ISARA-Lyon
Première année

le 11/06/2009

41ième promotion

durée : 2 heures

STATISTIQUE
EXAMEN
Pascale NEYRAN

Conditions d'examen

Documents		Autorisés
	X	**Non autorisés**
Calculatrice		Non autorisée
	X	**Collège autorisée**
		Tout type autorisée

Remarques particulières : **Répondre directement sur la feuille du sujet.**

Nom : .... Prénom : ....
Note sur 20 : ....

ISARA 1ère année Pascale **NEYRAN**

Nom : ....Prénom : ....

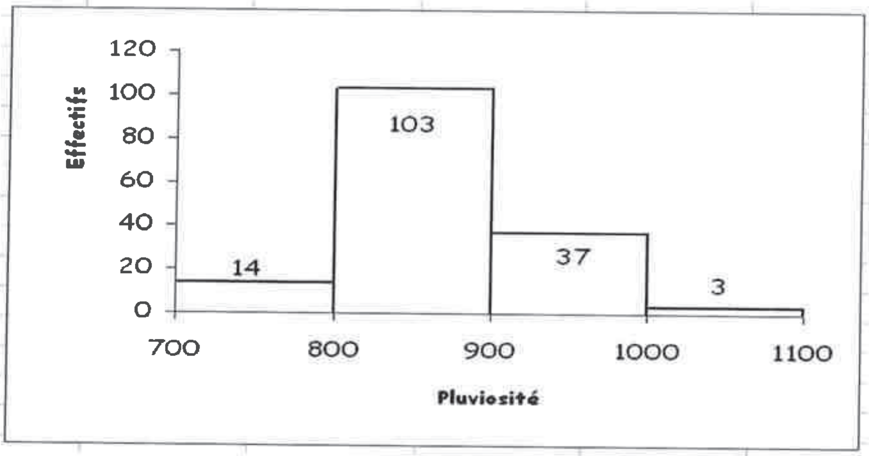
STATISTIQUE (durée 2 heures)

Une calculatrice collège est autorisée. Les tables statistiques distribuées au début de l'épreuve sont à rendre aux surveillant(e)s.
Le devoir est noté sur 50 points (barème entre parenthèses).

Exercice I (9,5 points)

Une étude est réalisée sur la répartition de la végétation en fonction de divers facteurs écologiques.

Partie A : Pour le chêne pubescent, les données concernant la pluviosité sont les suivantes :



Déterminer au centième, les paramètres suivants :

Moyenne	
$C_{(60)}$	
Q_3	
Médiane	
$\sum n_i(x_i - \overline{x})^2$	

Partie B : Pour le chêne rouge, la pluviosité se répartit ainsi :

pluviosité	[600 ; 800[	[800 ; 850[	[850 ; 900[	[900 ; 1100[	[1100 ; 1500[
effectifs	22	45	28	30	15

On nous précise que :

- la moyenne est égale à 903,75
- l'écart type est égal à 165,80
- $\sum n_i(x_i - \bar{x})^3 = 751\ 273\ 359{,}4$
- $\mu_4 = 2\ 943\ 149\ 337$.

1) Comparer la dispersion des pluviosités des 2 types de chênes (pubescent et rouge).

2) Etudier l'asymétrie de cette série.

3) Calculer le coefficient d'aplatissement de Fisher et conclure.

Exercice II (5 points)

Un nouveau vaccin a été testé sur un échantillon de 120 personnes pour juger de son efficacité. Un groupe témoin de 120 personnes non vaccinées a été suivi parallèlement. Les résultats sont les suivants :

	ont contracté la maladie	n'ont pas contracté la maladie
Vaccinés	13	107
Non vaccinés	26	94

On souhaiterait savoir si la vaccination et la survenue de la maladie sont 2 caractères indépendants ?

Exercice III (7 points)

Un moyen de tester que le sexe des enfants n'est pas déterminé (ou influencé) par un facteur génétique est de vérifier que la répartition des sexes dans les familles suit parfaitement une loi binomiale.
Une étude portant sur 220 familles de 5 enfants a donné les résultats suivants où X est la variable mesurant le nombre de garçons :

X	0	1	2	3	4	5
Nombre d'observations	5	30	60	80	35	10

1) Donner la raison pour laquelle le paramètre p de la loi binomiale est estimé à 0,53 (au centième).

2) Compléter le tableau suivant :

X	p_i	effectif théorique	effectif observé
0			
1			
2	0,2916		
3	0,3289		
4	0,1854		
5			

3) Tester au seuil de 5 %, l'hypothèse selon laquelle la distribution est binomiale.

Exercice IV (7,5 points)

On a mesuré chez 10 sujets la tension maximale d'un muscle donné avant dénervation (X) et après dénervation (Y). On dispose des données suivantes :

$\sum xi = 500$ $\sum yi = 330$ SPE(X ; Y) = 3 064

$\sum xi^2 = 29062$ V(Y) = 245,40

1) Déterminer l'équation de la droite de régression de Y par rapport à X.

2) Calculer le coefficient de corrélation linéaire.

3) Déterminer la variance résiduelle et la variance expliquée par la régression.

4) Quelle est la somme des carrés des résidus ?

5) A quoi sert le calcul des résidus standardisés ?

Exercice V (17 points)

PARTIE A

On étudie la variable aléatoire X qui mesure le taux des retours du rayon jardinage d'un magasin.
On suppose que la variable aléatoire X suit une loi normale de moyenne 25% et d'écart type 8%.

Question 1
Quel est le taux des retours qui est dépassé avec une probabilité de 0,10 ?

Question 2
Quelle est la probabilité que le taux des retours dépasse 40% ?

Question 3
A quel intervalle centré sur la moyenne peut-on s'attendre pour que le taux des retours soit connu avec une probabilité de 0,95 ?

Question 4
Déterminer le premier quartile.

PARTIE B
Compte tenu de la politique d'approvisionnement d'un certain détaillant, on estime que la probabilité qu'un client donné puisse trouver le produit qu'il cherche est de 0,95.
Soit Y la variable aléatoire donnant le nombre de clients **non satisfaits** parmi 15 clients qui se présentent au magasin. (Un client est non satisfait s'il ne trouve pas le produit qu'il cherche)

Question 5
Quelle est la loi suivie par la variable aléatoire Y ?

Question 6
Quelle est la probabilité qu'il y ait plus d'un client non satisfait ?

Question 7
Quelle est la loi suivie par la variable aléatoire Z donnant le nombre de clients non satisfaits parmi les 180 clients d'une journée.

Question 8
Par quelle loi discrète peut-on approcher la loi de Z ? Justifier.

Question 9
En utilisant cette approximation, déterminer la probabilité d'avoir plus de 15 clients non satisfaits ?

PARTIE C
Cette société fait étudier sur un mois par un de ses stagiaires la fréquence des sinistres liés au transport des commandes. Un échantillon aléatoire de 200 colis est examiné. On obtient la distribution suivante :

Nombre de produits abîmés	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Nombre De colis	25	53	58	37	19	5	2	1

Question 10
On appelle X la variable aléatoire correspondant au nombre de produits abîmés par colis. Calculer la moyenne et la variance arrondie à l'unité de cette variable.

Question 11
On suppose que X suit une loi de Poisson. Quel est le paramètre de cette loi ?

Question 12

On expédie **4 colis** au même magasin dans le mois et on en profite pour compter le nombre V de produits abîmés.

a) Quelle est la loi suivie par la variable aléatoire V ?

b) Quelle est la probabilité pour que V soit comprise entre 8 et 10 au sens large ?

Question 13

Dans une année, **50 colis** sont expédiés. Le nombre de produits abîmés W est une variable aléatoire qui suit une loi de poisson de paramètre m.

a) Calculer m.

b) Par quelle loi peut-on approcher la loi de W ?

c) Quelle est la probabilité pour que W soit strictement inférieure à 80 ?

Exercice VI (4 points)

Sélectionner la réponse correcte :

Question	*Réponse A*	*Réponse B*	*Réponse C*	*Choix*
SCE_x	$s^2(x)/n$	$\sum n_i x_i^2 - n\overline{x}^2$	$\sum n_i(x_i - \overline{x}^2)$	
La variance inter =	$\dfrac{\sum n_i s_i^2}{N}$	$\dfrac{\sum n_i (m_i - \overline{X})^2}{N}$	$\dfrac{\sum n_i m_i^2 - \overline{X}^2}{N}$	
Soit L(T) = N(0 ; 1), $P(-t \leq T \leq t) =$	$P(T \leq t) + P(T \geq -t)$	$2P(T \leq t)$	$2P(T \leq t) - 1$	
Si L(X) = N(10 ; 3)	$P(7 \leq X \leq 13) = 1$	$Q_1 \neq$ mode	$P(X < 10) = 0,5$	
Si L(X) = N(6 ; 2) alors L(-3X + 8) =	N(-10 ; 6)	N(-18 ; 2)	N(-10 ; - 2)	
Si X suit une loi de Poisson de variance 4	E(X) = 2	E(X) = 4	E(X) = 16	
$t_{0,001}(10)$ vaut	-2,764	2,764	2,326	
$s^2(\hat{y}_i) =$	$(1-r^2)s^2(y_i)$	$(1-r^2)s^2(e_i)$	$r^2 s^2(y_i)$	