

La variance et l'écart type

Julie Milot

Professeure de mathématique
Département de mathématiques
Collège de Maisonneuve
jmilot@cmaisonneuve.qc.ca



Ressource développée dans le cadre du projet Mathéma-TIC
Financé par le ministère de l'Enseignement supérieur, de la Recherche et de la Science (MESRS)
du Québec dans le cadre du Programme d'arrimage universités-collèges

Mise en situation



Examen 1 Notes x_i	Écarts à la moyenne $x_i - \mu$	Carrés des écarts à la moyenne $(x_i - \mu)^2$
62	-14	196
67	-9	81
68	-8	64
68	-8	64
69	-7	49
70	-6	36
72	-4	16
73	-3	9
73	-3	9
74	-2	4
75	-1	1
77	1	1
77	1	1
78	2	4
78	2	4
79	3	9
79	3	9
79	3	9
80	4	16
80	4	16
81	5	25
83	7	49
89	13	169
93	17	289
	0 (moyenne des écarts)	47,08 (moyenne des carrés des écarts) (variance)

$\sqrt{47,08} \approx 6,86$ (écart type)

Variance

Moyenne des carrés des écarts entre chacune des valeurs d'une variable et la valeur moyenne de cette variable.

Pour les valeurs provenant d'une population :

$$\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + (x_3 - \mu)^2 + \dots + (x_N - \mu)^2}{N} = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

où les x_i représentent les valeurs de la variable étudiée.

Pour les valeurs provenant d'un échantillon :

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

où les x_i représentent les valeurs de la variable étudiée.

Écart type

Racine carrée de la variance.

Pour les valeurs provenant d'une population :

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

où les x_i représentent les valeurs de la variable étudiée.

Pour les valeurs provenant d'un échantillon :

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

où les x_i représentent les valeurs de la variable étudiée.

Calcul d'un écart type à partir de données brutes

Données provenant d'un échantillon de 17 adultes canadiens, janvier 2015.

Identifiant du répondant	Sexe	Âge	Revenu annuel brut	Niveau de scolarité	Nombre d'enfants
1	H	39	63751	universitaire	4
2	F	24	28604	collégial	1
3	F	57	81379	universitaire	3
4	H	54	16641	secondaire	2
5	F	40	72865	collégial	3
6	H	68	69532	collégial	3
7	H	59	17735	secondaire	1
8	H	61	34369	universitaire	2
9	F	21	17206	secondaire	0
10	F	26	15195	collégial	1
11	H	53	54268	universitaire	4
12	F	65	63743	secondaire	3
13	F	70	84963	collégial	0
14	F	45	54040	universitaire	2
15	H	34	56388	collégial	1
16	F	27	74159	collégial	0
17	H	70	24819	collégial	2

$$\bar{x} \approx 1,88 \text{ enfant}$$

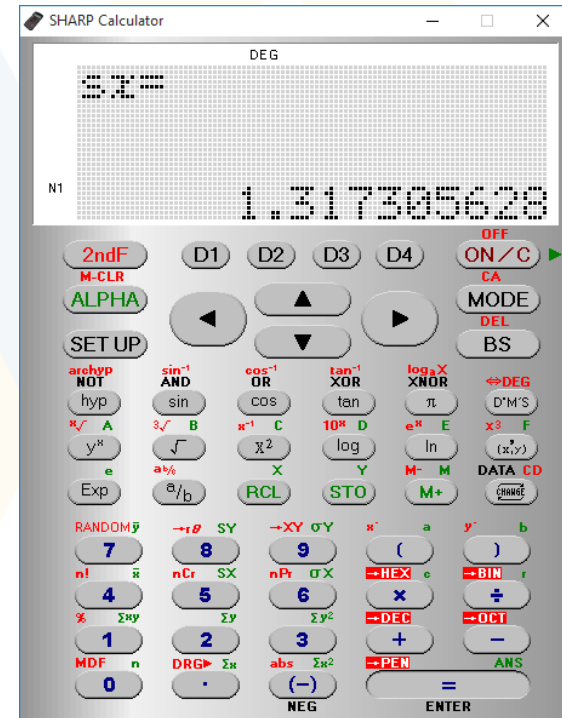
$$s \approx \sqrt{\frac{(4-1,88)^2 + (1-1,88)^2 + \dots + (2-1,88)^2}{17-1}}$$

$$s \approx 1,3173 \text{ enfant}$$

Calcul d'un écart type à partir de données brutes en mode statistique

Données provenant d'un échantillon de 17 adultes canadiens, janvier 2015.

Identifiant du répondant	Sexe	Âge	Revenu annuel brut	Niveau de scolarité	Nombre d'enfants
1	H	39	63751	universitaire	4
2	F	24	28604	collégial	1
3	F	57	81379	universitaire	3
4	H	54	16641	secondaire	2
5	F	40	72865	collégial	3
6	H	68	69532	collégial	3
7	H	59	17735	secondaire	1
8	H	61	34369	universitaire	2
9	F	21	17206	secondaire	0
10	F	26	15195	collégial	1
11	H	53	54268	universitaire	4
12	F	65	63743	secondaire	3
13	F	70	84963	collégial	0
14	F	45	54040	universitaire	2
15	H	34	56388	collégial	1
16	F	27	74159	collégial	0
17	H	70	24819	collégial	2



Calcul d'un écart type à partir d'un tableau de distribution de fréquences absolues

Répartition des 17 adultes canadiens interrogés selon le nombre d'enfants qu'ils ont, janvier 2015.

Nombre d'enfants	Nombre d'adultes canadiens
0	3
1	4
2	4
3	4
4	2
Total	17

Source : données fictives.

$$s \approx \sqrt{\frac{\overbrace{(0-1,88)^2 + (0-1,88)^2 + (0-1,88)^2 + \dots + (4-1,88)^2}^{3 \text{ fois}}}{17-1}}$$

$$s \approx \sqrt{\frac{(0-1,88)^2 \times 3 + (1-1,88)^2 \times 4 + \dots + (4-1,88)^2 \times 2}{17-1}}$$

$$s \approx 1,3173 \text{ enfant}$$

Pour les valeurs provenant d'un échantillon ou d'une population:

$$s = \sqrt{\frac{\sum_{i=1}^k (x_i - \bar{x})^2 n_i}{n-1}} \quad \text{ou} \quad \sigma = \sqrt{\frac{\sum_{i=1}^k (x_i - \mu)^2 n_i}{N}}$$

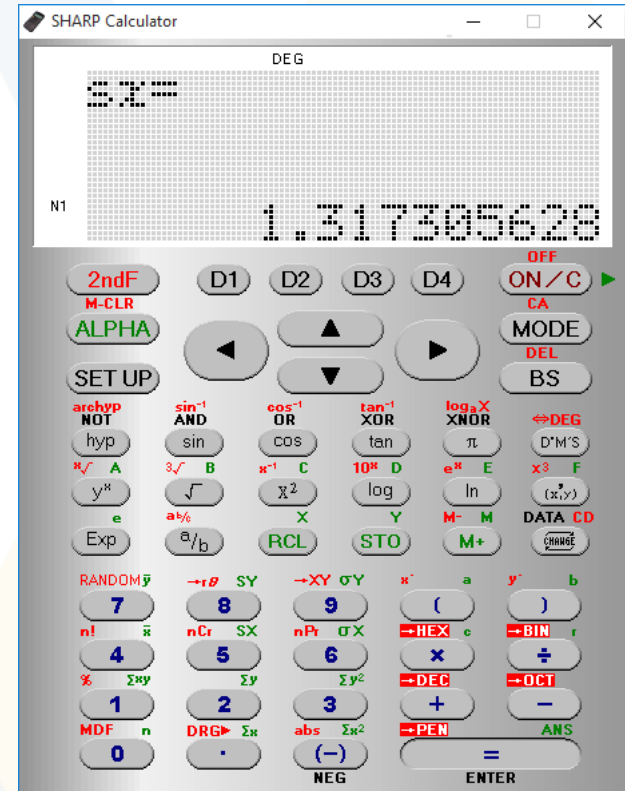
où les x_i représentent les valeurs de la variable étudiée, les n_i représentent les fréquences absolues et k représente le nombre de valeurs différentes.

Calcul d'un écart type à partir d'un tableau de distribution de fréquences absolues en mode statistique

Répartition des 17 adultes canadiens interrogés selon le nombre d'enfants qu'ils ont, janvier 2015.

Nombre d'enfants	Nombre d'adultes canadiens
0	3
1	4
2	4
3	4
4	2
Total	17

Source : données fictives.



Estimation d'un écart type à partir de valeurs classées et de fréquences absolues

Distribution des 17 adultes canadiens interrogés selon leur revenu annuel, janvier 2015.

Revenu annuel (\$)	Nombre d'adultes canadiens
[15 000; 30 000[6
[30 000; 45 000[1
[45 000; 60 000[3
[60 000; 75 000[5
[75 000; 90 000[2
Total	17

Source : données fictives.

$$s \approx \sqrt{\frac{(22500 - 48970,59)^2 \times 6 + \dots + (82500 - 48970,59)^2 \times 2}{17 - 1}}$$

$$s \approx 22828,48 \$$$

Pour les valeurs provenant d'un échantillon ou d'une population:

$$s = \sqrt{\frac{\sum_{i=1}^k (m_i - \bar{x})^2 n_i}{n - 1}} \quad \text{ou} \quad \sigma = \sqrt{\frac{\sum_{i=1}^k (m_i - \mu)^2 n_i}{N}}$$

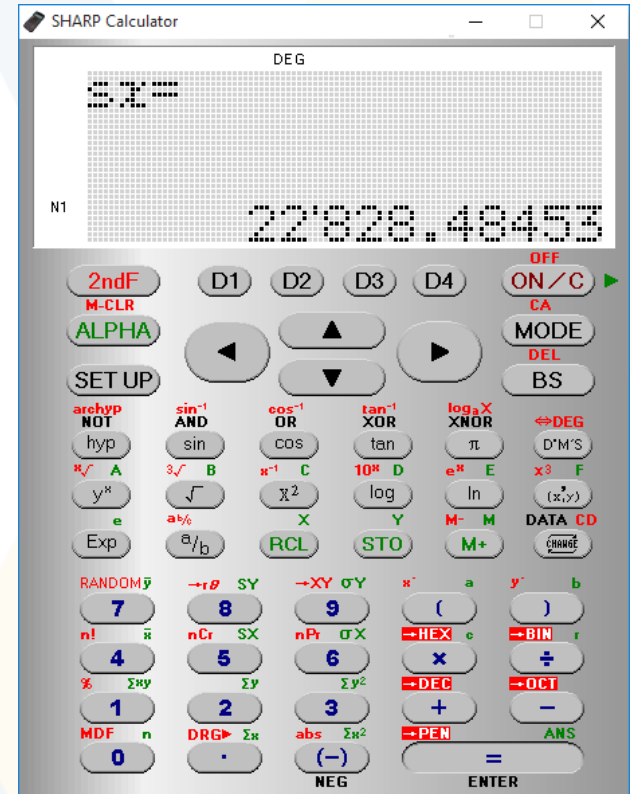
où les m_i représentent les milieux de classes, les n_i représentent les fréquences absolues et k représente le nombre de valeurs différentes.

Estimation d'un écart type à partir de valeurs classées, de fréquences absolues et du mode statistique

Distribution des 17 adultes canadiens interrogés selon leur revenu annuel, janvier 2015.

Revenu annuel (\$)	Nombre d'adultes canadiens
[15 000; 30 000[6
[30 000; 45 000[1
[45 000; 60 000[3
[60 000; 75 000[5
[75 000; 90 000[2
Total	17

Source : données fictives.



Avantages et inconvénients de l'utilisation de l'écart type comme mesure de dispersion

Avantages

- Il se prête facilement aux opérations algébriques.
- Sa valeur est relativement stable d'un échantillon à l'autre.

Inconvénients

- Son calcul est lourd.
- Il est sensible aux valeurs extrêmes.
- Son interprétation n'est pas immédiate.
- Il doit être utilisé avec d'autres mesures pour prendre un sens.

Résumé

L'écart type :

- a pour symbole s (échantillon) ou σ (population);
- se calcule en faisant la racine de la moyenne des carrés des écarts entre chaque valeur d'une variable et la moyenne de la variable ou en utilisant tout simplement le mode statistique d'une calculatrice ou la fonction écart type d'un logiciel;
- s'estime en substituant les valeurs par les milieux de classes si seules les valeurs classées sont disponibles.

Conception du contenu

Julie Milot

Collège de Maisonneuve
jmilot@cmaisonneuve.qc.ca

Révision du contenu

Samuel Bernard et Hélène Lambert

samuel.bernard@collanaud.qc.ca
hlambert@cmaisonneuve.qc.ca

Direction de projet

**Samuel Bernard
Bruno Poellhuber**

Postproduction

Marie-Ève Lanthier

Musique

Sébastien Belleudy

sebe.bandcamp.com

Crédit images

Wikipedia

en.wikipedia.org

Calculatrice virtuelle

PC-Simulator

sharp-world.com

Conception graphique

Christine Blais

Production des modèles en LaTeX

Nicolas Beauchemin

nicolas.beauchemin@bdeb.qc.ca

Production

Samuel Bernard



Bruno Poellhuber



Vidéo mise à disposition selon les termes de la licence

Creative Commons internationale 4.0

Paternité / Pas d'utilisation commerciale / Partage dans les mêmes conditions

Les autorisations au-delà du champ de cette licence peuvent être obtenues à

Mathema-TIC.ca