Statistiques descriptives univariées

Ce qu'il faut connaître :
-le vocabulaire : population, individu, taille, caractère, variable quantitative, variable qualitative, modalité
-les définitions de : effectif, fréquence et pourcentage
-les différentes représentations graphiques (diagramme en bâtons, histogramme)
-les caractéristiques de position : mode, moyenne, médiane
-les caractéristiques de dispersion : étendue, variance, écart-type, quartile, décile

1. Comment calculer la moyenne d'une série statistique

a. si on a les valeurs $a_1, a_2, \cdots, a_n$ de la série statistique, $\bar{x} = \frac{1}{n}\sum_{k=1}^{n} a_k$

b. si on a les modalités $x_1, x_2, \cdots, x_p$ et les effectifs $\bar{x} = \frac{n_1x_1 + \cdots + n_px_p}{n_1 + \cdots + n_p} = \frac{1}{n}\sum_{i=1}^{p} n_ix_i$

2. Comment calculer la variance ou l'écart-type d'une série statistique

L'écart-type est la racine carrée de la variance : $\sigma_x = \sqrt{\sigma_x^2}$

a. si on a les valeurs $a_1, a_2, \cdots, a_n$ de la série statistique $\sigma_x^2 = \frac{1}{n}\sum_{k=1}^{n}(a_k - \bar{x})^2 = \frac{1}{n}\sum_{k=1}^{n} {a_k}^2 - \bar{x}^2$

b. si on a les modalités $x_1, x_2, \cdots, x_p$ et les effectifs $\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{p} n_i(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{p} n_i{x_i}^2 - \bar{x}^2$

Statistiques descriptives bivariées

Ce qu'il faut connaître :
-la représentation par nuage de points
-la définition du point moyen
-la définition de la covariance
-la définition du coefficient de corrélation
-procéder à un ajutement affine

4. Comment déterminer le point moyen d'une série double $((x_i, y_i))_{i\in[|1,n]|}$

on calcule les moyennes $\bar{x}$ et $\bar{y}$ des séries $(x_i)_{i\in[|1,n]|}$ et $(y_i)_{i\in[|1,n]|}$ alors $G(\bar{x}, \bar{y})$

5. Comment calculer la covariance d'une série double $((x_i, y_i))_{i\in[|1,n]|}$

a. on calcule les moyennes $\bar{x}$ et $\bar{y}$ des séries $(x_i)_{i\in[|1,n]|}$ et $(y_i)_{i\in[|1,n]|}$ alors $s_{x,y} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

b. on calcule les moyennes $\bar{x}$, $\bar{y}$ et $\overline{xy}$ des séries $(x_i)_{i\in[|1,n]|}$, $(y_i)_{i\in[|1,n]|}$ et $(x_iy_i)_{i\in[|1,n]|}$, $s_{x,y} = \overline{xy} - \bar{x}.\bar{y}$

5. Comment calculer le coefficient de corrélation linéaire d'une série double $((x_i, y_i))_{i\in[|1,n]|}$

-on calcule la covariance $s_{x,y}$ de la série double $((x_i, y_i))_{i\in[|1,n]|}$
-on calcule les écart-types σ_x et σ_y des séries $(x_i)_{i\in[|1,n]|}$ et $(y_i)_{i\in[|1,n]|}$
-on calcule $r_{x,y} = \frac{s_{x,y}}{\sigma_x\sigma_y}$

6. Comment déterminer la droite de régression de y par rapport à x pour la série $((x_i, y_i))_{i\in[|1,n]|}$

-on calcule la covariance $s_{x,y}$ de la série double $((x_i, y_i))_{i\in[|1,n]|}$
-on calcule la variance σ_x^2 de la série $(x_i)_{i\in[|1,n]|}$
-on calcule les moyennes $\bar{x}$ et $\bar{y}$ des séries $(x_i)_{i\in[|1,n]|}$, $(y_i)_{i\in[|1,n]|}$
la droite de régression de y par rapport à x est la droite passant par$G(\bar{x}, \bar{y})$, de coefficient directeur $\frac{s_{x,y}}{\sigma_x^2}$
elle a donc pour équation $y = ax + b$ avec $a = \frac{s_{x,y}}{\sigma_x^2}$ et $b = \bar{y} - a\bar{x}$