

Travaux dirigés Statistique descriptive

Prof. Armel YODE

Exercice 1 La répartition des femmes à la recherche d'un emploi à Abidjan d'octobre 1994 à mars 1995 selon leur situation matrimoniale et selon le diplôme le plus élevé obtenu dans l'enseignement général:

Situation matrimoniale	Effectif	Diplôme	Effectif
Célibataire	4221	Sans	2041
Mariée	620	CEPE	3354
Divorcée	188	BEPC	693
Union libre	2258	BAC	22
Non déclarée	56	Supérieur	9
Total	7501	Total	7501

Source: Département Emploi de l'AGEPE, 1996.

- (1) Quelle est la taille de la population des femmes à la recherche d'un emploi à Abidjan d'octobre 1994 à mars 1995?
- (2) Quelles sont les caractères étudiés? Donner leur nature.
- (3) Donner le mode de chaque caractère.
- (4) Donner la distribution des fréquences de chaque caractère.
- (5) Donner une représentation graphique qui fasse apparaître l'importance relative de chaque modalité pour chaque caractère.

Exercice 2 *Une affaire de moyennes !*

- (1) *Moyenne géométrique et moyenne arithmétique:* Supposons que les taux d'intérêt pour 4 années consécutives soient respectivement de 5, 10, 15, et 10%.
 - (a) Que va-t-on obtenir après 4 ans si je place 100 francs ?
 - (b) Calculer la moyenne arithmétique des taux?
 - (c) Calculer la moyenne géométrique des taux?
 - (d) Quel est le bon taux moyen? Justifier.
- (2) Soit un commerçant qui augmente le prix d'un produit de 4%. A la suite de cette augmentation, la quantité vendue baisse de 3%. Calculer le taux de croissance de la recette totale.
- (3) *Moyenne harmonique et moyenne arithmétique:*

Un cycliste parcourt 4 étapes de 100 km. Les vitesses respectives pour ces étapes sont de 10 km/h, 30 km/h, 40 km/h, 20 km/h.

- (a) Par un raisonnement simple, calculer sa vitesse moyenne.
- (b) Calculer la moyenne arithmétique des vitesses.

- (c) Calculer la moyenne harmonique des vitesses.
 (d) Quelle est la bonne vitesse moyenne? Justifier.
- (4) Pour répondre aux questions suivantes, vous devrez trouver quelle moyenne doit être utilisée.
- Une entreprise comporte trois salariés dont les salaires nets sont respectivement de 2300000, 1100000 et 800000 francs. Quel est le salaire moyen dans cette entreprise ?
 - Deux départements ont la même population. Dans l'un, il y a une voiture pour 4 habitants et dans l'autre une voiture pour 12 habitants. Quel est le taux d'équipement moyen de la région formée par ces deux départements ?
 - En Côte d'Ivoire, la croissance du PIB a été de 8.84% en 2015, 8.34% en 2016 et 7.80% en 2017. Quel a été le taux de croissance annuel moyen sur ces trois années?

Exercice 3 Etude de la symétrie.

Voici trois séries:

$$A = \{1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5\}$$

$$B = \{1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5\}$$

$$C = \{1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5\}$$

- Pour chaque série, calculer la moyenne, la médiane et le mode. Comparer les valeurs obtenues.
- Pour chaque série, représenter le diagramme en bâtons des fréquences.
- Pour chaque série, représenter la boîte à moustaches.
- Pour chaque série, étudier la symétrie.
- Pour chaque série, étudier la forme.

Exercice 4 Au poste de péage, on compte le nombre de voitures se présentant sur une période de 5 mn. Sur 100 observations de 5 mn, on obtient les résultats suivants:

Nombre de voitures	1	2	3	4	5	6	7	8	9	10	11	12
Nombre d'observations	2	8	14	20	19	15	9	6	2	3	1	1

- Calculer la moyenne et l'écart-type de cette série.
- Déterminer analytiquement et graphiquement les quartiles.
- Etudier la forme et la symétrie de la série.

Exercice 5 Histogrammes

- Soit 100 ménages distribués selon leur revenu mensuel en euros. On définit des classes d'amplitudes égales à 1500 euros:

Revenu Mensuel	Effectif
[0, 1500[20
[1500, 3000[40
[3000, 4500[30

Représenter l'histogramme des fréquences ainsi que le polygone des fréquences.

- (2) Dans cette question, les classes sont d'amplitudes inégales:

Revenu Mensuel	Effectif
[0, 1500[20
[1500, 4500[40
[4500, 6000[30

Représenter l'histogramme des fréquences ainsi que le polygone des fréquences.

Exercice 6 Coefficient de variation

On connaît les salaires mensuels bruts en milliers de francs CFA des 200 employés de la même entreprise, à 10 ans d'intervalle. Les données sont groupées par classe. Le nombre d'employés est passé de 200 en 1994 à 280 en 2004. La dispersion des salaires a-t-elle augmenté?

Salaires	Effectif 1994	Effectif 2004
[1000, 2000[40	56
[2000, 3000[70	118
[3000, 4000[80	92
[4000, 5000[5	10
[5000, 10000[5	4

Exercice 7 Considérons le tableau de données de l'exercice précédent.

- (1) Identifier la population étudié, le caractère et sa nature.
- (2) Pour l'année 2004, représenter sur le même graphique les fréquences cumulées croissantes et décroissantes. Que représente l'intersection des deux courbes cumulatives? Justifier votre réponse.
- (3) Déterminer analytiquement et graphiquement les quartiles. Interpréter.
- (4) Déterminer la médiane. Interpréter.
- (5) Tracer la courbe de Lorenz. Interpréter.
- (6) L'indice de Gini confirme-t-il votre réponse à la question précédente?

Exercice 8

Le tableau suivant donne les prix et les quantités de deux produits 1 et 2, aux périodes 0 et t .

	Période 0		Période t	
	Prix	Quantité	Prix	Quantité
Produit 1	10	4	14	8
Produit 2	6	12	5	9

- (1) Calculer l'indice de Laspeyres des prix et l'indice de Laspeyres des quantités. Interpréter.
- (2) Calculer l'indice de Paasche des prix et l'indice de Paasche des quantités. Interpréter.
- (3) Comparer les résultats obtenus aux questions précédentes.
- (4) Calculer l'indice de Fisher des prix et l'indice de Fisher des quantités

Exercice 9 PRIX ET QUANTITES DES TROIS GRANDES CATEGORIES DE PRODUITS LAITIERS VENDUES DANS UN HYPERMARCHÉ ENTRE 1995 ET 2015.

	Prix en franc CFA			Quantité en milliers de kg		
	1995	2005	2015	1995	2005	2015
Yaourts	1100	1900	2500	1715	1835	1930
Lait	500	900	1100	2012	2128	2365
Fromage	3000	5500	9500	2478	2253	1992

- (1) Calculer et interpréter :
 - (a) La valeur globale du lait en 1995, 2005 et 2015
 - (b) La valeur globale du fromage en 1995, 2005 et 2015
- (2) Calculer l'indice de valeur I_v entre 1995 et 2005 puis entre 2005 et 2015 (2005 sera la date de référence) pour :
 - (a) Le lait
 - (b) Le fromage
- (3) Vérifier que

$$I_v = L_{pt_0} \times P_{qt_0}$$

$$I_v = L_{qt_0} \times P_{pt_0}$$

L_p et P_p correspondent respectivement aux indices de Laspeyres et de Paasche des prix. L_q et P_q correspondent respectivement aux indices de Laspeyres et de Paasche des quantités.

- (4) Peut-on dire que l'augmentation du prix est responsable de l'augmentation de la valeur du lait ? Vous expliquerez votre réponse.
- (5) Calculer la valeur globale de l'ensemble des produits laitiers en 2005 et 2015.
- (6) Calculer l'indice de Laspeyres "produits laitiers" des prix entre 2005 et 2015.
- (7) Calculer l'indice de Laspeyres "produits laitiers" des quantités entre 2005 et 2015.
- (8) Calculer les indices de Paasche et de Fisher des prix entre 2005 et 2015.

Exercice 10 Le tableau de contingence suivant est entre le salaire mensuel X et l'ancienneté Y des ouvriers d'une entreprise.

X	Y			
		[20, 25[[25, 30[[30, 35[
[0, 6[10	8	5	
[6, 12[8	9	4	
[12, 18[15	11	9	
[18, 24[3	6	2	

- (1) Quel est le nombre d'individus qui ont une ancienneté de chômage moins d'un an ?
- (2) Déterminer les deux distributions marginales.
- (3) Déterminer la distribution de X conditionnelle à $Y = [25, 30[$.
- (4) Les variables X et Y sont-elles indépendantes? Justifier.
- (5) Calculer le coefficient de corrélation linéaire. Commenter.
- (6) Donner l'équation de la droite de régression de Y en fonction de X .
- (7) Quel sera l'âge d'une personne ayant une ancienneté de chômage de 15 mois.

Exercice 11 Un site internet reçoit 113 457 visiteurs durant un mois. On désigne par X le navigateur internet utilisé et Y le système d'exploitation utilisé.

X \ Y	Windows	Mac	Linux
Chrome	14103	1186	427
Firefox	30853	4392	3234
Internet explorer	47389	23	0
Safari	668	6416	0
Autres	2974	40	1752

- (1) Identifier la population, sa taille ainsi que les variables étudiées en précisant leur type.
- (2) Quelle est la proportion de visiteurs sous Windows ?
- (3) Quelle proportion de visiteurs utilisent le navigateur Safari ?
- (4) Parmi les utilisateurs de Mac, quelle proportion utilise Chrome ?
- (5) Parmi les utilisateurs de Safari, quelle proportion est sous Windows ?
- (6) Représenter graphiquement la distribution des proportions par Navigateur pour chaque système d'exploitation. Les variables X et Y sont-elles indépendantes ?
- (7) Calculer le Coefficient de Cramer. Interpréter.

Exercice 12 Espérance de vie moyenne des personnes vulnérables au sein des dix villes les plus riches et des villes les plus pauvres aux Etats-Unis (richesse urbaine évaluée sur la base du revenu moyen par habitant).

Ville	Richesse	Espérance de vie	Ville	Richesse	Espérance de vie
New York City	R	79.5	Gary, Ind	P	74.2
Santa Barbara	R	79.4	Indianapolis	P	74.6
Santa Rosa	R	79	Detroit	P	74.8
Los Angeles	R	79	Louisville, Ky	P	74.9
San Francisco	R	78.8	Tulsa, Okla	P	74.9
San Diego	R	78.8	Toledo Ohio	P	74.9
Portland, Me	R	78.2	Oklahoma City	P	75
Boston	R	78.1	Dayton, Ohio	P	75,1
Miami	R	78.3	Knoxville, Ten	P	75.1
Newmark	R	78.2	Little Rock, Ark	P	75.1

- (1) Calculer la moyenne et la variance marginale de l'espérance de vie à la naissance.
- (2) Déterminer la moyenne et la variance de l'espérance de vie conditionnellement au niveau de richesse de la ville.
- (3) Calculer le rapport de corrélation.
- (4) Commenter le résultat obtenu.

Exercice 13 Le tableau suivant nous donne l'évolution du chiffre d'affaire trimestriel (en millions) d'un auto-entrepreneur au cours de quatre années consécutives.

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Y_t	8	10	7	30	13	14	8	40	16	18	11	50	20	20	14	60

- (1) Représenter graphiquement cette série chronologique (faire les deux types de représentation).
- (2) Déterminez la tendance de cette série en utilisant les moyennes mobiles (dont vous choisirez correctement l'ordre). La tendance est-elle linéaire?
- (3) Calculer l'équation de la droite de la tendance et tracer cette droite sur le graphique du correlogramme.
- (4) Le modèle est-il additif ou multiplicatif?
- (5) Déterminez les variations saisonnières.
- (6) Déterminer les séries corrigées des variations saisonnières.
- (7) Quelle est la prévision du chiffre d'affaire à la date 17?

Exercice 14 On dispose d'une statistique de vente d'un certain produit dans une entreprise (l'unité retenue ci-dessous étant la centaine d'articles la plus proche du nombre exact) :

	Du 01/01 au 30/01	Du 01/05 au 31/08	Du 01/09 au 31/12
2009	43	41	48
2010	49	47	57
2011	55	53	

- (1) Déterminer le trend T_t par la méthode de la moyenne mobile pour éliminer les mouvements saisonniers puis faire un ajustement linéaire du trend par la méthode des moindres carrés ordinaires. Afin d'estimer la qualité de l'ajustement, calculer la valeur du coefficient de détermination.
- (2) Après avoir estimé les composantes saisonnières, donner une estimation des ventes pour la période allant du 01/09/2011 au 31/12/2011.