TD n°6

EX 1 1°) Le programme est correct si X suit la loi normale $N(109, (0,5)^2)$ c-a-d si $U = \frac{X - 109}{0,5}$ suit la loi normale $N(0,1)$. On choisit comme hypothèse nulle H_0: U a pour fonction de répartition F celle de la loi normale centrée réduite qu'on note Π et on fabrique le test :

$H_0 : F = \Pi$ contre $H_1 : F \neq \Pi$.

on calcule les bornes centrées réduites des classes; pour tout i de 1 à 12 on a $p_i = P(b_{i-1} < U \leq b_i) = \Pi(b_i) - \Pi(b_{i-1})$ la loi normale prenant les valeurs de $-\infty$ à $+\infty$, il faut penser à ajouter les 2 classes extremes $]-\infty, b_0]$ et $]b_{12}, +\infty[$

on calcule ensuite les effectifs théoriques $n'_i = 500 \times p_i$.

comme on ne peut pas appliquer le test du chi-deux si des classes présentent des effectifs théoriques inférieurs à 5, il faut dans ce cas, regrouper des classes voisines et prendre en compte les nouveaux effectifs théoriques n''_i. Les calculs s'organisent dans le tableau suivant :

Classe	n_i	$[b_{i-1} ; b_i[$	p_i	$n'_i \to n''_i$	$\frac{(n_i - np_i)^2}{np_i}$
$]-\infty ; 107,8[$	0	$]-\infty ; -2,4[$	0,0082	4,1 |	
$[107,8 ; 108[$	5	$[-2,4 ; -2[$	0,0146	7,3 | 11,4	3,5930
$[108 ; 108,2[$	5	$[-2 ; -1,6[$	0,0320	16,0	7,5625
$[108,2 ; 108,4[$	30	$[-1,6 ; -1,2[$	0,0603	30,1	0,0003
$[108,4 ; 108,6[$	40	$[-1,2 ; -0,8[$	0,0968	48,4	1,4579
$[108,6 ; 108,8[$	70	$[-0,8 ; -0,4[$	0,1327	66,4	0,1952
$[108,8 ; 109[$	85	$[-0,4 ; 0[$	0,1554	77,7	0,6858
$[109 ; 109,2[$	75	$[0 ; 0,4[$	0,1554	77,7	0,0938
$[109,2 ; 109,4[$	85	$[0,4 ; 0,8[$	0,1327	66,4	5,2102
$[109,4 ; 109,6[$	55	$[0,8 ; 1,2[$	0,0968	48,4	0,9000
$[109,6 ; 109,8[$	30	$[1,2 ; 1,6[$	0,0603	30,1	0,0003
$[109,8 ; 110[$	10	$[1,6 ; 2[$	0,0320	16,0	2,2500
$[110 ; 110,2[$	10	$[2 ; 2,4[$	0,0146	7,3 | 11,4	0,1719
$[110,2 ; +\infty[$	0	$[2,4 ; +\infty[$	0,0082	4,1 |	
Somme	500		1	500	22,121

on a $\delta^2_{100} = \sum_{j=1}^{12} \frac{(n_j - np_j)^2}{np_j} \simeq 22,121$

la table du χ^2_{11} donne $\chi^2_{0,95}$ à 11 degrés de liberté $\simeq 19,7$

et $\chi^2_{0,99}$ " " " " $\simeq 24,7$

puisque $19,7 < 22,121 < 24,7$ alors on rejette H_0 au risque 5%, mais pas au risque 1%.

2.) on calcule la moyenne et l'ecart type à partir à partir des valeurs observés on a alors

$m = \bar{X}_n = \frac{1}{500} \sum_{i=1}^{12} n_i \frac{a_i + a_{i-1}}{2}$ (a_i et a_{i-1} sont les bornes des classes

$\simeq 109,04$

$\sigma = \sqrt{S_n^2}$ où $S_n^2 = \frac{1}{500-1} \sum_{i=1}^{12} n_i (x_i - \bar{x}_n)^2$

$\Rightarrow \sigma \simeq 0,44$

dans cette question on prend $U = \frac{X - 109,04}{0,44}$

et on recalcule le nouveau tableau et on trouve:

Classe	n_i	$[b_{i-1}\,;\,b_i[$	p_i	$n_i' \to n_i''$		$\frac{(n_i - np_i)^2}{np_i}$
$]-\infty\,;\,107,8[$	0	$]-\infty\,;-2,82[$	0,0024	1,21		
$[107,8\,;\,108[$	5	$[-2,82\,;-2,36[$	0,0066	3,32		
$[108\,;\,108,2[$	5 | 10	$[-2,36\,;-1,91[$	0,0191	9,54	14,1	1,1922
$[108,2\,;\,108,4[$	30	$[-1,91\,;-1,45[$	0,0448	22,4	2,5786	
$[108,4\,;\,108,6[$	40	$[-1,45\,;-1,00[$	0,0858	42,9	0,1960	
$[108,6\,;\,108,8[$	70	$[-1,00\,;-0,55[$	0,1341	67,0	0,1343	
$[108,8\,;\,109[$	85	$[-0,55\,;-0,09[$	0,1711	85,5	0,0029	
$[109\,;\,109,2[$	75	$[-0,09\,;0,36[$	0,1782	89,1	2,2313	
$[109,2\,;\,109,4[$	85	$[0,36\,;0,82[$	0,1514	75,7	1,1425	
$[109,4\,;\,109,6[$	55	$[0,82\,;1,27[$	0,1051	52,5	0,1190	
$[109,6\,;\,109,8[$	30	$[1,27\,;1,73[$	0,0595	29,8	0,0013	
$[109,8\,;\,110[$	10	$[1,73\,;2,18[$	0,0275	13,7	0,9993	
$[110\,;\,110,2[$	10	$[2,18\,;2,64[$	0,0104	5,19	7,3	0,9986
$[110,2\,;+\infty[$	0	$[2,64\,;+\infty[$	0,0042	2,10		
Somme	500		1	500		9,5962

on a $\int_{500}^2 = \sum_{j=1}^{12} \frac{(n_i - np_i)^2}{np_i} \simeq 9,6$

Rq Δ_n^2 suit une loi de χ^2 à $n-k-1$ degrés de liberté

ici $k=2$ car la loi normale inconnue avait 2 paramètres m et σ qui ont été estimé à partir des observations

ainsi $\chi^2_{0,95}$ à 9 degrés de liberté vaut 16,9

comme $9,6 < 16,9$ on ne rejette pas l'Hyp. nulle équivalente

à X suit la loi normale $N(109{,}04 ; (0{,}44)^2)$.

EX2 désignons par X le caractère qui spécifie si le médicament guérit ou non et Y le caractère qui spécifie si le médicament est cher ou pas cher.

calculons les effectifs marginaux :

$n_{1\cdot} = 156 + 44 = 200$; $n_{2\cdot} = 44 + 6 = 50$

$n_{\cdot 1} = 156 + 44 = 200$; $n_{\cdot 2} = 44 + 6 = 50$

puis les effectifs théoriques pour $1 \le i \le 2$ et $1 \le j \le 2$ par $n'_{ij} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{n}$

	médicament cher	médicament pas cher	$n_{i\cdot}$
Guérisons	$n'_{11} \frac{200 \times 200}{250} = 160$	$\frac{50 \times 200}{250} = n'_{12} = 40$	200
Non-guérisons	$n'_{21} = \frac{50 \times 200}{250} = 40$	$n'_{22} = \frac{50 \times 50}{250} = 10$	50
$n_{\cdot j}$	200	50	$n = 250$

on calcule la valeur de la statistique $d_n^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(n_{ij} - n'_{ij})^2}{n'_{ij}}$

d'où $d_n^2 = \frac{(156-160)^2}{160} + \frac{(44-40)^2}{40} + \frac{(44-40)^2}{40} + \frac{(6-10)^2}{10}$

$= 2{,}5$

pour un risque de 1ère espèce $\alpha = 0{,}05$ le fractile

d'ordre $1-\alpha$ de la loi χ^2_1 a pour valeur 3,84.
puisque $d_n^2 < 3{,}84$, on accepte l'hypothèse nulle d'indépendance du taux de guérison et du coût du médicament.

EX3 on désigne par X le caractère taille et Y le caractère de la catégorie socio-professionnelle CSP.

nous commençons par calculer les effectifs théoriques pour $1 \le i \le 4$, $1 \le j \le 3$

par exemple $n'_{11} = \frac{1900 \times 413}{2700} = 290{,}6$ arrondi à 291 puisqu'il s'agit d'effectifs entiers, on obtient ainsi au tableau suivant :

	Ouvriers	Employés	Cadres	Total
Moins de 165 cm	291	73	49	413
De 165 à moins de 170 cm	457	115	77	649
De 170 à moins de 175 cm	645	163	109	917
175 cm et plus	507	128	86	721
Total	1900	479	321	2700

$r = 4$
$s = 3$

on calcule la valeur de la statistique

$d_M^2 = \sum_{i=1}^{4} \sum_{j=1}^{3} \frac{(n_{ij} - n'_{ij})^2}{n'_{ij}}$ et on trouve que $d_M^2 \simeq 5{,}812$

Pour un risque de 1ère espèce $\alpha = 5\%$ le fractile d'ordre $1-\alpha$ de la loi du χ^2 à $6 = (r-1)\times(s-1)$ degrés de liberté a pour valeur 12,6. Puisque $d_M^2 > 12{,}6$ alors on rejette l'Hyp. Ho d'indépendance de la taille et de la CSP.