

**Objectifs du chapitre 2
d'Howell sur les
statistiques descriptives**

- ❖ Savoir calculer les caractéristiques d'une distribution de cotes dans un échantillon à l'aide des indices de tendance centrale et de variabilité
- ❖ Savoir définir les concepts suivants
 - suffisance
 - biais
 - efficacité
 - robustesse
 - degré de liberté
- ❖ Savoir discuter de la pertinence des indices de tendance centrale et de variabilité

**2 types de
statistique
descriptive**

- Indice (estimateur) de
tendance centrale
- Indice (estimateur) de
variabilité

**3 mesures de
tendance centrale**

- 1) Mode
- 2) Médiane
- 3) Moyenne
point d'équilibre de toutes les
valeurs d'une distribution

Comment trouver les mesures de tendance centrale

- 1) Mode

- 2) Médiane

- 3) Moyenne

Comparaison des mesures de tendance centrale (1)

	Mode	Médi-ane	Moy-enne
Vrai chiffre	Oui	Peut-être	Peut-être
Représente	majorité	Point central	Peut-être
Échelle de mesure	Nominale +	Ordinale +	Intervalle+
Dépend regroupement	Oui	Non	Non

Comparaison des mesures de tendance centrale (2)

	Mode	Médi-ane	Moy-enne
Influence des Valeurs extrêmes	Non	Non	Oui
Opérations mathématiques	Non	Non	Oui

**6 mesures de variabilité
(1)**

1) Étendue

2) Espace interquartile

**6 mesures de variabilité
(2)**

3) Écart moyen (pas utile)
somme de toutes les différences
entre les valeurs et la moyenne

4) Écart absolu moyen
(pas utile)
somme de toutes les différences
absolues entre les valeurs et la
moyenne

**6 mesures de variabilité
(3)**

5) Variance
somme des carrés des
différences entre les valeurs et
la moyenne

6) Écart type

Formule de calcul de la variance

La variance (et l'écart type) se calcule plus facilement avec la formule suivante

$$\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N-1}$$

(et l'écart type)

$$\sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N-1}}$$

D'où vient la formule de calcul de la variance ?

Développement de $\sum (X_i - M)^2$

$$\begin{aligned} \sum (X_i - M)^2 &= \sum (X_i - M) (X_i - M) \\ &= \sum (X_i^2 - 2 X_i M + M^2) \\ &= \sum X_i^2 - 2M \sum X_i + \sum M^2 \\ &= \sum X_i^2 - 2M \sum X_i + N M^2 \\ &= \sum X_i^2 - 2 \frac{\sum X_i}{N} \sum X_i + N \left(\frac{\sum X_i}{N} \right)^2 \\ &= \sum X_i^2 - 2 \left(\frac{\sum X_i}{N} \right)^2 + \left(\frac{\sum X_i}{N} \right)^2 \\ &= \sum X_i^2 - \left(\frac{\sum X_i}{N} \right)^2 \quad \text{où } M = \bar{X} \end{aligned}$$

Avantages et inconvénient de l'écart type

Avantages

- 1) sur la même échelle que les valeurs originales
- 2) Moyenne ± 1 (É.T.) = 66%
Moyenne ± 2 (É.T.) = 95%
Moyenne ± 3 (É.T.) = 99%
des valeurs de l'échantillon
- 3) É.T. = 1/5 ou 1/6 de l'étendue

Inconvénient

biais dû aux valeurs extrêmes

Pour comparer entre eux des écarts-types d'échantillons différents

Calculer le coefficient de variation (CV)

$$CV = \frac{\sigma}{\mu} = \frac{ET.}{\bar{X}}$$

La moyenne est un bon estimateur de la tendance centrale de la population

	Mode	Médiane	Moyenne
Suffisance	Non	Non	
sans biais	Non		Oui
Efficacité		Non	Oui
Robuste	Oui	Oui	Non

La variance est un bon estimateur de la variation dans la population

	Étendue	Écart inter-quartile	Variance
Suffisance		Non	Oui
sans biais	Non	Non	
Efficacité	Non	Non	Oui
Robustesse	Oui		
