Exo 1

On est dans le cadre où l'on connait l'écart-type de la population qui vaut 0.6.

On ne connait pas la loi mais nous sommes avec un échantillon de taille 100, donc le théorème Central limit peut nous aider au besoin.

i) On veut tester :

H_0 : la machine est bien réglé

H_a : la machine n'est pas bien réglé

Concrètement, cela signifie :

H_0 : $\mu = 8.3$

H_a : $\mu \neq 8.3$ où μ est l'espérance de la variable X

avec X : longueur des tiges métalliques.

On considère donc un test bilatéral.

Un estimateur de μ est $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

Par le théorème central limit, on a :

$$Z_n = \sqrt{n} \, \frac{\overline{X}_n - \mu}{\sigma} \rightsquigarrow \mathcal{N}(0;1)$$

Et intuitivement la région de rejet est

$$\{ \overline{X}_n > \text{seuil}_1 \} \cup \{ \overline{X}_n < \text{seuil}_2 \}$$

avec $\text{seuil}_1 > 8.3$ et $\text{seuil}_2 < 8.3$

Autrement dit on va rejeter $\mathcal{H}_0$ lorsque $\overline{X}_n$ s'écarte trop de la valeur de référence ~~8~~ 8.3.

Cela s'écrit encore :

$\{ |Z_n| > \text{seuil} \}$ (on prend le même seuil car on utilise la symétrie de la loi normale)

la détermination du seuil est liée au fait que

$$P_{\mathcal{H}_0} (|Z_n| > \text{seuil}) \leq \alpha$$

Or sous $\mathcal{H}_0$, on a $Z_n = \sqrt{n} \, \dfrac{\overline{X}_n - 8.3}{0.6} \rightsquigarrow \mathcal{N}(0;1)$

donc seuil est tel que :

$$P (|\mathcal{N}(0;1)| > \text{seuil}) \leq \alpha$$

soit $P(\mathcal{N}(0;1) < \text{seuil}) \geq 1 - \alpha/2$

soit seuil est la fractile d'ordre $1 - \frac{\alpha}{2}$ d'une normale centrée réduite.

Ainsi si $\alpha = 5\%$ seuil $= 1.96$

$\alpha = 1\%$ seuil $= 2.575$

$\alpha = 0.1\%$ seuil $= 3.3$

Maintenant que vaut Z_n sur notre échantillon?

$$Z_n = \sqrt{100} \times \frac{8.57 - 8.3}{0.6} = 10 \times \frac{8.57 - 8.3}{0.6} = 4.5$$

donc dans les trois cas, on a $|Z_n| > $ seuil.

Autrement dit, on rejette que la machine est bien réglée à 5%, 1% et 0.1%.

2) Pour $N = 20$, on ne peut plus utiliser le théorème de la limite centrale. C'est la raison pour laquelle il faut avoir l'information du caractère normal des longueurs.

Cependant, la région de rejet demeure la même et les seuils aussi.

Par contre $$Z_n = \sqrt{20} \times \frac{8.57 - 8.3}{0.6} \approx 2.01$$

Donc on rejette H_0 au seuil de 5%, mais on ne peut pas rejeter au seuil de 1% et 0.1%.

3) Cette fois, on ne connait pas l'écart-type de la population. (4)

Mais on connait une estimation.

Par ailleurs, on a l'information que les variables aléatoires sont de loi normale.

Donc, on peut considérer :

$$T_n = \sqrt{n}\,\frac{\overline{X}_n - \mu}{\hat{\sigma}_n} \quad \text{avec} \quad \hat{\sigma}_n^2 = \frac{1}{n-1}\sum (X_i - \overline{X}_n)^2$$

on sait que $T_n \sim T(n-1)$

la région de rejet va être $\{ |T_n| > \text{seuil} \}$

avec $P_{H_0}\left(|T_n| > \text{seuil} \right) \leq \alpha$

a sous H_0, $T_n = \sqrt{n}\,\frac{\overline{X}_n - 8.3}{0.6} \sim T(n-1)$

donc seuil est tel que :

$$P\left(|T(n-1)| > \text{seuil} \right) \leq \alpha$$

Si $\alpha = 5\%$ seuil $= 2.093$

$\alpha = 1\%$ seuil $= 2.861$

$\alpha = 0.1\%$ seuil $= 3.884$

et $t_n = \sqrt{20} \times \frac{8.57 - 8.3}{0.6} = 2.01$

donc ici, pour aucun des seuils on ne rejette H_0.

4) Selon que l'on connait l'écart-type de la population ou non, selon que l'on connait le caractère normal ou non, la décision ne sera pas forcément la même pour une même valeur de α.

De même, selon la valeur de α, la décision peut changer. Par contre on sait que si on rejette pour α_0, alors on rejettera pour tout $\alpha \geqslant \alpha_0$.