

Année universitaire 2015–2016

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 : jeudi 19 mai 2016

Durée de l'épreuve : 1h30

Liste des documents autorisés : aucun.

Liste des matériels autorisés : calculatrice de type Casio FX 92.

Nombre de pages : 8 (7 pages de sujet et une grille 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 grille de réponse située en dernière page, où vous collerez votre étiquette d'anonymat. **Seule cette grille 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 7 feuilles du sujet. **Seule la grille de réponses doit être dégrafée soigneusement à la fin pour y transcrire vos réponses puis être rendue aux surveillants.***

Exercice A

Une compagnie d'assurance dispose de données concernant 2220 jeunes conducteurs en 2009. Les caractéristiques recueillies sont les suivantes :

- **formule** : formule du contrat d'assurance en 3 modalités (C : minimum, B : moyenne, A : élevée),
- **valeur** : valeur marchande du véhicule assuré en 4 modalités (1 : faible valeur, 2 : valeur moyenne, 3 : valeur élevée, 4 : valeur très élevée),
- **anciennete** : ancienneté du véhicule assuré (en années).

Partie 1

Dans cette partie, on s'intéresse à la liaison entre l'ancienneté du véhicule et la formule d'assurance.

On donne les sorties du logiciel R suivantes :

```
> table(formule)
formule
  A   B   C
665 881 674

> tapply(anciennete, formule, mean)
  A     B     C
2.33 8.87 14.83
```

La sortie de R ci-dessous fournit les sommes des carrés des anciennetés des véhicules pour chaque formule d'assurance.

```
> tapply(anciennete^2, formule, sum)
  A         B         C
6332      80651     158056
```

Question 1 Quel est le graphique adapté pour représenter la liaison entre l'ancienneté et la formule d'assurance ?

- A des diagrammes en tuyaux d'orgues empilés.
 B des diagrammes en colonnes juxtaposés.
 C des boîtes à moustaches juxtaposées.
 D aucune des autre réponses n'est correcte.

Question 2 Que vaut l'ancienneté moyenne des véhicules pour l'ensemble des assurés (arrondie à 2 décimales) ?

- A 8,72 années. C 8,87 années.
 B 8,68 années. D aucune des autres réponses n'est correcte.

Zone de brouillon

Question 3 Quelle est la valeur de la variance conditionnelle de l'ancienneté sachant la formule d'assurance C (arrondie à 2 décimales) ?

- A 14,83. B 234,50. C 71,20. D 14,58.

Zone de brouillon

Question 4 Quelle est la valeur de la variance intra-groupes (arrondie à 2 décimales) ?

- A 34,34. B 10,76. C 110,38. D 23,58.

Zone de brouillon

Question 5 Quelle est la valeur du rapport de corrélation (arrondie à 2 décimales) ?

- A 0,31. B 0,46. C 0,50. D 0,69.

Zone de brouillon

Partie 2

Dans cette partie, on s'intéresse à la liaison entre la valeur du véhicule et la formule d'assurance.

Question 6 La variable valeur est

- A qualitative nominale. C quantitative continue.
 B qualitative ordinale. D quantitative discrète.

Question 7 Le tableau suivant, qui a été obtenu à partir des données, est

	formule		
valeur	A	B	C
1	0.00	0.52	0.48
2	0.28	0.40	0.32
3	0.38	0.41	0.21
4	0.63	0.20	0.17

- A le tableau des effectifs théoriques. C le tableau des profils-colonnes.
 B le tableau de contingence. D le tableau des profils-lignes.

Question 8 La commande du logiciel R qui permet d'obtenir le tableau de la question précédente est

- A `table(valeur, formule)`
 B `prop.table(table(valeur, formule), 1)`
 C `prop.table(table(valeur, formule), 2)`
 D `chisq.test(table(valeur, formule))$expected`

Question 9 On dispose du tableau des contributions au χ^2 suivant (arrondies à deux décimales) :

	formule		
valeur	A	B	C
1	46.73	5.89	16.13
2	2.28	0.04	1.63
3	7.80	0.19	10.72
4	?	14.75	8.63

Sachant que $n_{41}=93$, $n_{4.}=147$ et $n_{.1}=665$, la valeur manquante vaut

- A 0,14. B 44,03. C 48,97. D 54,45.

Zone de brouillon

Question 10 La valeur du C de Cramer (arrondie à 3 décimales) est

- A 0,195. B 0,192. C 0,189. D 0,276.

Zone de brouillon

Exercice B

On considère les salaires mensuels nets en euros en 2001 (salaire01 noté X) et en 2002 (salaire02 noté Y) pour 1000 salariés français.

Partie 3

Question 11 Un coefficient de corrélation linéaire

- A est toujours dans l'intervalle $[0; 1]$.
- B permet de mesurer le lien entre deux variables qualitatives.
- C est multiplié par 2 quand les deux variables sont multipliées par $\sqrt{2}$.
- D peut être négatif.

Question 12 L'inégalité de Cauchy-Schwarz est

- A $Var(X + Y) \leq Var(X) + Var(Y)$.
- B $Cov(X, Y) \geq \sigma_X \sigma_Y$.
- C $|Cov(X, Y)| \geq \sigma_X \sigma_Y$.
- D $|Cov(X, Y)| \leq \sigma_X \sigma_Y$.

Question 13 Le coefficient de détermination est

- A le carré du coefficient de corrélation linéaire.
- B la valeur absolue du coefficient de corrélation linéaire.
- C le rapport entre la variance des résidus et la variance totale.
- D une mesure de la liaison entre une variable quantitative et une variable qualitative.

Question 14 Lorsque les coefficients de régression de Y sur X sont estimés par moindres carrés,

- A la moyenne des valeurs ajustées est égale à la moyenne des résidus.
- B la moyenne des résidus est égale à la moyenne des valeurs de X .
- C la moyenne des résidus est égale à la moyenne des valeurs de Y .
- D la moyenne des valeurs de Y est égale à la somme de la moyenne des résidus et de la moyenne des valeurs ajustées.

Pour vous permettre de répondre aux deux questions suivantes, on vous donne l'écart-type de la variable X qui vaut 672,26 euros (arrondi à deux décimales) et celui de Y qui vaut 699,85 euros (arrondi à deux décimales) ainsi que la sortie du logiciel R ci-dessous.

Call:

```
lm(formula = salaire02 ~ salaire01)
```

Coefficients:

(Intercept)	salaire01
177.76	0.94

Call:

```
lm(formula = salaire01 ~ salaire02)
```

Coefficients:

(Intercept)	salaire02
103.96	0.87

Question 15 L'équation de la droite de régression de Y sur X est

- A $y = 177,76x + 0,94$.
- B $y = 0,94x + 177,76$.
- C $y = 103,96x + 0,87$.
- D $y = 0,87x + 103,96$.

Question 16 Le coefficient de corrélation linéaire entre X et Y (arrondi à deux décimales) vaut

- A 0,90. B 0,94. C 0,87. D 0,98.

Zone de brouillon

Partie 4

Les données de la variable X ont été regroupées en classes dans le tableau suivant et on s'intéresse à la concentration des salaires.

Salaire en 2001	[0,1000[[1000,1500[[1500,2000[[2000,2500[[2500,6100]
Effectif n_j	299	384	190	76	51
Proportion cumulée ϕ_j	0,30	0,68	0,87	0,95	1
Masse $n_j c_j$	149 500	?	332 500	171 000	219 300
Masse relative cumulée M_j	0,11	0,47	0,71	0,84	1
$f_j * (M_{j-1} + M_j)$	0,03	0,22	0,22	0,12	0,09

Question 17 Que vaut la masse salariale totale M ?

- A on ne peut pas la trouver avec les données fournies. C 1 352 300 euros.
 B 219 300 euros. D aucune des autres réponses n'est correcte.

Zone de brouillon

Question 18 Laquelle de ces affirmations est fausse ?

- A les 68% des salariés les moins payés gagnent 47% de la masse salariale totale.
 B les 5 % des salariés les mieux payés gagnent 16 % de la masse salariale totale.
 C 71 % de la masse salariale est gagnée par 87 % des salariés.
 D les 29 % des salaires les plus élevés sont gagnés par 23 % des salariés.

Question 19 Quelle est la valeur du salaire médial (arrondi à deux décimales) ?

A 1500,88 euros.

C 1236,84 euros.

B 1750 euros.

D aucune des autres réponses n'est correcte.

Zone de brouillon



Question 20 Combien vaut l'indice de Gini (arrondi à deux décimales) ?

A il est compris entre 0,30 et 0,32.

B il est compris entre 0,64 et 0,66.

C il est compris entre 0,33 et 0,36.

D aucune des autres réponses n'est correcte.

Zone de brouillon



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

GRILLE DE RÉPONSES

← 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 ↓

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

Espace dédié à l’étiquette d’anonymat

Pour chacune de ces 20 questions, vous devez noircir **une seule** des 4 réponses proposées, si possible au feutre noir, **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