Année universitaire 2015/2016

CONFERENCE DE STATISTIQUES

TUTORAT DE LA FACULTE DES SCIENCES ECONOMIQUES

Louis SIRUGUE : louis.sirugue@etudiant.univ-rennes1.fr

Maxime BESRET : maxime.besret@etudiant.univ-rennes1.fr

Quentin LE CARRER : quentin.lecarrer@etudiant.univ-rennes1.fr

Pierre MAYOT : pierre.mayot@etudiant.univ-rennes1.fr

Ceci est le sujet de la conférence. Il est possible de le retrouver – ainsi que sa correction – sur l'espace Tutorat de Moodle.

La conférence dure deux heures. Vous composerez pendant quarante-cinq minutes puis les tuteurs réaliseront la correction. N'hésitez pas à leur poser des questions.

Pour nous contacter, veuillez utiliser les adresses mail étudiantes des tuteurs. Le Forum en ligne peut également vous servir à poser vos questions.

Facebook : Tutorat Eco Rennes I

Site internet : http://tutoecogestionl1.forumactif.org/

Mail : associationarte@gmail.com

Partie I : Taux et indices

Exercice 1 : Taux de croissance appliqué à une seule variable

Le 1er septembre 2016, compte tenu de sa réussite fulgurante, l'Association des Tuteurs en Economie de Rennes devient une société et entre en bourse. Elle met en vente des actions de l'entreprise ARTE nouvellement créée au prix initial de 8,00€. On constate que quatre ans plus tard, l'action atteint la valeur de 11,50€.

a. Quelle a été la variation absolue de l'action ARTE sur la période 2016 – 2020 ?

La variation absolue est la différence entre la valeur d'arrivée et la valeur de départ. On a donc :

$V_t - V_0$=Variation absolue= 11,50-8,00=3,5 → Le prix de l'action ARTE a augmenté de 3,50€.

b. Quel a été le taux de croissance global de l'action ARTE sur cette même période ?

On applique la formule du taux de croissance global : $g_{Global} = \frac{V_t - V_0}{V_0}$; ce qui nous donne $\frac{11,5-8}{8} = 0,4375$ → Le taux de croissance global de la valeur de l'action a été de 43,75% sur la période

c. Quel a été le taux de croissance annuel moyen de l'entreprise sur cette période ?

Formule du taux de croissance annuel moyen (TCAM) = $\overline{g} = \left(\frac{V_t}{V_0}\right)^{\frac{1}{t}}-1$; sachant que t=4, on a donc :

$\overline{g} = \left(\frac{11,5}{8}\right)^{\frac{1}{4}}-1 = 0,095$ → Le TCAM sur la période 2016 – 2020 a été de 9,5%. C'est-à-dire que tous les ans, la valeur de l'action a augmenté de 9,5% si on considère que l'augmentation a été linéaire.

d. Si le taux de croissance restait le même, au bout de combien de temps l'entreprise aurait-elle doublé le cours de son action ?

Pour répondre, on doit poser l'équation. On cherche t, tel que $V_t = 2 \times V_0$.

On a donc $(1 + 0,095)^t \times V_0 = 2 \times V_0 \Leftrightarrow (1,095)^t \times 8 = 2 \times 8 \Leftrightarrow (1,095)^t = 2$.

En utilisant les logarithmes, on obtient $t \times ln(1,095) = ln(2) \Leftrightarrow t = \frac{ln\,2}{ln\,1,095} = 7,64$.

A partir de t=8 - la huitième année -, le cours de l'action sera supérieur à deux fois sa valeur initiale.

Finalement, en 2020, la présidente de l'association - Maëlle Le Gall -, lance un nouveau système de tutorat par vidéoconférences dans le monde entier. L'effet se fait sentir sur le cours de l'action et son taux de croissance atteint 15% en 2021 et 20% en 2022.

e. Quel est le multiplicateur composite sur la période 2016 – 2022 ?

Le multiplicateur composite est le produit des multiplicateurs. Il est donc égal à :

$$(1 + 0,095) \times (1 + 0,095) \times (1 + 0,095) \times (1 + 0,095) \times (1 + 0,15) \times (1 + 0,20)$$

Ce qui revient à $(1 + 0,095)^4 \times (1 + 0,15) \times (1 + 0,20) \approx 1,98$.

f. Quel a donc été le taux de croissance global ?

Taux de croissance : $g = (Multiplicateur) - 1 = 0,98$.

Exercice 2 : Taux de croissance appliqué aux grandeurs liées

En 2022, compte tenu de l'expansion internationale de la société, la direction prend la décision de rémunérer ses tuteurs à hauteur de 100€ par mois. Sachant que le seul bien de consommation des tuteurs est le café de la faculté (0,40€) :

a. Quelle est la capacité de consommation des tuteurs (en nombre de cafés) en 2022 ?

$Capacité\ de\ consommation = \frac{salaire}{prix\ du\ bien\ consommé}$. Donc, la CC des tuteurs est de $\frac{100}{0,40} = 250$.

Compte tenu de leur salaire, les tuteurs peuvent consommer 250 cafés par mois.

Voyant bien que cela ne suffisait pas pour maintenir les tuteurs éveillés pendant la journée, Maëlle décide d'augmenter les salaires de 20% pour l'année 2023. En parallèle, Rennes 1 augmente le prix du café à 0,50€.

b. Quel est le taux de croissance du prix du café ?

Le prix du café passe de 0,40€ à 0,50€, son taux de croissance est donc $\frac{0,50-0,40}{0,40} = 0,25$.

c. Quel aura été le taux de croissance de la capacité de consommation des tuteurs entre 2022 et 2023 ?

La formule du taux de croissance appliqué aux grandeurs liées doit être utilisée. On sait que $g_{GL} = \frac{(1+g_A)}{(1+g_B)} - 1$ donc le taux de croissance de la capacité de consommation $g_{cc} = \frac{(1+g_{salaire})}{(1+g_{café})} - 1 = \frac{1,2}{1,25} - 1 = -0,04$. Compte tenu de la hausse des salaires et du prix du café, la capacité de consommation des tuteurs a baissé de 4%.

d. Quelle est la capacité de consommation des tuteurs en 2023 ?

La capacité de consommation initiale était de 250. Après une baisse de 4% de leur CC, cette dernière est donc de $250 \times \left(1 + (-0,04)\right) = 250 \times 0,96 = 240$.

(Question bonus : Quel salaire faudrait-il donner aux tuteurs pour que leur capacité de consommation augmente de 20%, compte tenu de la hausse du prix du café ?)

Pour répondre à cette question, il faut poser la formule du taux de croissance appliqué aux grandeurs liées :

$g_{GL} = \frac{(1+g_A)}{(1+g_B)} - 1$; dans notre cas, on a $g_{cc} = \frac{(1+g_{salaire})}{(1+g_{café})} - 1$. On cherche donc la valeur de $g_{salaire}$ telle que $g_{cc} = 0,2$; pour que la capacité de consommation des tuteurs augmente de 20%. On pose donc

$0,2 = \frac{(1+g_{salaire})}{(1+0,25)} - 1 \Leftrightarrow 1,2 = \frac{(1+g_{salaire})}{(1+0,25)} \Leftrightarrow 1 + g_{salaire} = \frac{1,2}{1,25}$. Donc $g_{salaire} = \frac{1,2}{1,25} - 1 = 0,5$. Il faudrait donc que les salaires augmentent de 50% lorsque le prix du café augmente de 25% pour que la capacité de consommation des tuteurs augmente de 20%. Le salaire qui permettrait une hausse de 20% de cette capacité est 150€ car $100 \times (1 + 0,5) = 150$.

Exercice 3 : Indices élémentaires

Dans le tableau suivant est présentée l'évolution de la demande de bières d'un étudiant en économie pour les mois de septembre à décembre. À l'aide des indices, calculez la variation de la demande pour chaque mois en utilisant le mois de septembre comme base 100.

Mois	Septembre V0	Octobre V1	Novembre V2	Décembre V3
Demande	40	50	60	80
Indice	100	125	150	200

Ici on calcule les indices en prenant comme base septembre de telle manière que l'indice d'octobre soit égal à :

octobre/septembre x 100 = 50/40 x 100 = 125.

On procède de la même manière pour les suivants :

novembre/septembre x 100 = 60/40 x 100 = 150.

décembre/septembre x 100 = 80/40 x 100 = 200.

Exercice 4 : Propriétés

Partie A : Réversibilité

a. Le tableau suivant représente l'évolution des prix du Doliprane. Les étudiants ayant tous trop bu décident de calculer les indices de prix en utilisant le mois d'août comme base 100.

Mois	Août V0	Septembre V1	Octobre V2	Novembre V3
Prix	5	6,25	7,50	8,50
Indice	100	125	150	170

septembre/août x 100 = 6,25/5 x 100 =125.

octobre/août x 100 = 7,50/5 x 100 =150.

novembre/août x 100 = 8,50/5 x 100 = 170.

b. Les tuteurs veulent connaître l'évolution du prix entre septembre et août. Ils demandent donc aux étudiants de calculer par réversibilité l'indice d'août avec septembre comme base (ne pas hésiter à regarder la boîte à outils !).

Dans cette question on cherche à appliquer la propriété de réversibilité. Cette propriété permet le changement de la base de calcul, c'est-à-dire le changement de base 100. Dans cet exemple nous utilisions août comme base 100 mais comme nous voulions connaître l'évolution de septembre à août, il nous fallait donc prendre septembre comme base de travail, comme base 100.
Pour ce faire nous appliquons la formule : Indice V0/V1 = 100^2/ Indice V1/V0.
Dans notre cas on a :

$$I_{\frac{\text{août}}{\text{septembre}}} = \frac{100^2}{I_{\frac{\text{septembre}}{\text{août}}}} = \frac{100^2}{125} = 80$$

L'indice août/septembre est donc égal à 80 avec septembre comme base 100.

Partie B : Circularité

c. Plus tard, nous apprenons que le Doliprane coûtera 20% de plus en février qu'en janvier. Quel sera l'indice correspondant à l'évolution de son prix ?

Ici on nous demande l'indice de février avec janvier en base 100. Nous n'avons pas de données pour calculer l'indice, en revanche nous savons que le prix a augmenté de 20%. Les indices nous permettent donc de déduire le calcul suivant :

janvier base 100 → 20% → février = 100 + 20%(de 100), Indice février/janvier = 120.

$$I_{\frac{\text{février}}{\text{janvier}}} = 100 + 100 \times 20\% = 120$$

d. En mars son prix augmente de 5%. Quelle est l'évolution par rapport à février ?

Même raisonnement que pour la question précédente :

février base 100 → 5% →mars = 100 + 5%(de 100), Indice mars/février = 105.

$$I_{\frac{mars}{février}} = 100 + 100 \times 5\% = 105$$

e. Enfin, les tuteurs de nature très curieuse leur demandent l'indice de janvier à mars (en utilisant la circularité, bien entendu !).

Ici on nous demande l'indice de janvier/mars. Pour ce calcul nous allons utiliser la propriété de circularité. Cette propriété permet de trouver l'indice entre deux périodes sans forcément connaître l'évolution entre ces deux points. Ici par exemple on ne connaît pas la variation du prix entre janvier et mars. En revanche nous connaissons les évolutions qu'il y a entre janvier et mars c'est-à-dire l'indice de janvier à février et de février à mars, avec ces données nous pouvons calculer l'indice par circularité avec la formule suivante : Indice V2/V0 = (Indice V2/V1 x Indice V1/V0) / 100.

Dans notre cas :
Indice mars/janvier = (Indice mars/février x Indice février/janvier) / 100
Indice mars/janvier = (120 x 105) / 100
Indice mars/janvier = 126

$$I_{\frac{mars}{janvier}} = \frac{I_{\frac{mars}{février}} \times I_{\frac{février}{janvier}}}{100} = \frac{105 \times 120}{100} = 126$$

L'indice de mars / janvier est égal à 126 ce qui est cohérent avec les données que nous avons eu.

Partie II : Distributions

Exercice 1 : Représentations graphiques

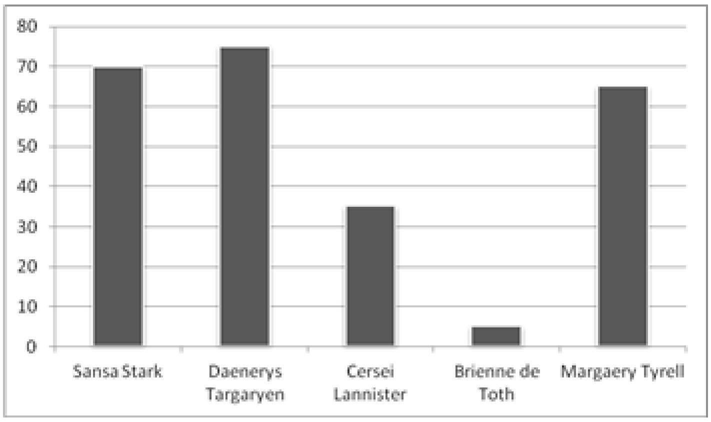
Une étude cherchant à déterminer la popularité des principaux personnages d'une série américaine a été menée auprès d'un groupe d'étudiants. Chacun d'entre eux devait choisir son personnage favori parmi cinq proposés. Les résultats sont les suivants :

Personnages	Sansa Stark	Daenerys Targaryen	Cersei Lannister	Brienne de Torth	Margaery Tyrell
Nombre de choix	70	75	35	5	65

a. Déterminez la population étudiée ainsi que le caractère étudié et sa nature.

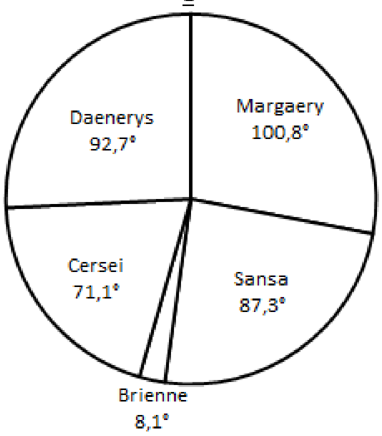
La population étudiée est un groupe d'étudiants. Le caractère étudié est le personnage choisi, de nature qualitative nominale. Le caractère de la variable est de nature qualitative car ce n'est pas la quantité qui varie (comme pour un logement qui pourrait être d'une, deux, trois ou quatre pièces), mais l'essence même du caractère, ici le personnage. Il en serait de même pour la couleur des yeux d'une population par exemple. Nominal ici fait référence au fait qu'il n'y ait pas d'ordre précis entre les variables, en opposition avec une variable ordinale qui peut être hiérarchisée (comme le degré de satisfaction : insatisfait < peu satisfait <satisfait < très satisfait).

b. Représentez cette distribution à l'aide d'un graphique à tuyaux d'orgue réalisé à main levée.

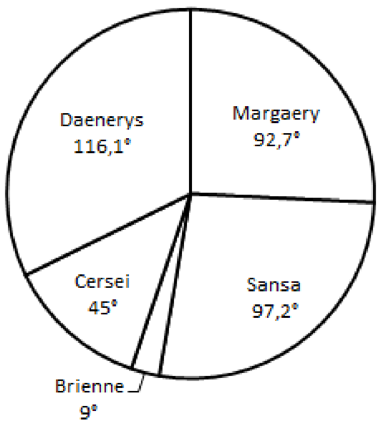


c. Déterminez la fréquence associée à chaque personnage. Déduisez–en le graphique à secteurs circulaires correspondant à cette distribution.

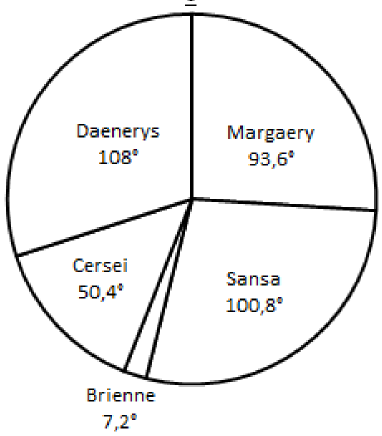
1



2



3



Dans un premier temps, il faut calculer les fréquences associées à chaque personnage pour déterminer la proportion qu'il occupe dans l'effectif. On divise donc l'effectif correspondant au personnage par l'effectif total. Ensuite, il faut multiplier cette fréquence par 360 pour obtenir un angle en degrés proportionnel à sa part de l'effectif total représenté ici par l'aire totale du disque.

Les fréquences :
Sansa Stark : 70/250=0.28
Daenerys Targaryen : 75/250=0.30
Cersei Lannister : 35/250=0.14
Brienne de Torth : 5/250=0.02
Margaery Tyrell : 65/250=0.26

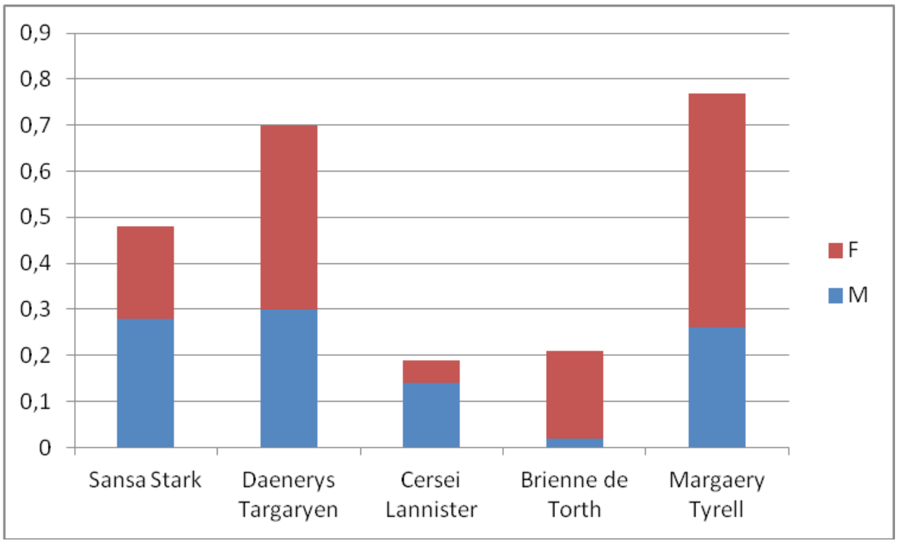
Les angles :
Sansa Stark : 0.28 x 360 = 100,8
Daenerys Targaryen : 0.30 x 360 = 108
Cersei Lannister : 0.14 x 360 = 50,4
Brienne de Torth : 0.02 x 360 = 7,2
Margaery Tyrell : 0.26 x 360 = 93,6

Ainsi, le graphique à secteurs circulaires correspondant à la distribution est le graphique n°3.

(Question bonus : L'étude en question a aussi recensé l'âge des étudiants interrogés. L'effectif a été divisé en deux classes d'âge dont la répartition est la suivante :

Personnages Classes d'âge	Sansa Stark	Daenerys Targaryen	Cersei Lannister	Brienne de Torth	Margaery Tyrell
[15 ; 25[	35	30	15	0	45
[25 ; 35[	35	45	20	5	20
∑	70	75	35	5	65

d. Réalisez un graphique en tuyaux d'orgue composés pour représenter cette répartition par tranche d'âge.)



Exercice 2 : Etude d'une distribution

Un sondage a été réalisé auprès de 400 étudiants concernant le temps hebdomadaire consacré au visionnage de séries pour chacun d'entre eux. Les résultats sont les suivants :

Temps hebdomadaire (en min)	Nombre d'étudiants
[0 ; 45[	64
[45 ; 90[	136
[90 ; 180[	128
[180 ; 360[	72

a. Déterminez la population étudiée ainsi que le caractère étudié et sa nature.

La population étudiée est un groupe de 400 étudiants. Le caractère étudié est le temps hebdomadaire passé à regarder des séries, de nature quantitative continue. Quantitative fait ici référence au fait que la variable est sommable, qu'elle concerne une quantité. La nature continue vient du fait que la variable est étudiée sur des intervalles (entre 0 et 45min, entre 180 et 360min…) et non sur des valeurs ponctuelles (comme pour l'exemple précédent sur le logement).

b. Complétez le tableau statistique standard suivant associé à ces données.

xi	ni	fi	F+	F-	ai	Ci	f'i
[0 ; 45[	64	0,16	0,16	1	45	1	0,16
[45 ; 90[	136	0,34	0,50	0,84	45	1	0,34
[90 ; 180[	128	0,32	0,82	0,50	90	2	0,16
[180 ; 360[	72	0,18	1	0,18	180	4	0,045

On détermine fi en divisant ni par l'effectif total, la somme de tous les ni.

F+ est la somme des fi en partant de l'intervalle le plus faible.

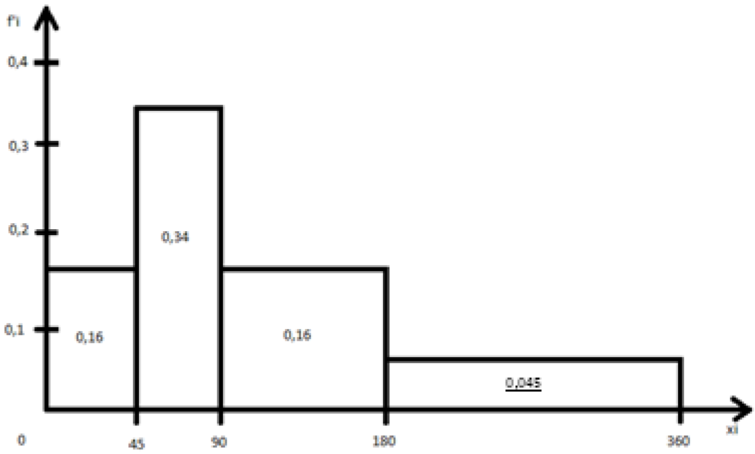
F- est le somme des fi en partant de l'intervalle le plus élevé.

Ai est l'amplitude de l'intervalle (pour [90 ; 180[ai = 180 – 90 = 90).

Ci est l'amplitude de l'intervalle divisée par l'amplitude de référence. Ici on choisit 45 comme amplitude de référence.

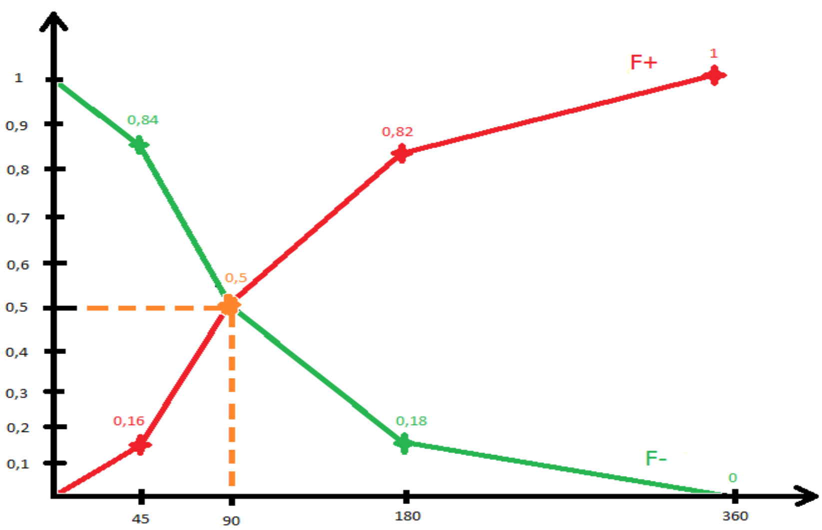
f'i correspond à fi divisé par Ci, cela sert à ajuster la fréquence en fonction des différences d'amplitude des intervalles.

c. Représentez à l'aide d'un histogramme la distribution observée.



Ici on utilise f 'i et non fi en ordonnées car on ajuste la fréquence à la taille des intervalles. En effet une fréquence étalée sur une amplitude valant trois fois celle de référence biaiserait la représentation graphique de la distribution en conférant au rectangle en question une aire plus importante que la place qu'il occupe réellement dans l'effectif total. Ainsi, la largeur des bases des rectangles compte pour que l'aire du rectangle soit bien proportionnelle à sa part de l'effectif total.

d. Représentez graphiquement la répartition de la série statistique.



On représente la répartition à l'aide des fonctions cumulées croissantes et décroissantes. Elles sont symétriques par rapport à 0,5 en ordonnées, c'est-à-dire que la valeur de F+ en un point est égale à la différence entre 1 et la valeur de F- en ce même point et inversement. F+ croît toujours de 0 à 1 tandis que F- décroît toujours de 1 à 0.

e. Déterminez graphiquement les coordonnées du point d'intersection de ces deux fonctions de répartition. Commentez.

Le point d'intersection de ces deux fonctions est de coordonnées (90 ; 0,5). Il correspond à la médiane qui divise la série statistique en deux groupes de même effectif. Ainsi, on peut dire qu'il y a autant d'étudiants qui passent moins de 90 minutes par semaine à regarder des séries que d'étudiants qui y passent entre 90 et 360 minutes.

f. Quel est le temps hebdomadaire minimum passé à regarder des séries pour les 84% d'étudiants qui en regardent le plus ?

Le temps hebdomadaire minimum passé à regarder des séries pour les 84% d'étudiants qui en regardent le plus est de 45 minutes. On peut déduire cela en se basant sur le graphique des fonctions de répartition ou sur le tableau des fréquences cumulées décroissantes.

g. Combien de temps maximum par semaine passent à regarder des séries les 16% d'étudiants qui en regardent le moins ?

Le temps hebdomadaire maximum consacré à regarder des séries par les 16% d'étudiants qui en regardent le moins est de 45 minutes. On peut déduire cela en se basant sur le graphique des fonctions de répartition ou sur le tableau des fréquences cumulées croissantes.

Partie III : Tendances centrales

Exercice 1 : QCM

a. La médiane d'une distribution :

- Est plus élevée que la moyenne.
  Faux
- Est inférieure à la moyenne.
  Faux
- Est égale à la moyenne.
  Faux
- Correspond à la valeur telle que la moitié de la distribution lui sera inférieure et l'autre moitié lui sera supérieure.
  Vrai. C'est exactement la définition d'une médiane. Mais il faut bien comprendre qu'on ne peut pas prévoir de quel côté de la médiane sera la moyenne ! Celle-ci peut très bien être égale, inférieure, ou supérieure. Par exemple le revenu MOYEN en France est de l'ordre de 2300€ mensuels alors que le revenu MEDIAN est de l'ordre de 1700€ mensuels. Cette différence considérable est due aux très hauts revenus.

b. A propos des quantiles :

- Ils permettent de diviser la distribution en parts égales.
  Vrai. C'est leur définition. Les quartiles divisent en quatre parts, les déciles en dix parts, les centiles en cent parts, etc.
- Ils sont utilisés pour réaliser une boîte à moustache.

Vrai. Une boîte à moustache se construit à partir de la médiane, et des quartiles 1 et 3 (ainsi que des valeurs extrêmes de la série et des déciles 1 et 9).

- Ils ne sont pas utilisés pour étudier une variable continue.
  Faux. On peut utiliser les quantiles pour étudier n'importe quelle variable.
- Il en existe seulement quatre : la médiane, le premier quartile, le deuxième quartile et le troisième quartile.
  Faux. En réalité il existe une infinité de quantiles : on peut choisir de diviser la distribution en 3, en 7 ou en 1664 si l'on veut. Mais les indicateurs le plus souvent utilisés sont la médiane, les quartiles, les déciles et les centiles.

c. A propos des différentes moyennes :

- La moyenne Arithmétique est supérieure à la moyenne Géométrique.
  Faux. C'est l'inverse.
- La moyenne Géométrique peut être utilisée pour calculer des vitesses moyennes.
  Faux. On utilisera plutôt une moyenne Harmonique.
- La moyenne Harmonique peut être utilisée pour calculer des taux de croissances moyens.
  Faux. On utilisera plutôt une moyenne Géométrique.
- Le paramètre associé à la moyenne géométrique est R=2.
  Faux. Ici R=e. Astuce : pour retenir « l'ordre des moyennes » on peut utiliser l'ordre alphabétique des deuxièmes lettres. Je m'explique. hArmonique<gEométrique<aRithmétique<qUadratique = A<E<R<U.

Questions bonus : Application

d. Question bonus : A propos du tableau suivant (Insee) :

Déciles de revenu disponible des ménages selon la configuration familiale en 2013

en euros

	D1	D2	D3	D4	D5	D6	D7	D8	D9
Personnes seules	10 270	12 860	14 870	16 570	18 250	20 290	22 940	26 770	33 720
Familles monoparentales	14 000	17 150	19 350	21 360	23 940	26 760	30 370	34 710	43 530
Couples sans enfant	20 380	25 040	28 450	31 920	35 420	39 520	45 340	52 920	67 360
Couples avec un enfant	24 140	29 940	34 550	38 380	42 330	46 900	52 450	59 950	75 580
Couples avec deux enfants	26 570	33 340	38 270	42 190	46 310	50 840	57 370	66 670	82 360
Couples avec trois enfants ou plus	26 160	31 870	36 590	41 220	46 340	51 920	58 140	68 740	88 110
Ménages complexes	17 250	22 080	26 680	29 950	34 400	39 100	45 070	53 590	66 800
Ensemble	**13 580**	**17 340**	**20 910**	**25 050**	**29 540**	**34 720**	**40 960**	**49 190**	**63 260**

Champ : France métropolitaine, ménages dont le revenu déclaré au fisc est positif ou nul et dont la personne de référence n'est pas étudiante.
Sources : CCMSA ; Cnaf ; Cnav ; DGFiP ; DGI ; Insee, enquête Revenus fiscaux et sociaux 2013.

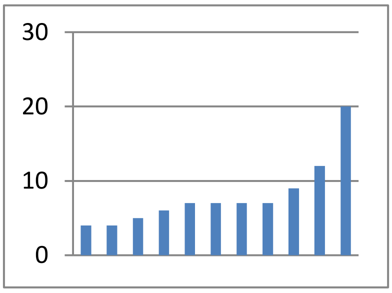
- On peut calculer le niveau de vie moyen des ménages complexes.
  Faux. Les quantiles comme seule information ne nous permettent en aucun cas de calculer une moyenne… Nous sommes tentés ici d'additionner tous les déciles et de diviser ensuite par neuf pour calculer une moyenne, mais le résultat – bien qu'approchant probablement la réalité – ne peut pas être exact.
  Prenons par exemple le Décile 1. Nous savons que 10% des ménages complexes ont un revenu disponible inférieur à 17 250€, mais nous ne savons pas si en moyenne ces ménages ont un revenu disponible de plutôt 15 000€ ou de plutôt 10 000€ ! Et le même problème se pose pour chaque décile.
- On peut calculer le revenu disponible médian des couples sans enfants.

Vrai. Ici D5=35 420 signifie que 50% des ménages sans enfants gagnent moins de 35 420€ (et que 50% gagnent plus).

- Le mode est D5.
  Faux. On ne peut pas déterminer de mode dans cette étude car on ne connaît pas les effectifs de chaque classe. On connait uniquement les valeurs qui séparent chaque dixième de la population.
- On peut représenter ce tableau avec un diagramme à secteurs circulaires.
  Faux. Un diagramme de ce type est adapté pour représenter des variables qualitatives et non quantitatives
- La moitié de la population étudiée a un revenu disponible inférieur à 29 540€.
  Vrai. L'information nous est donnée en ligne « Ensemble » et en colonne « D5 ». Ensemble signifie que l'on agrège les résultats des catégories utilisées précédemment, et D5 correspond à la médiane, donc à la valeur qui sépare la population en deux parts égales.

e. Question bonus : A propos de la série suivante : (4 ; 4 ; 5 ; 6 ; 7 ; 7 ; 7 ; 7 ; 9 ; 12 ; 20)

- La médiane est 8.
  Faux. Il y a 11 valeurs dans cette série, la médiane est donc la 6e. Il s'agit de 7. En effet : 5 valeurs sont inférieures ou égales à 7 et 5 valeurs sont supérieures ou égales à 7.
- La moyenne est 7.
  Faux : 4+4+5+6+7+7+7+7+9+12+20=88 et 88/11=8. La moyenne de cette série est donc 8.
- Le mode est 7.
  Vrai. Le mode est la valeur la plus fréquente de la série, il s'agit bien de 7, présente quatre fois.
- L'étendue est de 11.
  Faux. L'étendue désigne l'écart entre la valeur la plus petite et la valeur la plus grande d'une série. Ici 20-4=16. L'étendue est donc de 16.
- On peut représenter cette série par un diagramme en bâtons.
  Vrai. Un diagramme en bâtons permet de représenter des variables quantitatives discrètes, cela correspond donc à notre série.
  (Sur Excel je n'ai pas réussi à faire de bâtons, donc ça ressemble un peu à graphique en tuyaux d'orgues… imaginez juste un peu !)



f. Question bonus : A propos de l'étude statistique suivante :

Age des Tuteurs (années)	**[18 ; 20[**	**[20 ; 22[**	**[22 ; 27[**
Effectif	10	5	8

- L'âge moyen est 21,35 ans.
  Vrai. Pour calculer la moyenne, on calcule d'abord le milieu de chaque classe. Ensuite on multiplie le milieu de chaque classe par l'effectif qui lui correspond, on somme les résultats, puis on divise par le nombre total d'individus.
  1. Calcul des milieux. 19 est le milieu de la classe [18 ; 20[car (18+20)/2=19. 21 est le milieu de la classe [20 ; 22[car (20+22)/2=21. 24.5 est le milieu de la classe [22 ; 27[car (22+27)/2=24.5.
  2. On effectue le produit « milieu x effectif » pour chaque classe. 19x10=190. 21x5=105. 24.5x8=196.
  3. On somme les résultats. 190+105+196=491.
  4. On divise par l'effectif total. 491/(10+5+8)=21.35. L'âge moyen des tuteurs est de 21.35 ans (soit 21 ans 4 mois 5 jours.)
- On doit faire une interpolation linéaire pour calculer la moyenne.
  Faux. Il faut faire une interpolation linéaire pour calculer la MEDIANE.
- La classe modale est [18 ; 20[

Vrai. C'est en effet cette classe qui abrite l'effectif le plus élevé.

- Le premier quartile se trouve dans la classe [18 ; 20[

Vrai. L'effectif est de 23. Pour obtenir un quartile on doit diviser la série en 4 : 23/4=5.75. Le premier quartile est donc la « 5.75ième » valeur de la série. Elle se trouve donc entre dans la classe [18 ; 20[puisque son effectif est de 10.

Boîte à outils

Partie I : Taux et indices

- Taux de croissance global : $g_{Global} = \frac{V_t - V_0}{V_0}$.
- Taux de croissance annuel moyen : $\overline{g} = \left(\frac{V_t}{V_0}\right)^{\frac{1}{t}} - 1$.
- Propriété : $\ln(a^n) = n \times \ln(a)$.
- Multiplicateur composite : produit des multiplicateurs.
- Taux de croissance appliqué aux grandeurs liées : Soient deux valeurs A et B ; et GL la grandeur qui les relie (Exemple : capacité de consommation= $\frac{Salaire}{Prix}$). Si GL = $\frac{A}{B}$, on a : $g_{GL} = \frac{(1+g_A)}{(1+g_B)} - 1$ → Le taux de croissance appliqué aux grandeurs liées est égal au multiplicateur de la valeur A divisé par le multiplicateur de la valeur B, moins 1.
- Calcul d'un indice élémentaire : $\text{Indice}_{V1/V0} = \frac{V1}{V0} \times 100$.
- Réversibilité : $\text{Indice}_{V0/V1} = \frac{100^2}{I_{V1/V0}}$.
- Circularité : $\text{Indice}_{V2/V0} = \frac{I_{\frac{V2}{V1}} \times I_{\frac{V1}{V0}}}{100}$.

Partie II : Distributions

- Fréquence : f $= \frac{population\ associée\ à\ un\ caractère}{population\ totale}$.
- Angle correspondant à une fréquence pour un graphique à secteurs circulaires : f × 360.
- Contenu d'un tableau statistique standard :
  - xi : abscisses
  - ni : populations associées à un caractère
  - fi : fréquences
  - F+ : fréquences cumulées croissantes
  - F- : fréquences cumulées décroissantes
  - ai : amplitude (quand les intervalles de classe ne sont pas équivalents, il faut corriger l'amplitude. Ainsi on choisit une amplitude de base a et on calculera le rapport ai/a)
  - Ci : ai/a soit amplitude de l'intervalle divisée par l'amplitude de référence
  - f'i : fi/Ci fréquence ajustée
- Histogramme : xi en abscisses et f'i en ordonnées.
- Fonctions de répartition : xi en abscisses et F+ et F- en ordonnées. Les deux fonctions se croisent en (xi ; 0,5).

Partie III : Tendances centrales

- Représentations graphique :

Variables qualitatives : - Diagramme à secteurs circulaires

- Graphique à tuyaux d'orgue

- Graphique à tuyaux d'orgue composés

Variables quantitatives discrètes : - Diagramme en bâtons

- Courbe cumulative

Variables quantitatives continues : - Histogramme

- Fonction de répartition

- Le mode est la valeur d'une série qui revient le plus souvent.
- La classe modale est la classe ayant le plus grand effectif.
- La médiane est la valeur qui coupe la série en deux sous-ensembles de tailles égales.
- Les quantiles divisent la série en sous-ensembles de tailles égales : 2 sous-ensembles pour la médiane, 4 pour les quartiles, 10 pour les déciles, 20 pour les vingtiles, 100 pour les centiles, etc.
- La boîte à moustaches (Box Plot) est un graphique synthétique nécessitant 7 valeurs : Min, D1, Q1, Médiane, Q3, D9, Max.
- On distingue 4 moyennes différentes : la moyenne arithmétique (paramètre R=1), la moyenne quadratique (paramètre R=2), la moyenne harmonique (paramètre R=-1) et la moyenne géométrique (paramètre R=e).
- La formule générale d'un moyenne de paramètre R est : $M=[(1/n)*\Sigma(n_i(x_i^R))]^{1/R}$
- L'étendue d'une série est la différence entre la valeur maximale et la valeur minimale.