COURS. STATISTIQUE DESCRIPTIVE
Chapitre 03. PARAMETRES D'UNE VARIABLE STATISTIQUE

Département Mathématiques.

PRESENTATION
Dr. André Souleye Diabang

25 novembre 2021

Plan

Plan

Introduction

.

L'objectif est de résumer à travers quelques indicateurs numériques ou paramètres caractéristiques la distribution d'une variable statistique.

Introduction

.

L'objectif est de résumer à travers quelques indicateurs numériques ou paramètres caractéristiques la distribution d'une variable statistique. On les appelle des indicateurs de synthèse d'une distribution statistique.

Introduction

.

L'objectif est de résumer à travers quelques indicateurs numériques ou paramètres caractéristiques la distribution d'une variable statistique. On les appelle des indicateurs de synthèse d'une distribution statistique. On utilise des paramètres de position (ou de tendance centrale), des paramètres de dispersion et des paramètres de forme.

Introduction

.

L'objectif est de résumer à travers quelques indicateurs numériques ou paramètres caractéristiques la distribution d'une variable statistique. On les appelle des indicateurs de synthèse d'une distribution statistique. On utilise des paramètres de position (ou de tendance centrale), des paramètres de dispersion et des paramètres de forme. L'analyse numérique et l'analyse graphique d'une distribution sont complémentaires et non exclusives.

Plan

Paramètres de position.

.

Les paramètres de position sont des valeurs numériques, calculées à partir d'une série statistique et qui permettent de déterminer la valeur typique ou l'ordre de grandeur de la distribution.

Paramètres de position.

.

Les paramètres de position sont des valeurs numériques, calculées à partir d'une série statistique et qui permettent de déterminer la valeur typique ou l'ordre de grandeur de la distribution. Les principales caractéristiques de position sont : le mode, la médiane et la moyenne.

Mode.

Définition.

Le mode est la valeur la plus fréquente dans une série d'observations. (c'est la valeur qui le plus grand effectif). On le note Mo.

Mode.

Définition.

Le mode est la valeur la plus fréquente dans une série d'observations. (c'est la valeur qui le plus grand effectif). On le note Mo.
Dans le cas d'une variable quantitative continue on appelle "classe modale "

Mode.

Définition.

Le mode est la valeur la plus fréquente dans une série d'observations. (c'est la valeur qui le plus grand effectif). On le note Mo.

Dans le cas d'une variable quantitative continue on appelle "classe modale " la classe qui présente l'effectif le plus élevé.

Mode.

Définition.

Le mode est la valeur la plus fréquente dans une série d'observations. (c'est la valeur qui le plus grand effectif). On le note Mo.
Dans le cas d'une variable quantitative continue on appelle "classe modale " la classe qui présente l'effectif le plus élevé.

Remarque.

Le mode d'une série n'est pas nécessairement unique. Il peut ne pas exister

Mode.

Exemple.

Déterminer le mode des séries A1 . . . A5.

Mode.

Exemple.

Déterminer le mode des séries A1 . . . A5.

$S = \{1; 7; 2; 4; 5; 3\}$

Mode.

Exemple.

Déterminer le mode des séries A1 ... A5.
$S = \{1; 7; 2; 4; 5; 3\}$

Avantages et inconvénients du mode.

Mode.

Exemple.

Déterminer le mode des séries A1 ... A5.
$S = \{1; 7; 2; 4; 5; 3\}$

Avantages et inconvénients du mode.

(1) La détermination du mode est aisée (graphiquement).

Mode.

Exemple.

Déterminer le mode des séries A1 . . . A5.
$S = \{1; 7; 2; 4; 5; 3\}$

Avantages et inconvénients du mode.

(1) La détermination du mode est aisée (graphiquement).
(2) Son intérêt est évident puisqu'il désigne la valeur de la variable qui est la plus observée sur l'échantillon.

Mode.

Exemple.

Déterminer le mode des séries A1 ... A5.
$S = \{1; 7; 2; 4; 5; 3\}$

Avantages et inconvénients du mode.

(1) La détermination du mode est aisée (graphiquement).
(2) Son intérêt est évident puisqu'il désigne la valeur de la variable qui est la plus observée sur l'échantillon.
(3) Le mode n'a de signification véritable que si l'effectif correspondant est nettement supérieur aux effectifs des autres modalités.

Mode.

Exemple.

Déterminer le mode des séries A1 . . . A5.
$S = \{1; 7; 2; 4; 5; 3\}$

Avantages et inconvénients du mode.

(1) La détermination du mode est aisée (graphiquement).
(2) Son intérêt est évident puisqu'il désigne la valeur de la variable qui est la plus observée sur l'échantillon.
(3) Le mode n'a de signification véritable que si l'effectif correspondant est nettement supérieur aux effectifs des autres modalités.
(4) Le mode n'est intéressant que lorsqu'il est unique.

Médiane.

Définition.

C'est la valeur qui sépare une série d'observations ordonnées en ordre croissant ou décroissant, en deux parties comportant le même nombre d'observations. On la désigne par la notation Me.

Médiane.

Définition.

C'est la valeur qui sépare une série d'observations ordonnées en ordre croissant ou décroissant, en deux parties comportant le même nombre d'observations. On la désigne par la notation Me.

Médiane.

Méthode de calcul-Cas des données non-groupées.

Médiane.

Méthode de calcul-Cas des données non-groupées.

–(1) Ordonner la série par ordre croissant ou décroissant.

Médiane.

Méthode de calcul-Cas des données non-groupées.

–(1) Ordonner la série par ordre croissant ou décroissant.
–(2) Déterminer l'effectif total N.

Médiane.

Méthode de calcul-Cas des données non-groupées.

–(1) Ordonner la série par ordre croissant ou décroissant.

–(2) Déterminer l'effectif total N.

–(3) Calculer $\dfrac{N}{2}$.

Médiane.

Méthode de calcul-Cas des données non-groupées.

–(1) Ordonner la série par ordre croissant ou décroissant.
–(2) Déterminer l'effectif total N.
–(3) Calculer $\frac{N}{2}$. on distingue 2 cas possibles :

Médiane.

Méthode de calcul-Cas des données non-groupées.

–(1) Ordonner la série par ordre croissant ou décroissant.
–(2) Déterminer l'effectif total N.
–(3) Calculer $\dfrac{N}{2}$. on distingue 2 cas possibles :
– (a) Cas 1 : Si $\dfrac{N}{2}$ est un entier. On l'occalise $x(\frac{N}{2})$ et $x_{(\frac{N}{2}+1)}$ dans le série. $Me = \frac{x_{(\frac{N}{2})}+x_{(\frac{N}{2}+1)}}{2}$.

Médiane.

Méthode de calcul-Cas des données non-groupées.

–(1) Ordonner la série par ordre croissant ou décroissant.
–(2) Déterminer l'effectif total N.
–(3) Calculer $\frac{N}{2}$. on distingue 2 cas possibles :
– (a) Cas 1 : Si $\frac{N}{2}$ est un entier. On l'occalise $x(\frac{N}{2})$ et $x_{(\frac{N}{2}+1)}$ dans le série. $Me = \frac{x_{(\frac{N}{2})}+x_{(\frac{N}{2}+1)}}{2}$.
– (b) Cas 2 : Si $\frac{N}{2}$ n'est pas un entier. On détermine $E(\frac{N}{2})$ et on l'occalise $x_{(E(\frac{N}{2})+1)}$ dans le série. Alors $Me = x_{(E(\frac{N}{2})+1)}$.

Médiane.

Méthode de calcul-Cas des données groupées.

Médiane.

Méthode de calcul-Cas des données groupées.

Si les données sont groupées par classes (cas des variables continues) il faut :

Médiane.

Méthode de calcul-Cas des données groupées.

Si les données sont groupées par classes (cas des variables continues) il faut :
–(1) localiser la classe médiane, c'est-à-dire celle qui contient la médiane.

Médiane.

Méthode de calcul-Cas des données groupées.

Si les données sont groupées par classes (cas des variables continues) il faut :
–(1) localiser la classe médiane, c'est-à-dire celle qui contient la médiane.
–(2) calculer par extrapolation linéaire la valeur de la médiane.

Médiane.

Méthode de calcul-Cas des données groupées.

Si les données sont groupées par classes (cas des variables continues) il faut :
–(1) localiser la classe médiane, c'est-à-dire celle qui contient la médiane.
–(2) calculer par extrapolation linéaire la valeur de la médiane.
–(3) ou déterminer la médiane par projection à partir du diagramme des fréquences cumulées.

Médiane.

Méthode de calcul-Cas des données groupées.

Si les données sont groupées par classes (cas des variables continues) il faut :
–(1) localiser la classe médiane, c'est-à-dire celle qui contient la médiane.
–(2) calculer par extrapolation linéaire la valeur de la médiane.
–(3) ou déterminer la médiane par projection à partir du diagramme des fréquences cumulées.

Médiane.

Méthode de calcul-Cas des données groupées.

Médiane.

Méthode de calcul-Cas des données groupées.

NB : La classe médiane est celle dont la fréquence cumulée est $\geq 50\%$ et dont la classe précédente à une fréquence cumulée $< 50\%$.

Médiane.

Méthode de calcul-Cas des données groupées.

NB : La classe médiane est celle dont la fréquence cumulée est $\geq 50\%$ et dont la classe précédente à une fréquence cumulée $< 50\%$.
Si on note Me la médiane, e_1 la borne inférieure de la classe médiane, e_2 la borne supérieure de la classe médiane, f_e la fréquence de la classe médiane, $F(e_1)$ est la fréquence cumulée à la classe précédant la classe médiane. Alors

$$Me = e_1 + \frac{0,5 - F(e_1)}{f_e} \times (e_2 - e_1)$$

Médiane.

Exemple.

Déterminer le médiane des séries A1, A4 et A5.

Médiane.

Exemple.

Déterminer le médiane des séries A1, A4 et A5.

Avantages et inconvénients de la médiane.

Médiane.

Exemple.

Déterminer le médiane des séries A1, A4 et A5.

Avantages et inconvénients de la médiane.

(1) Son calcul est facile.

Médiane.

Exemple.

Déterminer le médiane des séries A1, A4 et A5.

Avantages et inconvénients de la médiane.

(1) Son calcul est facile.
(2) Donne une idée satisfaisante de la tendance centrale de la distribution.

Médiane.

Exemple.

Déterminer le médiane des séries A1, A4 et A5.

Avantages et inconvénients de la médiane.

(1) Son calcul est facile.
(2) Donne une idée satisfaisante de la tendance centrale de la distribution.
(3) Elle ne tient pas compte des valeurs prises par la variable mais seulement de leurs ordres de grandeur.

Médiane.

Exemple.

Déterminer le médiane des séries A1, A4 et A5.

Avantages et inconvénients de la médiane.

(1) Son calcul est facile.
(2) Donne une idée satisfaisante de la tendance centrale de la distribution.
(3) Elle ne tient pas compte des valeurs prises par la variable mais seulement de leurs ordres de grandeur.
(4)Elle concerne uniquement les variables quantitatives.

Moyenne.

Définition.

La moyenne arithmétique d'un ensemble de données est la somme des valeurs obtenues divisée par le nombre d'observations. Elle est notée $\overline{x}$ pour une variable notée X.

Moyenne.

Définition.

La moyenne arithmétique d'un ensemble de données est la somme des valeurs obtenues divisée par le nombre d'observations. Elle est notée $\overline{x}$ pour une variable notée X.

$$\overline{x} = \frac{\sum n_i x_i}{N}$$

Moyenne.

Définition.

La moyenne arithmétique d'un ensemble de données est la somme des valeurs obtenues divisée par le nombre d'observations. Elle est notée $\overline{x}$ pour une variable notée X.

$$\overline{x} = \frac{\sum n_i x_i}{N}$$

Moyenne.

Remarque

Dans le cas où les données sont groupées par classes, on remplace les x_i par les c_i.

Moyenne.

Remarque

Dans le cas où les données sont groupées par classes, on remplace les x_i par les c_i.

Exemple

Calculer la moyenne des séries A_1, A_4 et A_5

Moyenne.

Avantages et inconvénients de la moyenne.

Moyenne.

Avantages et inconvénients de la moyenne.

(1) Du fait qu'elle utilise pour son calcul toutes les valeurs prises par la variable, la moyenne arithmétique est la meilleure des caractéristiques de position.

Moyenne.

Avantages et inconvénients de la moyenne.

(1) Du fait qu'elle utilise pour son calcul toutes les valeurs prises par la variable, la moyenne arithmétique est la meilleure des caractéristiques de position.
(2) La moyenne n'a de sens que pour des variables quantitatives.

Moyenne.

Avantages et inconvénients de la moyenne.

(1) Du fait qu'elle utilise pour son calcul toutes les valeurs prises par la variable, la moyenne arithmétique est la meilleure des caractéristiques de position.
(2) La moyenne n'a de sens que pour des variables quantitatives.
(3)La moyenne arithmétique présente l'inconvénient d'être sensible aux valeurs extrêmes (valeurs aberrantes).

Plan

Paramètres de dispersion.

.

Elles servent à mesurer la variabilité de la variable statistique et de juger de la pertinence (représentativité) de la caractéristique de tendance centrale. On distingue :

Paramètres de dispersion.

.

Elles servent à mesurer la variabilité de la variable statistique et de juger de la pertinence (représentativité) de la caractéristique de tendance centrale. On distingue : l'étendue,

Paramètres de dispersion.

.

Elles servent à mesurer la variabilité de la variable statistique et de juger de la pertinence (représentativité) de la caractéristique de tendance centrale. On distingue : l'étendue,Intervalle interquartile,

Paramètres de dispersion.

.

Elles servent à mesurer la variabilité de la variable statistique et de juger de la pertinence (représentativité) de la caractéristique de tendance centrale. On distingue : l'étendue,Intervalle interquartile,l'écart absolu moyen,

Paramètres de dispersion.

.

Elles servent à mesurer la variabilité de la variable statistique et de juger de la pertinence (représentativité) de la caractéristique de tendance centrale. On distingue : l'étendue,Intervalle interquartile,l'écart absolu moyen,variance et écart-type,

Paramètres de dispersion.

.

Elles servent à mesurer la variabilité de la variable statistique et de juger de la pertinence (représentativité) de la caractéristique de tendance centrale. On distingue : l'étendue,Intervalle interquartile,l'écart absolu moyen,variance et écart-type,les coefficients de variation

Etendue.

Définition.

C'est la différence entre la plus grande et la plus petite valeur observée.

Etendue.

Définition.

C'est la différence entre la plus grande et la plus petite valeur observée.

Exemple

Calculer l'étendue des séries A_1, A_4 et A_5

Etendue.

Interprétation, avantages et inconvénients.

La signification de l'étendue est claire et sa détermination facile. Cependant, elle présente des inconvénients sérieux.

Etendue.

Interprétation, avantages et inconvénients.

La signification de l'étendue est claire et sa détermination facile. Cependant, elle présente des inconvénients sérieux. En effet, ne dépendant que des valeurs extrêmes qui sont souvent exceptionnelles

Etendue.

Interprétation, avantages et inconvénients.

La signification de l'étendue est claire et sa détermination facile. Cependant, elle présente des inconvénients sérieux. En effet, ne dépendant que des valeurs extrêmes qui sont souvent exceptionnelles voire aberrantes et non pas de tous les termes,

Etendue.

Interprétation, avantages et inconvénients.

La signification de l'étendue est claire et sa détermination facile. Cependant, elle présente des inconvénients sérieux. En effet, ne dépendant que des valeurs extrêmes qui sont souvent exceptionnelles voire aberrantes et non pas de tous les termes, elle est sujette à des fluctuations considérables d'un échantillon à un autre.

Les Quartiles.

Définition.

Ce sont les valeurs du caractère qui partagent la série en quatre sous-ensembles de tailles égales.

Les Quartiles.

Définition.

Ce sont les valeurs du caractère qui partagent la série en quatre sous-ensembles de tailles égales. Ils sont au nombre de 3 : Q_1, Q_2 et Q_3

Les Quartiles.

Définition.

Ce sont les valeurs du caractère qui partagent la série en quatre sous-ensembles de tailles égales. Ils sont au nombre de 3 : Q_1, Q_2 et Q_3
Q_1 : 25% de valeurs inférieures et 75% de valeurs supérieures.

Les Quartiles.

Définition.

Ce sont les valeurs du caractère qui partagent la série en quatre sous-ensembles de tailles égales. Ils sont au nombre de 3 : Q_1, Q_2 et Q_3

Q_1 : 25% de valeurs inférieures et 75% de valeurs supérieures.

Q_2 : 50% de valeurs inférieures et 50% de valeurs supérieures, Q_2 est la médiane.

Les Quartiles.

Définition.

Ce sont les valeurs du caractère qui partagent la série en quatre sous-ensembles de tailles égales. Ils sont au nombre de 3 : Q_1, Q_2 et Q_3

Q_1 : 25% de valeurs inférieures et 75% de valeurs supérieures.

Q_2 : 50% de valeurs inférieures et 50% de valeurs supérieures, Q_2 est la médiane.

Q_3 : 75% des valeurs inférieures et 25% des valeurs supérieures.

Les Quintiles, Déciles, Centiles.

Généralisation

Quintiles : Ils divisent la série en cinq sous-ensembles de tailles égales, soit 20%. Ils sont au nombre de cinq.

Les Quintiles, Déciles, Centiles.

Généralisation

Quintiles : Ils divisent la série en cinq sous-ensembles de tailles égales, soit 20%. Ils sont au nombre de cinq.
Déciles : Ils divisent la série en dix sous-ensembles de tailles égales, soit 10%. Ils sont au nombre de dix.

Les Quintiles, Déciles, Centiles.

Généralisation

Quintiles : Ils divisent la série en cinq sous-ensembles de tailles égales, soit 20%. Ils sont au nombre de cinq.
Déciles : Ils divisent la série en dix sous-ensembles de tailles égales, soit 10%. Ils sont au nombre de dix.
Centiles : Ils divisent la série en cent sous-ensembles de tailles égales, soit 1%. Ils sont au nombre de cent.

Les Quintiles, Déciles, Centiles.

Généralisation

Quintiles : Ils divisent la série en cinq sous-ensembles de tailles égales, soit 20%. Ils sont au nombre de cinq.
Déciles : Ils divisent la série en dix sous-ensembles de tailles égales, soit 10%. Ils sont au nombre de dix.
Centiles : Ils divisent la série en cent sous-ensembles de tailles égales, soit 1%. Ils sont au nombre de cent.

Intervalle interquartile.

Définition.

C'est la différence entre le 3^{me} et le $1^{e}r$ quartile. $I_Q = Q_3 - Q_1$.

Intervalle interquartile.

Définition.

C'est la différence entre le 3^{me} et le $1^{e}r$ quartile. $I_Q = Q_3 - Q_1$.
On définit de la même façon l'intervalle inter-décile
$(I_D = D_9 - D_1)$

Intervalle interquartile.

Définition.

C'est la différence entre le 3^{me} et le $1^{e}r$ quartile. $I_Q = Q_3 - Q_1$.
On définit de la même façon l'intervalle inter-décile
$(I_D = D_9 - D_1)$ et l'intervalle intercentile $(I_C = C_{99} - C_1)$.

Intervalle interquartile.

Interprétation, avantages et inconvénients.

L'utilisation de ces intervalles permet d'éliminer l'influence des valeurs extrêmes qui sont des valeurs rares ou aberrantes.

Intervalle interquartile.

Interprétation, avantages et inconvénients.

L'utilisation de ces intervalles permet d'éliminer l'influence des valeurs extrêmes qui sont des valeurs rares ou aberrantes.
La perte de l'information du fait de la diminution des observations qu'elle entraîne est compensée par l'homogénéité des données dans l'intervalle interquartile.

Ecart absolu moyen.

Définition.

C'est la moyenne des écarts absolus entre chaque observation et la moyenne.

$$EM = \frac{\sum_{i=1}^{n} n_i \mid x_i - \overline{x} \mid}{n}$$

Intervalle interquartile.

Interprétation, avantages et inconvénients.

• L'écart absolu moyen mesure la dispersion des valeurs observées d'une variable statistique autour d'une valeur centrale.

Intervalle interquartile.

Interprétation, avantages et inconvénients.

• L'écart absolu moyen mesure la dispersion des valeurs observées d'une variable statistique autour d'une valeur centrale.
• Une valeur faible de l'écart absolu moyen traduit une faible dispersion des valeurs autour de la valeur centrale.

Intervalle interquartile.

Interprétation, avantages et inconvénients.

• L'écart absolu moyen mesure la dispersion des valeurs observées d'une variable statistique autour d'une valeur centrale.

• Une valeur faible de l'écart absolu moyen traduit une faible dispersion des valeurs autour de la valeur centrale.

• Cependant la comparaison de cette caractéristique pour deux séries est difficile car sa valeur dépend de l'ordre de grandeur (échelle ou unité de mesure) des observations.

Variance.

Définition.

La variance est la moyenne des écarts (élevés au carré) des valeurs observées par rapport à la moyenne arithmétique de la série. On la note $V(X)$ pour une variable notée X.

$$V(X) = \frac{\sum_{i=1}^{n} n_i(x_i - \overline{x})^2}{N}$$

Variance.

Définition.

La variance est la moyenne des écarts (élevés au carré) des valeurs observées par rapport à la moyenne arithmétique de la série. On la note $V(X)$ pour une variable notée X.

$$V(X) = \frac{\sum_{i=1}^{n} n_i(x_i - \overline{x})^2}{N}$$

Ecart-type

L'écart-type est la racine carrée de la variance. On le note $\sigma(X)$ sa formule :

$$\sigma(X) = \sqrt{V(X)}$$

Variance et écart-type.

Interprétation, avantages et inconvénients.

• L'écart-type et la variance mesurent la dispersion de la variable autour de la moyenne. Ainsi, des valeurs élevées (respectivement faibles) de ces caractéristiques traduisent une grande (respectivement faible) dispersion des valeurs autour de la moyenne.

Variance et écart-type.

Interprétation, avantages et inconvénients.

• L'écart-type et la variance mesurent la dispersion de la variable autour de la moyenne. Ainsi, des valeurs élevées (respectivement faibles) de ces caractéristiques traduisent une grande (respectivement faible) dispersion des valeurs autour de la moyenne.

Variance et écart-type.

Interprétation, avantages et inconvénients.

• L'écart-type et la variance mesurent la dispersion de la variable autour de la moyenne. Ainsi, des valeurs élevées (respectivement faibles) de ces caractéristiques traduisent une grande (respectivement faible) dispersion des valeurs autour de la moyenne.

• La variance est calculée à partir des valeurs de la série élevées au carré. Ainsi l'unité (de mesure) de la variance est le carré de celle de la variable.

Variance et écart-type.

Interprétation, avantages et inconvénients.

• L'écart-type et la variance mesurent la dispersion de la variable autour de la moyenne. Ainsi, des valeurs élevées (respectivement faibles) de ces caractéristiques traduisent une grande (respectivement faible) dispersion des valeurs autour de la moyenne.

• La variance est calculée à partir des valeurs de la série élevées au carré. Ainsi l'unité (de mesure) de la variance est le carré de celle de la variable.

Par exemple, si la variable est mesurée en francs, en kg ou en mètre, la variance sera mesurée en francs au carré, en kg au carré ou en mètres au carré. Par contre l'écart-type a la même unité de mesure que la variable.

Coefficients de variation.

Coefficient de variation de l'écart-type.

Le coefficient de variation de l'écart-type est le rapport entre l'écart-type et la moyenne de la distribution. On le note CV_σ.

$$\boxed{CV_\sigma = \frac{\sigma(X)}{\overline{X}}}$$

Coefficients de variation.

Coefficient de variation de l'écart-type.

Le coefficient de variation de l'écart-type est le rapport entre l'écart-type et la moyenne de la distribution. On le note CV_σ.

$$CV_\sigma = \frac{\sigma(X)}{\overline{X}}$$

Coefficient de variation de l'intervalle interquartile.

Le coefficient de variation de l'intervalle interquartile est le rapport entre l'intervalle interquartile I_Q et la médiane M_e. On le note CV_Q.

$$CV_Q = \frac{Q_3 - Q_1}{Q_2}$$

Coefficients de variation.

Interprétation, avantages et inconvénients.

• Contrairement aux autres indicateurs de dispersion, le coefficient de variation est sans unité de mesure. On l'exprime souvent en pourcentage.

Coefficients de variation.

Interprétation, avantages et inconvénients.

• Contrairement aux autres indicateurs de dispersion, le coefficient de variation est sans unité de mesure. On l'exprime souvent en pourcentage.

• Du fait qu'elle est sans unité, le coefficient de variation présente l'avantage de ne pas être sensible à l'ordre de grandeur (ou à l'unité de mesure) de la variable mais seulement à la dispersion des valeurs autour de la moyenne. Ainsi on peut l'utiliser pour comparer la dispersion de deux séries dont les ordres de grandeur (ou les unités de mesure) sont différents.

Coefficients de variation.

Interprétation, avantages et inconvénients.

• Un coefficient de variation élevé (respectivement faible) traduit une grande (respectivement faible) dispersion de la variable autour de la moyenne.

Coefficients de variation.

Interprétation, avantages et inconvénients.

- Un coefficient de variation élevé (respectivement faible) traduit une grande (respectivement faible) dispersion de la variable autour de la moyenne.
- L'appréciation du niveau (faible ou élevé) du coefficient de variation est laissée aux soins de l'utilisateur. Cependant une valeur du CV supérieure à 10% doit susciter des questions quant à la représentativité de la moyenne comme caractéristique de tendance centrale.

Plan

Coefficient d'asymétrie de Fisher.

Définition

Le coefficient d'asymétrie de fisher est

$$g_1 = \frac{m_3}{\sigma(x)^3}$$

Coefficient d'asymétrie de Fisher.

Définition

Le coefficient d'asymétrie de fisher est

$$g_1 = \frac{m_3}{\sigma(x)^3}$$

avec $m_3 = \frac{\sum_{i=1}^{n} n_i(x_i - \overline{x})^3}{N}$ appelé le moment centré d'ordre 3.

Coefficient d'asymétrie de Pearson.

Définition

Le coefficient d'asymétrie de Pearson est basé sur une comparaison de la moyenne et du mode, et est normalisé par l'écart-type :

Coefficient d'asymétrie de Pearson.

Définition

Le coefficient d'asymétrie de Pearson est basé sur une comparaison de la moyenne et du mode, et est normalisé par l'écart-type :

$$A_P = \frac{\overline{x} - Mode}{\sigma(x)}$$

Coefficient d'asymétrie de Yule.

Définition

Le coefficient d'asymétrie de Yule est basé sur les positions des 3 quartiles (1er quartile, médiane et troisième quartile), et est normalisé par la distance interquartile :

Coefficient d'asymétrie de Yule.

Définition

Le coefficient d'asymétrie de Yule est basé sur les positions des 3 quartiles (1er quartile, médiane et troisième quartile), et est normalisé par la distance interquartile :

$$A_y = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

Interprétation des coefficients d'asymétries.

Interprétation.

Tous les coefficients d'asymétrie ont les mêmes propriétés, ils sont nuls si la série est symétrique, négatifs si la série est asymétrie à gauche et positifs si la série est asymétrie à droite.

Coefficient d'aplatissement de Pearsons.

Définition.

L'aplatissement de Pearson est :

$$\beta_2 = \frac{m_4}{\sigma(x)^4}$$

Coefficient d'aplatissement de Pearsons.

Définition.

L'aplatissement de Pearson est :

$$\boxed{\beta_2 = \frac{m_4}{\sigma(x)^4}}$$

avec $\boxed{m_4 = \frac{\sum_{i=1}^{n} n_i (x_i - \overline{x})^4}{N}}$ appelé le moment centré d'ordre 4.

Coefficient d'aplatissement de Fisher.

Définition.

L'aplatissement de Fisher est :

$$g_2 = \beta_2 - 3$$

Interprétation des coefficients d'aplatissements.

Interprétation.

Si $\beta_2 \cong 3$ ou $g_2 \cong 0$, alors la courbe du polygône statistique est mésokurtique.

Interprétation des coefficients d'aplatissements.

Interprétation.

Si $\beta_2 \cong 3$ ou $g_2 \cong 0$, alors la courbe du polygône statistique est mésokurtique.
Si $\beta_2 > 3$ ou $g_2 > 0$, alors la courbe du polygône statistique est leptokurtique.

Interprétation des coefficients d'aplatissements.

Interprétation.

Si $\beta_2 \cong 3$ ou $g_2 \cong 0$, alors la courbe du polygône statistique est mésokurtique.
Si $\beta_2 > 3$ ou $g_2 > 0$, alors la courbe du polygône statistique est leptokurtique.
Si $\beta_2 < 3$ ou $g_2 < 0$, alors la courbe du polygône statistique est platykurtique.