Chapitre 4

Tests d'hypothèse

Contenu

Définition 1. *(Hypothèse statistique)*

est une affirmation concernant les valeurs des paramètres d'une population

Hypothèse nulle H_0 : est l'hypothèse que l'on désire contrôler.

Hypothèse alternative H_1 : est la négation de H_0.

Définition 2. *(Test d'hypothèse)*

est une démarche qui a pour but de fournir une règle de décision permettant de faire un choix entre deux hypothèses statistiques sur la base de résultats d'échantillon.

Définition 3. *(Tests Khi-deux)*

Les tests de Khi-deux sont basés sur la statistique de χ^2 proposée par Karl Pearson. L'objectif de ces tests est principalement de comparer des distributions entre elles. Ces tests peuvent être appliqués à des variables qualitatives ou quantitatives.

Trois types de test de khi-deux : Test de conformité, Test d'homogénéité, Test d'indépendance.

4.1 Test de conformité

L'objectif :

comparer une distribution observée sur un échantillon à une distribution théorique.

Notations :

T_i : les effectifs théoriques

O_i : les effectifs observés

n : l'effectif total.

k : le nombre de modalité.

Les étapes du test :

❶ Formulation des hypothèses

H_0 : la distribution observée est conforme à la distribution théorique.

H_1 : la distribution observée est différente de la distribution théorique.

❷ Calcul des effectifs théoriques

condition de validité du test : tous les $T_i \geq 5$.

❸ La statistique de test :

$$\chi_c^2 = \sum_{i=1}^{k} \frac{(O_i - T_i)^2}{T_i}$$

❹ Le seuil critique :se lit dans la table (loi de Khi-deux) avec $ddl = k - 1$

$$\chi_T^2 = \chi^2_{(ddl, 1-\alpha)}$$

❺ Décision

Si $\chi_c^2 \leq \chi_T^2$ On accepte H_0

Si $\chi_c^2 > \chi_T^2$ On rejette H_0

Example 4.1.

Dans une maternité sur 100 *naissances on a observé* 44 *garçons et* 56 *filles. cette observation est-elle compatible avec la statistique nationale donnant les proportions de naissances de garçons et de filles respectivement* 53% *et* 47%.

Solution

Choix du test : test khi-deux de conformité

❶ *Formulation des hypothèses*

H_0 *: la distribution observée est conforme à la distribution nationale.*

H_1 *: la distribution observée est différente de la distribution nationale.*

❷ *Calcul des effectifs théoriques*

sexe	effectifs observés O_i	statistique nationale	effectifs théoriques T_i
garçons	44	0.53	53
filles	56	0.47	47
total	100	1	100

tous les $T_i \geq 0$ *donc le test est valide.*

❸ *La statistique du test :*

$$\chi_c^2 = \sum_{i=1}^{k} \frac{(O_i - T_i)^2}{T_i} = \frac{(44-53)^2}{53} + \frac{(56-47)^2}{47} = 3.25$$

❹ *Le seuil critique :*

$$ona \begin{cases} ddl = k-1 = 2-1 = 1 \\ 1-\alpha = 1-0.05 = 0.95 \end{cases} \quad donc \quad \chi_T^2 = \chi^2_{(1;0.95)} = 3.84$$

❺ *Décision : On a* $\chi_c^2 < \chi_T^2$

Décision statistique : On accepte H_0

Décision pratique : la distribution observée est conforme à la distribution nationale.

4.2 Test d'homogénéité

L'objectif :

comparer deux ou plusieurs distributions observées sur des échantillons.

Notations :

l : nombre de lignes

c : nombre de colonnes

n : l'effectif total.

Les étapes du test :

❶ Formulation des hypothèses

H_0 : Les distributions observées sont équivalentes.

H_1 : Les distributions observées ne sont pas équivalentes.

❷ Calcul des effectifs théoriques

$$n_i = \sum_{j=1}^{c} O_{ij} \qquad n_j = \sum_{i=1}^{l} O_{ij} \qquad T_{ij} = \frac{n_i n_j}{n}$$

Condition de validité du test : tous les $T_{ij} \geq 5$.

❸ La statistique du test :

$$\chi_c^2 = \sum_{j=1}^{c} \sum_{i=1}^{l} \frac{(O_{ij} - T_{ij})^2}{T_{ij}}$$

❹ Le seuil critique :se lit dans la table (loi de Khi-deux) avec

$ddl = (c-1)(l-1)$

$$\chi_T^2 = \chi^2_{(ddl, 1-\alpha)}$$

❺ Décision

Si $\chi_c^2 \leq \chi_T^2$ On accepte H_0

Si $\chi_c^2 > \chi_T^2$ On rejette H_0

Example 4.2.
Deux médicaments A et B ont été testés sur deux groupes de malades, on a observé les résultats suivants

	disparition des symptômes	persistance des symptômes	aggravation	réaction secondaire
A	100	40	20	30
B	220	80	70	40

Peut-on dire que ces deux traitements ont les mêmes effets ? on prendra $\alpha = 5\%$

Solution

Choix du test : test khi-deux d'homogénéité

❶ *Formulation des hypothèses*

H_0 *: les deux traitements sont équivalentes.*

H_1 *: les deux traitements ne sont pas équivalentes.*

❷ *Calcul des effectifs théoriques : tableau de contingence*

	disparition des symptômes		persistance des symptômes		aggravation		réaction secondaire		total n_i
	O_{ij}	T_{ij}	O_{ij}	T_{ij}	O_{ij}	T_{ij}	O_{ij}	T_{ij}	
A	100	101.33	40	38	20	28.5	30	22.16	190
B	220	218.66	80	82	70	61.5	40	47.83	410
total n_j	320		120		90		70		n=600

tous les $T_{ij} \geq 5$ *donc le test est valide.*

❸ *La statistique du test :*

$$\chi_c^2 = \sum_{i=1}^{l}\sum_{j=1}^{c}\frac{(O_{ij}-T_{ij})^2}{T_{ij}} = \sum_{j=1}^{4}\sum_{i=1}^{2}\frac{(O_{ij}-T_{ij})^2}{T_{ij}} = 7.94$$

❹ *Le seuil critique :*

$$ona \begin{cases} ddl = (c-1)(l-1) = (4-1)(2-1) = 3 \\ 1-\alpha = 1-0.05 = 0.95 \end{cases} \quad donc \quad \chi_T^2 = \chi^2_{(3;0.95)} = 7.815$$

❺ *Décision : On a* $\chi_c^2 > \chi_T^2$

Décision statistique : On rejette H_0

Décision pratique : les deux traitements ne sont pas équivalentes.

4.3 Test d'indépendance

L'objectif : étudier sur un même échantillon la liaison entre deux variables

Example 4.3.

Un épisode d'intoxication alimentaire collective est survenu parmi les éléves d'un école primaire, un docteur fut chargé de l'enquête et il a dressé le tableau suivant

	malades	*sains*
glace au chocolat	*69*	*83*
pas de glace	*31*	*17*

L'intoxication alimentaire est-elle liée à la consommation de glace au chocolat ?

Solution

Choix du test : test khi-deux d'indépendance

❶ Formulation des hypothèses

H_0 : il n' ya pas de liaison entre l'intoxication alimentaire et la consommation de glace au chocolat.

H_1 : il existe une liaison entre l'intoxication alimentaire et la consommation de glace au chocolat.

❷ Calcul des effectifs théoriques : tableau de contingence

	malades		sains		total n_i
	O_{ij}	T_{ij}	O_{ij}	T_{ij}	
glace au chocolat	69	76	83	76	152
pas de glace	31	24	17	24	48
total n_j	100		100		n=200

tous les $T_{ij} \geq 5$ donc le test est valide.

❸ La statistique de test :

$$\chi_c^2 = \sum_{i=1}^{l}\sum_{j=1}^{c}\frac{(O_{ij}-T_{ij})^2}{T_{ij}} = \sum_{j=1}^{2}\sum_{i=1}^{2}\frac{(O_{ij}-T_{ij})^2}{T_{ij}} = 5.4$$

❹ Le seuil critique :

$$ona \begin{cases} ddl = (c-1)(l-1) = (2-1)(2-1) = 1 \\ 1-\alpha = 1-0.05 = 0.95 \end{cases} \qquad donc \quad \chi_T^2 = \chi^2_{(1;0.95)} = 3.841$$

❺ Décision : On a $\chi_c^2 > \chi_T^2$

Décision statistique : On rejette H_0

Décision pratique : il existe une liaison entre l'intoxication alimentaire et la consommation de glace au chocolat.