E.S.S.T.I.N. **Vendredi 29 mars 2013**
UNIVERSITE DE LORRAINE

3e année
Durée : 1 heure

Eléments de correction

DEVOIR SURVEILLE n°1

Statistiques

Barème indicatif sur 20 : Partie A : 8 points Partie B : 12 points

PARTIE A

1. *A partir de la loi de probabilité suivie par la variable X, déterminer* $E(X)$, $V(X)$ *et* $E(X^2)$.

X suit une loi normale de moyenne μ et de variance σ^2 donc $E(X)=\mu$, $V(X)=\sigma^2$ et $E(X^2)=V(X)+\left[E(X)\right]^2=\sigma^2+\mu^2$.

2. *Etudier la qualité des quatre estimateurs* S_1, S_2, S_3 et S_4 *en déterminant le biais, la variance et le Carré Moyen de l'Erreur (CME) de chaque estimateur.*

Le paramètre que l'on cherche à estimer est la superficie du champ : μ^2.

$\underline{S_1=X_1^2}$

$$\text{Biais}(S_1)=E(X_1^2)-\mu^2=\sigma^2+\mu^2-\mu^2=\sigma^2$$
$$V(S_1)=V(X_1^2)=2(\sigma^4+2\,\mu^2\sigma^2)$$
$$\text{CME}(S_1)=V(S_1)+\left[\text{Biais}(S_1)\right]^2=3\sigma^4+4\mu^2\sigma^2$$

$\underline{S_2=X_2^2}$

$$\text{Biais}(S_2)=E(X_2^2)-\mu^2=\sigma^2+\mu^2-\mu^2=\sigma^2$$
$$V(S_2)=V(X_2^2)=2(\sigma^4+2\,\mu^2\sigma^2)$$
$$\text{CME}(S_2)=V(S_2)+\left[\text{Biais}(S_2)\right]^2=3\sigma^4+4\mu^2\sigma^2$$

$\underline{S_3=X_1.X_2}$

En supposant l'indépendance de X_1 et X_2,

$$\text{Biais}(S_3)=E(X_1.X_2)-\mu^2=E(X_1).E(X_2)-\mu^2=\mu^2-\mu^2=0$$
$$V(S_3)=V(X_1.X_2)=E\left[(X_1.X_2)^2\right]-\left[E(X_1.X_2)\right]^2$$
$$=E(X_1^2).E(X_2^2)-\left[E(X_1)\right]^2\left[E(X_2)\right]^2=\sigma^4+2\,\mu^2\sigma^2$$
$$\text{CME}(S_3)=V(S_3)+\left[\text{Biais}(S_3)\right]^2=\sigma^4+2\,\mu^2\sigma^2$$

$\underline{S_4=\dfrac{S_1+S_2}{2}}$

En supposant l'indépendance de S_1 et S_2,

$$\text{Biais}(S_4)=E\left(\frac{S_1+S_2}{2}\right)-\mu^2=\frac{1}{2}E(S_1)+\frac{1}{2}E(S_2)-\mu^2=\sigma^2+\mu^2-\mu^2=\sigma^2$$
$$V(S_4)=V\left(\frac{S_1+S_2}{2}\right)=\frac{1}{4}V(S_1)+\frac{1}{4}V(S_2)=\sigma^4+2\,\mu^2\sigma^2$$
$$\text{CME}(S_4)=V(S_4)+\left[\text{Biais}(S_4)\right]^2=2\,\sigma^4+2\,\mu^2\sigma^2$$

3. *A l'aide des résultats de la question 2, aider l'agriculteur à choisir l'estimateur parmi les quatre qui vous semble préférable aux autres.*

On choisit l'estimateur qui a le CME le plus petit donc S_3.

PARTIE B

4. *Préciser le caractère étudié (nom et nature).*

X : mesure du côté du champ (en m). Variable quantitative continue.

5. *Calculer la moyenne et la médiane des mesures.*

$\bar{x}=\dfrac{\sum_{i=1}^{10}x_i}{10}=500{,}7\text{ m}$ $\quad$ $Me=502\text{ m}$ (valeur moyenne entre la 5e et la 6e mesure)

6. *Calculer l'étendue et l'écart type des mesures.*

$E=x_{Max}-x_{min}=512-485=27\text{ m}$ $\quad$ $s^2=\dfrac{\sum_{i=1}^{10}x_i^2}{10}-\bar{x}^2$ d'où $s=8{,}79\text{ m}$

7. *Calculer le coefficient de variation et conclure.*

$CV=\dfrac{s}{\bar{x}}=1{,}75\ \%$. CV est très petit, la distribution est donc homogène.

8. Quelle est l'estimation ponctuelle sans biais de la moyenne de X ?

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 500{,}7 \text{ m}$$

9. Déterminer la marge d'erreur associée à cette estimation pour α = 5 %.

X suit une loi normale de variance population inconnue et la taille de l'échantillon est inférieure à 30 donc la marge d'erreur k est égale à :

$$k = t_{1-\alpha/2\,;\,n-1} \frac{s^*}{\sqrt{n}} = 2{,}262 \frac{9{,}26}{\sqrt{10}} = 6{,}62 \text{ m} \quad \text{avec} \quad s^* = \sqrt{\frac{n}{n-1}}.s$$

10. En déduire l'intervalle de confiance pour μ (α = 5 %). Proposer une interprétation pour cet intervalle de confiance.

$$\bar{x} - k \le \mu \le \bar{x} + k \quad \text{soit} \quad 494{,}1 \text{ m} \le \mu \le 507{,}3 \text{ m}$$

avec un niveau de confiance de 95 %

L'intervalle $\left[494{,}1 \text{ m}\,;507{,}3 \text{ m}\right]$ a 95 % de chances d'encadrer la vraie valeur μ.

11. Calculer la taille de l'échantillon qu'il faudrait prélever pour diminuer la marge d'erreur sur la moyenne de 25 % (α = 5 %).

On veut diminuer la marge d'erreur k de 25 % soit une nouvelle marge d'erreur $k = 6{,}62\,.\,0{,}75 = 4{,}97 \text{ m}$

Pour le calcul de n, comme t dépend de n, on approxime $t_{1-\alpha/2\,;\,n-1}$ par $u_{1-\alpha/2}$

Cette approximation est valide pour n supérieur ou égal à 30.

$$n = \left[\frac{u_{1-\alpha/2}\, s^*}{4{,}97}\right]^2 = 13{,}3$$

On obtient une valeur de n inférieure à 30, donc il est nécessaire de calculer la marge d'erreur k avec la loi de Student et n = 14.

On obtient : $k = t_{1-\alpha/2\,;\,13} \frac{s^*}{\sqrt{14}} = 5{,}34$

La marge d'erreur k obtenue est plus grande que la marge d'erreur voulue donc il faut augmenter n et recalculer k.
Pour n = 16, on trouve k = 4,93. Il faut donc réaliser 16 mesures de X pour obtenir une marge d'erreur inférieure à 4,97 m avec un risque α égal à 5 %.