2. ESTIMATION PONCTUELLE ET ESTIMATION PAR INTERVALLE

EXERCICE 2.1

La production moyenne d'un quartier de mine est de 400 berlines par jour. On veut connaître le rapport $\frac{\text{net}}{\text{brut}}$ de la production du quartier.

On a déterminé, pour chaque berline d'un échantillon de 20, le rapport $\frac{\text{net}}{\text{brut}}$.

On a trouvé :

0,60	0,71	0,63	0,71	0,56
0,71	0,46	0,45	0,55	0,75
0,58	0,79	0,79	0,74	0,62
0,57	0,63	0,68	0,64	0,58

1. Déterminer l'estimation ponctuelle de la moyenne population

Utiliser la formule du polycopié Chapitre 2 – Tr 28.

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{n} X_i}{n} = 0{,}6375$$

2. Déterminer un intervalle de confiance à 95 % pour la moyenne.

Choisir, dans le polycopié Chapitre 2 – Tr 28, la formule la plus proche du contexte de l'exercice (en ajoutant éventuellement des hypothèses).

Données échantillon : $n = 20$ $\quad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 0{,}6375 \quad s = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2} = 0{,}0950$

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

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

avec $\quad s^* = \sqrt{\frac{n}{n-1}}\,s = 0{,}0975 \quad$ et

$$t_{1-\frac{\alpha}{2};\,n-1} = t_{0{,}975\,;\,19} = 2{,}093$$

valeur lue dans la table de Student bilatérale pour la colonne $\alpha = 0{,}05$ et la ligne 19

OU lue dans la table de Student unilatérale pour la colonne $P = 0{,}975$ et la ligne 19

Intervalle de confiance à 95 % pour μ

$$0{,}6375 - 0{,}0456 \leq \mu \leq 0{,}6375 + 0{,}0456$$

$$0{,}5919 \leq \mu \leq 0{,}6831$$

L'intervalle de confiance $[0{,}59\,;0{,}68]$ a 95 % de chances de contenir la valeur μ de la moyenne population.

3. Déterminer la taille de l'échantillon pour ramener la marge d'erreur sur la moyenne à $\pm\,0{,}03$ ($\alpha = 5$ %).

Utiliser, dans un premier temps, l'approximation de u par t pour déterminer n.

Vérifier (puis ajuster éventuellement) le résultat obtenu en recalculant la marge d'erreur k avec t et la valeur de n obtenue précédemment.

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

$$k = |\bar{x} - \mu| = 0{,}03 = t_{1-\frac{\alpha}{2};\,n-1}\frac{s^*}{\sqrt{n}} \qquad \text{avec} \qquad \widehat{\sigma^2} = s^{*2} = 0{,}0975^{\,2}$$

On conserve la valeur estimée pour la variance (calculée à la question précédente) car, faute d'information supplémentaire, elle représente toujours la meilleure estimation pour la variance population.

On en déduit l'expression : $\quad n = \frac{(t_{0{,}975;\,n-1})^2}{k^2} s^{*2}$

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 :

$$n \cong \frac{(u_{0{,}975})^2}{k^2} s^{*2} = \frac{(1{,}96)^2}{0{,}03^2} 0{,}0975^2 = 40{,}6 \cong 41$$

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

Pour vérifier le résultat, on calcule la marge d'erreur k pour $n = 41$ avec la loi de Student :

$$k = t_{0{,}975;\,41-1=40}\frac{0{,}0975}{\sqrt{41}} = 2{,}021\frac{0{,}0975}{\sqrt{41}} = 0{,}0308$$

On trouve une marge d'erreur légèrement supérieure à la marge d'erreur voulue.

Pour affiner, on peut augmenter n d'une unité par itération et à chaque nouvelle valeur on peut recalculer k avec la loi de Student.

Le résultat le plus précis est obtenu pour $n = 43$:

$$k = t_{0{,}975;\,43-1=42}\frac{0{,}0975}{\sqrt{43}} = 2{,}018\frac{0{,}0975}{\sqrt{43}} = 0{,}0300$$

Remarque : la valeur de t pour 42 DL ne figure pas dans la table distribuée ; elle est obtenue avec une fonction informatique, par exemple une fonction Excel (cf. tableau des fonctions Excel disponible sur Arche).

4. Déterminer un intervalle de confiance à 95 % pour la variance (n = 20).

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

<u>Intervalle de confiance à 95 % pour</u> σ^2

$$\frac{(n-1)s^{*2}}{\chi^2_{0{,}975\,;\,n-1=19}} \leq \sigma^2 \leq \frac{(n-1)s^{*2}}{\chi^2_{0{,}025\,;\,n-1=19}}$$

$$\frac{(n-1)s^{*2}}{32{,}85} \leq \sigma^2 \leq \frac{(n-1)s^{*2}}{8{,}91}$$

$$0{,}005 \leq \sigma^2 \leq 0{,}020$$

L'intervalle de confiance $[0{,}005\,;0{,}020]$ a 95 % de chances de contenir la valeur σ^2 de la variance population.

On peut vérifier que l'estimation ponctuelle $\widehat{\sigma^2} = s^{*2} = 0{,}0095$ est comprise entre les deux bornes de l'intervalle de confiance.