Licence 3 - SDV 2008-2009

UE Statistiques de base

TD5 – Corrigé

Exercice 1 :

Hypothèse H_0 : l'échantillon est extrait de la population parente (les deux distributions de type de connexion ne différent pas significativement).

	NET			Total
	NET1	NET2	NET3	
Population parente	2248787	2893279	5033316	10175382
% théorique	0,22	0,28	0,50	1
Effectifs observés (o_i)	48	37	80	165
% observés	0,29	0,22	0,49	1
Effectifs théoriques (c_i)	36,47	46,92	81,62	165
$\chi^2 = \frac{(o_i - c_i)^2}{c_i}$	3,65	2,10	0,03	5,78

Tous les c_i étant plus grand que 5, la condition d'application sur les effectifs est respectée et le test réalisé sur ces bases.

$\chi^2 = 5{,}78$; ddl=k-1=2 ; $\chi^2_{0,05} = 5{,}991$; $\chi^2 < \chi^2_{0,05}$

On ne peut pas rejeter H_0 au seuil de 0,05 (on pourrait par contre la rejeter au seuil de 0,10).
L'échantillon est représentatif de la population parente quant à la distribution du type de connexion Internet des foyers.

Exercice 2 :

Hypothèse H_0 : l'échantillon est extrait de la population parente (les deux distributions de niveaux d'études ne différent pas significativement). On ne peut traiter le problème avec des pourcentages, il faut donc remplacer les pourcentages par des effectifs théoriques.

Niveau de diplôme	c_i	o_i	χ^2
Pas de diplôme	85	72	1,98
BEPC	20	24	0,8
CAP BEP	62,5	65	0,1
BAC ou BP	27,5	25	0,23
1er cycle universitaire ou BTS ou DUT	25	29	0,64
2ème ou 3ème cycle universitaire	17,5	19	0,13
Autre	12,5	16	0,98
Total	250	250	4,86

Tous les c_i étant plus grand que 5, la condition d'application sur les effectifs est respectée et le test réalisé sur ces bases.

On a 7-1=6 dd. Donc $\chi^2_{0,05} = 12{,}59$

$\chi^2 < \chi^2_{0,05}$ donc on conserve H_0. L'échantillon est bien représentatif au seuil 0 ,05.

Exercice 3 :

Hypothèse H_0: La répartition des groupes sanguins dans la ville n'est pas différente de celle de la France.
<=> π1=0,45, π2=0,44, π3=0,08, π4=0,03 où π_i (i = 1, ..., 4) représente la proportion de sujets dans le groupes sanguin n°i au niveau national.

Groupe sanguin	O	A	B	AB
o_i	485	392	74	49
π_i	0,45	0,44	0,08	0,03
c_i	450	440	80	30
χ^2	2,72	5,24	0,45	20,44

Tous les c_i étant plus grand que 5, la condition d'application sur les effectifs est respectée et le test réalisé sur ces bases :

$$\chi^2=\sum_{ji}\frac{(n_j-N\pi_j)^2}{N\pi_j}=\approx 20{,}44$$

On a 3 degrés de liberté : ddl=4-1. $\chi^2_{0,001}=16{,}27$

$\chi^2 > \chi^2_{0,001}$ donc on rejette H_0 au seuil de risque de 0,1%.

Il est possible de conclure, à ce seuil, que la répartition des groupes sanguins dans la ville est différente de celle de la France.

Exercice 4 :

Hypothèse H_0 : les deux variables sont indépendantes.

On évalue les c_i avec la formule $c_{ij}=\frac{N_i \times N_j}{N}$

Puis on calcule pour chaque CSP et chaque SEF : $\chi^2=\sum_{ij}\frac{(o_{ij}-c_{ij})^2}{c_{ij}}$

		CSP1	CSP2	CSP3	Total
SEF faible	oi	9	8	6	23
	ci	8,75	7,57	6,68	
SEF souple	oi	40	21	8	69
	ci	26,26	22,7	20,03	
SEF rigide	oi	10	22	31	63
	ci	23,98	20,73	18,29	
Total		59	51	45	155
χ^2		15,34645	0,229543	16,1268	31,7028

Tous les c_i étant plus grand que 5, la condition d'application sur les effectifs est respectée et le test réalisé sur ces bases.

On a 4 degrés de liberté : ddl=(3-1)(3-1). $\chi^2_{0,001}=18{,}47$

$\chi^2 > \chi^2_{0,001}$ donc on rejette H_0 au seuil de risque de 0,1%.

On peut considérer que les deux distributions sont liées significativement.

Exercice 5 :

1. L'hypothèse à tester (H_0) est l'indépendance entre l'opérateur et son rôle comme opérateur principal ou secondaire.

2. Les deux variables qualitatives nominales sont :
• d'une part le code de l'opérateur (1 à 8)
• d'autre part le statut comme opérateur principal (1 - principal, 2 - secondaire).
3. 4.

		1er ou 2ème opérateur		Total
Opérateur		1er	2ème	
1	oi	967	20	987
	ci	818,50	168,50	
2	oi	38	12	50
	ci	41,46	8,54	
3	oi	52	13	65
	ci	53,90	11,10	
4	oi	47	32	79
	ci	65,51	13,49	
5	oi	145	36	181
	ci	150,10	30,90	
6	oi	83	57	140
	ci	116,10	23,90	
7	oi	79	27	106
	ci	87,90	18,10	
8	oi	22	98	120
	ci	99,51	20,49	
Total		1433	295	1728
χ^2		103,42	502,37	605,79

On évalue les c_i avec la formule $c_{ij} = \frac{N_i \times N_j}{N}$

Tous les c_i étant plus grand que 5, la condition d'application sur les effectifs est respectée et le test réalisé sur ces bases.

On a 7 degrés de liberté : ddl=(8-1)(2-1). $\chi^2_{0,001} = 24,32$

$\chi^2 > \chi^2_{0,001}$ donc on rejette H_0 au seuil de risque de 0,1%.

Il existe donc une différence significative de rôle entre les chirurgiens de l'équipe.

Exercice 6 :

L'hypothèse à tester (H_0) est que le mélange est homogène. Si le mélange est homogène (même quantité des 4 variétés) les effectifs doivent s'ajuster à une loi uniforme discontinue (On peu appeler les variétés 1, 2, 3, 4 ou A, B, C, D). Avec 4 valeurs de la variable aléatoire, la probabilité théorique est de ¼, donc l'effectif théorique est de 25 (1/4x100). Tous les c_i étant plus grand que 5, la condition d'application sur les effectifs est respectée et le test réalisé sur ces bases. On fait un test de χ2 par rapport à l'effectif théorique de $\chi^2 = \sum_i \frac{(o_i - 25)^2}{25} = 4\sum_i \frac{(o_i - 25)^2}{100}$

On trouve :

Variété	A	B	C	D	Total
o_i	18	27	35	20	100
c_i	25	25	25	25	100
χ^2	1,96	0,16	4,00	1,00	7,12

On trouve χ^2 sur toutes les classes =7,12 La loi uniforme est définie par 1 paramètre, donc ddl = 4-1-1 = 2

$\chi^2_{0,05} = 5{,}99$ et $\chi^2_{0,01} = 9{,}21$ donc $\chi^2_{0,05} < \chi^2 < \chi^2_{0,01}$
L'hypothèse (H_0) de l'homogénéité est rejetée au risque de 5% et acceptée au risque de1%.

Exercice 7 :

L'hypothèse à tester (H_0) est que le nombre d'impulsions suit une loi de Poisson.
Paramètres statistiques de la distribution : n=28, μ=3,61 et σ=2,16
On calcule les ci par $n \times P(X=i) = n \times e^{-\lambda} \frac{\lambda^i}{i!} = 28 \times e^{-3,61} \frac{3{,}61^i}{i!}$ en considérant λ=μ

Mesure	0	1	2	3	4	5	6	7	8	9	Total
oi	1	4	5	4	5	4	2	2	0	1	28
ci	0,76	2,73	4,94	5,94	5,36	3,87	2,33	1,20	0,54	0,22	28
χ^2	0,08	0,59	0,00	0,63	0,02	0,00	0,05	0,53	0,54	2,82	5,26

La loi de Poisson est définie par 1 paramètre, donc on a 8 degrés de liberté : ddl=(10-1)-1. $\chi^2_{0,05} = 15{,}5$
$\chi^2 < \chi^2_{0,05}$ donc on accepte H_0 au seuil de risque de 5 %.
L'hypothèse à tester (H_0) est retenue. La loi de Poisson est acceptée.
En toute rigueur, pour avoir $o_i > 4$, il faudrait regrouper les classes 0 et 1 ainsi que 6, 7, 8 et 9. Avec ce regroupement, la loi est toujours acceptée.