ISARA 1ère année

Pascale **NEYRAN**

Corrigé épreuve STATISTIQUE juin 2007

Exercice I (4 points)

Réponses: B / C / B / B / A / C / A / C.

Exercice II (11 points)

Intéressement en k€	*ni*	*ni⁺*	*hi*
[0 ; 10[	6	6	2,4
[10 ; 20[	30	36	12
[20 ; 24[	50	86	50
[24 ; 30[	40	126	26,67
[30 ; 40[	20	146	8
[40 ; 50[	4	150	1,6

1) Déterminer au centième, les paramètres suivants :

Classe modale	[20 ; 24[elle correspond à la hauteur (hi) la plus forte sur l'histogramme.
Moyenne	3540/150 = 23,60
C_{25}	$C_{(25)} = 37{,}75 \Rightarrow C_{25} = 20{,}14$
Q_3	$Q_{(3)} = 113{,}25 \Rightarrow Q_3 = 28{,}09$
Médiane	Le rang de la me = 75,50 $\Rightarrow$ me = 23,16
$\sum n_i x_i^2 - n\bar{x}^2$	SCE = 9316
Coefficient de Yule	$\dfrac{(28{,}09-23{,}16)-(23{,}16-20{,}14)}{(28{,}09-20{,}14)} = 0{,}24$

2) Le coefficient de Yule est strictement positif : cette série n'est pas symétrique, elle est étalée à droite.

3) $\gamma_2 = \dfrac{2064607/150}{7{,}88^4} - 3 = 0{,}57$. La série est leptocurtique. Elle est moins aplatie que la distribution de la loi normale de même moyenne et de même écart type.

4) Si on appelle X la variable mesurant l'ancien intéressement et Y la variable correspondant au nouvel intéressement après augmentation :

Y = 1,015X + 1.

La moyenne de Y est donc égale à $1{,}015 \times 23{,}6 + 1 = 24{,}95$.

L'écart type de Y est égal à $|1{,}015| \times 7{,}881 = 8$.

Exercice III (6 points)

Hauteur / origine	*Faible*	*Moyenne*	*Forte*
Région 1	27 / 48,90	189 / 186,60	54 / 34,50
Région 2	33 / 32,60	124 / 124,40	23 / 23
Région 3	103 / 81,50	309 / 311	38 / 57,50

Hypothèse nulle : Il y a indépendance entre la taille des arbres et leur origine géographique.

Critère statistique calculé :	33,16
Nombre de degrés de liberté :	(3-1)(3-1) = 4
Critère statistique théorique :	9,49
Conclusion :	On rejette H_0 avec un risque de se tromper inférieur à 5 %
Risque réel ?	Le risque de se tromper en refusant H_0 est inférieur à 0,1%.

Exercice IV (10 points)

Partie A :

1) X peut s'écrire comme la somme de 180 variables aléatoires indépendantes :

$X = \sum_{i=1}^{180} X_i$ où X_i est la variable aléatoire mesurant le nombre d'heures de stationnement payées par le client numéro i $\Rightarrow$ X suit une loi de Poisson de paramètre $180 \times 3 = 540$.

2) Le paramètre de la loi de Poisson étant > 18, la loi de X peut être approximée par une loi normale $N(540 ; \sqrt{540})$.

3) Il faut trouver x tel que $P(540 - x \leq X \leq 540 + x) = 0,92$.

Or $P(T \leq \frac{x}{\sqrt{540}}) - P(T < -\frac{x}{\sqrt{540}}) = 2P(T \leq \frac{x}{\sqrt{540}}) - 1 \Rightarrow P(T \leq \frac{x}{\sqrt{540}}) = 0,96$.

Dans la table 2 de la loi normale on lit la valeur $1,7507 = \frac{x}{\sqrt{540}}$. On obtient x = 40,68 et l'intervalle centré sur la moyenne [499,32 ; 580,68].

Partie B :

1) Pour chaque client qui se présente aux caisses, il y a 2 issues possibles :

- il choisit la caisse B avec une probabilité p = 0,3
- ou non avec la probabilité q = 0,7.

Cette épreuve se répète de façon indépendante pour les 180 clients.
L(Y) = B(180 ; 0,3).

2) $E(Y) = 180 \times 0,3 = 54$ et $V(Y) = 180 \times 0,3 \times 0,7 = 37,8 \Rightarrow \sigma(Y) = 6,15$.

3) $n > 5$ et $\left| \sqrt{\frac{0,3}{0,7}} - \sqrt{\frac{0,7}{0,3}} \right| \times \frac{1}{\sqrt{180}} = 0,065$ donc inférieur à 0,34.

On peut approcher la loi Binomiale par la loi normale N(54 ;6,15).

4) $P(Y \leq 62) - P(Y < 38) = P(T \leq 1,30) - P(T < -2,60) = 0,9032 - (1 - 0,9953) = 0,8985$.

5) On cherche x tel que : $P(Y \leq x) = 0,97 \Rightarrow P(T \leq \frac{x-54}{6,15}) = 0,97$.

On en déduit que $(x - 54)/6,15 = 1,8808$. Il faut donc 66 reçus.

Exercice V (11 points)

1) $\text{cov}(X\,;\,Y) = \frac{1623,10}{8} - \frac{24,4}{8} \times \frac{515}{8} = 6,5438$.

$V(X) = \frac{78,2}{8} - \left(\frac{24,4}{8}\right)^2 = 0,4725$

$\Rightarrow b = 6,5438/0,4725 = 13,8493$ et $a = \frac{515}{8} - b \times \frac{24,4}{8} = 22,1346$

$V(Y) = \frac{34025}{8} - \left(\frac{515}{8}\right)^2 = 108,9844$

$r^2 = \frac{6,5438^2}{0,4725 \times 108,9844} = 0,8316 \Rightarrow$ 83,16 % des variations du chiffre d'affaires sont expliquées par les frais de publicité.

$SPE(X\,;\,Y) = n \times \text{cov}(X\,;\,Y) = 52,35$.

2) L'écart type des résidus $s(e_i) = \sqrt{\frac{146,87}{8}} = 4,28$.

N° observation	Yi estimé	résidu	e_i^2	Résidu standardisé
1	**49,83**	**8,83**	**77,97**	**2,06**
8	**78,92**	**3,92**	**15,37**	**0,91**
Total	515	**0**	146,87	**0**

3) La première est une donnée OUT car $\left| \frac{e_i}{s(e_i)} \right| > 2$.

4) $\text{cov}(x_i ; e_i) = \dfrac{\sum x_i e_i}{n} \approx 0$

5) $s^2(\hat{y}_i) = r^2 \times s^2(y_i) = 90{,}63$.

Exercice VI (8 points)

Hypothèse nulle **:** La distribution observée suit une loi Normale N(25,5 ; 10).

Calculons :

- P(X < 10) = P(T < (10-25,5)/10) = P(T < -1,55) = 1 - 0,9394 = 0,0606
- P(25 ≤ X < 30) = P(T < 0,45) - P(T < - 0,05) = 0,6736 - (1 - 0,5199) = 0,1935
- P(X ≥ 40) = P(T ≥ 1,45) = 1 - 0,9265 = 0,0735

Valeurs de X	probabilité	Effectifs théoriques	Effectifs observés
X < 10	**0,0606**	**12,12**	**10**
10≤X<20	0,2306	46,12	**50**
20≤X<25	0,1889	37,78	**38**
25≤X<30	**0,1935**	**38,70**	**40**
30≤X<40	0,2528	50,57	**52**
X ≥ 40	**0,0735**	**14,70**	**10**
Total	**1**	**200**	**200**

Il n'y a pas d'effectif théorique < 5.

Critère statistique calculé : 2,29

Critère statistique théorique : 11,10 pour un ddl de 5

Conclusion : On accepte H_o, car si on la refuse on prend un risque de se tromper supérieur à 5% (risque réel compris entre 50% et 90%).