

Année universitaire 2017–2018

Session 1 – Semestre 2

Licence 1 mention Economie, parcours Economie et Gestion
Licence 1 mention Economie, parcours Economie et Droit

Epreuve : Statistique descriptive

Date de l'épreuve : lundi 7 mai 2018

Durée de l'épreuve : 1h30

*Liste des documents autorisés : aucun.**Liste des matériels autorisés : calculatrice de type Casio FX 92.**Nombre de pages : 7 (6 feuilles de sujet et une feuille de réponses).**Chaque question attend **une seule** bonne réponse. Chaque réponse juste rapporte un point. Chaque réponse fausse enlève 0,25 point. L'absence de réponse n'enlève pas de point.**Les quatre parties du sujet sont indépendantes.*

*Vos réponses définitives doivent être transcrites dans la feuille de réponse située en dernière page, où vous collerez votre étiquette d'anonymat. **Seule cette feuille de réponse sera rendue.** Des zones de brouillon sont disposées dans le sujet pour faire vos calculs. Vous pouvez déborder de ces zones ou écrire au verso des feuilles du sujet. Ces calculs ne seront pas du tout considérés pour la notation étant donné que vous garderez le sujet.*

Vous ne devez pas dégrafer entre elles les feuilles du sujet. Seule la feuille de réponses doit être dégrafée soigneusement juste avant la fin de l'épreuve pour y transcrire vos réponses puis être rendue aux surveillants.

Exercice A

On dispose de données sur 31 appartements en location situés dans une ville américaine. Les variables considérées dans cet exercice sont les suivantes :

- **loyer** : loyer mensuel (en dollars),
- **surface** : surface de l'appartement (en m^2).

Partie 1

Dans cette partie, les valeurs de la variable **loyer** ont été regroupées en classes dans le tableau suivant (une des cases a été volontairement laissée vide) :

Loyer	[500 ; 600[[600 ; 700[[700 ; 800[[800 ; 900[[900 ; 1000[[1000 ; 1300[
Effectif n_j	3	7	6	6	5	4
Fréquence relative cumulée ϕ_j	0,10	0,33	0,52	0,71	0,87	1
Masse $n_j c_j$	1650	4550	4500	5100	4750	4600
Masse relative cumulée M_j	0,07	0,25		0,63	0,82	1
$f_j \times (M_{j-1} + M_j)$	0,01	0,07	0,13	0,20	0,23	0,24

Question 1 La médiale du loyer (arrondie à 2 décimales) vaut

- A aucune des autres réponses n'est correcte.
- B 850,00 dollars.
- C 835,00 dollars.
- D 789,47 dollars.

Question 2 Laquelle de ces affirmations est **fausse** ?

- A 63 % de la masse totale des loyers est concentrée par 71 % des appartements.
- B Les 29 % des appartements dont le loyer est supérieur à 800 dollars concentrent 37 % de la masse totale des loyers.
- C Les 33 % des appartements dont le loyer est le moins élevé représentent 25 % de la masse totale des loyers.
- D Les 13 % des appartements dont le loyer est le plus élevé représentent 18 % de la masse totale des loyers.

Question 3 L'indice de Gini (arrondi à 2 décimales) est égal à

- A 0,12. B 0,56. C 0,88. D 0,44.

Question 4 D'après le résultat de la question précédente, comment peut-on qualifier la répartition des loyers dans cet échantillon ?

- A plutôt égalitaire.
- B aucune des autres réponses n'est correcte.
- C plutôt inégalitaire.
- D parfaitement égalitaire.

Zone de brouillon

Partie 2

Dans cette partie, on s'intéresse à la liaison entre la surface (notée X) et le loyer (noté Y). On dispose des quantités suivantes :

$$\sum_{i=1}^n x_i = 1\,170, \quad \sum_{i=1}^n y_i = 25\,250, \quad \sum_{i=1}^n x_i y_i = 982\,677, \quad \sum_{i=1}^n x_i^2 = 45\,236, \quad \sum_{i=1}^n y_i^2 = 21\,585\,338.$$

Question 5 Soient X et Y deux variables quantitatives. On a la propriété suivante :

- A $Var(X + Y) = Var(X) + Var(Y) - Cov(X, Y)$.
 B $Var(X + Y) = Var(X) + Var(Y) + Cov(X, Y)$.
 C $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$.
 D $Var(X + Y) = Var(X) + Var(Y) - 2Cov(X, Y)$.

Question 6 Soient a, b, c et d des constantes réelles. On note $r(X, Y)$ le coefficient de corrélation linéaire entre X et Y . On a la propriété suivante :

- A $r(aX + b, cY + d) = r(X, Y)$.
 B $r(aX + b, cY + d) = \left| \frac{ac}{ac} \right| r(X, Y)$.
 C $r(aX + b, cY + d) = \frac{ac}{|ac|} r(X, Y)$.
 D $r(aX + b, cY + d) = ac r(X, Y)$.

Question 7 Lorsque les coefficients de la régression de Y sur X sont estimés par les moindres carrés,

- A la somme des carrés des résidus est nulle.
 B tous les résidus sont nuls.
 C aucune des autres réponses n'est correcte.
 D la variance des résidus est nulle.

Question 8 La covariance entre X et Y (arrondie à 2 décimales) est égale à

- A 922 242,46. B -921 284,61. C 957,84. D 31 699,26.

Question 9 En arrondissant les calculs intermédiaires et les résultats obtenus à 2 décimales, l'équation de la droite de régression de Y sur X est

- A $y = 27,43 x - 220,69$.
 B $y = 0,03 x + 813,39$.
 C $y = -220,69 x + 27,43$.
 D $y = 0,89 x + 780,93$.

Zone de brouillon

Question 10 Le coefficient de détermination associé à la régression de Y sur X (arrondi à 2 décimales) est compris entre

A 0,72 et 0,75.

C 0,93 et 0,96.

B 0,88 et 0,91.

D 0,78 et 0,81.

Zone de brouillon

Exercice B

On considère 342 clients d'un magasin de chaussures pour lesquels on a observé les variables suivantes :

- **montant** : montant dépensé au mois de mars 2018 dans le magasin (en euros),
- **formation** : niveau d'études (en 3 modalités : licence, master et doctorat),
- **revenu** : revenu annuel en euros (en 4 modalités : $< 50\ 000$, entre 50 000 et 100 000, entre 100 000 et 150 000, $\geq 150\ 000$.)

Toutes les valeurs proposées comme réponses dans cet exercice sont obtenues en arrondissant les calculs intermédiaires et le résultat final à 2 décimales.

Partie 3

Dans cette partie, on s'intéresse à la liaison entre le niveau d'études (noté X) et le revenu (noté Y). On obtient le tableau de contingence suivant :

	$< 50\ 000$	entre 50 000 et 100 000	entre 100 000 et 150 000	$\geq 150\ 000$
licence	46	126	32	43
master	12	37	21	16
doctorat	0	1	1	7

Question 11 Laquelle de ces affirmations est **fausse** ?

A 72,22 % des clients ont le niveau licence.

B 79,31 % des clients dont les revenus sont inférieurs à 50 000 euros ont le niveau licence.

C 18,62 % des clients de niveau licence ont des revenus inférieurs à 50 000 euros.

D 15,25 % des clients ont le niveau licence et des revenus inférieurs à 50 000 euros.

Question 12 Pour le couple de modalités (doctorat, $\geq 150\ 000$), l'effectif théorique attendu en cas d'indépendance parfaite des deux variables est égal à

- A 1,74. B 1,82. C 1,79. D 1,86.

Question 13 Laquelle de ces affirmations est **exacte** ?

- A La contribution au χ^2 du couple de modalités (doctorat, $\geq 150\ 000$) est égale à 15,90 et l'association entre les deux modalités est positive.
- B La contribution au χ^2 du couple de modalités (doctorat, $\geq 150\ 000$) est égale à 3,03 et l'association entre les deux modalités est négative.
- C La contribution au χ^2 du couple de modalités (doctorat, $\geq 150\ 000$) est égale à 3,03 et l'association entre les deux modalités est positive.
- D La contribution au χ^2 du couple de modalités (doctorat, $\geq 150\ 000$) est égale à 15,90 et l'association entre les deux modalités est négative.

Question 14 Sachant que $\chi^2 = 27,72$, la valeur du C de Cramer est

- A 0,08. B 0,20. C 0,28 D 0,04.

Question 15 Sous Excel, on peut réaliser un tableau de contingence à l'aide

- A d'un tableau croisé évolutif. C de la fonction `table()`.
- B d'un tableau croisé actif. D d'un tableau croisé dynamique.

Zone de brouillon

Partie 4

Dans cette partie, on s'intéresse à la liaison entre le montant dépensé (noté Y) et le revenu (noté X). On dispose du tableau suivant, où les colonnes Moyenne et Variance contiennent des indicateurs relatifs au montant dépensé :

Modalité du revenu	Effectif	Moyenne	Variance
< 50 000	58	118,83	2742,97
Entre 50 000 et 100 000	164	159,20	2766,93
Entre 100 000 et 150 000	54	222,67	1859,85
$\geq 150\ 000$	66	248,15	4710,46

Question 16 Le type de la variable **revenu** est

- A quantitatif continu. C qualitatif ordinal.
 B quantitatif discret. D qualitatif nominal.

Question 17 La colonne Moyenne du tableau ci-dessus représente

- A les moyennes conditionnelles de X sachant Y .
 B les moyennes conditionnelles de Y sachant X .
 C les moyennes marginales.
 D aucune des autres réponses n'est correcte.

Question 18 La variance INTRA est égale à

- A la moyenne des variances conditionnelles pondérée par les effectifs.
 B la variance des moyennes conditionnelles.
 C la somme de la variance totale et de la variance INTER.
 D la variance des variances conditionnelles.

Question 19 La valeur de la variance INTER est

- A 3 020,05. B 2 607,87. C 2 994,71. D 2 025,60.

Question 20 La valeur du rapport de corrélation η^2 est

- A 0,46. B 0,40. C 0,68. D 0,60.

Zone de brouillon



Licence 1 mention Economie, parcours Economie et Gestion
 Licence 1 mention Economie, parcours Economie et Droit
 Epreuve : Statistique descriptive

Feuille de réponses

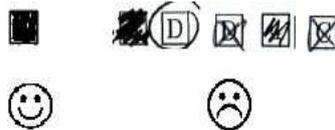
Ne rien écrire ni dessiner sur cette feuille en dehors de ce qui vous est demandé. Sinon votre copie ne sera pas corrigée.

0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9

← Ecrivez ci-contre votre numéro d’anonymat puis noircissez dans l’ordre (de gauche à droite) les 8 cases correspondant aux **8 derniers chiffres de votre numéro d’anonymat** (1ère colonne = 1er chiffre, 2ème colonne = 2ème chiffre, etc.), **en remplissant complètement chaque case**. Collez ci-dessous votre étiquette ↓

Espace dédié à l’étiquette d’anonymat

Pour chaque question, vous devez noircir **une seule** réponse, **en remplissant bien toute la case** pour que la réponse soit prise en considération lors de sa lecture automatique. **Utilisez le correcteur blanc** pour effacer une réponse erronée mais dans ce cas-là, ne redessinez pas la case.



- QUESTION 1 : A B C D
- QUESTION 2 : A B C D
- QUESTION 3 : A B C D
- QUESTION 4 : A B C D
- QUESTION 5 : A B C D
- QUESTION 6 : A B C D
- QUESTION 7 : A B C D
- QUESTION 8 : A B C D
- QUESTION 9 : A B C D
- QUESTION 10 : A B C D

- QUESTION 11 : A B C D
- QUESTION 12 : A B C D
- QUESTION 13 : A B C D
- QUESTION 14 : A B C D
- QUESTION 15 : A B C D
- QUESTION 16 : A B C D
- QUESTION 17 : A B C D
- QUESTION 18 : A B C D
- QUESTION 19 : A B C D
- QUESTION 20 : A B C D