DUT TC2 - Module OS 01 - PROBABILITÉS ET STATISTIQUE INFÉRENTIELLE

CORRECTION Exercices Chapitre 4 - Le test du χ^2.

Exercice 27 *Correction* :
On pose l'hypothèse H_0 :"Le dé n'est pas truqué", ce qui revient à poser l'hypothèse que chaque face a la même probabilité d'apparaître (équiprobabilité). Si tel est le cas, en 60 jets, les effectifs théoriques attendus devraient être :

$$e_1 = e_2 = e_3 = \ldots = e_6 = 60 \times \frac{1}{6} = 10.$$

On peut dresser le tableau suivant :

Face x_i	1	2	3	4	5	6	Total
Effectif empirique n_i	15	7	4	11	6	17	60
Effectif théorique e_i	10	10	10	10	10	10	60
$\frac{(n_i - e_i)^2}{e_i}$	2,5	0,9	3,6	0,1	1,6	4,9	13,6

Tous les effectifs théoriques sont supérieurs ou égaux à 5 et la taille de l'échantillon ($N = 60$) est suffisante pour que nous puissions appliquer le test d'ajustement du χ^2. La quantité $\sum_{i=1}^{6} \frac{(n_i - e_i)^2}{e_i}$ est un χ^2 à $\nu = k - 1 = 5$ degrés de liberté. Sa valeur pour l'échantillon est $\chi^2_{obs} = 13,6$. Pour $\alpha = 0,05$, la valeur critique lue dans la table est $\chi^2_{lu} = 11,07$. Puisque $\chi^2_{obs} > \chi^2_{lu}$, on rejette l'hypothèse H_0, ce qui signifie qu'on peut considérer que le dé est truqué (ou que le joueur triche).

Exercice 28 *Correction* :
On pose l'hypothèse H_0 : "Les arrivées sont régies par une loi de Poisson de moyenne $m = 2$". On en déduit les effectifs théoriques $e_i = Np_i$ avec $N = 264$, les probabilités $p_i = p(X = x_i)$ étant calculées à l'aide de la loi de Poisson de paramètre $m = 2$.

Nombre de clients x_i	Effectif empirique n_i	Probabilité $p_i = P(X = x_i)$	Effectif théorique Np_i	$\frac{(n_i - Np_i)^2}{Np_i}$
0	23	0,1353	35,7	4,53
1	75	0,2707	71,5	0,17
2	68	0,2707	71,5	0,17
3	51	0,1804	47,6	0,24
4	30	0,0902	23,8	1,61
Total	264	1	264	7,4
$\left.\begin{array}{c}5\\ \text{plus de } 5\end{array}\right\}$ plus de 4	$\left.\begin{array}{c}10\\ 7\end{array}\right\}$ 17	$\left.\begin{array}{c}0,0361\\ 0,0166\end{array}\right\}$ 0,0527	$\left.\begin{array}{c}9,5\\ 4,4\end{array}\right\}$ 13,9	0,69

L'effectif théorique associé à un nombre de clients inférieur à 5 étant insuffisant relativement aux conditions de validité du test, il convient de fusionner les deux dernières lignes du tableau en une seule $x_i > 4$. La valeur du χ^2 observé calculé sur l'échantillon s'élève alors à 7,4 tandis que la valeur χ^2_{lu} lue dans dans la table au seuil $\alpha = 0,05$ et $\nu = k - 1 = 5$ ddl est égale à 11,07. Puisque $\chi^2_{obs} < \chi^2_{lu}$, on ne peut rejeter l'hypothèse H_0. Au seuil de 5%, on peut admettre que les arrivées des clients sont régies par la loi de Poisson de moyenne $m = 2$.

Exercice 29 Pour tester l'ajustement d'une loi normale aux données de l'échantillon, il nous faut, pour en estimer les paramètres, calculer la moyenne et l'écart-type de l'échantillon. D'où le tableau de calculs suivant :

Classe i	Centre de classe x_i	Effectif n_i	n_ix_i	$n_ix_i^2$
$[5,5;6,5[$	6	2	12	72
$[6,5;7,5[$	7	3	21	147
$[7,5;8,5[$	8	12	96	768
$[8,5;9,5[$	9	27	243	2187
$[9,5;10,5[$	10	23	230	2300
$[10,5;11,5[$	11	15	165	1815
$[11,5;12,5[$	12	12	144	1728
$[12,5;13,5[$	13	5	65	845
$[13,5;14,5[$	14	2	28	392
$[14,5;15,5[$	15	2	30	450
Total	–	103	1034	10704

On a alors :

- $m = \dfrac{1304}{103} = 10,038$
- $s^2 = \dfrac{10704}{103} - (10,038)^2 = 3,144 \Rightarrow s \simeq 1,773$

On pose l'hypothèse H_0 : "La distribution des chiffres d'affaires mensuels des magasins est régie par une loi normale". On estimera ses paramètres par leurs estimations sans biais :

- μ_x est estimé par $m = 10,038$
- σ_x est estimé par $\hat{s} = \sqrt{\dfrac{N}{N-1}} \times s \simeq 1,78$

Il reste à calculer les probabilités associées à chaque classe. Pour cela, il suffit de calculer les valeurs de l'écart centré réduit correspondant aux limites de classes, puis de déterminer les probabilités des intervalles obtenus à l'aide de la loi normale centrée réduite.

Classe en X	n_i	Classe en T	p_i	$e_i = Np_i$	$\dfrac{(n_i - e_i)^2}{e_i}$
$\left.\begin{array}{l}[5,5;6,5[\\ [6,5;7,5[\end{array}\right\} [5,5;7,5[$	$\left.\begin{array}{c}2\\ 3\end{array}\right\} 5$	$\left.\begin{array}{r}]-\infty;-1,99[\\ [-1,99;-1,43[\end{array}\right\}]-\infty;-1,43[$	0,0233 0,0531	$\left.\begin{array}{c}\boxed{2{,}3999}\\ 5,4693\end{array}\right\} 7,8692$	1,0461
[7,5 ;8,5[	12	[-1,43 ;-0,86[	0,1185	12,2055	0,0035
[8,5 ;9,5[	27	[-0,86 ;-0,30[	0,1872	19,2816	3,0897
[9,5 ;10,5[	23	[-0,30 ;0,26[	0,2205	22,7115	0,0037
[10,5 ;11,5[	15	[0,26 ;0,82[	0,1913	19,7039	1,1230
[11,5 ;12,5[	12	[0,82 ;1,38[	0,1223	12,5969	0,0283
$\left.\begin{array}{l}[12,5;13,5[\\ [13,5;14,5[\\ [14,5;15,5[\end{array}\right\} [12,5;15,5[$	$\left.\begin{array}{c}5\\ 2\\ 2\end{array}\right\} 9$	$\left.\begin{array}{l}]1,38;1,94[\\ [1,94;2,51[\\ [2,51;+\infty[\end{array}\right\}]1,38;+\infty[$	0,0576 0,0202 0,006	$\left.\begin{array}{c}5,9328\\ \boxed{2{,}0806}\\ \boxed{0{,}618}\end{array}\right\} 8,6314$	0,0157
Total	103	–	1	103	5,31

Explications des calculs pour les deux premières lignes du tableau : sous H_0, $X \rightsquigarrow \mathcal{N}(10,038;1,78)$. On pose alors $T = \dfrac{X - 10,038}{1,78}$ et $T \rightsquigarrow \mathcal{N}(0,1)$. On s'intéresse à la classe $[5,5;6,5[$. On cherche donc

$$p_1 = p(X \in [5,5;6,5[) = p(5,5 \leq X < 6,5) = p\left(\frac{5,5-10,038}{1,78} \leq X < \frac{6,5-10,038}{1,78}\right) = p\left(-2,55 \leq T < -1,99\right).$$

On assimile la classe $]-2,55;-1,99[$ à $]-\infty;-1,99]$ pour deux raisons :

— l'aire additionnelle, c'est-à-dire $p(-\infty < T < -2,55) = 1 - 0,9946 = 0,0054$, est négligeable. Donc on peut affirmer que $p(-2,55 \leq T < -1,99) \simeq p(-\infty < T < -1,99)$.

— La somme des probabilités p_i devant être égale à 1, il faut que la variable T parcoure $\mathbb{R} =]-\infty;\infty[$ entièrement. On aura de fait $\sum_i Np_i = \sum_i e_i = N = 103$.

On cherche donc $p_1 = p(X \in [5,5;6,5[) \simeq p(-\infty < T < -1,99)$ qui vaut d'après la table de la loi normale centrée réduite $1 - 0,9767 = 0,0233$. On en déduit l'effectif théorique associé $e_1 = Np_1 = 103 \times 0,0233 = 2,3999$. Comme cet effectif est inférieur à 5 (la valeur a été encadrée dans le tableau), il faudra adjoindre (au moins) deux classes pour que la condition sur les effectifs soit vérifiée.
On s'intéresse maintenant à la classe $[6,5;7,5[$. On cherche

$$p_2 = p(X \in [6,5;7,5[) = p(6,5 \leq X < 7,5) = p\left(\frac{6,5-10,038}{1,78} \leq X < \frac{7,5-10,038}{1,78}\right) = p(-1,99 \leq T < -1,43).$$

Donc $p_2 = \Pi_T(-1,43) = \Pi_T(-1,99) = (1-\Pi_T(1,43))-(1-\Pi_T(1,99)) = \Pi_T(1,99)-\Pi_T(1,43) = 0,9767-0,9236 = 0,0531$. L'effectif théorique associé est égal à $e_2 = Np_2 = 103 \times 0,0531 = 5,4693 \geq 5$. Il suffit donc d'adjoindre les classes $[5,5;6,5[$ et $[6,5;7,5[$ pour que la condition sur les effectifs théoriques à ce niveau des calculs (attention, on n'a travaillé que sur les deux premiŕes lignes...) soit vérifiée. On additionne alors les effectifs observés n_1 et n_2, et les effectifs théoriques e_1 et e_2, et on calcule la "première erreur" $\dfrac{(5-7,8692)^2}{7,8692} = 1,0461$. On continue le travail pour chaque classe en veillant systématiquement à ce que la condition sur les effectifs théoriques soit observée. On précise pour conclure que la dernière classe en X $[14,5;15,5[$ admet pour équivalent en T $[2,51;3,07[$ et est assimilée à $[2,51;+\infty[$ pour les mêmes raisons invoquées plus haut.

Après regroupement des deux premières lignes et des trois dernières, on trouve $\chi^2_{obs} = 5,31$. La valeur χ^2_{lu} lue dans la table au seuil de 5% et avec un nombre de degrés de liberté $\nu = k - r - 1 = 7 - 1 - 2 = 4$ est égale à 9,488. Comme $\chi^2_{obs} < \chi^2_{lu}$, on ne peut rejeter l'hypothèse H_0. On peut donc considérer que les chiffres d'affaires sont normalement distribués.

Exercice 30 *Correction* :

On pose l'hypothèse H_0 : "Le nombre de garçons par famille (de 7 enfants) obéit à une loi binomiale". En toute rigueur, la proportion p des garçons doit être estimée par la fréquence observée sur l'échantillon. En effectuant la somme $\sum_i n_i x_i$, on constate que l'échantillon compte au total 6650 garçons pour $1883 \times 7 = 13181$ enfants, soit une fréquence égale à $0,504$, qu'on arrondira, pour simplifier les calculs et permettre l'utilisation de la table, à $0,5$. Donc la loi (hypothétique) de proababilité régissant le nombre X de garçons par famille de 7 enfants serait une loi binomiale de paramètres $n = 7$ et $p = 0,5$. Elle s'écrit :

$$P(X = x_i) = C_7^{x_i}\left(\frac{1}{2}\right)^{x_i}\left(1-\frac{1}{2}\right)^{7-x_i} = C_7^{x_i}\left(\frac{1}{2}\right)^7.$$

On peut alors établir les distributions des effectifs théoriques :

Nombre de garçons x_i	Effectif n_i	Probabilité p_i	$e_i = Np_i$	$\dfrac{(n_i-e_i)^2}{e_i}$
0	27	0,0078	14,7	10,32
1	111	0,0547	103	0,62
2	287	0,1641	309	1,57
3	480	0,2734	514,8	2,35
4	529	0,2734	514,8	0,39
5	304	0,1641	309	0,08
6	126	0,0547	103	5,14
7	19	0,0078	14,7	1,27
Total	1883	1	1877	21,74

On a $\chi^2_{obs} = 21,74$. Cette valeur est sensiblement plus élevée que la valeur limite du χ^2 à $8 - 1 - 1 = 6$ ddl (un paramètre, la probabilité p, ayant été estimée), lue dans la table au seuil de 5% : $\chi^2_{lu} = 12,59$. L'hypothèse doit donc être rejetée à ce seuil. Deux interprétations pourraient être données de ce rejet :

— On pourrait en premier lieu supposer que la probabilité de naissance d'un garçon qu'on a arrondie à 0,5 est incorrecte (minorée), la loi binomiale pouvant néanmoins s'appliquer. Toutefois, une aussi faible minoration de p ne saurait expliquer un tel écart entre l'échantillon et la loi d'ajustement.

— En réalité, on constate que la source de divergence entre les effectifs empiriques et les effectifs théoriques est étroitement localisée aux deux extrémités de la distribution ($X = 0$, $X = 6$ et 7). Autrement dit, la répartition du sexe des enfants dans une famille n'est pas une véritable loterie. Le fait de n'avoir eu aucun garçon (ou aucune fille) sur les 6 premières naissances n'est pas sans influence sur les probabilités relatives à la 7e naissance. L'indépendance entre les épreuves n'étant pas respectée, la loi binomiale ne peut pas être appliquée.

Exercice 31 On pose l'hypothèse :

"Les caractéristiques de la machine correspondent à l'assertion du fournisseur"

Cette hypothèse équivaut à appliquer au caractère "conforme" ou "hors-norme" de chaque pièce les probabilités 0,95 et 0,05 respectivement. D'où le tableau des effectifs empiriques et théoriques :

	Effectif empirique n_i	Probabilité p_i	Effectif théorique $e_i = Np_i$
Conformes	91	0,95	95
Hors-normes	9	0,05	5
Total	$N = 100$	1	100

La taille de l'échantillon est suffisamment grande ($N \geq 30$) et $\forall i$, $e_i \geq 5$. Les conditions d'application du test du χ^2 sont réunies. On a

$$\chi^2_{obs} = \sum_{i=1}^{2} \frac{(n_i - e_i)^2}{e_i} = \frac{(9-5)^2}{5} + \frac{(91-95)^2}{95} \simeq 3,37.$$

Au seuil de 5%, la valeur limite de χ^2_{lu} à $k-1 = 1$ ddl est égale à $3,841$. Puisque $\chi^2_{obs} < \chi^2_{lu}$, on ne peut rejeter l'hypothèse selon laquelle 95% des pièces fabriquées par la machine seraient conformes aux normes exigées.

Exercice 32 On pose l'hypothèse :

H_0 : "Les résultats sont les mêmes pour les deux populations d'où l'on a extrait les deux échantillons"

n_i	Succès	Échec	Total
1^er^ groupe	42	8	50
2^ème^ groupe	18	12	30
Total	60	20	80

Si les deux groupes répondent de la même façon au questionnaire (hypothèse H_0), les effectifs sont respectivement proportionnels aux totaux horizontaux et verticaux. Il est facile de les calculer en multipliant les totaux de la ligne et de la colonne correspondantes puis en divisant par l'effectif total ; on obtient de cette manière les valeurs suivantes :

$$\begin{array}{c|c} 50 \times \dfrac{60}{80} = 37,5 & 50 \times \dfrac{20}{80} = 12,5 \\ \hline 30 \times \dfrac{60}{80} = 22,5 & 30 \times \dfrac{20}{80} = 7,5 \end{array}$$

Les effectifs ainsi calculés fournissent les valeurs des effectifs théoriques :

n_i	Succès	Échec	Total
1^er^ groupe	37,5	12,5	50
2^ème^ groupe	22,5	7,5	30
Total	60	20	80

La détermination du nombre de ddl donne $\nu = (a-1) \times (b-1) = (2-1) \times (2-1) = 1$. L'applicayion de la fomule du χ^2 donne :

$$\chi^2_{obs} = \frac{(42-37,5)^2}{37,5} + \frac{(8-12,5)^2}{12,5} + \frac{(18-22,5)^2}{22,5} + \frac{(12-7,5)^2}{7,5} \simeq 5,76.$$

Le seuil choisi est $\alpha = 0,05$ et la valeur lue dans la table est $\chi^2_{lu} = 3,841$. Comme $\chi^2_{obs} > \chi^2_{lu}$, on rejette l'hypothèse H_0 à ce seuil. On considère que les résultats ne sont pas les mêmes pour les deux populations d'où l'on a extrait les deux échantillons.

Remarque 0.1 Si le seuil choisi avait été $\alpha = 0,01$, la valeur correspondante dans la table étant $\chi^2_{lu} = 6,635$, on accepterait l'hypothèse H_0 puisque dans ce cas, $\chi^2_{obs} < \chi^2_{lu}$.

Exercice 33 On fait l'hypothèse :

H_0 : "L'intensité de l'asthme et la présence d'eczéma sont indépendantes".

Asthme / Eczéma	Présent	Passé	Jamais	Total
Fort	22	28	28	78
Faible	12	33	37	82
Total	34	61	65	160

1. S'il n'existe pas de relation entre l'intensité de l'asthme et la présence d'eczéma, cela se traduira pas une répartition au hasard des résultats, c'est-à-dire que les nombres de chaque case du tableau précédent seront voisins des nombres théoriques du tableau à venir, en faisant le produit des deux extrêmes (ligne et colonne), divisépar l'effectif total. Plus le χ sera grand, moins les répartitions seront imputables au hasard : on pourra alors faire l'hypothèse d'une liaison entre les variables. On obtient ainsi :

$$\begin{array}{c|c|c}
78 \times \dfrac{34}{160} \simeq 16,6 & 78 \times \dfrac{61}{160} \simeq 29,7 & 78 \times \dfrac{65}{160} \simeq 31,7 \\
\hline
82 \times \dfrac{34}{160} \simeq 17,4 & 82 \times \dfrac{61}{160} \simeq 31,3 & 82 \times \dfrac{65}{160} \simeq 33,3
\end{array}$$

ce qui donne

Asthme / Eczéma	Présent	Passé	Jamais	Total
Fort	16,6	29,7	31,7	78
Faible	17,4	31,3	33,3	82
Total	34	61	65	160

2. La détermination du nombre de degrés de liberté se fait par la formule $\nu = (a-1)\times(b-1) = (3-1)\times(3-1) = 4$.
3. Le calcul du χ^2 donne :

$$\chi^2_{obs} = \sum_{i=1}^{6} \frac{(n_i - n'_i)^2}{n'_i} \frac{(22-16,6)^2}{16,6} + \frac{(28-29,7)^2}{29,7} + \ldots + \frac{(37-33,3)^2}{33,3} \simeq 4,465.$$

4. Au seuil de signification $0,95$, on trouve $\alpha = 0,05$, la valeur lue dans la table est $\chi^2_{lu} = 5,991$. Comme $\chi^2_{obs} < \chi^2_{lu}$, on ne rejette pas l'hypothèse H_0 selon laquelle l'instensité de l'asthme et la présence d'eczéma sont indépendantes.

Exercice 34 On fait l'hyptohèse :

H_0 : "Les résultats électoraux du village du Sud-Est de la France correspondent aux résultats nationaux pour cette élection".

1. Construction du tableau :

	A	B	C	D	E	Totaux
Nombre de votants du village	475	364	1968	1633	560	5000
Résultats nationaux en % (p_i)	10,3	7,1	40,8	32,1	9,7	100
Résultats théoriques sous H_0 ($5000 \times sp_i$)	515	355	2040	1605	485	5000

2. Calcul du χ^2_{obs} :

$$\chi^2_{obs} = \frac{(475-515)^2}{515} + \frac{(364-355)^2}{355} + \frac{(1968-2040)^2}{2040} + \frac{(1633-1605)^2}{1605} + \frac{(560-485)^2}{485} \simeq 17,963.$$

3. Le nombre de degrés de liberté ν est égal à la différence entre

. d'une part, le nombre de quantités de la forme $\dfrac{(n_i - n'_i)^2}{n'_i}$ dont la somme est le χ^2_{obs},

. d'autre part, le nombre des relations qui unissent les quantités en question.

Dans notre exemple, on connaît l'effectif total ($N = 5000$), donc le nombre de ddl est égal à $\nu = 5 - 1 - 4$. Au risque de 1% ($\alpha = 0,01$), la valeur lue dans la table vaut : $\chi^2_{lu} = 13,277$. Comme $\chi^2_{obs} > \chi^2_{lu}$, on rejette l'hypothèse H_0. On considère par conséquent que les résultats électoraux du village du Sud-Est de la France ne correspondent pas aux résultats nationaux pour cette sélection.

Exercice 35 On fait l'hypothèse :

H_0 : "Il y a analogie des choix des hommes et des femmes".

1. Construction du tableau des distributions théoriques :

Candidats	A	B	C	D	E	Totaux
Hommes	25	75	105	130	75	410
Femmes	147	116	17	80	105	465
Totaux	172	191	122	210	180	875

Principe du raisonnemment : si les hommes et les femmes votent de manière analogue (hypothèse H_0), les effectifs sont respectivement proportionnels aux totaux horizontaux et verticaux. Il est facile de les calculer en multipliant les totaux de la ligne et de la colonne correspondantes et en divisant par l'effectif total. On obtient ainsi :

$410 \times \dfrac{172}{875} \simeq 80,59$	$410 \times \dfrac{191}{875} \simeq 89,5$	$410 \times \dfrac{122}{875} \simeq 57,17$	$410 \times \dfrac{210}{875} \simeq 98,4$	$410 \times \dfrac{180}{875} \simeq 84,34$
$465 \times \dfrac{172}{875} \simeq 91,41$	$465 \times \dfrac{191}{875} \simeq 101,5$	$465 \times \dfrac{122}{875} \simeq 64,83$	$465 \times \dfrac{210}{875} \simeq 111,6$	$465 \times \dfrac{180}{875} \simeq 95,66$

On construit avec ces résultats le tableau des effectifs théoriques :

Candidats	A	B	C	D	E	Totaux
Hommes	80,59	89,50	57,17	98,40	84,34	410
Femmes	91,41	101,50	64,83	111,60	95,66	465
Totaux	172	191	122	210	180	875

2. Nombre de degrés de liberté : leur détermination se fait via la formule $\nu = (a-1)\times(b-1) = (5-1)(2-1) = 4$.
3. Calcul du χ^2_{obs} : l'application de la formule du χ^2 donne :

$$\chi^2_{obs} = \frac{(25 - 80,59)^2}{80,59} + \frac{(75 - 89,5)^2}{89,5} + \ldots + \frac{(80 - 111,6)^2}{111,6} + \frac{(105 - 95,66)^2}{95,66} \simeq 172,92.$$

4. Au risque de 1% ($\alpha = 0,01$), la valeur lue dans la table vaut $\chi^2_{lu} = 13,277$. Comme $\chi^2_{obs} > \chi^2_{lu}$, on rejette l'hypothèse H_0. On considère qu'il n'y a pas d'analogie entre les hommes et les femmes.

Exercice 36

1. On fait l'hypothèse :

H_0 : "Les deux caractères rhtyme cardiaque et vie urbaine ou rural sont indépendants".

Population / Rythme	$R < 65$	$[65, 70[$	$[70, 75[$	$[75, 80[$	$80 \leq R$	Totaux
Urbaine	6	21	53	60	10	150
Rurale	32	80	163	112	13	400
Totaux	38	101	216	172	23	550

Principe du raisonnement : s'il n'existe pas de relation entre le rythme cardiaque et le type de vie urbaine ou rurale (hypothèse H_0), les effectifs sont respectivement proportionnels aux totaux horizontaux et verticaux. Il est facile de les calculer en multipliant les totaux de la ligne et de la colonne correspondantes et en divisant pas l'effectif total. On obtient ainsi :

$150 \times \dfrac{38}{550} \simeq 10,36$	$150 \times \dfrac{101}{550} \simeq 27,55$	$150 \times \dfrac{101}{550} \simeq 58,91$	$150 \times \dfrac{172}{550} \simeq 46,91$	$150 \times \dfrac{23}{550} \simeq 6,27$
$400 \times \dfrac{38}{550} \simeq 27,64$	$400 \times \dfrac{101}{550} \simeq 73,45$	$400 \times \dfrac{216}{550} \simeq 157,09$	$400 \times \dfrac{172}{550} \simeq 125,09$	$400 \times \dfrac{23}{550} \simeq 16,6$

On construit avec ces résultats le tableau des effectifs théoriques :

Rythme \ Population	$R < 65$	$[65, 70[$	$[70, 75[$	$[75, 80[$	$80 \leq R$	Totaux
Urbaine	10,36	27,55	58,91	46,91	6,27	150
Rurale	27,64	73,45	157,09	125,09	16,73	400
Totaux	38	101	216	172	23	550

2. Nombre de degrés de liberté : leur détermination se fait via la formule $\nu = (a-1)\times(b-1) = (5-1)(2-1) = 4$.
3. Calcul du χ^2_{obs} : l'application de la formule du χ^2 donne :
$$\chi^2_{obs} = \frac{(6-10,36)^2}{10,36} + \frac{(21-27,55)^2}{27,55} + \ldots + \frac{(112-125,09)^2}{125,09} + \frac{(13-16,73)^2}{16,73} \simeq 13,552.$$
4. Au risque de 5% ($\alpha = 0,05$), la valeur lue dans la table vaut $\chi^2_{lu} = 9,488$. Comme $\chi^2_{obs} > \chi^2_{lu}$, on rejette l'hypothèse H_0. On considère qu'il y a un lien entre le rythme cardiaque et le type de vie urbaine ou rurale.

Exercice 37 On pose l'hypothèse :

H_0 : "Il n'y a pas de différence significative entre les proportions de réussite des groupes selon les diverses méthodes pédagogiques".

Soit H_0 : $p_1 = p_2 = p_3 = p$ (p_1, p_2 et p_3 désignant les propositions afférentes aux trois populations, toutes trois égales à une même valeur inconnue p). La remise en cause de cette hypothèse impliquerait donc, ces proportions n'étant pas toutes égales, qu'au moins deux d'entre elles soient significativement différentes. Dressons le tableau des effectifs empiriques :

	Méthode traditionnelle	Méthode traditionnelle renforcée	Nouvelle méthode	Totaux
Admis	$K_1 = 82$	$K_2 = 71$	$K_3 = 26$	179
Ajournés	$N_1 - K_1 = 33$	$N_2 - K_2 = 19$	$N_3 - K_3 = 9$	61
Totaux	$N_1 = 115$	$N_2 = 90$	$N_3 = 35$	240

La proportion hypothétique p des réussites commune aux 3 populations est estimée par la fréquence observée sur l'ensemble des 3 échantillons :
$$\hat{p} = \frac{K_1 + K_2 + K_3}{N_1 + N_2 + N_3} = \frac{179}{240} = 0,746$$
et $1 - \hat{p} = \frac{61}{240} \simeq 0,254$ d'où le tableau des effectifs théoriques :

	Méthode traditionnelle	Méthode traditionnelle renforcée	Nouvelle méthode	Totaux $N_{i.}$
Admis	$N_{.1}\hat{p} = 85,77$	$N_{.2}\hat{p} = 67,13$	$N_{.3}\hat{p} = 26,1$	$N_{1.} = 179$
Ajournés	$N_{.1}(1-\hat{p}) = 29,23$	$N_{.2}(1-\hat{p}) = 22,87$	$N_{.3}(1-\hat{p}) = 8,9$	61
Totaux	$N_{.1} = 115$	$N_{.2} = 90$	$N_{.3} = 35$	240

On obtient :
$$\chi^2_{obs} = \frac{(82-85,77)^2}{85,77} + \frac{(71-67,13)^2}{67,13} + \frac{(26-26,1)^2}{26,1} + \frac{(33-29,23)^2}{29,23} + \frac{(19-22,87)^2}{22,87} + \frac{(9-8,9)^2}{8,9} \simeq 1,53$$
La valeur critique du χ^2 à $(r-1)\times(s-1) = 1\times 2 = 2$ ddl est égale à $5,991$ pour $\alpha = 0,05$. Puisque $\chi^2_{obs} < \chi^2_{lu}$, on ne peut rejeter l'hypothèse H_0. Les différences de fréquences constatées sur les 3 échantillons n'apparaissent pas significatives au seuil de 5% : aucune méthode d'apparaît significativement plus efficace que les autres. Cette

conclusion ne signifie pourtant pas qu'il n'y ait aucune relation entre la méthode pédagogique et la probabilité de succès, mais seulement que l'échantillon ne permet pas de révéler cette (éventuelle) corrélation.

$\boxed{\textbf{Exercice 38}}$ On teste l'hypothèse selon laquelle les deux variables X (niveau de revenu) et Y (niveau d'épargne) sont indépendantes. À cet effet, il faut construire le tableau des effectifs théoriques $e_{ij} = \dfrac{N_{i.} \times N_{.j}}{N}$:

Revenus \ Épargne	y_1 : taux faible	y_2 : taux intermédiaire	y_3 : taux élevé	Total $N_{i.}$
x_1 : faibles	$27,375$	$29,93$	$15,695$	$N_{1.} = 73$
x_2 : intermédiaires	$30,375$	$33,21$	$17,415$	$N_{2.} = 81$
x_3 : élevés	$17,25$	$18,86$	$9,89$	$N_{3.} = 46$
Total $N_{.j}$	$N_{.1} = 75$	$N_{.2} = 82$	$N_{.3} = 43$	$N = 200$

Toutes les valeurs e_{ij} étant supérieures à 5, aucun regroupement n'est nécessaire. On a

$$\chi^2_{obs} = \frac{(53 - 27,375)^2}{27,375} + \ldots + \frac{(29 - 9,89)^2}{9,89} \simeq 117.$$

χ^2_{obs} étant considérablement plus élevé que χ^2_{lu}, valeur critique lue dans la table pour $(3-1) \times (3-1) = 2 \times 2 = 4$ ddl (sauf pour un seuil infinitésimal), on a donc la quasi-certitude que l'hypothèse testée est inexacte. Il existe donc une relation hautement significative entre le niveau de revenu et la propoension moyenne à épargner (le coefficient de corrélation entre les deux variables est significativement différent de zéro).