3. TESTS STATISTIQUES (UNE POPULATION)

Tests paramétriques

EXERCICE 3.4

La durée de vie moyenne d'un échantillon de 100 tubes fluorescents a été établie à 1570 heures avec un écart type de 120 heures.

Si μ est la moyenne de la population, vérifier l'hypothèse μ = 1600 heures relativement à l'hypothèse μ ≠ 1600 heures, avec un niveau de signification de 0,05 et de 0,01.

- Repérer si les données de l'énoncé concernent la population ou l'échantillon.
- En déduire la procédure de test à mettre en œuvre.
- Les hypothèses H0 et H1 sont imposées dans cet exercice. Chercher à les interpréter.
- Mettre en œuvre la procédure de test en respectant les 7 étapes. Pour simplifier la mise en œuvre de cette procédure, quelle approximation peut-on faire quand n est supérieur à 30 ?
- Conclure en répondant à la question posée ; la conclusion peut être différente en fonction du niveau de signification (c'est-à-dire α) choisi.

On va tester la durée de vie moyenne de la population de tubes.

La variance de la population n'est pas connue, on choisit donc la procédure du Tr 41.

1) H0 $\mu = 1600\text{ h}$ Durée de vie moyenne égale à 1600 h

H1 $\mu \neq 1600\text{ h}$ Durée de vie moyenne différente de 1600 h

On choisit l'alternative les hypothèses H0 et H1 conformément à ce qui est demandé dans l'énoncé. Une durée de vie moyenne trop courte pénalise le client, une durée de vie moyenne trop grande coûte plus cher au fabricant (utilisation de composants plus fiables donc plus onéreux).

2) $\alpha = 0{,}05$ puis $\alpha = 0{,}01$

3) La variance de la population σ^2 est inconnue, l'échantillon a une taille supérieure ou égale à 30.

4) $T = \left.\dfrac{\overline{X}-1600}{\frac{s^*}{\sqrt{n}}}\right|_{H0\ \mu=1600\text{ h}} \rightarrow$ Loi de Student $\upsilon = n-1$ DL

Comme $n \geq 30$, on peut approximer la loi de Student par la loi normale centrée réduite (mais on peut conserver la loi de Student pour plus de précision) :

$$U = \left.\frac{\overline{X}-1600}{\frac{s^*}{\sqrt{n}}}\right|_{H0\ \mu=1600\text{ h}} \rightarrow N(0\,;1)$$

5) Rejeter H0 $\mu = 1600\text{ h}$

$$\text{si} \quad U = \frac{\overline{X}-1600}{\frac{s^*}{\sqrt{n}}} < -u_{1-\frac{\alpha}{2}} = -u_{0,975} = -1{,}96$$

$$\text{ou si} \quad U = \frac{\overline{X}-1600}{\frac{s^*}{\sqrt{n}}} > u_{1-\frac{\alpha}{2}} = 1{,}96$$

6) Échantillon : $n = 100$ $\bar{x} = 1570\text{ h}$ et $s = 120\text{ h}$

$$\text{d'où} \quad u = \frac{\bar{x}-1600}{\frac{s^*}{\sqrt{n}}} = \frac{\bar{x}-1600}{\frac{s}{\sqrt{n-1}}} = \frac{1570-1600}{\frac{120}{\sqrt{99}}} = -2{,}49 \quad \text{avec} \quad s^* = \sqrt{\frac{n}{n-1}}\,s$$

7) $\alpha = 0{,}05$ $u = -2{,}49 < -u_{0,975} = -1{,}96$ d'où Rejet H0

Sur la base des résultats de cet échantillon, on peut conclure avec un risque α égal à 0,05 que la durée de vie moyenne de 1600 h n'est pas respectée.

$\alpha = 0{,}01$

$-u_{0,995} = -2{,}575 \leq u = -2{,}49 \leq u_{0,995} = 2{,}575$ d'où Non rejet H0

Sur la base des résultats de cet échantillon, on peut conclure avec un risque α égal à 0,01 que la durée de vie moyenne de 1600 h est respectée.

Dans ce cas, l'échantillon mène à une valeur très proche de la région de rejet, pour un risque égal à 0,01, on privilégiera donc la conclusion « Rejet de H0 ».

EXERCICE 3.5

Le fabricant d'un médicament annonce qu'un de ses produits est efficace à 90 %, en supprimant une allergie dans un délai de 8 heures. Dans un échantillon de 200 personnes, le résultat a été effectif pour 160 d'entre elles. Déterminer si l'affirmation du fabricant est légitime.

- Déduire de la question posée si le test doit porter sur une moyenne, une variance ou une proportion.
- Repérer si les données de l'énoncé concernent la population ou l'échantillon.
- En déduire la procédure de test à mettre en œuvre.
- Choisir l'hypothèse H0 et son alternative H1 ; ces deux hypothèses doivent porter sur le paramètre population à étudier et correspondre aux deux réponses possibles à la question posée (affirmation du fabricant correcte ou incorrecte) ; les deux hypothèses ne doivent pas mentionner les résultats obtenus sur l'échantillon. La compréhension du phénomène étudié impose le choix de l'alternative H1.
- Mettre en œuvre la procédure de test en respectant les 7 étapes.
- Conclure en répondant à la question posée ; la conclusion peut être différente en fonction du niveau de signification (c'est-à-dire α) choisi.

Il s'agit de tester un pourcentage d'efficacité d'un médicament donc on va mettre en œuvre un test sur une proportion.
On choisit la procédure du Tr 44 à condition de vérifier que l'on peut utiliser l'approximation de la loi binomiale par la loi normale.

1) H0 $p = 0{,}90$ Affirmation du fabricant correcte

H1 $p < 0{,}90$ Affirmation du fabricant incorrecte : efficacité du médicament plus faible qu'annoncée

On choisit l'alternative « H1 Proportion population inférieure à 0,90 » car on cherche à mettre en évidence une efficacité plus faible que celle annoncée par le fabricant.

2) $\alpha = 0{,}05$

3) Approximation de la loi binomiale par la loi normale vérifiée car

$200\,.\,0{,}90 = 180 \ \geq 5$ et $200\,(1 - 0{,}90) = 20 \geq 5$

4) $U = \dfrac{\widehat{P} - 0{,}90}{\sqrt{\dfrac{0{,}90\,(1-0{,}90)}{n}}} \rightarrow N(0\,;\,1)$

5) Rejeter H0 $p = 0{,}90$

$$\text{si} \quad U = \frac{\widehat{P} - 0{,}90}{\sqrt{\dfrac{0{,}90\,(1 - 0{,}90)}{n}}} < -u_{1-\alpha} = -u_{0,95} = -1{,}645$$

6) Échantillon : $n = 200$ $f = \frac{160}{200} = 0{,}80$

$$u = \frac{0{,}80 - 0{,}90}{\sqrt{\dfrac{0{,}90\,(1 - 0{,}90)}{n}}} = -4{,}71$$

7) $\alpha = 0{,}05$ $u = -4{,}71 < -u_{0,95} = -1{,}645$ d'où Rejet H0

Sur la base des résultats de cet échantillon, on peut conclure avec un risque α égal à 0,05 que l'affirmation du fabricant est incorrecte : l'efficacité du médicament est inférieure à 90 %.

Pour confirmer cette conclusion, on refait le test avec $\alpha = 0{,}01$.

$\alpha = 0{,}01$ $u = -4{,}71 < -u_{0,99} = -2{,}33$ d'où Rejet H0

Sur la base des résultats de cet échantillon, on peut conclure avec un risque α égal à 0,01 que l'affirmation du fabricant est incorrecte : l'efficacité du médicament est inférieure à 90 %.

Les résultats sont donc hautement significatifs.

EXERCICE 3.6

Sur 4000 naissances, une enquête relève 2065 garçons.

Tester l'hypothèse selon laquelle la probabilité d'avoir un garçon à la naissance est 1/2 avec un seuil de 0,05 et avec un seuil de 0,01. Quel est le seuil descriptif du test ?

- Déduire de la question posée si le test doit porter sur une moyenne, une variance ou une proportion.
- Repérer si les données de l'énoncé concernent la population ou l'échantillon.
- En déduire la procédure de test à mettre en œuvre.
- Choisir l'hypothèse H0 et son alternative H1 ; ces deux hypothèses doivent porter sur le paramètre population à étudier et correspondre aux deux réponses possibles à la question posée (probabilité théorique d'avoir un garçon à la naissance respectée ou non) ; les deux hypothèses ne doivent pas mentionner les résultats obtenus sur l'échantillon.
- Pour cet exercice, deux alternatives H1 peuvent être acceptées. Toutefois, les deux alternatives ne mènent pas à la même phrase de conclusion.
- Mettre en œuvre la procédure de test en respectant les 7 étapes.
- Conclure en répondant à la question posée ; la conclusion peut être différente en fonction du niveau de signification (c'est-à-dire α) choisi.
- Déterminer le seuil descriptif défini au Tr 45, en suivant la procédure complémentaire disponible sur Arche. Interpréter la valeur obtenue.

Il s'agit de tester la loi de la génétique c'est-à-dire la probabilité d'avoir un garçon à la naissance égale à 50 %. On va mettre en œuvre un test sur une proportion.
On choisit la procédure du Tr 44 à condition de vérifier que l'on peut utiliser l'approximation de la loi binomiale par la loi normale.

1) H0 $p = 0{,}50$ Loi de la génétique respectée

H1 $p > 0{,}50$ Loi de la génétique non respectée : surplus de garçons

Si on choisit l'alternative « H1 Proportion population supérieure à 0,50 », on cherche à mettre en évidence un surplus de garçons par rapport à la loi de la génétique.

<u>Pour cet exercice, on pourrait également choisir l'alternative « H1 Proportion population différente de 0,50 »</u>. Dans ce cas, on chercherait à mettre en évidence un respect ou non de la loi de la génétique. La phrase de conclusion, en cas de rejet de H0, serait différente.

2) $\alpha = 0{,}05$ puis $\alpha = 0{,}01$

3) Approximation de la loi binomiale par la loi normale vérifiée car

$4000 \,.\, 0{,}50 = 2000 \geq 5$ et $4000\,(1 - 0{,}50) = 2000 \geq 5$

4) $U = \dfrac{\hat{P} - 0{,}50}{\sqrt{\dfrac{0{,}50\,(1-0{,}50)}{n}}} \rightarrow N(0\,;1)$

5) Rejeter H0 $p = 0{,}50$

$$\text{si} \quad U = \frac{\hat{P} - 0{,}50}{\sqrt{\dfrac{0{,}50\,(1-0{,}50)}{n}}} > u_{1-\alpha} = u_{0{,}95} = 1{,}645$$

6) Échantillon : $n = 4000$ $f = \frac{2065}{4000} = 0{,}51625$

$$u = \frac{0{,}51625 - 0{,}50}{\sqrt{\dfrac{0{,}50\,(1-0{,}50)}{n}}} = 2{,}055$$

7) $\alpha = 0{,}05$ $u = 2{,}055 > u_{0{,}95} = 1{,}645$ d'où Rejet H0

Sur la base des résultats de cet échantillon, on peut conclure avec un risque α égal à 0,05 que la loi de la génétique n'est pas respectée : on observe un surplus de garçons.

$\alpha = 0{,}01$ $u = 2{,}055 \leq u_{0{,}99} = 2{,}33$ d'où Non rejet H0

Sur la base des résultats de cet échantillon, on peut conclure avec une confiance de 99 % (ou un risque α égal à 0,01) que la loi de la génétique est respectée : on observe autant de garçons que de filles à la naissance.

Une petite variation du risque α modifie la conclusion. Il faut donc poursuivre l'étude pour conclure avec une très grande confiance sur le respect ou non de la loi de la génétique.

<u>**Calcul du seuil descriptif du test (ou p-value) :**</u>

$\alpha_P = Pr\,(U > u = 2{,}055) = 1 - Pr\,(U \leq u = 2{,}055) = 1 - 0{,}98 = 0{,}02$

Dans la table U, on se place à la valeur $u = 2{,}055$ (ligne 2,0 et colonne 0,05 ou 0,06) et on en déduit que la probabilité correspondant à 2,055 est égale à 0,98.
La valeur du seuil descriptif du test recherchée est égale au complément de la probabilité lue dans la table U.
On trouve une valeur du seuil descriptif entre 0,01 et 0,10, on n'accorde pas une très grande confiance à la conclusion. Toutefois, la valeur étant tout de même faible, on privilégie donc la décision « Rejet H0 ».