

Test d'ajustement Test d'indépendance

Dr. Sacha Varone

Objectif

Savoir effectuer

- un test d'ajustement
- un test d'indépendance

Rappels

Test d'ajustement

Test
d'indépendance

h e g



Rappels

Test d'ajustement

Test
d'indépendance

Rappels

Test de la médiane

Si la médiane hypothétique de la population est trop éloignée de la médiane de l'échantillon, alors H_0 rejeté.

-> Test des rangs signés de Wilcoxon

La statistique de test W suit la loi tabulée pour le test des rangs signés de Wilcoxon.

1. Calculer les différences d_i entre chaque valeur et la médiane postulée $\tilde{\mu}$
2. Calculer la valeur absolue des différences précédentes : $|d_i|$
3. Déterminer le rang pour chacune des valeurs $|d_i|$, en ne tenant pas compte des valeurs nulles.
Si des observations ont la même valeur $|d_i|$, alors affecter le rang moyen de ces observations.
4. Calculer la statistique W qui est la somme des rangs dont les d_i sont positifs.

Rejet de H_0 si $W > W_\alpha =$ valeur critique.

Rappels

Test d'ajustement

Test
d'indépendance

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

Test d'ajustement

Utilité

Supposition nécessaire pour certains tests
population suit une loi spécifique (ex loi normale).

Vérification ?

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

Supposition nécessaire pour certains tests
population suit une loi spécifique (ex loi normale).

Vérification ? Test d'ajustement du χ^2

- Adéquation entre distribution hypothétique et empirique -

Principe

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

1. Acquisition d'un échantillon de taille suffisamment grande.
2. Classement en k différentes catégories des données.
3. Calcul des fréquences absolues observées.
4. Comparaison des fréquences absolues théoriques e_i et fréquences observées.

Rejet de H_0 = distribution théorique
si une trop grande différence existe.

RappelsTest d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

La statistique à utiliser suit une loi du χ^2 à $k - 1$ degrés de liberté et est calculée ainsi :

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$$

où

o_i = fréquence observée pour la catégorie i

e_i = fréquence théorique pour la catégorie i

k = nombre de catégories

Supposition suivante : LA TAILLE DE L'ÉCHANTILLON EST SUFFISAMMENT GRANDE

Exemple

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

Un nouveau directeur d'un centre d'appels pour FAI constate que le personnel est réduit de 20% les mercredi, jeudi et dimanche. Son prédécesseur avait procédé ainsi car le nombre d'appels était 20% moins élevé ces jours-là. Afin de savoir si tel est toujours le cas, il fait relever le nombre d'appels sur 1 mois pour chaque jour de la semaine et obtient les données agrégées suivantes :

Jours	Lu	Ma	Me	Je	Ve	Sa	Di
Nombre d'appels	1000	1200	900	1000	1200	1100	800

Il souhaite savoir s'il y a effectivement une baisse de 20% les mercredi, jeudi et dimanche, avec un niveau de signification de 0.05.

Exemple (suite)

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

1. Hypothèses nulle et alternative

Exemple (suite)

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

1. Hypothèses nulle et alternative

H_0 La distribution des appels est identique
les lu, ma, ve et sa,
et 20% moins élevée les me, je et di.

H_1 La distribution des appels n'est pas celle décrite en H_0

2. Le niveau de signification

Exemple (suite)

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

1. Hypothèses nulle et alternative

H_0 La distribution des appels est identique
les lu, ma, ve et sa,
et 20% moins élevée les me, je et di.

H_1 La distribution des appels n'est pas celle décrite en H_0

2. Le niveau de signification 0.05

3. La valeur critique

Exemple (suite)

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

1. Hypothèses nulle et alternative

H_0 La distribution des appels est identique
les lu, ma, ve et sa,
et 20% moins élevée les me, je et di.

H_1 La distribution des appels n'est pas celle décrite en H_0

2. Le niveau de signification 0.05

3. La valeur critique est celle d'une distribution du χ^2 à 6 degrés de libertés.

$$\chi_{0.05}^2 = 12.5916$$

Exemple (fin)

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

4. Le nombre total d'appels sur la période observée est 7200

Jours	Lu	Ma	Me	Je	Ve	Sa	Di
o_i	1000	1200	900	1000	1200	1100	800
e_i	1125	1125	900	900	1125	1125	900

La statistique est alors

$$\chi^2 = 46.6$$

5.

Exemple (fin)

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

4. Le nombre total d'appels sur la période observée est 7200

Jours	Lu	Ma	Me	Je	Ve	Sa	Di
o_i	1000	1200	900	1000	1200	1100	800
e_i	1125	1125	900	900	1125	1125	900

La statistique est alors

$$\chi^2 = 46.6$$

5. Comme $46.6 > 12.5916$, l'hypothèse H_0 est rejetée.
6. La conclusion est

Exemple (fin)

Rappels

Test d'ajustement

Utilité

Principe

Exemple

Test
d'indépendance

4. Le nombre total d'appels sur la période observée est 7200

Jours	Lu	Ma	Me	Je	Ve	Sa	Di
o_i	1000	1200	900	1000	1200	1100	800
e_i	1125	1125	900	900	1125	1125	900

La statistique est alors

$$\chi^2 = 46.6$$

5. Comme $46.6 > 12.5916$, l'hypothèse H_0 est rejetée.
6. La conclusion est que la distribution du nombre d'appels n'est plus telle qu'indiquée par le précédent directeur.

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

Test d'indépendance

Rappel

Table de contingence = tableau croisé = crosstable
Principe : 2 variables qualitatives, données groupées

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

Principe du test d'indépendance

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

But : Tester l'indépendance de deux variables de type catégoriel
Moyen : Table de contingence.

Principe du test

Comparaison des fréquences observées avec les fréquences théoriques e_{ij} en cas d'indépendance.

Si une trop grande différence existe, alors l'hypothèse d'indépendance des variables est rejetée.

$$e_{ij} = \frac{(\text{Total ligne } i)(\text{Total colonne } j)}{\text{Taille de l'échantillon}}$$

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

où

 o_{ij} = fréquence observée de la cellule (i, j) e_{ij} = fréquence théorique de la cellule (i, j) r = nombre de lignes c = nombre de colonnes

Supposition : LA TAILLE DE L'ÉCHANTILLON EST
SUFFISAMMENT GRANDE

En pratique, effectif observé par cellule ≥ 5

Exemple

Une recherche est effectuée afin de savoir si le nombre de sorties le week-end est indépendant des résultats aux examens, avec un niveau de signification de 0.05. La table de contingence suivante résume les données récoltées :

	Résultat				<i>Total</i>
	Insuffisant	Acquis	Bien	Excellent	
Sortie WE					
Jamais	84	50	50	16	200
Occasionnel	82	64	34	20	200
Fréquent	34	36	16	14	100
<i>Total</i>	200	150	100	50	500

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

Exemple (suite)

1. Hypothèses nulle et alternative

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

Exemple (suite)

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

1. Hypothèses nulle et alternative

H_0 Sorties du week-end indépendantes
des résultats aux examens.

H_1 Sorties du week-end ne sont PAS indépendantes
des résultats aux examens

2. Niveau de signification

Exemple (suite)

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

1. Hypothèses nulle et alternative

H_0 Sorties du week-end indépendantes
des résultats aux examens.

H_1 Sorties du week-end ne sont PAS indépendantes
des résultats aux examens

2. Niveau de signification 0.05

3. Valeur critique est

Exemple (suite)

Rappels

Test d'ajustement

Test
d'indépendance

Table contigence

Principe

Statistique

Exemple

1. Hypothèses nulle et alternative

H_0 Sorties du week-end indépendantes
des résultats aux examens.

H_1 Sorties du week-end ne sont PAS indépendantes
des résultats aux examens

2. Niveau de signification 0.05

3. Valeur critique est celle d'une distribution du χ^2 à
 $(3 - 1)(4 - 1) = 6$ degrés de libertés.

$$\chi_{0.05}^2 = 12.5916$$

Exemple (suite)

4. Les fréquences théoriques de la table de contingence sont

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

Exemple (suite)

4. Les fréquences théoriques de la table de contingence sont

	Résultat				<i>Total</i>
	Insuffisant	Acquis	Bien	Excellent	
Sortie WE					
Jamais	80	60	40	20	200
Occasionnel	80	60	40	20	200
Fréquent	40	30	20	10	100
<i>Total</i>	200	150	100	50	500

La statistique est alors

$$\chi^2 = 10.88$$

Exemple (fin)

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

5. Comme $10.88 < 12.5916$, l'hypothèse H_0 ne peut pas être rejetée.
6. La conclusion est

Exemple (fin)

Rappels

Test d'ajustement

Test
d'indépendance

Table contingence

Principe

Statistique

Exemple

5. Comme $10.88 < 12.5916$, l'hypothèse H_0 ne peut pas être rejetée.
6. La conclusion est qu'il n'y a pas suffisamment d'évidence pour conclure que les sorties du week-end et le résultat aux examens ne sont pas indépendants.

... et ceci n'est aucunement une incitation à faire la fête le week-end !