coefficient de corrélation linéaire

Le coefficient de corrélation linéaire, noté r, est une valeur numérique ($-1 \leq r \leq 1$) qui permet d'évaluer le lien existant entre deux variables statistiques quantitatives.

corrélation linéaire

La corrélation est le lien qui existe entre deux variables statistiques quantitatives. On parle de corrélation linéaire lorsqu'on peut représenter approximativement une distribution à l'aide d'une droite.

cote standard

La cote standard d'une donnée dans une distribution statistique à un caractère quantitatif, qu'on appelle également sa cote z, est un indice qui nous permet de déterminer à combien d'écarts-types au-dessous ou au-dessus de la moyenne se situe une donnée. On la calcule à l'aide de la formule $z = \frac{x - \bar{x}}{s}$, où x représente la donnée numérique, $\bar{x}$ la moyenne de toutes les données de la distribution, et s l'écart-type de l'ensemble de toutes les données de la distribution.

droite de régression

On appelle droite de régression la droite qui offre la meilleure approximation de l'ensemble des points d'une distribution statistique à deux caractères quantitatifs.

écart-type

L'écart-type d'une distribution, noté s, est une mesure de dispersion de cette distribution à un caractère quantitatif. On le calcule à l'aide de la formule $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$, où x_i représente la i^{ème} donnée, $\bar{x}$ la moyenne, et n le nombre de données.

méthode des moyennes

La méthode des moyennes permet de donner une approximation de l'équation de la droite de régression. Elle consiste à partager en deux groupes à peu près égaux en nombre l'ensemble des points d'un nuage à l'aide d'une droite qui passe par le point ($\bar{x}$, $\bar{y}$) et par le plus grand nombre possible de points de la distribution.

© 2006 KINÉSIS ÉDUCATION INC.

méthode du rectangle

La méthode du rectangle permet d'approximer le coefficient de corrélation existant entre deux variables statistiques. Elle consiste à partager en deux groupes à peu près égaux en nombre l'ensemble des points d'un nuage à l'aide d'une droite, puis à tracer autour de cette droite un rectangle parfaitement symétrique encadrant tous les points de la distribution. On calcule ensuite la valeur approximative du coefficient de corrélation en mesurant le rectangle ainsi formé et la formule :

$$r \approx \pm \left(1 - \frac{\text{mesure du petit côté}}{\text{mesure du grand côté}}\right).$$

nuage de points

Un nuage de points est la représentation graphique des couples d'une distribution à deux caractères quantitatifs dans un plan cartésien.

tableau de distribution à deux variables

Un tableau de distribution à deux variables est un tableau dans lequel apparaissent les valeurs des deux variables reliées à une même observation dans une distribution statistique.

variance

La variance, notée s^2, est une mesure de dispersion d'une distribution à un caractère quantitatif. On la calcule à l'aide de la formule $s^2 = \frac{\sum (x_i - \overline{x})^2}{n - 1}$, où x_i représente la i^{ème} donnée, $\overline{x}$ la moyenne, et n le nombre de données.

© 2006 KINÉSIS ÉDUCATION INC.