

ISARA-Lyon
Première année

le 12/06/2008

40^{ième} promotion

durée : 2 heures

<p>STATISTIQUE EXAMEN Pascale NEYRAN</p>

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.

Répondre sur la feuille du sujet.

Le devoir est noté sur 50 points (barème entre parenthèses).

Exercice I (4 points)

Sélectionner la réponse correcte :

Question	Réponse A	Réponse B	Réponse C	Choix
$s(x) =$	$\frac{\sum n_i x_i^2}{n} - \bar{x}^2$	$\frac{SCE}{\sqrt{n}}$	$(\sum f_i x_i^2 - \bar{x}^2)^{0,5}$	
La moyenne vérifie :	$\sum n_i(x_i - \bar{x}) = 0$	$\sum n_i x_i - \bar{x} = 0$	$\sum f_i x_i - n\bar{x} = 0$	
Le coefficient d'aplatissement de Fisher =	$\frac{\mu_4}{s(x)^4}$	$\frac{\mu_3}{\mu_2^{1,5}} - 3$	$\frac{\mu_4}{\mu_2^2} - 3$	
Si $Y = bX + a$	$s^2(y) = bs^2(x) + a$	$s^2(y) = bs^2(x)$	$s(y) = b s(x)$	
$P(X \leq 2) =$	$1 - P(X \geq 2)$	$1 - P(X > 2)$	$P(X=1) + P(X=2)$	
Si $L(X) = B(60 ; 0,02)$ $L(X) \approx$	$N(60 ; 0,02)$	$N(1,2 ; \sqrt{0,02})$	$P(1,2)$	
Si $L(X) = B(20 ; 0,45)$ $L(X) \approx$	$N(9 ; \sqrt{4,95})$	$N(9 ; 4,95^2)$	$N(20 ; \sqrt{4,95})$	
Si $L(X) = P(144)$ $L(X) \approx$	$B(144 ; 12)$	$N(144 ; 12)$	$N(144 ; 144)$	

Exercice II (16,5 points)

Partie A : (11 points)

On considère la série statistique consignée dans le tableau ci-dessous donnant la distribution des exploitations agricoles de la zone Alpha suivant leur SAU (Surface Agricole Utilisable) en 2007 :

Superficies (en ha SAU)	Nombre d'exploitations
[0 ; 5[56
[5 ; 10[100
[10 ; 12[123
[12 ; 14[150
[14 ; 18[200
[18 ; 23[130
[23 ; 30[250
[30 ; 40[300
[40 ; 100[98
[100 ; 150[26
[150 ; 300[12

On nous précise que :

$$\sum n_i x_i = 39993$$

$$\sum n_i x_i^2 = 2189053$$

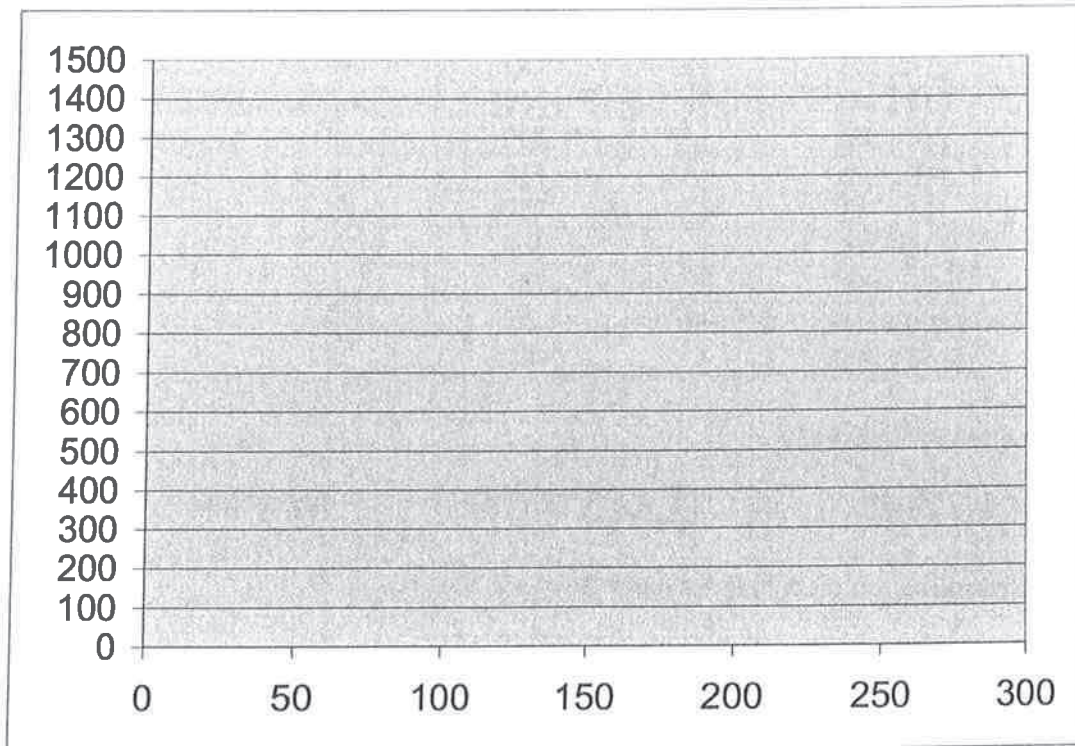
$$Q_3 = 32,52$$

1) Déterminer au centième, les paramètres suivants :

Classe modale	
Moyenne	
$D_{(4)}$	
Q_1	
Médiane	
SCE	
Coefficient de variation	

2) Calculer le coefficient de Yule de cette série. Que peut-on en déduire ?

3) Représenter les effectifs cumulés croissants et vérifier graphiquement la valeur de la médiane.



Partie B : (5,5 points)

Une étude similaire est réalisée dans la zone Béta comprenant 1430 exploitations. La Surface Agricole Utilisable moyenne dans cette zone est égale à 35 ha avec un écart type de 30 ha (on prendra 27,37 ha comme écart type dans la zone Alpha).

1) Comparer la dispersion des surfaces agricoles utilisables des 2 zones Alpha et Béta.

2) On définit la région Gama formée par les 2 zones Alpha et Béta.

a) Déterminer au niveau de la région Gama :

- la moyenne
- la variance intra
- la variance inter

b) Commenter les résultats obtenus dans la région Gama.

Exercice III (6 points)

On cherche à savoir si la fréquence d'une maladie est liée au groupe sanguin.

Sur 200 malades observés :

Groupe sanguin	O	A	B	AB	total
Effectifs observés	104	76	18	2	200

Sur la population générale la répartition entre les groupes est :

- groupe O : 47 %
- groupe A : 43 %
- groupe B : 7 %
- groupe AB : 3 %.

La répartition des groupes sanguins dans l'échantillon des 200 malades est-elle conforme à celle observée dans la population générale ? On prendra un niveau de confiance de 95 %.

Hypothèse nulle :

Critère statistique calculé :	
Nombre de degrés de liberté :	
Critère statistique théorique :	
Conclusion :	

Exercice IV (12 points)

Les parties A, B et C sont indépendantes.

Partie A : (6 points)

Le personnel d'un laboratoire pharmaceutique ayant suivi un stage de remise à niveau, doit passer un test d'évaluation.

On appelle X , la variable aléatoire mesurant le résultat à ce test.

La loi suivie par X est une loi normale de moyenne m et d'écart type σ .

On sait que 16 % des salariés ont obtenu une note inférieure à 10 et que 63 % des salariés ont obtenu un score plus petit que 14.

1) Déterminer à 10^{-2} près, les paramètres m et σ .

Pour la suite de l'exercice on prendra une moyenne de 13 et un écart type de 3.

2) Au-dessus de quelle note, trouve t'on 75 % des salariés ?

3) Déterminer un intervalle $[a ; b]$ centré sur la moyenne ayant une probabilité de 95 % de contenir X .

Partie B : (2,5 points)

Le directeur des Ressources Humaines souhaite connaître l'avis des salariés sur la création éventuelle d'une cafétéria. Il réalise une enquête auprès de 20 personnes. Parmi elles, 90 % sont favorables à la création de la cafétéria.

On note Y la variable aléatoire qui, à tout ensemble de 20 employés choisis au hasard, associe le nombre de personnes ayant répondu négativement.

1) Quelle est la loi suivie par Y ?

2) Calculer la moyenne et l'écart type de Y , à 10^{-2} près.

3) Calculer la probabilité pour que plus de 14 personnes soient favorables à cette proposition.

Partie C : (3,5 points)

La commission hygiène et sécurité de cette entreprise a recensé le nombre mensuel d'accidents du travail sur les 10 dernières années.

1) On appelle Z la variable aléatoire mesurant le nombre d'accidents du travail **pour un mois donné**. On sait que Z suit une loi de Poisson de paramètre 4.

a) Quelle est la probabilité pour qu'il y ait plus de 2 accidents dans le prochain mois ?

b) Calculer la probabilité pour que Z soit comprise entre 2 et 7 (bornes incluses).

2) On appelle U la variable aléatoire mesurant le nombre d'accidents du travail **pour un semestre donné**.

a) Quelle est la loi suivie par U ?

b) Quelle est la probabilité pour que le nombre d'accidents dans les six prochains mois soit inférieur à 30 ?

Exercice V (5,5 points)

On veut étudier la relation linéaire simple qui pourrait exister entre 2 variables X et Y . On dispose des données suivantes, pour les 10 couples recensés :

$$\sum x_i = 2260$$

$$\sum y_i = 2800$$

$$\sum x_i y_i = 472000$$

$$\sum x_i^2 = 624400$$

$$SCE(y_i) = 230800$$

Déterminer :

- les coordonnées du centre de gravité du nuage :
- la SPE de X et Y :
- la pente de la droite de régression de Y en X :
- le coefficient de corrélation :
- la variance résiduelle :

Exercice VI (6 points)

Une entreprise de distribution de matériel agricole a recensé le nombre de produits A vendus chaque jour, sur une période de 50 jours. Elle dispose des données suivantes :

x_i	0	1	2	3	4	5	6	7
n_i	3	7	11	11	8	5	3	2

avec :

- x_i = nombre de produits vendus, un jour donné
- n_i = nombre de jours correspondant

Montrer que la répartition donnée peut-être ajustée par une loi de Poisson de paramètre 3.