

**centile**

Les centiles divisent une distribution statistique ordonnée en 100 groupes comportant chacun 1 % des données de la distribution.

**coefficient de corrélation linéaire**

Le coefficient de corrélation linéaire est un nombre compris entre -1 et 1 qui quantifie l'intensité de la corrélation linéaire associée à un nuage de points.

**corrélacion**

La corrélation entre deux variables statistiques numériques est l'intensité de la liaison qui existe entre ces variables. Le type le plus simple de liaison est la relation affine. Lorsque la liaison entre les variables se traduit par une droite, on dit que la corrélation est linéaire.

**diagramme à tige et à feuilles**

Un diagramme à tige et à feuilles permet de classer des données quantitatives rapidement et efficacement au fur et à mesure de leur collecte.

**droite de régression**

La droite de régression est une droite qui représente le mieux possible le nuage de points.

**écart moyen**

L'écart moyen, noté  $EM$  est une mesure de dispersion qui mesure la moyenne des écarts des données d'une distribution statistique par rapport à la moyenne des données:

$$EM = \frac{\text{somme des écarts à la moyenne}}{\text{nombre total de données}}$$

**étendue**

L'étendue d'une distribution statistique notée habituellement  $E$  est définie comme la différence entre la valeur maximale et la valeur minimale de cette distribution.

**extrapolation**

L'extrapolation consiste à déterminer la valeur d'une variable à partir d'une valeur connue de l'autre variable à l'extérieur d'un nuage de points.

**feuille**

Les feuilles d'un diagramme à tige et à feuilles contiennent le dernier chiffre de chacune des données.

**interpolation**

L'interpolation consiste à déterminer la valeur d'une variable à partir d'une valeur connue de l'autre variable à l'intérieur d'un nuage de points.

**médiane**

La médiane, symbolisée par les lettres  $Md$ , est la valeur située au centre de la distribution lorsque les données sont classées en ordre croissant.

**méthode de la droite médiane-médiane**

La méthode de la droite médiane-médiane permet d'obtenir une approximation de la droite de régression en utilisant les valeurs médianes de trois sous-groupes des données d'une distribution à deux caractères.

**méthode de l'ellipse**

La méthode de l'ellipse est une méthode graphique qui permet d'estimer le coefficient de corrélation entre deux variables.

Cette méthode consiste à tracer une ellipse autour d'un nuage de points et à appliquer la formule:  $r \approx \pm \left( 1 - \frac{\text{mesure du petit axe}}{\text{mesure du grand axe}} \right)$ .

**méthode de Mayer**

La méthode de Mayer permet d'obtenir une approximation de la droite de régression à partir de deux points dont les coordonnées sont les moyennes des valeurs en  $x$  et en  $y$  de deux sous-groupes des données d'une distribution à deux caractères.

**méthode des moyennes**

La méthode des moyennes permet de déterminer une approximation de l'équation de la droite de régression en utilisant un point dont les coordonnées sont la valeur moyenne des valeurs de  $x$  et la valeur moyenne des valeurs de  $y$  des points d'une distribution, et un autre point de la distribution.

**méthode du rectangle**

La méthode du rectangle est une méthode graphique qui permet d'estimer le coefficient de corrélation entre deux variables.

Cette méthode consiste à tracer un rectangle autour d'un nuage de points et à appliquer la formule:  $r \approx \pm \left( 1 - \frac{\text{mesure du petit côté}}{\text{mesure du grand côté}} \right)$ .

**mode**

Le mode, noté  $Mo$ , est la donnée la plus fréquente de la distribution.

**moyenne**

La moyenne d'une distribution, notée  $\bar{x}$ , est égale à la somme de toutes les données d'une distribution, divisée par le nombre total de données de cette distribution. Pour calculer la moyenne, on utilise la formule suivante:  $\bar{x} = \frac{\sum x_i}{n}$ , où  $x_i$  représente la  $i^{\text{e}}$  donnée et  $n$ , le nombre total de données.

**nuage de points**

Un nuage de points est l'ensemble des points représentant une distribution de données quantitatives dépendant de deux variables. Il permet de mettre en évidence le degré de corrélation entre deux variables liées.

**rang centile**

Le rang centile est une mesure de position qui permet de positionner une valeur par rapport aux autres valeurs d'une distribution statistique ordonnée. Le rang centile d'une donnée indique le pourcentage des données ayant une valeur inférieure ou égale à cette valeur.

**tableau à double entrée**

Un tableau à double entrée ou tableau de distribution à deux caractères, comporte des lignes et des colonnes. Les deux caractères étudiés apparaissent dans chacune des moitiés de la première case du tableau. Dans un tableau de distribution à deux caractères, la lecture se fait par croisement des lignes et des colonnes.

**tableau de distribution à deux caractères**

Un tableau de distribution à deux caractères appelé aussi tableau à double entrée comporte des lignes et des colonnes. Les deux caractères étudiés apparaissent dans chacune des moitiés de la première case du tableau. Dans un tableau de distribution à deux caractères, la lecture se fait par croisement des lignes et des colonnes.

**tige**

La tige d'un diagramme à tige et à feuilles comporte tous les chiffres de chacune des données, sauf le dernier.

**valeur absolue**

La valeur absolue d'un nombre est: le nombre lui-même s'il s'agit d'un nombre plus grand ou égal à 0; le nombre multiplié par -1 s'il s'agit d'un nombre plus petit que 0.