mathcal{T}\,,\,P)$, mutuellement indépendantes et suivant la même loi uniforme sur l'intervalle $[0\,,\,1]$.

On pose $\quad Y_n = \max\{X_1\,,\,X_2\,,\,\ldots\,,\,X_n\}$

1. Rappeler la fonction de répartition F de chacune des variables aléatoires X_k .
2. (a) Soit x un nombre réel. Exprimer l'événement $[Y_n \leqslant x]$ en fonction des événements $[X_k \leqslant x]$.
   (b) Déterminer la fonction de répartition G_n de Y_n.
3. En déduire que Y_n est une variable aléatoire à densité et donner une densité g_n de Y_n.
4. Calculer l'espérance mathématique de Y_n.
5. Montrer que $V(Y_n) = \dfrac{n}{(n+1)^2(n+2)}$.

Pensez à changer de copie

Exercice 2 : deux questions indépendantes (6 points)

1. Soit n un entier supérieur à 2 et $q \in]0\,,\,1[$.

   Un jury de concours pose une question dont la réponse est connue par une proportion q d'étudiants. Les étudiants ont le choix entre n réponses dont une seule est la bonne.
   Un étudiant choisi au hasard, répond correctement à la question posée.

   Calculer en fonction de n et de q, la probabilité qu'il ait répondu en connaissant réellement la réponse.

2. Un QCM d'examen comporte 10 questions indépendantes. 1600 étudiants passent cet examen. Chaque étudiant a 1 chance sur 3 de répondre correctement à chacune des 10 questions posées et répond à toutes les questions.
   Soit X la variable aléatoire égale au nombre de bonnes réponses obtenues pour un étudiant à ce QCM. X représente donc la note sur 10, obtenue par un étudiant à cet examen (les mauvaises réponses ne sont pas pénalisées).
   (a) Reconnaître la loi de probabilité de X.
   (b) On désigne par par $\overline{X}$ la moyenne des notes obtenues par l'ensemble des 1600 étudiants.
   Par quelle loi peut-on approcher la loi de $\overline{X}$? *On précisera ses paramètres.*
   En déduire une valeur approchée du réel a vérifiant $P\big(\overline{X} \leqslant a\big) = 5\%$.

Pensez à changer de copie

Exercice 3 : estimation par intervalle de confiance (9 points)

λ désigne un paramètre réel strictement positif, λ étant inconnu.

Soit n un entier naturel tel que $n \geqslant 3$. On considère un n-échantillon $(X_1, X_2, ..., X_n)$ de variables aléatoires à valeurs strictement positives, indépendantes, de même loi exponentielle de paramètre λ.

On pose pour tout entier $n \geqslant 3$, $S_n = \sum_{k=1}^{n} X_k$ et $J_n = \lambda S_n$.

1. Calculer pour tout entier $n \geqslant 3$, $E(S_n)$, $V(S_n)$, $E(J_n)$ et $V(J_n)$.
2. On admet que $E\left(\dfrac{1}{J_n}\right) = \dfrac{1}{n-1}$ et que $E\left(\dfrac{1}{J_n^2}\right) = \dfrac{1}{(n-1)(n-2)}$

   On pose pour tout n supérieur ou égal à 3, $T_n = \dfrac{n}{S_n}$.

   (a) Montrer que T_n est un estimateur asymptotiquement sans biais de λ.

   (b) Calculer l'erreur quadratique moyenne associée à T_n, que l'on notera $r(T_n)$.

   (c) En déduire la limite, lorsque n tend vers $+\infty$, de l'erreur quadratique moyenne de T_n en λ.
3. Dans cette question, on veut construire un intervalle de confiance du paramètre λ au risque α. On note Φ la fonction de répartition de la loi normale centrée réduite, et t_α le réel strictement positif tel que $\Phi(t_\alpha) = 1 - \dfrac{\alpha}{2}$.

   (a) En appliquant le théorème de la limite centrée, justifier que la variable aléatoire S_n^* définie par $S_n^* = \lambda \dfrac{S_n}{\sqrt{n}} - \sqrt{n}$ converge en loi vers une variable aléatoire U de loi normale centrée réduite.

   (b) En déduire que pour n assez grand, on a approximativement : $P([-t_\alpha \leqslant S_n^* \leqslant t_\alpha]) \approx 1 - \alpha$.

   (c) Montrer que pour n assez grand, l'intervalle $\left[\left(1 - \dfrac{t_\alpha}{\sqrt{n}}\right) T_n \,,\, \left(1 + \dfrac{t_\alpha}{\sqrt{n}}\right) T_n\right]$ est un intervalle de confiance aléatoire de λ au risque α.
4. Avec le logiciel Scilab, on simule 100 réalisations indépendantes $x_1, x_2, \ldots, x_{100}$ d'une variable aléatoire de loi exponentielle de paramètre λ.

   La valeur affectée à λ est cachée. La commande `sum(grand(1,100,"'exp"',1/lambda)` fournit la valeur $14,646$

   (a) Donner une valeur approchée de t_α lorsque $\alpha = 5\%$.

   (b) Proposer un intervalle de confiance observé pour λ au niveau de confiance 95%.