Partiel Statistique

Première année GEA

Année 2014 2015

# Partie 1

## Application 1 (5 points)

A partir des données du questionnaire promo GEA les individus aux yeux marron ont été sélectionnés. Voici le poids des 48 étudiants concernés :

|  |
| --- |
| 48 |
| 50 |
| 50 |
| 50 |
| 50 |
| 50 |
| 50 |
| 51 |
| 52 |
| 52 |
| 52 |
| 54 |
| 54 |
| 54 |
| 55 |
| 55 |
| 57 |
| 57 |
| 58 |
| 58 |
| 58 |
| 59 |

|  |
| --- |
| 70 |
| 72 |
| 74 |
| 75 |
| 75 |
| 76 |
| 78 |
| 80 |

|  |
| --- |
| 60 |
| 60 |
| 60 |
| 60 |
| 60 |
| 62 |
| 62 |
| 63 |
| 63 |
| 65 |
| 65 |
| 65 |
| 65 |
| 65 |
| 65 |
| 68 |
| 68 |
| 69 |

**Travail à faire**

1. **Calculez avec la formule de Yule, l’intervalle des classes (0,5 points)**
2. **Etablissez le tableau statistique en regroupant les données poids en classe d’amplitude 5 (0,5 points)**
3. **Calculez les fréquences et les fréquences cumulées (1 points)**
4. **Dressez l’histogramme des fréquences et le polygone des fréquences cumulées (0,5 points)**
5. **Déterminez graphiquement la médiane et les quartiles Q1 et Q3 (1 point)**
6. **Calculez par interpolation linéaire la médiane et les quartiles Q1 et Q3 (1,5 points)**

Vous pourrez utiliser le tableau ci après

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Poids | ni | fi | ficum |  |  |  |  |
| [ 48 - 53 [  | 11 | 0,23 | 0,23 |  |  |  |  |
| [ 53 - 58 [ | 7 | 0,15 | 0,38 |  |  |  |  |
| [ 58 - 63 [ | 11 | 0,23 | 0,60 |  |  |  |  |
| [ 63 - 68 [  | 8 | 0,17 | 0,77 |  |  |  |  |
| [ 68 - 73 [  | 5 | 0,10 | 0,88 |  |  |  |  |
| [ 73 - 78 [ | 4 | 0,08 | 0,96 |  |  |  |  |
| [ 78 - 83 [ | 2 | 0,04 | 1,00 |  |  |  |  |
| total | 48 |  |  |  |  |  |  |

1)

 Sachant que Nombre de classes = $2,5\sqrt[4]{n}$

N étant de 48

$$2,5\sqrt[4]{48=}6,58$$

$$int=\frac{(x\_{max}-x\_{min})}{nb classes}$$

80-48/6,58=4,86

6) Q1=0,25 ( Q1-52)/(57-52)= (0,25-0,23)/(0,38-0,23)

52=0,23 Q1= 0,67+52= 52,67 On retient 53

57=0,38

Me=0,5 ( Me-57)/(62-57)= (0,5-0,38)/(0,61-0,38)

57=0,38 Q1= 2,61+57= 59,61 On retient 59

62=0,61

Q3=0,75 ( Q3-62)/(67-62)= (0,75-0,6)/(0,77-0,6)

62=0,6 Q3= 4,41+62= 66,41 on retient 66

67=0,77

## Application 2 (3 points)

A partir des données du questionnaire promo GEA les individus aux yeux marron ont été sélectionnés. Voici la taille des 48 étudiants concernés regroupés en classe :

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Taille | ni | fi | ficum | ci | fici | Fi(ci-xbar)2 | densité |
| [157-162[  | 5 | 0,10 | 0,10 | 159 | 15 ,9 | 15,876 | 0,02 |
| [162-167[ | 9 | 0,19 | 0,29 | 164 | 30,75 | 10,9744 | 0,038 |
| [167-172[ | 10 | 0,21 | 0,50 | 169 | 35,21 | 1,4196 | 0,042 |
| [172-177[ | 11 | 0,23 | 0,73 | 174 | 39,88 | 1,3248 | 0,046 |
| [177-182[ | 8 | 0,17 | 0,90 | 179 | 29,83 | 9,3092 | 0,034 |
| [182-191[ | 5 | 0,10 | 1 | 186 | 19,38 | 20,736 | 0,011 |
| total | 48 |  |  |  | 171,6 | 59,64 var7,72 et |  |

**Travail à faire**

1. **Calculez les fréquences et les fréquences cumulées (0,5 points)**
2. **Etablissez l’histogramme de cette série (1 point)**
3. **Calculez la moyenne et l’écart type de la série (1,5 points)**

# Partie 2

## Application 3(5,5 points)

A partir des extraits de données du questionnaire promo GEA

**Travail à faire**

1. **Calculer la moyenne et l’écart type pour les variables nombre d’heures sur les RS et nombre de textos envoyés (1,5 points)**
2. **Calculez la covariance et le coefficient de corrélation linéaire entre ces deux variables (2,5 points)**
3. **Que peut -on conclure ? (1,5 points)**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Nombre d’heure RS | NB textos | (x-x bar) | (y-ybar) | (x-xbar)2 | (y-ybar)2 | (x-x bar) (y-ybar) |
| une fille | 1 | 15 | -2,63 | -48,13 | 6,92 | 2316,50 | 126,58 |
| une fille | 1 | 10 | -2,63 | -53,13 | 6,92 | 2822,80 | 139,73 |
| une fille | 1 | 40 | -2,63 | -23,13 | 6,92 | 535,00 | 60,83 |
| une fille | 2 | 5 | -1,63 | -58,13 | 2,66 | 3379,10 | 94,75 |
| une fille | 2 | 150 | -1,63 | 86,87 | 2,66 | 7546,40 | -141,60 |
| une fille | 2 | 15 | -1,63 | -48,13 | 2,66 | 2316,50 | 78,45 |
| une fille | 2 | 50 | -1,63 | -13,13 | 2,66 | 172,40 | 21,40 |
| une fille | 2 | 300 | -1,63 | 236,87 | 2,66 | 56107,40 | -386,10 |
| une fille | 3 | 30 | -0,63 | -33,13 | 0,40 | 1097,60 | 20,87 |
| une fille | 3 | 30 | -0,63 | -33,13 | 0,40 | 1097,60 | 20,87 |
| une fille | 4 | 15 | 0,37 | -48,13 | 0,14 | 2316,50 | -17,81 |
| une fille | 4 | 50 | 0,37 | -13,13 | 0,14 | 172,40 | -4,86 |
| une fille | 5 | 40 | 1,37 | -23,13 | 1,88 | 535,00 | -31,69 |
| une fille | 5 | 30 | 1,37 | -33,13 | 1,88 | 1097,60 | -45,39 |
| une fille | 6 | 30 | 2,37 | -33,13 | 5,62 | 1097,60 | -78,52 |
| une fille | 15 | 200 | 11,37 | 136,87 | 129,28 | 18733,40 | 1556,21 |
|  | 3,63 | 63,13 |  |  | 10,86 var | 6333,98 var | 88,36 cov |
|  |  |  |  |  | 3,30 et | 79,59 et | 0,34 r |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Nombre d’heure RS | NB textos | (x-x bar) | (y-ybar) | (x-xbar)2 | (y-ybar)2 | (x-x bar) (y-ybar) |
| un garçon | 0 | 3 | -4,44 | -37,94 | 19,71 | 1439,44 | 168,45 |
| un garçon | 1 | 10 | -3,44 | -30,94 | 11,83 | 957,28 | 106,43 |
| un garçon | 1 | 2 | -3,44 | -38,94 | 11,83 | 1516,32 | 133,95 |
| un garçon | 1 | 15 | -3,44 | -25,94 | 11,83 | 672,88 | 89,23 |
| un garçon | 2 | 20 | -2,44 | -20,94 | 5,95 | 438,48 | 51,09 |
| un garçon | 2 | 10 | -2,44 | -30,94 | 5,95 | 957,28 | 75,49 |
| un garçon | 3 | 70 | -1,44 | 29,06 | 2,07 | 844,48 | -41,85 |
| un garçon | 3 | 10 | -1,44 | -30,94 | 2,07 | 957,28 | 44,55 |
| un garçon | 3 | 40 | -1,44 | -0,94 | 2,07 | 0,88 | 1,35 |
| un garçon | 4 | 20 | -0,44 | -20,94 | 0,19 | 438,48 | 9,21 |
| un garçon | 4 | 35 | -0,44 | -5,94 | 0,19 | 35,28 | 2,61 |
| un garçon | 4 | 30 | -0,44 | -10,94 | 0,19 | 119,68 | 4,81 |
| un garçon | 5 | 30 | 0,56 | -10,94 | 0,31 | 119,68 | -6,13 |
| un garçon | 6 | 50 | 1,56 | 9,06 | 2,43 | 82,08 | 14,13 |
| un garçon | 7 | 10 | 2,56 | -30,94 | 6,55 | 957,28 | -79,21 |
| un garçon | 25 | 300 | 20,56 | 259,06 | 422,71 | 67112,08 | 5326,27 |
|  | 4,44 | 40,94 |  |  | 31,62 var | 4790,56 var | 368,78 cov |
|  |  |  |  |  | 5,62 et | 69,21 et | 0,95 r |

## Application 4 (3 points)

**Travail à faire**

Sachant que les calculs effectués sur **toute** population garçon de la promo première année GEA donne les résultats suivants :

* Var (x) = 17,43 Var (y)=3889,28
* Ecart type (x)= 4,14 Ecart type (y)= 62,36
* Covariance (xy)= 164,34
* R(xy) = 0,63

Sachant que les calculs effectués sur **toute** population fille de la promo première année GEA donne les résultats suivants :

* Var (x)= 14,53 V(y)= 3635,02
* Ecart type (x)= 3,81 Ecart type (y)= 60,29
* Covariance (xy)= 12,18
* R(xy)= 0,05

Que peut-on conclure ?

1. Vous comparerez les résultats que vous avez trouvez sur votre échantillon avec ceux qui concerne la population étudiée dans son ensemble. (1,5 points)

Population Garçon et échantillon garçon : Les écarts types de la population sont plus faibles que sur l’échantillon. La population est donc plus homogène que l’échantillon. Le coefficient de corrélation est positif dans les deux cas. Il existe donc une relation positive entre les deux variables considérées. Plus on passe de temps sur les réseaux sociaux plus on envoie de texto. Cette relation est très forte dans le cas de l’échantillon.

Ces différences peuvent provenir d’une moyenne relativement différente entre l’échantillon et la population. Or la moyenne est la valeur qui minimise l’écart dans le calcul de la variance. Cela expliquerait les différences d’écart type.

On peut supposer que l’échantillon n’est pas représentatif de la population, peut être la sélection n’a pas été aléatoire.

1. Vous comparerez les résultats de la population fille et ceux de la population garçon (1,5 points)

Les populations sont différentes. La population fille est plus homogène. Le fait remarquable est l’absence de corrélation entre les deux variables dans le cas des filles alors qu’il existe une corrélation pour le cas des garçons. Ce résultat est interpellant sur les conclusions que l’on peut avoir. Doit-on considérer qu’il existe une relation linéaire entre ces deux variables ou sont elles indépendantes ? A y réfléchir on ne voit pas en quoi passer du temps sur les réseaux sociaux ait un quelconque lien avec le nombre de textos envoyés sauf à penser qu’il existe un profil que l’on pourrait appelé technologique ie adepte de la relation à distance. Envoyer des textos et passer du temps sur les réseaux sociaux c’est une même façon de communiquer et d’être en relation avec les autres. C’est une relation qui est médiée par la technologie.

Peut être faut il considérer l’utilisation du réseau social…l’utilisation est elle pour le tchat ? ou les images… la question est que fait on quant on passe du temps sur les réseaux sociaux ? c’est cette utilisation qui déterminera s’il existe un lien ou non entre les deux variables étudiées.

## Application 5 (3,5 points)

**Travail à faire**

1. A partir des résultats effectués sur toute la population garçon vous justifierez l’existence d’une relation linéaire entre les variables (1,5 points)

Lorsque deux variables quantitatives sont correctement correlées (|rxy | voisin de 1) et que l’on peut considérer, a priori, que l’une (nous supposerons qu’il s’agit de x) est cause de l’autre (il s’agira donc de Y ), il est alors assez naturel de chercher une fonction de x approchant y Le coefficient de corrélation est positif. Il existe à priori une relation linéaire entre les deux variables

1. Vous calculerez l’équation de la droite par la méthode des moindres carrés (2 points)

Pour calculer à partir de la population garçon totale il aurait fallu avoir la moyenne… je ne vous l’avais pas donné donc vous ne pouviez que calculer a et écrire l’équation avec b générique. Ceux qui ont trouvé b et écrit l’équation ont eu 1,75/2 et ceux qui en sus ont mentionné le manque des données sur la moyenne on eu 2/2

$$\hat{a}=\frac{s\_{xy}}{s^{2}\_{x}}$$

a=164,34/17,43 =9,43

$$\hat{a}=\frac{\sum\_{i+1}^{n}\left(x\_{i}-\overbar{x}\right)\left(y\_{i}-\overbar{y}\right)}{\sum\_{I=1}^{n}\left(x\_{i}-\overbar{x}\right)^{2}}$$

b=67,86 – 9,43\*2,91=

$$\hat{b}=\overbar{y}-\hat{a}\overbar{x}$$

y= 9,43 x+ 27

Calcul à partir des données de l’échantillon garçon

a=368,78/ 31,62=11,66

b =40,94 – 11,66\*4,44 =-10,83

y=11,66x – 10,83