TD6 : test du χ^2 (conformité en loi)

Exercice 1 (Exo 11 du TD 1)**.** Trouver la valeur de z telle que

1. $P(\chi^2(5) \leq z) = 0.9$, $z = 9.236$.
2. $P(\chi^2(13) \leq z) = 0.95$, $z = 22.362$.
3. $P(\chi^2(20)) = 0.99$, $z = 37.566$.
4. $P(\chi^2(11) \leq z) = 0.9$, $z = 17.275$.
5. $P(\chi^2(3) \leq z) = 0.95$, $z = 7.815$.
6. $P(\chi^2(2) \leq z) = 0.99$, $z = 9.210$.

1 Exercices

Exercice 2. Les ventes automobiles de l'année 2000 était

Type de moteur	Essence	Diesel	Électrique ou Hybride
Pourcentage des ventes	60%	30%	10%

.

Un concessionnaire regarde ces ventes sur le dernier semestre de l'année 2022 et constate la répartition suivante sur les 150 véhicules vendus :

Type de moteur	Essence	Diesel	Électrique ou Hybride
Volume des ventes	72	37	41

.

1. Au risque 10%, la répartition des ventes a-t-elle significativement changée ?

Correction. La loi initial $\mathcal{L}_0$ est celle de 2000 représentée par le tableau

M	E	D	H
% des ventes	60%	30%	10%

.

On a la loi des données inconnue $\mathcal{L}$.
On teste
— l'hypothèse $H_0 : \mathcal{L} = \mathcal{L}_0$, le marché n'a pas changé *contre*
— l'hypothèse $H_1 : \mathcal{L} \neq \mathcal{L}_0$, le marché a changé.
Pour le faire, on construit le tableau

M	E	D	H	Total
% des ventes	60%	30%	10%	100%
Ventes actuelles	72	37	41	150
Ventes théoriques	$150 \times 0.6 = 90$	45	15	150
χ^2	$\frac{(90-72)^2}{90} = 3.6$	$\frac{(45-37)^2}{45} = 1.4$	$\frac{(15-41)^2}{15} = 45.1$	

Chaque élément de la ligne χ^2 est $\frac{(n_{\text{théo}} - n_{\text{vrai}})^2}{n_{\text{théo}}}$.
Ne pas oublier de vérifier que le total > 30 et chaque vente théorique ≥ 5.

La statistique Z que l'on obtient est $Z = 3.6 + 1.4 + 45.1 = 50.1$. Cette statistique suit une loi du χ^2 à $2 = 3 - 1$ degrés de libertés. Le 3, c'est le nombre de colonnes/choix possibles.

La zone de rejet est toujours de la forme $[z, \infty[$. Pour trouver z,
— ligne 2 car 2 degrés de liberté et
— le risque étant de 10%, colonne 0.1.
Donc $z = 4.605$. Comme $50.1 \in [4.605; +\infty[$, on rejette H_0 et on en conclut que le marché a changé.

Exercice 3 (Rattrapage 2021-2022)**.** En vu de la période de solde, un magasin de vêtements étudie son stock pour savoir s'il va mettre la même démarque sur toutes les tailles. Il sait qu'en moyenne il vend 22% de vêtements en taille S, 46% en taille M, 26% en taille L et 6% en taille XL. Il mettra la même démarque sur toutes les tailles si son stock correspond à la moyenne de ses ventes habituelles. L'inventaire du stock donne

Taille	S	M	L	XL
Stock	141	187	141	61

.

1. Au risque 5%, le magasin va-t-il faire des démarques différenciées en fonction de la taille ?

Correction. On teste

— l'hypothèse H_0 : le stock correspond aux ventes, pas de démarques différenciées *contre*

— l'hypothèse H_1 : le stock ne correspond pas aux ventes, on fait des démarques différenciés.

Pour le faire, on construit le tableau

Taille	S	M	L	XL	Total
% des ventes	22%	46%	26%	6%	100%
Stocks actuels	141	187	141	61	530
Stocks théo	116.6	243.8	137.8	31.8	530
χ^2	$\frac{(116.6-141)^2}{116.6} = 5.1$	$\frac{(243.8-187)^2}{243.8} = 13.2$	$\frac{(137.8-141)^2}{137.8} = 0.07$	$\frac{(31.8-61)^2}{31.8} = 26.8$	

La statistique Z vaut $Z = 5.1 + 13.2 + 0.07 + 26.8 = 45.17$. Cette statistique suit une loi du χ^2 à 3 degrés de libertés.

La zone de rejet est de la forme $[z, \infty[$. Pour trouver z,

— ligne 3

— le risque étant de 5%, colonne 0.05.

Donc $z = 7.815$. Comme $45.17 > 7.815$, on rejette H_0 et on fait des démarques différenciées.

2 Exercices d'entrainement

Exercice 4. Une parfumerie qui produit trois parfums – Eaux du Nil (50% des ventes), Fleur de Lune (30% des ventes), et Mina (20% des ventes) – décide de refaire le design des flacons. Un an plus tard, afin d'en évaluer les effets, elle prélève les données suivantes sur la vente de 300 flacons de parfums :

Parfum	Eaux du Nil	Fleur de Lune	Mina
Ventes	129	90	81

.

Au risque 1%, les nouveaux designs ont-ils changés la répartition des ventes ?

Correction. On teste

— l'hypothèse H_0 : le nouveau design n'a pas changé la répartition des ventes *contre*

— l'hypothèse H_1 : le nouveau design a eu un effet sur la répartition des ventes.

Pour le faire, on construit le tableau

Parfum	Eaux du Nil	Fleur de Lune	Mina	Total
% des ventes	50%	30%	20%	100%
Ventes actuelles	129	90	81	300
Ventes théo	150	90	60	300
χ^2	$\frac{(150-129)^2}{150} = 2.94$	$\frac{(90-90)^2}{90} = 0$	$\frac{(60-81)^2}{60} = 7.35$	

La statistique Z vaut $Z = 2.94 + 0 + 7.35 = 10.29$. Cette statistique suit une loi du χ^2 à 2 degrés de libertés.

La zone de rejet est de la forme $[z, \infty[$. Pour trouver z,

— ligne 2

— le risque étant de 1%, colonne 0.01.

Donc $z = 9.210$. Comme $10.29 > 9.210$, on rejette H_0, les nouveaux design ont bien changé la répartition des ventes.

Exercice 5. Pepsi se lance dans une grande campagne publicitaire. Avant cette campagne, la répartition des ventes de soda était de 60% pour Coca-Cola, 20% pour Pepsi et 20% pour les autres marques de cola. Pendant la campagne, dans un supermarché, sur 234 bouteilles vendus : 130 sont de Coca-Cola, 59 sont de Pepsi et 45 d'autres marques.

Au risque 10%, la campagne publicitaire a-t-elle un impact ?

Correction. On teste

— l'hypothèse H_0 : la campagne n'a pas d'impact *contre*

— l'hypothèse H_1 : la campagne a un impact.

Pour le faire, on construit le tableau

Marques	Caco-Cola	Pepsi	Autres	Total
% des ventes	60%	20%	20%	100%
Ventes actuelles	130	59	45	234
Ventes théo	140.4	46.8	46.8	234
χ^2	$\frac{(140.4-130)^2}{140.4} = 0.77$	$\frac{(46.8-59)^2}{46.8} = 3.18$	$\frac{(46.8-45)^2}{46.8} = 0,07$	

La statistique Z vaut $Z = 0.77 + 3.18 + 0.07 = 4.02$. Cette statistique suit une loi du χ^2 à 2 degrés de libertés.

La zone de rejet est de la forme $[z, \infty[$. Pour trouver z,

— ligne 2

— le risque étant de 10%, colonne 0.1.

Donc $z = 4.605$. Comme $4.02 < 4.605$, on ne rejette pas H_0, la campagne publicitaire n'a pas eu d'impact significatif.

Exercice 6. Une entreprise mène une enquête sur ses 5 sites français : Lyon, Marseille, Lille, Bordeaux et Rennes. Sur les 3000 salariés auxquels le questionnaire a été envoyé, 800 ont répondu. Les effectifs et le nombre de gens qui ont répondus en fonction du site sont représentés dans le tableau suivant :

Site	Lyon	Marseille	Lille	Bordeaux	Rennes
Effectifs	1500	500	400	400	200
Réponses	400	110	105	125	60

.

Au risque 1%, l'échantillon des réponses reçues est-il représentatif de la répartition des effectifs ? [Indication : calculer le pourcentage des effectifs en premier.]

Correction. On teste

- l'hypothèse H_0 : l'échantillon des réponses réçues est représentatif de la répartition des effectifs *contre*
- l'hypothèse H_1 : il ne l'est pas.

Pour le faire, on construit le tableau

Site	Lyon	Marseille	Lille	Bordeaux	Rennes	Total
% des effectifs	50%	16.67%	13.33%	13.33%	6.67%	100%
Réponses	400	110	105	125	60	800
Réponses théo	400	133.36	106.64	106.64	53.36	800
χ^2	0	4.09	0.03	3.16	0.83	

La statistique Z vaut $Z = 0 + 4.09 + 0.03 + 3.16 + 0.83 = 8.11$. Cette statistique suit une loi du χ^2 à 4 degrés de libertés.

La zone de rejet est de la forme $[z, \infty[$. Pour trouver z,

- ligne 4
- le risque étant de 1%, colonne 0.01.

Donc $z = 13.277$. Comme $8.11 < 13.277$, on ne rejette pas H_0, l'échantillon de réponses réçues est représentatif de la répartition des effectifs.