

STATISTIQUE DESCRIPTIVE

DOSSIER DES APPLI_COURS DU SECOND SEMESTRE 2017

S2_APPLI_COURS_1 : *fréquences simples – effectifs catégoriels*

SUJET : bibliothécaire

S2_APPLI_COURS_2 : *variables qualitatives et quantitatives et diagrammes*

Sujet : imaginer des variables et leurs diagrammes

S2_APPLI_COURS_2bis : VARIABLE QUALITATIVE : tuyaux d'orgue et diagramme empilé :

Sujet : Population active par CSP

S2_APPLI_COURS_3 : Variable discrète, fréquence simple, fréquences cumulées, et diagrammes

Sujet : nombre de frères et soeurs

S2_APPLI_COURS_4 : Variable discrète : x_M , la médiane

Sujet : Portefeuille d'actions

S2_APPLI_COURS_4bis: VARIABLE CONTINUE : Mode et Médiane -Méthode du tableau et méthode du diagramme

Sujet : primes salariales

S2_APPLI_COURS_5_ : Série continue : centre de classe, amplitude de classe

Sujet : distribution de salaires en Euro

S2_APPLI_COURS_6 : Variable discrète : \bar{x} , la moyenne pondérée

Sujet : mouvements de la population

S2_APPLI_COURS_7_ : Série continue : construction de classes et Diagramme différentiel - Moyenne, Médiane

Sujet : échantillon d'entreprises selon le chiffre d'affaires

S2_APPLI_COURS_8_ : Série continue : interpréter la courbe cumulative

Sujet : Contribuables A,B,C

S2_APPLI_COURS_9_ : La moyenne « \bar{x} » : la règle du réel

Sujet : nombre de pièces moyen par logement

S2_APPLI_COURS_10_ : Ecart absolu moyen (e_{am}) par rapport à la moyenne

Sujet : La Moyenne des distances à la moyenne *est nulle !*

S2_APPLI_COURS_11_ : Ecart absolu moyen (e_{am}) par rapport à la médiane
Sujet : *L'écart absolu par rapport à la médiane minimise l'écart absolu moyen (e_{am})*

S2_APPLI_COURS_12_ : Les PHI moyennes
Sujet : appliquer les formules des Phi Moyennes

S2_APPLI_COURS_13_ : Fonction de répartition et recherche de Quantiles (*définition : voir APPLI 14*)
Sujet : Lecture d'une courbe cumulative

S2_APPLI_COURS_14_ : Quantiles, classes, et courbe cumulative
Sujet : Etrences

S2_APPLI_COURS_15_ : La VARIANCE ($V(x)$ ou σ^2) et la déduction de l'écart type et du coefficient de variation
Sujet : Formule développée (ou simplifiée), ou moyenne des carrés des écarts $(x_i - \bar{x})^2$, il faut choisir !

S2_APPLI_COURS_16_ : MOYENNE, VARIANCE ($V(x)$ ou σ^2) ET ECART TYPE ($(V(x))^{1/2}$ ou σ)- Technique du Changement de variable
Sujet : salaires

S2_APPLI_COURS_17_ : CARACTERISTIQUES DE FORME (KURTOSIS et APLATISSEMENT)
Sujet : - 2 exemples -

S2_APPLI_COURS_18_ : CARACTERISTIQUES DE CONCENTRATION -Indice de Gini et courbe de Lorenz Gini
Sujet : inégalités de répartition salariale

S2_APPLI_COURS_19_ : EFFET DE STRUCTURE- - METHODE SHIFT SHARE
Sujet : comparaison salariale *entreprise-branche*

S2_APPLI_COURS_20_ : INDICES SYNTHETIQUES - (Laspeyres, Paasche, fisher), et égalité fondamentale-
Sujet : dépenses de consommation d'un ménage

S2_APPLI_COURS_1 : fréquences simples – effectifs catégoriels

SUJET : bibliothécaire

Matière	Code	fi%	ni
histoire générale	1	9,0%	
XVIème siècle	2	2,1%	
XVIIème siècle	3	3,8%	
XVIIIème siècle	4	7,6%	
XIXème siècle	5	7,0%	
XXème siècle	6	5,0%	
géographie générale	7	10,0%	
géographie orientale	8		1240
physique	9	2,1%	
chimie	10		
sciences naturelles	11	1,0%	
biologie	12	2,0%	
mathématique	13	9,9%	
musique	14	3,0%	
français	15	11,2%	
langues	16	7,8%	
Ensemble		100%	8602

Vous disposez ci-dessus d'un inventaire incomplet réalisé par un bibliothécaire, qui souhaite le compléter en réalisant le moins de calculs possibles.

A l'aide des relations entre fréquences et effectifs, compléter cet inventaire (calculs arrondis sans décimales).

Répondez d'abord aux deux questions ci-dessous :

- De quel type est la variable du tableau ?
- Que représente la valeur 8602 ?

Dans quel ordre faut-il réaliser les calculs ?

S2_APPLI_COURS_2 : variables qualitatives et quantitatives et diagrammes

SUJET : imaginer les variables et leur diagramme respectif

**Cocher le (ou les) diagrammes au(x)quel(s) peut donner lieu chaque type de variable
(les variables sont celles du S1, et dont le corrigé a déjà été réalisé)**

Lignes 16, 17 et 18 : à compléter par une variable de chaque type.

Variable	Qualitative	discrète	continue	Diagramme différentiel (*)			Diagramme intégral	
				secteur	bandeaux	bâtons	histogramme	courbe cumulative en escaliers
Population active selon l'âge	1							
Résidents de la région Nord Pas de Calais selon la nationalité	2							
Personnel d'une entreprise selon le nombre d'années d'étude	3							
élèves d'une classe maternelle selon la couleur des yeux	4							
Ouvrages de la bibliothèque selon le support	5							
Fleuves du monde suivant le débit	6							
Population des ménages lillois selon le nombre de personnes du ménage	7							
Entreprises du secteur de la distribution selon le chiffre d'affaires annuel	8							
Pays de l'Union Européenne selon le salaire minimum moyen	9							
Etats de l'OCDE selon le montant de la dette publique	10							
Population des animaux d'un zoo selon le genre	11							
Arbres de la forêt tropicale selon la taille exacte	12							
Employés du secteur public selon la qualification	13							
Flux journalier d'automobiles selon la vitesse	14							
Films européens réalisés dans l'année selon le coût de production	15							
	16							
	17							
	18							

(*) plusieurs réponses possibles parfois

S2_APPLI_COURS_2bis : SUJET VARIABLE QUALITATIVE : tuyaux d'orgue et diagramme empilé : Population active par CSP

Les données du tableau ci-dessous diffusées par



Institut national de la statistique
et des études économiques

Mesurer pour comprendre

sont celles de la distribution de la population active française en 2011, par Professions et catégories socio professionnelles. On connaît l' effectif (n_i) de chaque PCS, et donc la population active totale (N =Effectif total). Les modalités (repérées par l'indice i), de la variables PCS sont QUALITATIVES et au nombre de 7.

Travail demandé :

- Représenter graphiquement cette série après avoir calculé les fréquences : SIMPLES ($f_i\%$) et CUMULEES ($F_i\%$). Le mode calcul est répertorié sous le tableau.

La représentation doit être réalisée sous forme de 2 graphiques :

- Graphique 1: Diagramme en tuyaux d'orgue (ou à défaut « circulaire »)
- Graphique 2: DIAGRAMME EMPILE 100% (appelé « en bandeaux » dans le Manuel).

Remarque : le travail peut être fait de différentes manières : manuellement sur papier millimétré, ou non ; par informatique, en disposant de la version 2010 et plus de « Excel » (étudiée au semestre 2).

i	Catégorie socioprofessionnelle (PCS)	effectifs en 2011 (n_i)	fréquences simples ($f_i\%$)	fréquences cumulées ($F_i\%$)
1	Agriculteurs exploitants	456 907		
2	Artisans, commerçants, chefs d'entreprise	1 625 414		
3	Cadres, professions intellectuelles supérieures	4 415 255		
4	Professions intermédiaires	6 728 283		
5	Employés	7 475 214		
6	Ouvriers (y compris agricoles)	5 725 587		
7	Chômeurs	3 880 232		
	Total = N = effectif global	30 306 892		

Note : les données ont été révisées suite à la rénovation du questionnaire de l'enquête emploi.

Champ : population des ménages de 15 ans ou plus, vivant en France métropolitaine. Résultats en moyenne annuelle.

Source : Insee, enquêtes Emploi.

MODE DE CALCULS

Une fréquence simple ($f_i\%$) = $(n_i/N) \times 100\%$ et $\sum f_i\% = 100\%$

Une fréquence cumulée ($F_i\%$) = $f_{(i-1)} + f_i$

S2_APPLI_COURS_3 : Variable discrète, fréquence simple, fréquences cumulées, et diagrammes

Sujet : nombre de frères et soeurs

Enoncé : Une étude réalisée par une cellule d'orientation de l'Université de Montréal (Quebec), portant sur un échantillon de 250 étudiants de première année de Lettres, a permis de construire le *tableau de distribution ci-dessous*. Les étudiants sont répartis selon le *nombre de frères et soeurs*.

Nombre de frères et soeurs	Effectifs					
0	25					
1	36					
2	43					
3	75					
4	40					
5	21					
6	10					
	250					

CALCULS et réalisation du tableau de distribution

- 1- Quelle est la variable et de quel type est-elle ?
- 2- Nommer par un symbole chacune des deux premières colonnes (en libellé).
- 3- Indiquer le symbole de la valeur 250. Comment se nomme cette valeur ?
- 4- Calculer les fréquences simples en les désignant correctement, et en songeant aux calculs ultérieurs qu'elles vous permettront de réaliser -(utiliser 2 décimales)
- 5- Continuer vos calculs en déterminant les fréquences cumulées que vous désignerez correctement par un symbole dans chaque colonne. (utiliser 2 décimales).
- 6- Choisissez dans chaque colonne $F(x_i)$ une valeur dont vous écrirez une interprétation en 1 ligne maximum.

REPRESENTATION GRAPHIQUE

- 7- Réaliser un diagramme différentiel de cette série en veillant à ne rien omettre dans les informations qu'il doit porter.
- 8- Réaliser un diagramme intégral avec la même attention.

S2_APPLI_COURS_4 : Variable discrète : xMé, la médiane

Sujet : Portefeuille d'actions

On veut appliquer à la série donnée ci-dessous les deux méthodes de détermination de la valeur de $x_{Mé}$, la médiane. C'est-à-dire :

- **La méthode des effectifs cumulés**
- **La méthode de la courbe cumulative.**

On dispose de la série du nombre d'actions acquises au cours d'une année par un échantillon de spéculateurs.

xi nombre d'actions	ni effectifs				
0	20				
10	40				
20	30				
30	70				
40	40				
	200				

NB : Le nombre de colonnes est indicatif

En utilisant le tableau de distribution, il est demandé de déterminer le nombre médian d'actions, soit $x_{Mé}$:

- Par la méthode $N(x_i+)$, en expliquant le résultat
- Par la méthode $F(x_i)$ et donc par le tableau et la courbe cumulative.

(NB : vérifier que les deux résultats convergent).

Puis, interpréter le résultat en une phrase.

APPLI_COURS_4bis: VARIABLE CONTINUE : Mode et Médiane -Méthode du tableau et méthode du diagramme
Sujet : primes salariales

La variable ci-dessous représente les primes versées au personnel d'une entreprise. On connaît les effectifs des bénéficiaires.

xi-	xi+	effectifs (ni)						
2	3	1						
3	5	5						
5	7	20						
7	10	15						
10	12	8						
12	15	2						

Travail demandé :

En utilisant les deux méthodes : du tableau et du diagramme déterminer (en veillant à utiliser les symboles appropriés)

- 1) Le mode de cette série**
- 2) La Médiane et le premier décile (ce dernier pouvant être calculé ultérieurement, voir APPLI 13 et 14)**
- 3) Donner la signification du nombre « 13 » calculé quelque part dans le tableau.**

S2_APPLI_COURS_5_ : Série continue : centre de classe, amplitude de classe

Sujet : distribution de salaires en Euro

La distribution des salaires mensuels en Euro, observée au 15 Janvier 2015 dans une usine, est partiellement donnée dans *le fichier excel de l'appli_5*. On dispose des classes de salaires en € (voir ci contre)

Il est demandé, le cas échéant en réalisant le travail au « tableur », de

1- calculer rapidement les centres de classe

2- calculer les amplitudes de classes

(Résultats arrondis aux entiers)

NB : Vous veillerez à faire figurer dans le tableau toutes les informations qu'il doit contenir.

Répondre aux questions ci-dessous :

1- Ecrire la formule du centre de classe et justifier qu'il puisse devenir une variable du même type que « xi » pour une variable discrète.

2- Ecrire la formule de l'amplitude, et dire quel est l'intérêt du calcul des amplitudes de classes ?

3- La somme des centres de classe a-t'elle un sens ?

4- La somme des amplitudes a-t'elle un sens (NB : cette question se réfère à une étape ultérieure du cours – voir le manuel ou compléter ultérieurement -) ?

5- Est-il possible de réaliser le diagramme différentiel de cette série ? (expliquez) ?

6- Est-il possible de progresser dans l'analyse de cette série (expliquez) ?

xi-	xi+
2500	3000
3000	3450
3450	3850
3850	4400
4400	5400
5400	7350
7350	10400
10400	15450
15450	20000
20000	50000

S2_APPLI_COURS_6 : Variable discrète : \bar{x} , la moyenne pondérée

Sujet : mouvements de la population

Le site de l'INSEE (www.insee.fr/fr/themes/) fournit pour un échantillon de pays, les mouvements de la population pour l'année 2012. On connaît (ci-contre) le **taux d'accroissement total de la population pour 1000 habitants** (*en pour 1000 ou taux/00*). De même nous donne t'on le nombre de pays vérifiant chaque taux.

Question 1 : Quelle est la variable x_i , de quel type est-elle ?

Question 2 : en complétant le tableau de distribution ci-dessous veuillez calculer la taux moyen d'accroissement pour ces pays. (Indiquer clairement l'intitulé de vos colonnes).

Pour rappel : le cours a donné les deux méthodes de calcul de la moyenne \bar{x} , qui peuvent être appliquées. Leurs résultats doivent converger.

Le tableau de distribution

*NB : Le nombre de colonnes du tableau n'est qu'indicatif, mais suffisant pour permettre l'application des deux modes de calculs de \bar{x} , la moyenne.
(Travailler avec deux décimales)*

x_i taux ₁₀₀	n_i nb pays						
3	0						
4	3						
5	2						
6	2						
7	1						
8	1						
9	1						
10	3						

(source : INSEE – les véritables taux ont été ici modifiés pour harmoniser les modalités)

Ecrivez ci-dessous les formules algébriques appliquées dans le tableau, et donner le résultat aussi à l'aide d'une phrase.

S2_APPLI_COURS_7_ : Série continue : construction de classes et Diagramme différentiel - Moyenne, Médiane

Sujet : échantillon d'entreprises selon le chiffre d'affaires

Une enquête auprès d'un échantillon d'entreprises, sondées dans l'ordre donné dans le tableau ci-contre, a permis de connaître leur chiffre d'affaires (CA) en millions d'Euro, réalisé sur leur marché respectif, en Décembre 2014.

Il est demandé de réaliser un *dénombrement* à partir du tableau pour **construire LE TABLEAU DE DISTRIBUTION de la série continue du chiffre d'affaire de ces entreprises.**

C'est-à-dire de construire d'abord les classes de CA, et les reporter dans le tableau de distribution (page suivante).

I) - **La construction des classes est laissée à votre attention(*)** . Mais Vous respecterez néanmoins :

- 1- Les valeurs extrêmes que sont $x_{Min} = 0,05$ et $x_{Max} = 12$
- 2- Un choix rationnel du nombre de classes, c'est-à-dire guidé par quelques principes :
 - **les effectifs catégoriels (ni) ne doivent pas être déséquilibrés**
 - **le nombre de classes ne doit ni diluer l'information (trop de classes), ni la concentrer (ou ramasser) (trop peu de classes) excessivement.**

II) Les classes étant constituées et reportées dans le tableau de distribution, on veut étudier la diversité de ces entreprises.

Il est donc demandé :

- 1- **de réaliser le diagramme différentiel, pour situer la classe modale (NB : les fréquences par unité d'amplitude seront calculées en « pourcentage » pour plus de clarté),**
- 2- **de calculer :**
 - a. **le chiffre d'affaire moyen (\bar{x})**
 - b. **le chiffre d'affaire médian ($x_{Mé}$)**

(les calculs seront réalisés au centième. Mais la précision est accrue avec 3 décimales)

(*) Le cours ayant exposé les deux règles applicables pour l'effectif global dénommé N, et k le nombre de classes à appliquer :

règle de Brooks-Carruthers : $k < 5 + \log N$

règle de Huntsberger : $k \approx 1 + (3,3 \log N)$

Entreprise	CA (millions €)
1	0,25
2	1
3	2,5
4	0,11
5	0,5
6	5
7	2,5
8	0,25
9	10
10	0,5
11	5
12	1
13	0,5
14	10
15	10
16	0,11
17	0,25
18	5
19	2,5
20	1
21	7
22	5
23	0,5
24	0,5
25	0,16
26	1
27	2,5
28	0,25
29	10
30	1
31	1
32	0,5
33	10
34	6
35	0,5
36	2,5
37	1
38	0,25
39	5
40	0,5

SUITE CI-DESSOUS

TABLEAU DE DISTRIBUTION DE L'APPLI 7

Indiquer un titre.....

x_i^-	x_i^+											

S2_APPLI_COURS_8 : Série continue : interpréter la courbe cumulative

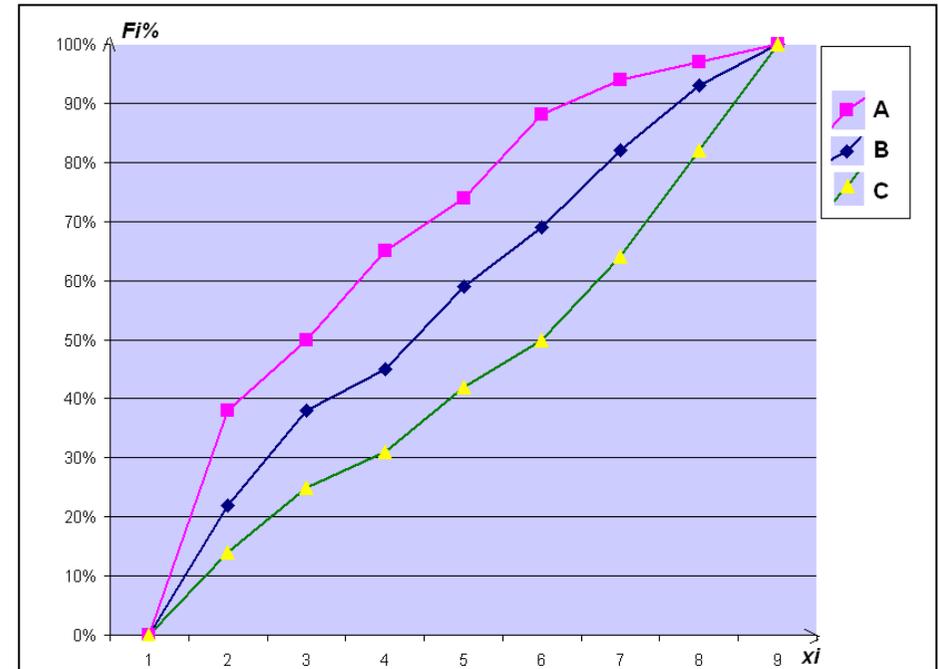
Sujet : Contribuables A,B,C

Soit trois distributions d'impôts sur le revenu, rangées en classes, de trois groupes de contribuables (A,B,C). Le montant de l'impôt payé est « xi » en milliers d'Euros.

L'analyse de cette distribution a déjà été réalisée dans 3 tableaux de distribution, dont on a retenu les fréquences cumulées et leur représentation par 3 fonctions de répartition.

Le diagramme intégral ci-contre représente donc les fréquences cumulées pour chacun des groupes (A,B et C).

L'exercice consiste à interpréter les courbes cumulatives en répondant aux questions.



QUESTIONS

- 1- A la seule observation du graphique, quel est le groupe qui paie les montants d'impôts les plus élevés ? Pourquoi ?
- 2- Est-il possible de répondre à la question 1 en comparant pour A, B et C, une seule fréquence $F_i\%$? Si oui, donner un exemple en situant le raisonnement dans le graphique par une horizontale ou une verticale. Vous donnerez ensuite l'ordre du classement des groupes.
- 3- Quel(s) autre(s) moyen(s) aurai(en)t facilité la comparaison des 3 groupes ?
- 4- Reconstruire les fréquences simples ($f_i\%$) à partir de la courbe (en estimant l'ordonnée), et recomposer le tableau de distribution.
- 5- Que peut on ajouter au tableau si on dispose de l'information : $N(A) = 1500$, $N(B) = 2000$ et $N(C) = 1740$? Réaliser les calculs en donnant la méthode.

S2_APPLI_COURS_9_ : La moyenne « xbar » : la règle du réel

Sujet : nombre de pièces moyen par logement

Le recensement de la population française réalisé par l'I.N.S.E.E. en 1990 donne les résultats suivants en ce qui concerne la répartition des logements par nombre de pièces principales :

x(i)	n(i) (milliers)					
Nombre de pièces	Nombre de logements					
1	1 303					
2	2 792					
3	5 063					
4	6 023					
5	3 937					
6	2 424					
ensemble	21 542					

Il vous est demandé de calculer *le nombre moyen de pièces par logement*, sachant qu'une enquête à dénombré exactement pour 1990, un nombre total de pièces égal à 100.000.

S2_APPLI_COURS_10_ : Ecart absolu moyen (*eam*) par rapport à la moyenne

Sujet : La Moyenne des distances à la moyenne *est nulle* !

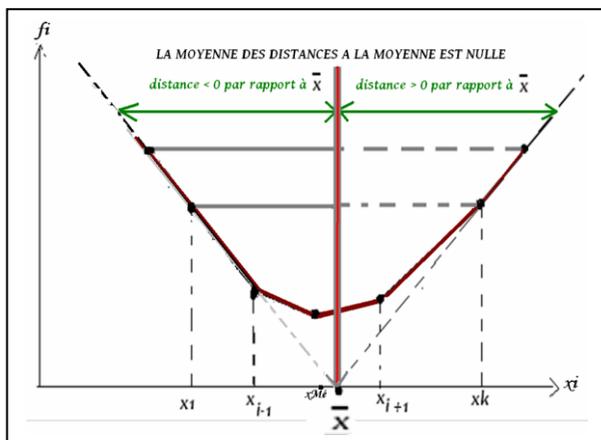
Les 70 exploitations d'une petite commune se répartissent suivant la superficie cultivée en milliers d'hectares (*m_ha*), de la manière suivante, donnée par le tableau de distribution :

Exploitations de la commune suivant la superficie

x_i^-	x_i^+	n_i					
0	20	7					
20	40	20					
40	50	18					
50	60	10					
60	80	15					
		70					

Question

Vérifier à l'aide de cet exemple la démonstration du cours (rappelée ci-dessous par le graphique et la formule algébrique), suivant laquelle *la moyenne des distances à la moyenne est nulle*. Vous utiliserez le tableau de distribution en réalisant vos calculs au centième.



$$\sum_{i=1}^n f_i (x_i - \bar{x}) = 0$$

S2_APPLI_COURS_11 _: Ecart absolu moyen (e_{am}) par rapport à la médiane

Sujet : L'écart absolu par rapport à la médiane minimise l'écart absolu moyen (e_{am})

Sur l'exemple de la variable x_i , discrète et quelconque, on veut vérifier la règle énoncée par le sujet (plus haut).

En utilisant le tableau de distribution, démontrer que $e_{xM\acute{e}} < e_{xbar}$, c'est-à-dire que l'écart absolu moyen par rapport à la médiane est inférieur à l'écart absolu moyen par rapport à la moyenne.

Il est donc nécessaire de déterminer préalablement les valeurs de la moyenne et de la médiane. Vous utiliserez deux décimales.

x_i	n_i	$n_i \cdot x_i$	$(x_i - \bar{x})$	$ x_i - \bar{x} $	$N(x_i)$	$(x_i - x_{M\acute{e}})$	$ x_i - x_{M\acute{e}} $		
2	1								
6	1								
15	1								
24	1								
38	1								
46	1								
53	1								
68	1								
85	1								

(deux colonnes vierges facultatives)

(NB : les n_i étant très faibles et égaux à 1, le détour par les fréquences est ici inutile)

S2_APPLI_COURS_12_ : Les PHI moyennes
Sujet : appliquer les formules des Phi Moyennes

En utilisant cette partie du cours

Phi moyennes ou « moyennes d'ordre « r »

$$M_r \text{ ou } (XM_r) = \left[\sum_{i=1}^k f_i \times xi^r \right]^{\frac{1}{r}}$$

sur l'exemple de ce tableau :

xi	ni										
1	20										
2	30										
3	15										
4	10										
5	5										
6	2										

1) Démontrer par application des formules des "Phi Moyennes" les valeurs suivantes (ajouter des colonnes ou utiliser plusieurs tableaux)

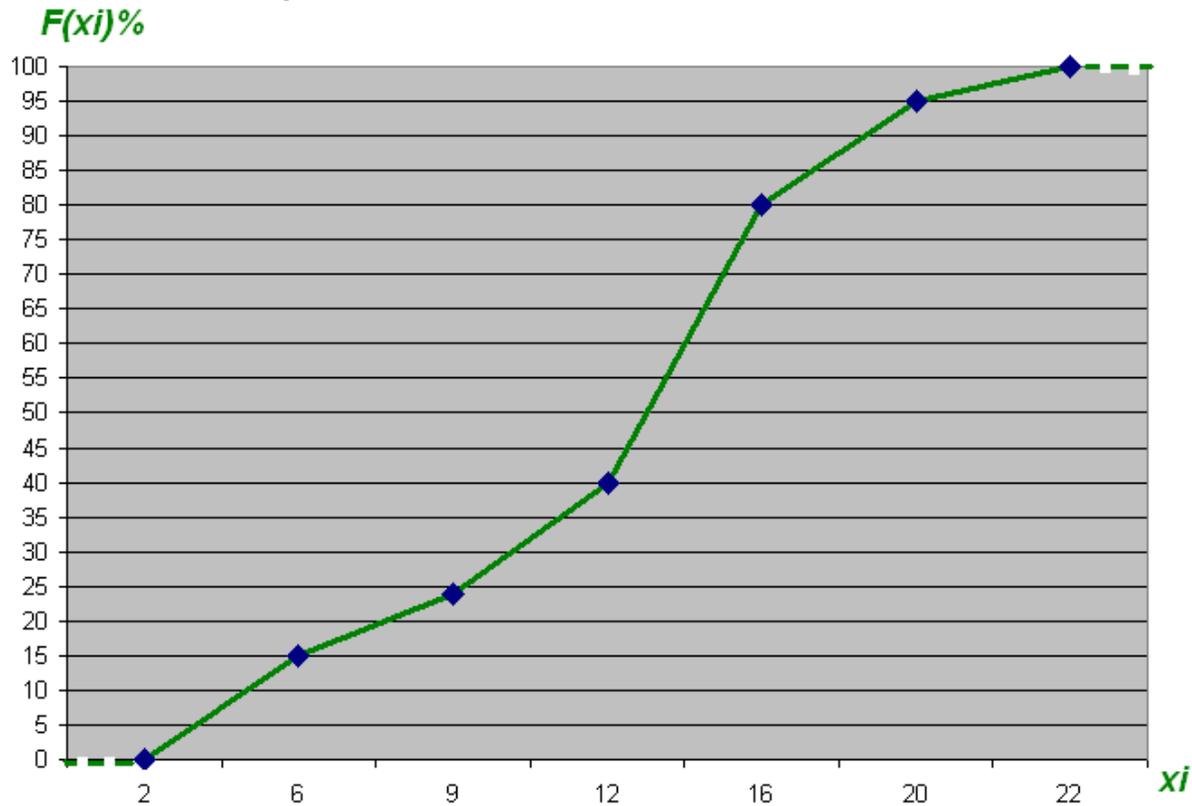
x	Moyenne arithmétique =	non donnée
Q ²	Moyenne quadratique =	2,78
H	moyenne harmonique =	1,87
G	Moyenne géométrique =	2,15

2) Montrer que la relation d'ordre entre moyennes est vérifiée.

3) De quel type est la variable ? Vos calculs conviennent ils dans le cas d'un autre type de variable ? laquelle ? et pourquoi ?

S2_APPLI_COURS_13_ : Fonction de répartition et recherche de Quantiles (définition : voir APPLI 14)
Sujet : Lecture d'une courbe cumulative

Soit la fonction de répartition ci-dessous :



Il est demandé : de situer deux quantiles au choix

- l'un dont la valeur est donnée par la courbe
- l'autre dont la valeur sera calculée par interpolation (avec plus ou moins de précision)

(en donnant clairement l'expression algébrique de chaque quantile, et la signification du résultat.)

APPLI_COURS_14 : Quantiles, classes, et courbe cumulative

Sujet : Etrennes

Un sondage réalisé auprès d'enfants d'un établissement secondaire au 15 Janvier 2015 a permis de dresser une distribution en quantiles des bénéficiaires selon le montant des Etrennes (x_i en €) reçues.

Les résultats donnés dans un ordre alléatoire sont les suivant :

9eme décile = 200€, 1^{er} quartile = 75€, 4eme décile = 100€, 1^{er} décile = 50€, 3eme quartile = 150€, 2eme quartile = 120€. On sait que le montant d'étrennes minimum est égal à 10€.

Il est demandé

- 1) de construire le tableau de distribution donnant :
 - les Fréquences cumulées
 - les fréquences simple ($f_i\%$) selon le montant reçu
- 2) de réaliser la courbe cumulative en situant la médiane

RAPPEL

- Un quantile est une modalité du caractère en-deça de laquelle sont strictement situés $F(x_i)\%$ des individus.

Soit x_i , tel que $(F_{x_i})\%$

- Les quantiles suivent un ordre croissant de 0% à 100%/ Leur position dans cet ordre est donnée par leur intitulé (ou nom) : centile, décile, quartile.

- Cet ordre est la seule manière pour retrouver les classes de la série, elles mêmes nécessaires pour la construction du tableau de distribution, et la courbe cumulative.

S2_APPLI_COURS_15_: La VARIANCE ($V(x)$ ou σ^2) et la déduction de l'écart type et du coefficient de variation

Sujet : Formule développée (ou simplifiée), ou moyenne des carrés des écarts $(x_i - \bar{x})^2$, il faut choisir !

Le tableau de distribution ci-dessous est celui d'une variable continue quelconque, dont on connaît les effectifs « ni »,

Il est demandé de calculer la variance de la distribution, pour en déduire l'écart type et le coefficient de variation.

Il est conseillé de vérifier sur cet exemple l'intérêt de la formule simplifiée de la variance. Pour cela le tableau de distribution comporte suffisamment de cellules vierges, vous permettant de réaliser les deux méthodes afin de les comparer, sachant que leur résultat doit être le même.

(Voirs cours)

Tableau de distribution

col 1	col 2	col 3	col 4	col 5	col 6	col 7	col 8		col 9	col 10	col 11	col 12	col 13		col 14
xi-	xi+	ni													
15	25	9													
25	35	15													
35	45	22													
45	55	29													
55	65	17													
65	75	6													
75	85	2													
		100													

(Les colonnes étroites indiquent le passage à une autre formule)

Technique du Changement de variable

Sujet : salaires

xi-	xi+	ni	cx _i	cx _i '	fi	fi.cxi'	(cx _i) ²	fi.cxi ²	fi.cxi	cx _i ²	fi.cxi ²
4400	5400	1271									
5400	7350	847									
7350	10400	296									
10400	15450	297									
		2711									

A l'aide d'une distribution de salaires (donnée ci-dessus) en 4 classes, on se propose de réduire la variable par un changement de variable. Les calculs de la moyenne, de la variance et de l'écart type, seront donc réalisés à l'aide la propriété de linéarité appliquée à chacun de ces indicateurs.

Soit cette distribution et le tableau dont les libellés de colonne sont donnés (l'ordre des calculs peut cependant être modifié) :

Il est demandé :

- 1) de réaliser les calculs de : la moyenne, de la variance et de l'écart type de *la variable intermédiaire (ou nouvelle variable)*, sachant que le changement de variable s'écrit :

$$cx'_i = \frac{cx_i - 6375}{10} \quad (\text{ou si on préfère alléger la formalisation } x'_i = \frac{x_i - 6375}{10}) - \text{Vous chercherez à justifier ce nombre 6375 -}$$

- 2) de vérifier la propriété de linéarité par le calcul dans le tableau, en donnant la relation vérifiée (sous le tableau).

(calculs au centième)

APPLI_COURS_17_ : CARACTERISTIQUES DE FORME (KURTOSIS et APLATISSEMENT)

Sujet : - 2 exemples -

Soit les deux distributions quelconques données ci-dessous avec leur tableau respectif de distribution usuel (des n_i aux $F(x_i)$). Ce tableau permet le calcul des caractéristiques de tendance centrales habituelles (x_{Mo} , $x_{Mé}$ et \bar{x}). On dénomme ces distributions : **exemple 1 et exemple 2.**

Travail demandé :

- 1) Etudier pour chaque distribution les caractéristiques de forme, en calculant les deux coefficients de Fisher : Gamma 1 (Γ_1), et Gamma 2 (Γ_2).
Il convient donc de compléter chaque tableau de distribution.
- 2) Conclure sur l'**asymétrie et l'aplatissement**, en vérifiant cette conclusion à l'aide des caractéristiques de tendance centrale
- 3) Réaliser le **polygone des fréquences** des deux distributions. Ce polygone doit normalement vérifier chaque conclusion.

Exemple 1									
Tableau usuel pour le Mode, la moyenne et la Médiane									
xi-	xi+	ni	fi%	cxi	fi,cxi	F(xi+)	F(xi-)	ai	fi/ai
50	55	10							
55	75	15							
75	90	9							
90	95	20							
95	110	96							
		150							

Exemple 2									
Tableau usuel pour le Mode, la moyenne et la Médiane									
xi-	xi+	ni	fi%	cxi	fi,cxi	F(xi+)	F(xi-)	ai	fi/ai
50	65	100							
65	80	350							
80	85	580							
85	90	435							
90	120	125							
		1590							

NB : ci-dessus, lire (fi.cxi) et (fi/ai)%

APPLI_COURS_18_ : CARACTERISTIQUES DE CONCENTRATION

Indice de Gini et courbe de Lorenz Gini

Sujet : inégalités de répartition salariale

Soit une distribution de salaires (prime comprise) en Euro, dans une entreprise, dont on connaît les effectifs salariés au cours d'une période.

On suppose, dans un premier temps inconnue la masse salariale (comprenant la prime) distribuée. C'est le tableau 1.

Puis dans un second temps (en fin de période), cette masse (**MS**) et **sa répartition réelle** (**Si**) sont connues (en K€). C'est le tableau 2.

Travail demandé

- 1) Évaluer dans chaque cas, à l'aide de l'indice de Gini, la concentration, pour dresser un constat sur les *inégalités de répartition*. Dans chaque tableau, on évaluera l'indice de Gini à l'aide des deux raisonnements : dans le triangle ($IG = 2(0,5-U)$), et dans le carré ($IG = 1 - A$). Les libellés des colonnes sont donnés à cet effet.
- 2) Déterminer pour chaque situation : la Médiane, et la Médiale. Calculer l'intervalle de variation (ΔM). Illustrer ces trois valeurs avec les deux courbes cumulatives $F(x_i^-)$ et $G(x_i^-)$ réalisées dans le même graphique. Vérifier la concentration constatée en 1) ci-dessus.
- 3) Réaliser les deux courbes de Lorenz Gini dans le même carré, et vérifier la concentration constatée en 1) ci-dessus.
- 4) Conclure en une phrase sur la différence entre les deux distributions : tableau et courbes des cas 1) 2).

Tableau 1 : Masse et répartition estimées
(NB : le tableau est à utiliser pour répondre aux trois question)

											dans le triangle	dans le carré
x_i^-	x_i^+	n_i	$f_i\%$	c_{x_i}	$f_i^* \cdot c_{x_i}$	$F(x_i^-)$	$F(x_i^+)$	g_i	$G(x_i^-)$	$G(x_i^+)$	$1/2 \cdot f_i((G_{x_i^-}) + G_{x_i^+})$	$f_i((G_{x_i^-}) + G_{x_i^+})$
1500	1700	180										
1700	1800	75										
1800	1950	26										
1950	2005	49										
		330										
											U : triangle	ou A : carré

(suite ci-dessous)

Tableau 2 : Masse et répartition réelles
(NB : le tableau est à utiliser pour répondre aux trois question)

xi-	xi+	ni	Si	fi%	cxi	fi*.cxi	F(xi-)	F(xi+)	gi	G(xi-)	G(xi+)	dans le triangle 1/2.fi((Gxi-)+G(xi+))	dans le carré fi((Gxi-)+G(xi+))
1500	1700	180	55										
1700	1800	75	150										
1800	1950	26	100										
1950	2005	49	200										
		330	505									0,247	0,494
			MS									U : triangle	ou A : carré

Fin de l'énoncé

APPLI_COURS_19_ : SUJET EFFET DE STRUCTURE- - METHODE SHIFT SHARE

Sujet : comparaison salariale *entreprise-branche*

On dispose ci-dessous d'un tableau donnant les effectifs salariés par catégorie, ainsi que les salaires catégoriels, pour deux niveaux d'analyse dénommés respectivement « E » pour une entreprise, située dans une branche « B ».

On utilise la convention de notation suivante : par exemple « $e n_i$ » désigne les effectifs (n) par catégorie (i) dans l'entreprise (E) .

Ou « $b f_i$ » désigne les fréquences simples (f_i) de chaque catégorie (i) dans la branche (B).

Le tableau comprend les informations suffisantes (et explicites), pour être complété en vue de l'application de la méthode shift share.

Le travail demandé est :

- 1) Compléter le tableau des données, pour déterminer les valeurs des moyennes nécessaires à la recherche d'un « effet de structure ».
- 2) Le but est de répondre à la question suivante : *la position relative de l'entreprise par rapport à la branche s'explique t'elle par un « effet de structure » ?* Vous disposez à cet effet du tableau à 4 cases qu'il convient de compléter pour répondre à la question

Tableau des données à compléter

catégorie	$e n_i$ nombre de salariés	$e f_i$ fréquences	$e X_i$ (en k€) salaires catégoriels	$e f_i \cdot e X_i$	$b f_i$ fréquences	$b f_i \cdot e X_i$	$b X_i$ (en k€) salaires catégoriels	$b f_i \cdot b X_i$	$e f_i \cdot b X_i$
cadres supérieurs	20		7,0		1%		6,7		
cadres moyens	40		4,0		9%		4,2		
employés qualifiés	80		3,0		10%		3,1		
employés non-qualifiés	60		2,0		15%		2,1		
ouvriers qualifiés	100		2,2		15%		2,4		
ouvriers non-qualifiés	200		1,6		50%		1,8		
Ensemble	500	100%			100%				
				$\bar{X}_{(E)}$		$\tilde{X}_{(B)}$		$\bar{X}_{(B)}$	\hat{X}

Suite ci-dessous

Tableau à 4 cases

		Structure catégorielle	
		Entreprise	Branche
Salaires catégoriels			

Fin de l'énoncé

APPLI_COURS_20_ : INDICES SYNTHETIQUES - (Laspeyres, Paasche, fisher), et égalité fondamentale-

Sujet : dépenses de consommation d'un ménage

Le tableau ci-dessous donne l'évolution de la consommation d'un ménage, résumée par trois postes de dépenses, entre deux période 0 et 1.

Période	période 0		Période 1					
Poste	p_0	q_0	p_1	q_1				
Aliments	1,20 €	96	1,85 €	110				
Boissons	2,50 €	48	3,10 €	64				
Divers	2,00 €	80	2,35 €	100				

Travail demandé :

En utilisant les colonnes du tableau prévues à cet effet, dont vous donnerez chaque libellé, calculer les Indices synthétiques suivants :

- 1) Laspeyres des prix
Formule utilisée :

Application :

Résultat :

- 2) Paasche des prix
Formule utilisée :

Application :

Résultat :

Suite 1/2 ci-dessous

3) Laspeyres des quantités

Formule utilisée :

Application :

Résultat :

4) Paasche des quantités

Formule utilisée :

Application :

Résultat :

5) En déduire le Fisher des prix

Formule utilisée :

Application :

Résultat :

6) de même, le Fisher des quantités

Formule utilisée :

Application :

Résultat :

Suite 2/2 ci-dessous

7) Ecrire et vérifier l'égalité fondamentale entre indices en y faisant figurer le Laspeyres des prix

Formule utilisée :

Application :

Résultat :

Vérification :

-Φ-fin du document : APPLI_COURS_ S2_17_ SUJETS