

Deux écarts types

Pour une série statistique dont les n valeurs sont $x_1, x_2, x_3, \dots, x_n$ et de moyenne $\bar{x}$.

L'**écart type** est $\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}}$.

Cela représente l'**écart-type** de la **population** concernée, c'est pour cela qu'il se détermine sur tableur par la fonction **ECARTYPE**. C'est celui à étudier dans le cadre du référentiel de 2^{nde} professionnelle.

Un autre **écart type**, appelé **écart type estimé**, est $s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1}}$

Il est plutôt vu en BTS car il « **estime** » (avec des erreurs d'estimation possibles) ce que serait l'écart type si la série statistique est un échantillon d'une population plus grande.

Plus les valeurs de n sont grandes, plus les deux valeurs des écart types sont proches, car diviser par une **valeur très grande** ou une **valeur très grande - 1** donnent des résultats très proches.

Exemple : Soit une population de 150 personnes dont on a mesuré les tailles en cm

172	180	157	185	221	184	188	168	177
194	188	194	164	173	173	184	168	193
184	169	175	154	185	140	164	202	165
174	193	179	198	194	174	181	189	196
191	163	199	181	179	166	180	184	192
184	166	164	158	157	180	181	167	186
190	181	180	159	176	192	170	191	172
191	168	185	185	160	190	175	166	188
171	193	177	172	191	164	199	176	166
189	185	177	185	173	176	182	184	191
172	185	183	180	178	171	184	187	181
162	190	169	179	187	176	185	171	168
189	167	186	177	177	169	190	178	197
182	181	174	184	193	189	199	182	159
168	191	185	189	185	206	184	177	177

	30 valeurs	30 valeurs	30 valeurs	Total valeurs
MOYENNE	180,43	179,47	177,80	179,78
ECARTYPE	10,23	10,54	11,63	11,86
ECARTYPEP	10,06	10,37	11,43	11,82

On note que l'**écart type estimé** sur 30 valeurs prises au hasard (1^{ère} et 2^e, puis 2^e et 3^e colonne, puis 3^e et 4^e colonne) est plus proche de la valeur de l'écart type de la population totale, bien qu'éloignée à cause du nombre faible de valeurs.

Les moyennes restent quant à elles d'assez bonnes estimations. Certes 177,8 est assez éloigné de 179,78, mais cela représente qu'une erreur d'environ 1%.

À noter que pour la **moyenne**, il n'y a pas d'autre formule pour la **moyenne estimée**. La moyenne usuelle est considérée comme une bonne estimation de la moyenne si la série statistique est un échantillon d'une population plus grande.