2. ESTIMATION PONCTUELLE ET ESTIMATION PAR INTERVALLE

EXERCICE 2.2

On extrait un échantillon de taille n = 16 dans une population dont l'écart type σ et la moyenne μ sont inconnus. L'intervalle de confiance à 95 % de la moyenne de la population, estimé à partir de l'échantillon, est égal à [24 ; 28].

1. Déterminer la meilleure estimation de la variance de la population.

A partir du contexte de l'exercice, déterminer parmi les deux estimations de la variance population Tr 30 l'expression de la valeur recherchée.

Déterminer alors la formule du Tr 28 qui a permis de calculer les deux bornes de l'intervalle de confiance pour la moyenne population puis en déduire la valeur recherchée.

Données échantillon : n = 16

La valeur de la moyenne population n'est pas connue, l'estimation de la variance recherchée est donc, d'après le Tr 30 :

$$\hat{\sigma}^2 = s^{*2}$$

La variance population est inconnue, l'échantillon est petit (n < 30) donc il faut faire l'hypothèse que X suit une loi normale pour pouvoir appliquer la formule suivante :

$$\bar{x} - t_{1-\frac{\alpha}{2};\, n-1}\frac{s^*}{\sqrt{n}} \le \mu \le \bar{x} + t_{1-\frac{\alpha}{2};\, n-1}\frac{s^*}{\sqrt{n}}$$

Cette formule a été utilisée pour le calcul de bornes de l'intervalle de confiance à 95 % de la moyenne population $[24; 28]$ d'où le système d'équations suivant :

$$\begin{cases} \bar{x} - t_{1-\frac{\alpha}{2};\, n-1}\dfrac{s^*}{\sqrt{n}} = 24 \\ \bar{x} + t_{1-\frac{\alpha}{2};\, n-1}\dfrac{s^*}{\sqrt{n}} = 28 \end{cases}$$

$$\text{d'où} \quad \begin{cases} \bar{x} = 26 \\ s^* = \dfrac{2\sqrt{n}}{t_{1-\frac{\alpha}{2};\, n-1}} = 3{,}754 \end{cases} \quad \text{avec} \quad t_{1-\frac{\alpha}{2};\, n-1} = t_{0{,}975\,;\,15} = 2{,}131$$

$$\text{d'où} \qquad \hat{\sigma}^2 = s^{*2} = 14{,}09$$

2. Déterminer un intervalle de confiance à 90 % de la variance de la population.

Il faut faire l'hypothèse que X suit une loi normale et, comme la moyenne population est toujours inconnue, on utilise la formule suivante :

Intervalle de confiance à 90 % pour σ^2

$$\frac{(n-1)s^{*2}}{\chi^2_{0{,}95\,;\, n-1=15}} \le \sigma^2 \le \frac{(n-1)s^{*2}}{\chi^2_{0{,}05\,;\, n-1=15}}$$

$$\frac{(n-1)s^{*2}}{25{,}00} \le \sigma^2 \le \frac{(n-1)s^{*2}}{7{,}26}$$

$$8{,}4 \le \sigma^2 \le 29{,}1$$

L'intervalle de confiance $[8{,}4\,; 29{,}1]$ a 90 % de chances de contenir la valeur σ^2 de la variance population.

3. Calculer la taille de l'échantillon à prélever pour déterminer la moyenne avec une précision de 10 % (le risque α sera pris égal à 1 %).

Déterminer la valeur de la marge d'erreur k à atteindre à partir de la précision mentionnée.

Procéder ensuite de la même façon qu'à la question 2 de l'exercice 2.4 jusqu'à atteindre cette marge d'erreur.

On ne connaît toujours pas la valeur de la variance population. On ne connaît toujours pas la moyenne population. En supposant que X suit une loi normale, l'expression de la marge d'erreur est donc :

$$k = 10\%\ \hat{\mu} = 0{,}10\,\bar{x} = 2{,}6 \qquad \text{avec} \qquad \bar{x} = 26 \quad (\text{voir question 1.})$$

$$k = t_{1-\frac{\alpha}{2};\, n-1}\frac{s^*}{\sqrt{n}} = 2{,}6 \qquad \text{avec} \qquad \widehat{\sigma^2} = s^{*2} = 14{,}09$$

On utilise pour exprimer k la valeur estimée pour la moyenne ainsi que la valeur estimée pour la variance calculées à la question 1. car, faute d'informations supplémentaires, elles représentent toujours les meilleures estimations pour les paramètres population.

On en déduit l'expression : $n = \frac{\left(t_{0{,}995;\, n-1}\right)^2}{k^2} s^{*2}$

Comme on ne connaît pas n et que la valeur à lire dans la table de Student dépend de n, on utilise l'approximation de la loi de Student par la loi normale centrée réduite, valable à partir de n = 30, ainsi :

$$n \cong \frac{\left(u_{0{,}995}\right)^2}{k^2} s^{*2} = \frac{(2{,}575)^2}{2{,}6^2} 14{,}09 = 13{,}8 \cong 14$$

La valeur de n obtenue est inférieure à 30 donc l'approximation n'est pas valide.

On calcule la marge d'erreur k pour n = 14 avec la loi de Student :

$$k = t_{0{,}995;\, 14-1=13}\frac{3{,}754}{\sqrt{14}} = 3{,}012\,\frac{3{,}754}{\sqrt{14}} = 3{,}02$$

On trouve une marge d'erreur nettement supérieure à la marge d'erreur voulue. Il faut donc augmenter n.
On augmente donc n de façon itérative et on recalcule, pour chaque valeur de n, k avec la table de Student jusqu'à atteindre la marge d'erreur voulue.
La valeur de n pour laquelle on obtient une marge d'erreur inférieure à 2,6 est n = 18.
La marge d'erreur obtenue pour n = 18 est :

$$k = t_{0,995;\,18-1=17}\frac{3,754}{\sqrt{18}} = 2,898\,\frac{3,754}{\sqrt{18}} = 2,56$$