

Faculté des sciences économique, commerciales, et sciences de gestion

Master 2 économie et gestion d'entreprise

Cours 05: Analyse bi-variée Tableau de contingence et le principe du test Khi 2 (Chi square)

DR BOUYACOU BRAHIM

**MAITRE DE CONFÉRENCES / DOCTEUR EN ÉCONOMIE
MONÉTAIRE ET FINANCIÈRE**

Analyse bi-variée

- L'objectif de cette partie est d'étudier deux variables différents X et Y et de rechercher s'il existe un lien entre ces deux variables.
- Chacune des deux variables peut être, soit quantitative, soit qualitative:
 - 1) Deux variables qualitatives (tableau de contingence).
 - 2) Une variable quantitative et une variable qualitative (Test ANOVA).
 - 3) Deux variables quantitatives (régression linéaire).

Deux variables qualitatives (tableau de contingence)

- ❑ La variable qualitative nominale est une information non mesurable.
- ❑ Elle présente des catégories que l'on nomme avec un nom codé qu'on appelle modalité (niveau d'étude, genre, salaire).
- ❑ La variable peut avoir plusieurs modalités codées. Par exemple : la variable couleur avec plusieurs modalités (1-rouge, 2-vert, 3-bleu, 4-...), variable âge plusieurs modalités, et variable logement avec plusieurs modalités (1- logement individuel, 2- logement semi collectif, 3- logement collectif).

Tableau de contingence (tableau croisé)

- ❑ tableau croisé (appelé aussi tableau de contingence) examine la relation entre deux variables catégorielles.
- ❑ C'est un arrangement dans lequel les données sont classées selon deux variables catégorielles. Les catégories d'une variable apparaissent dans les lignes et les catégories de l'autre variable apparaissent dans les colonnes.

Hypothèses (H_0) et (H_1)

L'hypothèse nulle (H_0) et l'hypothèse alternative (H_1) du test Chi-2 peuvent être exprimées de la manière suivante :

- 1) H_0 : « (Variable 1) n'est pas associée à (Variable 2) » (Absence de relation entre les 2 variables catégorielles)
- 2) H_1 : «(Variable 1) est associée à (Variable 2) » (Existence de relation entre les 2 variables).

Le principe du test Khi 2

- analyse bi-variée consiste à déterminer s'il existe une association entre deux variables qualitatives nominales.
- Significativité globale de la relation : d.d.l. = $(l - 1) * (c - 1)$
- Si $\chi^2 > 5\%$ la différence n'est pas significative au seuil de 5%
- Si $\chi^2 < \text{ou} = 5\%$ la différence est significative au seuil de 5%

Le principe du test Khi 2

- Le test du khi-deux consiste à tester la signification statistique d'une association de deux variables qualitatives (nominales ou ordinales). Plus précisément, il a pour objet de tester l'indépendance des variables dans un tableau croisé.

Le principe du test Khi 2

Le test de Chi-deux est utilisé pour tester l'hypothèse nulle d'absence de relation entre deux variables catégorielles. On peut également dire que ce test vérifie l'hypothèse d'indépendance de ces variables.

H0 : Il n'existe pas de relation entre les deux variables.

H1 : Il existe une relation entre les deux variables.

Le principe du test Khi 2

- 1) Si la valeur du khi deux de Person est supérieure à 0,05, on accepte H0.
- 2) Si la valeur du khi deux de Person est inférieure à 0,05, on rejette H0 et on accepte H1.

Exemple :

On s'intéresse à une éventuelle relation entre la variable $X = (A, B)$ de $n = 200$ et la variable $Y = (C, D, E)$.

Tableau 1. Effectifs observés n_{ij}

X/Y	C	D	E	total
A	$n_{11} = 10$	$n_{12} = 50$	$n_{13} = 20$	$n_{1.} = 80$
B	$n_{21} = 20$	$n_{22} = 60$	$n_{23} = 40$	$n_{2.} = 120$
total	$N_{.1} = 30$	$n_{.2} = 110$	$n_{.3} = 60$	$n = 200$

Les nombres $n_{1.}$; $n_{2.}$ Et $n_{.1}$; $n_{.2}$; $n_{.3}$ sont appelés effectifs marginaux.

Tableau 1. Fréquence observés n_{ij}

X/Y	C	D	E	total
A	$f_{11} = 0,05$	$f_{12} = 0,25$	$f_{13} = 0,10$	$f_{1.} = 0,40$
B	$f_{21} = 0,10$	$f_{22} = 0,30$	$f_{23} = 0,20$	$f_{2.} = 0,60$
total	$f_{.1} = 0,15$	$f_{.2} = 0,55$	$f_{.3} = 0,30$	1

Les nombres $f_{1.}$; $f_{2.}$ et $f_{.1}$; $f_{.2}$; $f_{.3}$ sont appelées fréquences marginales.

Tableau des effectifs théoriques

A partir des effectifs marginaux L_i et C_j , on peut calculer les effectifs attendus lorsque X et Y sont indépendants.

Ces effectifs théoriques, notés e_{ij} , sont donnés par la formule

$$e_{ij} = \frac{L_i C_j}{n}.$$

Tableau 2. Effectifs théoriques e_{ij}

X/Y	C	D	E	total
A	$n_{11} = 80 \cdot 30 / 200$	$n_{12} = 80 \cdot 110 / 200$	$n_{13} = 80 \cdot 60 / 200$	$n_{1.} = 80$
B	$n_{21} = 120 \cdot 30 / 200$	$n_{22} = 120 \cdot 110 / 200$	$n_{23} = 120 \cdot 60 / 200$	$n_{2.} = 120$
total	$N_{.1} = 30$	$n_{.2} = 110$	$n_{.3} = 60$	$n = 200$

Tests du khi-carré

	Valeur	ddl	Signification asymptotique (bilatérale)	Sig. exacte (bilatérale)	Sig. exacte (unilatérale)
khi-carré de Pearson	2,039 ^a	1	,153		
Correction pour continuité ^b	1,310	1	,252		
Rapport de vraisemblance	2,054	1	,152		
Test exact de Fisher				,171	,126
Association linéaire par linéaire	1,998	1	,158		
N d'observations valides	50				

a. 0 cellules (0,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 11,52.

b. Calculée uniquement pour une table 2x2



Signification du khi-deux de Person est SUPERIEUR à 0,05 (Seuil de signification).

On rejette H1 (dépendance) et on accepte H0 (Indépendance)

Interprétation : il n'existe pas d'une dépendance entre la variable du genre et celle de l'employabilité.

Exercice:

Pour cibler la clientèle d'un nouveau produit de consommation, une entreprise fait un sondage auprès de 321 personnes. L'intérêt dans le produit est noté par "aucun intérêt", "un intérêt mineur" ou un "intérêt important". La situation familiale (au moins un enfant à charge : oui ou non) est notée également. On cherche à vérifier si l'intérêt dans le produit dépend de la situation familiale. Les résultats sont les suivants :



<i>Enfant</i>	<i>aucun</i>	<i>mineur</i>	<i>important</i>
<i>oui</i>	10	12	3
<i>non</i>	7	38	9

On a donc 79 personnes qui répondent. On veut vérifier s'il y a un lien entre les deux mesures au niveau 5%.

Questions :

1. Déterminer les hypothèses du modèle bi-varie (tableau de contingence).
2. Calculer les effectifs marginaux, les fréquences observées n_{ij} et les fréquences marginaux.
3. calcule les effectifs théoriques.

Après utilisation du logiciel SPSS, interpréter les résultats suivants :

Tests du khi-carré

	Valeur	ddl	Signification asymptotique (bilatérale)
khi-carré de Pearson	6,162 ^a	2	,046
Rapport de vraisemblance	5,839	2	,054
Association linéaire par linéaire	4,020	1	,045
N d'observations valides	79		

a. 1 cellules (16,7%) ont un effectif théorique inférieur à 5.
L'effectif théorique minimum est de 3,80.

Faculté des sciences économique, commerciales, et sciences de gestion

Master 2 économie et gestion d'entreprise

Cours 05: Analyse bi-variée Tableau de contingence et le principe du test Khi 2 (Chi square)

DR BOUYACOU BRAHIM

MAITRE DE CONFÉRENCES / DOCTEUR EN ÉCONOMIE
MONÉTAIRE ET FINANCIÈRE