

TESTS D'HYPOTHÈSES

DIFFÉRENCE DE DEUX MOYENNES

(ÉCHANTILLONS INDÉPENDANTS)

Échantillon 2015

n_A	Âge
1	18,2
2	22,1
3	18,0
4	20,0
5	20,8
...	...
...	...
...	...
30	20,5
31	19,1
32	20,5
33	23,7
34	22,5
35	19,3
36	22,3

Échantillon 2020

n_B	Âge
1	18,3
2	23,7
3	25,5
4	20,7
5	24,2
...	...
...	...
...	...
43	21,3
44	25,2
45	23,8
46	19,2
47	21,6
48	22,4
49	22,0

OBJECTIF

Utiliser le logiciel Excel pour comparer une variable quantitative continue de deux populations à l'aide d'échantillons prélevés dans ces populations.

Mise en situation

Le Ministère de l'éducation veut savoir si la proportion de femmes de la clientèle qui s'inscrit à un programme universitaire a changé entre 2015 et 2020. On a prélevé un échantillon de taille 36 parmi les inscrits de 2015 et un échantillon de taille 49 parmi les inscrits de 2020.

Les données recueillies sont accessibles dans le fichier Excel «Hypo2ProportionsA». Effectuer un test bilatéral sur la différence des proportions avec un seuil de 1%.

CALCUL DES MESURES STATISTIQUES

1. Ouvrir le fichier Excel «Hypo2ProportionsA» et enregistrer sous un nom personnalisé.

2. En C11, définir le test logique :

«=SI(B11="F";1;0)».

Valider et incrémenter jusqu'en C46.

3. Dans la cellule C47, faire effectuer :

«=SOMME(C11:C46)».

4. Donner le nom «nA» à la cellule A46 et «nB» à la cellule D59.

5. En C48, faire calculer la proportion de femmes dans l'échantillon, soit :

«=C47/nA».

Donner le nom «PrA» à cette cellule.

6. Procéder de façon analogue pour calculer la proportion de femmes dans l'échantillon de 2015 et donner le nom PrB à la cellule affichant la proportion.

Remarque

La variable qualitative étudiée est le sexe lors de l'inscription au programme et les deux populations à comparer sont les inscrits au programme en 2015 et en 2020. C'est une variable binomiale, mais on peut utiliser une loi normale si la taille des échantillons est supérieure à 30.

Remarque

Les proportions réelles de femmes dans chacune des populations sont inconnues, il faut utiliser les proportions dans les échantillons pour effectuer le test.

Remarque

Après avoir calculé les proportions, on vérifie que les conditions pour utiliser une loi normale sont satisfaites. Ce qui est le cas ici puisque $n_A = 36 > 30$, $n_B = 49 > 30$, $n_A \times Pr_A = 15,99... \geq 5$, $n_B \times (1 - Pr_B) = 21,0... \geq 5$.

ÉNONCÉ DES HYPOTHÈSES

7. En A51, écrire «H0:» et en B51, «PrA = PrB».

8. En A52, écrire «H1:» et en B52, «PrA ≠ PrB».

RÈGLE DE DÉCISION

8. En A55, écrire «alpha =» et en B55 écrire «0,01». Donner le nom «alphaD» à cette cellule.

Remarque

Selon l'hypothèse nulle, les proportions des deux échantillons ont statistiquement la même valeur. Selon l'hypothèse alternative, elles sont statistiquement différentes.

9. En A56 écrire «zd =» et en M55, définir :
«=LOI.NORMALE.STANDARD.INVERSE(1- α /2)». Donner le nom «zd» à cette cellule.
10. En A57, écrire «Écart-ÉchA». En B57, définir :
«=(PrA*(1-PrA)/nA)^(1/2)».
Donner le nom «EDA» à cette cellule.
11. En A58, écrire «Écart-ÉchB». En B58, définir :
«=(PrB*(1-PrB)/nB)^(1/2)».
Donner le nom «EDB» à cette cellule.
12. En A59 écrire «Écart-Diff» et en B59 définir :
«=(EDA^2+EDB^2)^(1/2)».
Donner le nom «EDD» à cette cellule.
13. En A61, écrire «Marge=» et en B61, définir :
«=zp*EDD».
Donner le nom «MaD» à cette cellule.
14. En A63 et en A64, écrire les décisions possibles.

Remarque

L'écart-type de la distribution d'échantillonnage de la population A est estimée

Remarque

À l'étape 10, on calcule l'écart-type de la distribution d'échantillonnage de la proportion.

APPLICATION ET CONCLUSION

15. En A67 définir les tests logiques imbriqués
«=SI(OU(PrB-PrA<MaD;
PrB-PrA>MaD);"Rejeter H0";
"Conserver H0")».

Remarque

En écriture ordinaire, les tests logiques imbriqués s'écrivent

Si (PrB - PrA < MaD ou PrB - PrA > MaD) alors on rejette H_0 , sinon on conserve H_0 .