

MIASS 231.1
 MATHÉMATIQUES
 (APPLIQUÉES AUX SCIENCES SOCIALES) 3

© El Hadj Touré, Ph D. Sociologie
 Section de sociologie, UGB de St-Louis

Révision
Statistiques descriptives

15.08 1

1

Consignes de l'examen
 Samedi 6 mai, 12h-14h, C16

- Tous les documents imprimés sont autorisés
 - Attention, piège!
 - L'examen peut se faire sans consultation de documents
- Apportez une calculatrice de base
 - Opérations arithmétiques
- L'ordinateur, la tablette et le téléphone sont interdits
- Répondez rapidement à la suite des questions du cahier
 - Vous pouvez utiliser un crayon effaçable
 - Soyez bref et précis
- La notation: justesse des calculs & rédaction de qualité

2

2

Au programme

<ul style="list-style-type: none"> ■ Description statistique <ul style="list-style-type: none"> ❖ Classification des variables ❖ Statistiques descriptives & questions de recherche ■ Distribution de % <ul style="list-style-type: none"> ❖ Ratio, taux, %, proportion ❖ % de variation ■ Mesures de tendance centrale <ul style="list-style-type: none"> ❖ Mode, médiane & moyenne ❖ Avantages & inconvénients 	<ul style="list-style-type: none"> ■ Mesures de variation <ul style="list-style-type: none"> ❖ Variance & écart-type ❖ Coefficient de variation ❖ Asymétrie & aplatissement ❖ Scores standardisés-z ❖ Interprétation statistique et interprétation sociologique ■ Aperçu de l'examen intra <ul style="list-style-type: none"> - Type de questions - Contenu de l'épreuve
--	---

3

Concepts statistiques de base

Variable qualitative vs variable quantitative

- Une variable est dite qualitative lorsque ses valeurs possibles sont des catégories (qualités)
 - Elle est nominale lorsque ses catégories servent simplement à identifier, nommer les individus. Ex: sexe (1.homme, 2.femme)
 - Elle est ordinaire lorsque ses catégories sont ordonnées. Ex: niveau d'éducation (1.primaire, 2.secondeire, 3.post-secondeire)
- Une variable quantitative est une variable d'intervalles/ ratio, dont les valeurs sont des nombres (quantités)
 - Elle est continue lorsque les nombres, arrondis, contiennent intrinsèquement des décimales. Ex: âge (20ans,21ans, 22ans...)
 - Elle est discrète lorsque les nombres, séparés, relèvent d'un comptage. Ex: nombre d'enfants dans une famille (0; 1; 2;...)

4

Concepts statistiques de base

Remue-méninges

→ Dans les questions suivantes, nommez les variables et dites laquelle est discrète, continue, nominale ou ordinaire. Justifiez!

- Depuis combien de temps êtes-vous marié(e)?
- Combien de membres compte-t-elle votre famille?
- La religion est-elle importante? (très, assez, peu, pas du tout)
- À quelle religion appartenez-vous?

→ Un chercheur veut étudier la consommation de cigarettes chez les élèves du cycle secondaire de Saint-Louis

1. Quelle est l'unité d'analyse de l'étude?
2. Quelles sont les questions qu'il peut poser pour obtenir des informations sur l'importance de la consommation?
 - Variable qualitative nominale (précisez les valeurs)
 - Variable qualitative ordinaire (précisez les valeurs)
 - Variable quantitative (précisez les valeurs)

15:03 5

5

Description statistique

Aperçu des techniques

→ Statistiques descriptives= l'objet est de décrire un ensemble de données, surtout d'échantillon, afin de les rendre plus intelligibles, communicables, utiles (réduction)

- Distribution de fréquences (f) et de pourcentages (%)
 - Quel est le degré d'importance ou l'ampleur d'un phénomène?
- Mesures de tendance centrale
 - Quel est le score typique, le plus représentatif d'un phénomène?
- Mesures de variation ou dispersion
 - Quel est le degré de variabilité d'un phénomène (homogénéité)?

→ Les statistiques inférentielles (2^e volet) seront abordées au second semestre dans le cadre du MIASS241.1

6

Distribution de pourcentages

Proportions, %, taux & ratios

- Une étude porte sur la représentativité féminine en socio à l'UGB: homme (79) ; femme (75); n = 154 (Maths 4)
 - Proportion de femmes = $\frac{75}{154} = 0,49$
 - Pourcentage de femmes = $0,49 \times 100 = 49\%$
 - Taux de féminité (100) = 49%
 - Ratio femmes: hommes = $\frac{75}{79} = 0.95$
- Selon la Banque mondiale, avec 15 420 000 habitants, le Sénégal compte 1 066 médecins (public) en 2017. Quel est le taux de médecins et le ratio au Sénégal?
 - Taux de médecins/100 000 = $(\frac{1066}{15\ 420\ 000}) \times 100\ 000 = 6,9$
 - Ratio d'habitants par médecin = $(\frac{15\ 420\ 000}{1\ 066}) = 14\ 465$

7

Distribution de pourcentages

% de variation

- Selon la Banque mondiale, le PIB/hbt du Sénégal est passé de 617\$ en 2000 à 1 318\$ en 2012. Quel est le % d'augmentation du PIB/hbt de 2000 à 2012?
- Formule du % de variation (augmentation ou diminution)

$$\% \text{ de variation} = \frac{\text{Valeur à un tps donné (T2)} - \text{Valeur au tps de référence (T1)}}{\text{Valeur au tps de référence (T1)}} \times 100$$
- Calcul $\% \text{ de variation} = \frac{1318 - 617}{617} \times 100 = 114\%$
- Quel est le % d'augmentation du PIB/hbt de 2012 à 2020, si en 2020 le PIB/hbt s'élève à 1 488\$? 13%

8

Mesures de tendance centrale

Le mode

- Valeur qui présente la fréquence la plus élevée dans une distribution donnée (Mo): «c'est la valeur à la mode»

Nombre d'enfants désiré chez les étudiants SOL1020

→ Le mode est égal à 2. Une pluralité d'étudiants, soit 37%, désirent avoir une famille constituée de deux enfants

9

Mesures de tendance centrale

La médiane

- Valeur qui divise en deux parties égales un ensemble ordonné de scores d'une variable quantitative (**Md**)
 - 50^e centile (C_{50}). Les centiles C_i divisent les scores en 100 parties =
 - 2^e quartile (Q_2). Les quartiles Q_i divisent les scores en 4 parties =

X_{\min} 25% 25% **Md** 25% 25% X_{\max}
 Q_1/C_{25} Q_2/C_{50} Q_3/C_{75}

- La médiane s'interprète ainsi : « au moins 50% des individus ont le score médian ou moins »
 - Ex. avec n impair (n=5): 31 35 **(39)** 45 50
 - Ex. avec n pair (n=6): 30 34 38 42 49 55
(40)

10

Mesures de tendance centrale

La moyenne

- La moyenne est la somme Σ de tous les scores X_i d'une variable quantitative, relativisée par le nombre d'individus

Données d'échantillon	Données de population
Statistique $\bar{X} = \frac{\Sigma X_i}{n}$	Paramètre $\mu = \frac{\Sigma X_i}{N}$

- Calcul avec données brutes
 - Nombre d'enfants désiré chez 10 cas: 0 4 1 2 3 4 4 1 2 4
 - $\bar{X} = (0+4+1+2+3+4+4+1+2+4) / 10 = 2,5$ enfants par femme
- Interprétation statistique
 - La moyenne est égale à 2,5. Autrement dit, le nombre moyen d'enfants par femme est de 2,5 dans l'échantillon.

11

Mesures de tendance centrale

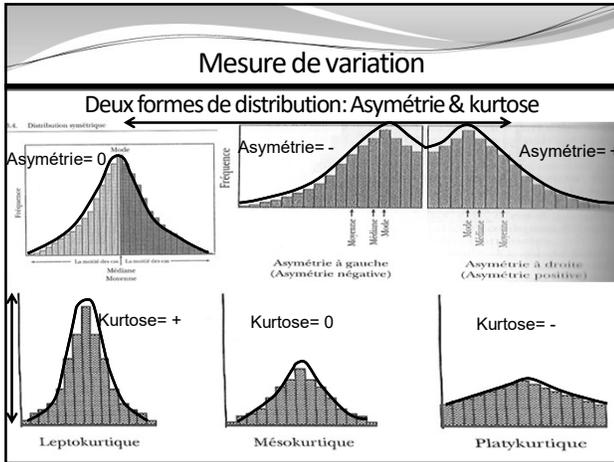
Moyenne, Médiane & Mode

→ Soit la distribution de revenu en \$ de 7 employés
20000, 22000, 34000, 34000, 35000, 36000, **500000**

- Moyenne = 96714\$;
- Médiane = 34000\$;
- Mode = 34000\$

- Au contraire de la médiane et du mode, la moyenne est affectée par la présence du cas déviant (500000)
 - Toutefois, elle intègre toute l'information contenue dans les scores
- La médiane est appropriée pour créer deux groupes, et le mode lorsqu'on veut le score réel le plus fréquent
- Les mesures de tendance centrale ne disent pas tout: elles n'informent pas sur la dispersion des scores

12



16

Mesures de variation

Scores standardisés ou scores-z

- Qui de ces deux étudiants de socio (Mamad) et géo (Dija) de l'UGB a la meilleure note par rapport à sa classe?

Étudiant	Note	Moyenne	Écart-type	Score-z
Mamad	17	14	3	1
Dija	16	12,5	2	1,75

- Note standardisée de Mamad: $(17 - 14) / 3 = 6/6 = 1,00$
- Note standardisée de Dija: $(16 - 12,5) / 2 = 1,75$

- Interprétation stat.:** Les cotes-z respectifs de Dija et Mamad sont de 1,75 et 1. La note de Dija est meilleure: elle se situe à 1,75 écart-type au-dessus de la moyenne de sa classe

17

Statistiques descriptives

Interprétation statistique

- Un sociologue s'intéresse au nombre d'heures passées sur Internet par jour chez 100 étudiants inscrits en L2 de sociologie à l'UGB (n=100). Interprétez statistiquement chacune des statistiques ci-dessous?

Moyenne	7,140
Médiane	5,000
Mode	4,0
Ecart type	5,7177
Asymétrie	1,450
Kurtose	1,479
Étendue	24,0
Minimum	,0
Maximum	24,0
Percentiles	
25	4,000
50	5,000
75	9,500

18

Aperçu de l'examen

- Questions objectives de compréhension
 - Identifier la variable clé et l'unité d'analyse d'une étude quanti.
 - Différencier les variables : nominale, ordinale, discrète, continue
- Questions de calculs et problèmes pratiques
 - Calculer les mesures suivantes: ratio, taux, % de variation, écart-type, coefficient de variation
 - Classer des individus selon leurs scores-z
- Analyse de statistiques descriptives
 - Savoir interpréter statistiquement: mode, médiane, moyenne, écart-type, coefficient d'asymétrie et d'aplatissement, quartiles

19

19

BON SUCCÈS 😊

20

20
