**PROBABILITÉS - STATISTIQUES - SQ20**

**TRONC COMMUN**

**FINAL - PRINTEMPS 2012**

DURÉE DE L'ÉPREUVE : 2 HEURES

**Une calculatrice de poche, les tables de statistiques ainsi qu'une feuille A4 manuscrite de notes personnelles sont les seuls documents autorisés. Ces outils sont strictement personnels, et ne peuvent faire l'objet d'un prêt entre deux étudiants. *Les téléphones portables doivent être éteints, et rangés au fond d'un sac.* Le non respect d'une de ces consignes pourra entraîner l'annulation de votre semestre à l'UTBM.**
**Les deux exercices sont à rédiger sur des copies différentes.**

**Exercice 1** (8 points) Dans cet exercice, on considère un réel strictement positif $R$ fixé, et on considère la fonction $f$ définie sur R par :

$$\begin{cases} f(t) = \dfrac{2t}{R^2} & \text{si } t \in [0, R] \\ f(t) = 0 & \text{sinon} \end{cases}$$

1. Montrer que $f$ est une densité de probabilité.
   Dans toute la suite de l'exercice, on note $X_1$ et $X_2$ deux variables aléatoires indépendantes de densité $f$.
2. Calculer l'espérance, et la variance de $X_i$ (pour $i \in \{1, 2\}$).
3. Calculer la fonction de répartition $F_{X_i}(x)$ (pour $i \in \{1, 2\}$).
4. On note $T_1 = \dfrac{3}{4}(X_1 + X_2)$.
   (a) Calculer l'espérance, et la variance de $T_1$.
   (b) $T_1$ est-il un estimateur sans biais de $R$ ?
5. On note $T_2 = \dfrac{5}{4}\max(X_1, X_2)$ : on voit ainsi que $T_2$ est à valeurs dans l'intervalle $\left[0, \dfrac{5}{4}R\right]$.
   (a) Calculer la fonction de répartition $F_{T_2}$ de $T_2$ sur l'intervalle $\left[0, \dfrac{5}{4}R\right]$.
   (b) Montrer que la densité $f_{T_2}$ de $T_2$ est la fonction
   $$\begin{cases} f_{T_2}(t) = \dfrac{4^5}{5^4}\dfrac{t^3}{R^4} & \text{si } t \in \left[0, \dfrac{5}{4}R\right] \\ f_{T_2}(t) = 0 & \text{sinon} \end{cases}$$
   (c) Calculer l'espérance, et la variance de $T_2$.
   (d) $T_2$ est-il un estimateur sans biais de $R$ ?
6. Des estimateurs $T_1$ et $T_2$, lequel est le meilleur ?

Pensez à changer de copie.

**Exercice 2** (12 points) Dans une production, on a une proportion $p$ inconnue d'éléments défectueux.

On tire un échantillon aléatoire de taille $n$ (le tirage est assimilé à un tirage avec remise).
On note $X_{n,p}$ la variable aléatoire comptant le nombre d'éléments défectueux dans un tel échantillon aléatoire et on pose $\overline{X_p} = \dfrac{X_{n,p}}{n}$.

1. (a) Quelle est la loi de $X_{n,p}$ ?
   (b) Quelles sont l'espérance et l'écart type de $\overline{X_p}$ ?
2. $n = 400$.
   (a) D'après le TCL (théorème centrale limite), par quelle loi continue peut-on approcher la loi de $\overline{X_p}$ ?
   (b) On observe sur l'échantillon tiré une fréquence d'éléments défectueux égale à 13%. Donner un intervalle de confiance pour $p$ à 95% (pour le calcul des bornes de l'intervalle, on considérera que la valeur du paramètre $p$ est égale à son estimation $\widehat{p}$).
3. $n = 50$.
   Pour un échantillon de taille moyenne, l'utilisation du TCL est encore valable, mais on préfère ne plus recourir à l'estimation ponctuelle $\widehat{p}$ pour calculer l'écart-type de $\overline{X_p}$.
   (a) Sachant que, d'après le TCL, on peut considérer que $\overline{X_p}$ suit une loi normale d'espérance $p$, et d'écart type $\sigma = 0,1\sqrt{2\,p(1-p)}$, déterminer le nombre $k$ tel que,
   $$\mathbb{P}\left((p - \overline{X_p})^2 \leqslant k\,p(1-p)\right) = 0,95. \quad (1)$$
   Autrement dit, dans 95% des cas, on a
   $$(p - \widehat{p})^2 \leqslant k\,p(1-p) \quad (2)$$
   où $\widehat{p}$ est la valeur prise par $\overline{X_p}$ dans le tirage réalisé.
   (b) Sur l'échantillon tiré, la variable aléatoire $\overline{X_p}$ prend la valeur $\widehat{p} = 0,18$. En prenant $k = 0,0768$, résoudre l'inéquation (2), d'inconnue $p$, pour obtenir l'intervalle de confiance pour $p$ à 95%.
4. $n = 10$.
   Sur l'échantillon tiré, la variable aléatoire $\overline{X_p}$ prend la valeur $\widehat{p} = 30\%$ On se propose de déterminer pour quelle valeur de $p$ la probabilité d'observer $\overline{X_p} = 0,3$ sur un échantillon est maximale.
   On note $L$ la fonction définie sur $[0,1]$ par
   $$L(p) = \mathbb{P}(\overline{X_p} = 0,3) = \mathbb{P}(X_{10,p} = 3).$$
   (a) Donner l'expression de $L(p)$ puis de $g(p) = \ln(L(p))$.
   (b) Étudier les variations de la fonction $g$. En déduire l'estimation ponctuelle de $p$ suivant le maximum de vraissemblance.
   (c) Le responsable de la production assure que $p = 10\%$. En utilisant la loi de la variable aléatoire $X_{10,p}$ déterminée à la question 1a, construire un test pour valider, ou infirmer, l'hypothèse $p = 10\%$ contre l'hypothése alternative $p > 10\%$ au seuil de 5% (la loi de $X_{10,p}$ étant discrète, on écrira la règle de décision de façon à ce que la probabilité de rejeter à tort l'hypothèse nulle soit **au maximum** de 5%).
   (d) Quelle est la décision du test pour l'échantillon prélevé ($\widehat{p} = 30\%$) ?