

6 Tests du Khi-deux d'indépendance et d'homogénéité

Exercice 6.1

Test du Khi-deux d'indépendance.

- **Contexte.** Deux variables qualitatives mesurées simultanément sur une population.

\mathcal{P} : candidats.

Variable X : « Résultat au test d'aptitude », qualitative à $l = 2$ modalités.

Variable Y : « Résultat au test psychologique », qualitative à $c = 2$ modalités.

- **Hypothèses :**

$$\begin{cases} H_0 : \text{les deux variables sont indépendantes} \\ H_1 : \text{les deux variables sont liées} \\ \alpha = 1\% \end{cases}$$

- **Observations :**

On dispose d'un échantillon de taille $n = 120$.

- **Effectifs observés** n_{ij} :

| X \ Y | introverti | extraverti | total ligne L_i |
|---------------------|------------|------------|-------------------|
| apte | 14 | 34 | 48 |
| inapte | 31 | 41 | 72 |
| total colonne C_j | 45 | 75 | $n = 120$ |

- **On calcule les effectifs théoriques** $\frac{L_i \times C_j}{n}$ **attendus sous** H_0 :

| X \ Y | introverti | extraverti |
|--------|---------------------------------|---------------------------------|
| apte | $\frac{48 \times 45}{120} = 18$ | $\frac{48 \times 75}{120} = 30$ |
| inapte | $\frac{72 \times 45}{120} = 27$ | $\frac{72 \times 75}{120} = 45$ |

- **Statistique du test :** $Q^2 = \sum \frac{\left(N_{ij} - \frac{L_i \times C_j}{n}\right)^2}{\frac{L_i \times C_j}{n}}$.

Valeur observée de Q^2 :

$$q_{\text{obs}}^2 = \frac{(14 - 18)^2}{18} + \frac{(34 - 30)^2}{30} + \frac{(31 - 27)^2}{27} + \frac{(41 - 45)^2}{45} = 2,37.$$

- **Loi de Q^2 sous H_0 :**

On vérifie les conditions d'application du test :

$n \geq 30$ et tous les effectifs théoriques sont supérieurs ou égaux à 5.

Donc $Q^2 \underset{\text{approx.}}{\sim} \chi^2$ à $(l - 1)(c - 1) = 1$ ddl.

- **IA et RC pour Q^2 au risque $\alpha = 1\%$:**

On rejette H_0 pour les valeurs de Q^2 trop grandes : RC à l'extrémité droite du domaine de Q^2 .

$IA = [0; 6,635]$ où 6,635 est le quantile d'ordre $1 - \alpha = 0,99$ de la loi χ^2 à 1 ddl.

- **Décision :**

$q_{\text{obs}}^2 = 2,37 \in IA$, donc on conserve H_0 avec un risque d'erreur β inconnu.

On ne peut pas affirmer, au risque $\alpha = 1\%$, qu'il existe un lien entre le type de personnalité et l'aptitude au pilotage.

Exercice 6.2

Test du Khi-deux d'indépendance.

« Couleur des cheveux » (3 modalités) et « Couleur des yeux » (2 modalités).

- **Echantillon** de taille $n = 528$.

- **Effectifs théoriques** :

| Couleur des cheveux \ Couleur des yeux | bleus | noirs | clairs | total ligne L_i |
|--|--------|-------|--------|-------------------|
| châtains ou noirs | 120,91 | 91,22 | 72,87 | 285 |
| blonds | 103,09 | 77,78 | 62,13 | 243 |
| total colonne C_j | 224 | 169 | 135 | $n = 528$ |

Les conditions d'application du test sont vérifiées.

- **Valeur observée** de la statistique du test : $q_{\text{obs}}^2 = 112,36$.

- **IA au risque** $\alpha = 5\%$: $IA = [0; 5,991]$ (loi χ^2 à 2 ddl) et $q_{\text{obs}}^2 \notin IA$.

Exercice 6.3

Test du Khi-deux d'indépendance.

« Sexe » (2 modalités) et « Développement de la maladie » (2 modalités).

- **Echantillon** de taille $n = 700$.

- **Effectifs théoriques** :

| Sexe \ Développement de la maladie | oui | non | Total ligne L_i |
|------------------------------------|-------|--------|-------------------|
| femme | 28,57 | 371,43 | 400 |
| homme | 21,43 | 278,57 | 300 |
| Total colonne C_j | 50 | 650 | $n = 700$ |

Les conditions d'application du test sont vérifiées.

- **Valeur observée** de la statistique du test : $q_{\text{obs}}^2 = 1,121$.

- **IA au risque** $\alpha = 1\%$: $IA = [0; 6,635]$ (loi χ^2 à 1 ddl) et $q_{\text{obs}}^2 \in IA$.

Exercice 6.4

Test du Khi-deux d'indépendance.

« Alimentation » (3 modalités) et « Cancer » (2 modalités).

- **Echantillon** de taille $n = 500$.

- **Effectifs théoriques** :

| Alimentation \ Cancer | cancer | pas de cancer | Total ligne L_i |
|-----------------------|--------|---------------|-------------------|
| rapide | 31,25 | 93,75 | 125 |
| végétarienne | 12,50 | 37,50 | 50 |
| traditionnelle | 81,25 | 243,75 | 325 |
| Total colonne C_j | 125 | 375 | $n = 500$ |

Les conditions d'application du test sont vérifiées.

- **Valeur observée** de la statistique du test : $q_{\text{obs}}^2 = 34,4$.

- **IA au risque** $\alpha = 1\%$: $IA = [0; 9,21]$ (loi χ^2 à 2 ddl) et $q_{\text{obs}}^2 \notin IA$.

Exercice 6.5

Test du Khi-deux d'indépendance.

« Trouble » (2 modalités) et « Rattrapage » (2 modalités), chez les enfants de dernière année.

- **Echantillon** de taille $n = 302$.
- **Effectifs théoriques** :

| X \ Y | cours | pas de cours | Total ligne L_i |
|------------------------|-------|--------------|-------------------|
| normaux | 28,37 | 180,63 | 209 |
| déficit de l'attention | 12,63 | 80,37 | 93 |
| Total colonne C_j | 41 | 261 | $n = 302$ |

Les conditions d'application du test sont vérifiées.

- **Valeur observée** de la statistique du test : $q_{\text{obs}}^2 = 5,373$
- **IA au risque** $\alpha = 5\%$: $IA = [0; 3,841]$, (loi χ^2 à 1ddl) et $q_{\text{obs}}^2 \notin IA$.

Exercice 6.6

« Produit consommé » (2 modalités) et « Crise cardiaque » (2 modalités).

- **Echantillon** de taille $n = 22071$.
- **Effectifs théoriques** :

| X \ Y | crise | pas de crise | Total ligne L_i |
|---------------------|--------|--------------|-------------------|
| aspirine | 146,52 | 10890,48 | 11037 |
| placebo | 146,48 | 10887,52 | 11034 |
| Total colonne C_j | 293 | 21778 | $n = 22071$ |

Les conditions d'application du test sont vérifiées.

- **Valeur observée** de la statistique du test : $q_{\text{obs}}^2 = 25,014$.
- **IA au risque** $\alpha = 5\%$: $IA = [0; 3,841]$ (loi χ^2 à 1ddl) et $q_{\text{obs}}^2 \notin IA$.

Exercice 6.7

Test du Khi-deux d'homogénéité.

Une variable qualitative mesurée sur plusieurs populations.

- **Contexte** :

$$c = 2 \text{ populations : } \begin{cases} \mathcal{P}_1 : \text{élèves du collège A} \\ \mathcal{P}_2 : \text{élèves du collège B.} \end{cases}$$

Variable X : « Participation à un club sportif », qualitative à $l = 2$ modalités « oui » et « non ».

- **Hypothèses** :

$$\begin{cases} H_0 : \text{les deux populations sont homogènes (même taux de participation)} \\ H_1 : \text{les deux populations ne sont pas homogènes (taux de participation différents)} \\ \alpha = 1\% \end{cases}$$

- **Observations** :

On dispose de 2 échantillons indépendants prélevés dans \mathcal{P}_1 et \mathcal{P}_2 , de tailles respectives $n_1 = 50$ et $n_2 = 60$.

Au total, on a tiré au sort $n = 110$ individus.

- **Effectifs observés** n_{ij} :

| Participation \ Echantillon | collège A | collège B | total ligne L_i |
|-----------------------------|-----------|-----------|-------------------|
| oui | 12 | 26 | 38 |
| non | 38 | 34 | 72 |
| total colonne C_j | 50 | 60 | $n = 110$ |

- **Effectifs théoriques** $\frac{L_i \times C_j}{n}$:

| Participation \ Echantillon | collège A | collège B |
|-----------------------------|-----------|-----------|
| oui | 17,27 | 20,73 |
| non | 32,73 | 39,27 |

- **Statistique du test** : $Q^2 = \sum \frac{\left(N_{ij} - \frac{L_i \times C_j}{n}\right)^2}{\frac{L_i \times C_j}{n}}$.

Valeur observée de Q^2 : $q_{\text{obs}}^2 = 4,504$.

- **Loi de Q^2 sous H_0** :

On vérifie les conditions d'application du test :

$n \geq 30$ et tous les effectifs théoriques sont supérieurs ou égaux à 5.

Donc $Q^2 \underset{\text{approx.}}{\sim} \chi^2$ à $(l-1)(c-1) = 1$ ddl.

- **IA et RC pour Q^2 au risque $\alpha = 1\%$** :

On rejette H_0 pour les valeurs de Q^2 trop grandes : RC à l'extrémité droite du domaine de Q^2 .

$IA = [0; 6,635]$ où 6,635 est le quantile d'ordre $1 - \alpha = 0,99$ de la loi du χ^2 à 1 ddl.

- **Décision** :

$q_{\text{obs}}^2 = 4,504 \in IA$, donc on conserve H_0 avec un risque d'erreur β inconnu. Au niveau 1%, les résultats ne sont pas significativement différents dans les deux collèges.

Exercice 6.8

Question 1. Test du Khi-deux d'homogénéité.

« Résultat au test des gnoses digitales » (2 modalités) dans 3 populations d'enfants : « Enfants de 8 ans », « Enfants de 10 ans », « Enfants de 12 ans ».

- **3 échantillons indépendants**, de même taille 75. Total $n = 225$.

- **Effectifs théoriques** :

| Résultat \ Echantillons | 8 ans | 10 ans | 12 ans | total ligne L_i |
|-------------------------|-------|--------|--------|-------------------|
| réussite | 40,33 | 40,33 | 40,33 | 121 |
| échec | 34,67 | 34,67 | 34,67 | 104 |
| total colonne C_j | 75 | 75 | 75 | 225 |

Les conditions d'application du test sont vérifiées.

- **Valeur observée** de la statistique du test : $q_{\text{obs}}^2 = 17,2$

- **IA au risque $\alpha = 1\%$** : $IA = [0; 9,21]$, (loi χ^2 à 2 ddl) et $q_{\text{obs}}^2 \notin IA$.

Question 2. Test sur une proportion.

p : taux de réussite chez les enfants de 8 ans.

- **Hypothèses** :
$$\begin{cases} H_0 : p = 0,5 \\ H_1 : p < 0,5 \text{ unilatérale gauche} \\ \alpha = 1\% \end{cases}$$

- **Observations** : $n = 75$; $f_{\text{obs}} = \frac{26}{75} = 0,347$ et $z_{\text{obs}} = \frac{0,347 - 0,5}{\sqrt{\frac{0,5 \times 0,5}{75}}} = -2,65$.

Les conditions d'application du test sont vérifiées : $n = 75 \geq 30$ et $75 \times 0,5 = 75 \times (1 - 0,5) \geq 5$.

- **IA pour F_n** : $IA = [0,366; 1]$ et $f_{\text{obs}} \notin IA$.

- **IA pour Z_n** : $IA = [-2,325; +\infty[$ et $z_{\text{obs}} \notin IA$.

- **p-valeur** $\alpha_{\text{obs}} = P_{H_0}(F_n \leq 0,347) = P_{H_0}(Z_n \leq -2,65) \simeq 0,4\%$.

Exercice 6.9

« Résultat au test Droite-Gauche » (2 modalités) dans 2 populations : « Enfants droitiers de 8 ans », « Enfants gauchers de 8 ans ».

- **2 échantillons indépendants**, de même taille 50. Total $n = 100$.

- **Effectifs théoriques** :

| Résultat \ Echantillons | droitiers | gauchers | total ligne L_i |
|-------------------------|-----------|----------|-------------------|
| réussite | 39 | 39 | 78 |
| échec | 11 | 11 | 22 |
| total colonne C_j | 50 | 50 | 100 |

Les conditions d'application du test sont vérifiées.

- **Valeur observée** de la statistique du test : $q_{\text{obs}}^2 = 8,392$.

- **IA au risque $\alpha = 1\%$** : $IA = [0; 6,635]$, (loi χ^2 à 1 ddl) et $q_{\text{obs}}^2 \notin IA$.