STREUMAßE

Zu 5 unterschiedlichen Tageszeiten wird die Höhe des Eiffelturms gemessen. (Alle Einheiten in m.)

x_i	x_1	x_2	x_3	x_4	x_5	Σ	$\frac{\Sigma}{n}$
Höhe	329,9	330,1	330,3	329,7	330	1650	330 ← $\bar{x}$
$x_i - \bar{x}$	-0,1	0,1	0,3	-0,3	0	0	0
$\lvert x_i - \bar{x} \rvert$	0,1	0,1	0,3	0,3	0	0,8	0,16 ← $\bar{d}$
$(x_i - \bar{x})^2$	0,01	0,01	0,09	0,09	0	0,2	0,04 ← $\sigma^2 = Var(X)$

arithmetisches Mittel

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \frac{1}{n} \cdot (x_1 + x_2 \ldots x_n)$$

i: Laufvariable (kann die Werte 1 - n annehmen, hier 1 - 5)

$$\bar{x} = \frac{1}{5} \cdot (329{,}9 + 330{,}1 \ldots + 330)$$

n: Anzahl Messwerte

$$= \frac{1}{5} \cdot 1650 = \underline{\underline{330\,[m]}}$$

x_i: Messwert

mittlere absolute Abweichung

(durchschnittliche Abweichung vom Mittelwert)

$$\bar{d} = \frac{1}{n} \cdot \sum_{i=1}^{n} \lvert x_i - \bar{x} \rvert$$

$$\bar{d} = \frac{1}{5} \cdot \left(\lvert 329{,}9 - 330 \rvert + \lvert 330{,}1 - 330 \rvert + \lvert 330{,}3 - 330 \rvert + \lvert 329{,}7 - 330 \rvert + \lvert 330 - 330 \rvert \right)$$

$$= \frac{1}{5} \cdot \left(\lvert -0{,}1 \rvert + \lvert 0{,}1 \rvert + \lvert 0{,}3 \rvert + \lvert -0{,}3 \rvert + \lvert 0 \rvert \right)$$

$$= \frac{1}{5} \cdot (0{,}1 + 0{,}1 + 0{,}3 + 0{,}3 + 0)$$

$$= \frac{1}{5} \cdot 0{,}8 = \underline{\underline{0{,}16\,[m]}}$$

Varianz

$$Var(X) = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2$$

(logischere Formel)

$$Var(X) = \frac{1}{5} \cdot \left((329{,}9 - 330)^2 + (330{,}1 - 330)^2 + (330{,}3 - 330)^2 + (329{,}7 - 330)^2 + (330 - 330)^2 \right)$$

$$= \frac{1}{5} \cdot \left((-0{,}1)^2 + 0{,}1^2 + 0{,}3^2 + (-0{,}3)^2 + 0^2 \right)$$

$$= \frac{1}{5} \cdot (0{,}01 + 0{,}01 + 0{,}09 + 0{,}09 + 0)$$

$$= \frac{1}{5} \cdot 0{,}2 = \underline{\underline{0{,}04\,[m^2]}}$$

Varianz

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

(leichtere Formel zum Rechnen)

$$Var(X) = \frac{1}{5} \left(329{,}9^2 + 330{,}1^2 + \ldots 330^2 \right) - 330^2$$

$$= \frac{1}{5} \cdot 544.500{,}2 - 330^2 = 0{,}04\,[m^2]$$

Standardabweichung

$$\sigma = \sqrt{Var(X)}$$

Maß für die Abweichung vom Mittelwert

$$\sigma = \sqrt{0{,}04} = 0{,}2\,[m]$$

Variationskoeffizient

$$V = \frac{\sigma}{\bar{x}}$$

$$V = \frac{0{,}2}{330} = 0{,}0006 = 0{,}06\,\%$$

⇒ setzt die Standardabweichung ins Verhältnis zum Mittelwert

z.B. wäre $\frac{0{,}2}{0{,}3} = 66{,}7\,\%$

eine verhältnismäßig hohe Abweichung vom Mittelwert

Beispiel zum Variationskoeffizient:

A: Unternehmen mit 5 Mitarbeitern | B: Freiberufler

$\bar{x}_A = 100.000\,€$ ← durchschnittl. Monatsumsatz → $\bar{x}_B = 10.000\,€$

$\sigma_A = 5.000\,€$ ← Abweichung vom Monatsumsatz → $\sigma_B = 5.000\,€$

$V_A = \frac{5.000}{100.000} = 0{,}05$ verhältnismäßig hohe Abweichung vom Monatsumsatz $V_B = \frac{5.000}{10.000} = 0{,}5$

⇒ A und B haben zwar die gleiche Standardabweichung $\sigma = 5000\,€$, aber V_A und V_B zeigen das Risiko des Freiberuflers.

Hinweis:

Bei einer Population (alle Werte sind bekannt) gelten die angegebenen Formeln.
Bei einer Stichprobe, mit deren Hilfe wir auf die Population schließen wollen, sollte man für eine gute Schätzung der Varianz mit n-1 statt mit n rechnen.
Dass das so ist, könnten wir mit Computersimulationen nachweisen.
⇓

Population

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

Stichprobe

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

↑

weitere Schreibweisen: s^2, s_x^2, s_{n-1}^2

Mobil & whatsapp 0173-3047 206 • Internet franko-mathe.de/SRH • E-Mail info@franko-mathe.de • Dozent Rawe Franko • Bereich SRH Fachschulen Heidelberg

Mobil & whatsapp 0173-3047 206 • Internet franko-mathe.de/SRH • E-Mail info@franko-mathe.de • Dozent Rawe Franko • Bereich SRH Fachschulen Heidelberg

STREUMAßE LÖSUNG

① $\bar{x} = \frac{1}{5}\cdot(90+100+\ldots 140) = \frac{575}{5} = 115\,[g]$

$$\sigma^2 = \frac{1}{n}\cdot\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{5}\cdot\left((90-115)^2+(100-115)^2\ldots(140-115)^2\right) = \frac{1700}{5} = 340\,[g^2]$$

$$\sigma = \sqrt{340} = 18{,}44\,[g]$$

$$v = \frac{\sigma}{\bar{x}} = \frac{18{,}44}{115} = 0{,}160\ (=16{,}0\,\%)$$

$$\bar{d} = \frac{1}{n}\cdot\sum_{i=1}^{n}|x_i-\bar{x}| = \frac{1}{5}\cdot\left(|90-115|+|100-115|\ldots|140-115|\right) = \frac{80}{5} = 16\,[g]$$

Vergleiche Varianz bei Stichproben und bei Wahrscheinlichkeiten

$$\sigma^2 = \frac{1}{n}\cdot\sum_{i=1}^{n}(x_i-\bar{x})^2 \qquad Var(X) = \sum\left(x_i-E(X)\right)^2\cdot P(X=x_i)$$

mit Werten aus Aufgabe 3

	1	2	3	4	5	
absolute Häufigkeit	10	20	35	30	5	10+20+…5=100
relative Häufigkeit	$\frac{10}{100}$	$\frac{20}{100}$	$\frac{35}{100}$	$\frac{30}{100}$	$\frac{5}{100}$	
in Prozent	10%	20%	35%	30%	5%	

$$= \frac{1}{100}\cdot\Big[\underbrace{(1-\bar{x})^2+(1-\bar{x})^2\ldots(1-\bar{x})^2}_{10\text{-mal}}+\underbrace{(2-\bar{x})^2\ldots}_{20\text{-mal}}\ldots\Big]$$

$$= \frac{1}{100}\cdot\left[(1-\bar{x})^2\cdot 10+(2-\bar{x})^2\cdot 20\ldots(5-\bar{x})^2\cdot 5\right]$$

$$= \left[(1-\bar{x})^2\cdot\frac{10}{100}+(2-\bar{x})^2\cdot\frac{20}{100}\ldots(5-\bar{x})^2\cdot\frac{5}{100}\right]$$

$$Var(X) = \sum_{i=1}^{K}\left(x_i-E(X)^2\right)\cdot P(X=x_i)$$

②

Tägl. Arbeitszeit [h]	7	7,5	8	8,5	9	
absolute Häufigkeit	5	9	3	2	1	Σ=20
relative Häufigkeit	$\frac{5}{20}$	$\frac{9}{20}$	$\frac{3}{20}$	$\frac{2}{20}$	$\frac{1}{20}$	
Wahrscheinlichkeit in %	25%	45%	15%	10%	5%	

absolute Häufigkeit: gibt an, wie oft etwas vorkommt (ganze Zahl), z.B. „5" - ob dies viel oder wenig ist lässt sich nicht sagen, dazu braucht man einen Vergleich

relative Häufigkeit: setzt die absolute Häufigkeit ins Verhältnis zur Gesamtzahl, z.B. „$\frac{5}{20}$"

Zufallsvariable X: nimmt in diesem Beispiel den Wert der täglichen Arbeitszeit an

x_i: Messwert/Datenwert $P(X=x_i)$: Wahrscheinlichkeit (probability) für das jeweilige Ereignis

K: Anzahl der verschiedenen Datenwerte (hier $K=5$ im Gegensatz zu $n=20$)

$$E(X) = \sum_{i=1}^{K} x_i\cdot P(X=x_i) = 7\cdot 0{,}25+7{,}5\cdot 0{,}45\ldots+9\cdot 0{,}05 = 7{,}625\,[h]$$

$$MAD = \sum_{i=1}^{K}|x_i-E(X)|\cdot P(X=x_i) = |7-7{,}625|\cdot 0{,}25+|7{,}5-7{,}625|\cdot 0{,}45\ldots = 0{,}425\,[h]$$

$$Var(X) = \sum_{i=1}^{K}(x_i-E(X))^2\cdot P(X=x_i) = \sum_{i=1}^{K}x_i^2\cdot P(X=x_i) - E(X)^2 = 0{,}2969\,h^2$$

$$\sigma(X) = 0{,}545\,h \qquad v = \frac{0{,}545}{7{,}625} = 0{,}071$$

③ $$E(X) = \sum_{i=1}^{K} x_i\cdot P(X=x_i)$$
$$= x_1\cdot P(X=x_1)+x_2\cdot P(X=x_2)+\ldots x_5\cdot P(X=x_5)$$
$$= 1\cdot P(X=1)+2\cdot P(X=2)+\ldots 5\cdot P(X=5)$$
$$= 1\cdot 10\% + 2\cdot 20\% + \ldots 5\cdot 5\%$$
$$= 1\cdot 0{,}1 + 2\cdot 0{,}2 + \ldots 5\cdot 0{,}05 = 3$$

$$Var(X) = \sum_{i=1}^{K}(x_i-E(X))^2\cdot P(X=x_i)$$
$$= (1-3)^2\cdot 0{,}1+(2-3)^2\cdot 0{,}2+\ldots(5-3)^2\cdot 0{,}05 = 1{,}1$$

$$\sigma = \sqrt{Var(X)} = \sqrt{1{,}1} = 1{,}049$$

$$v = \frac{\sigma}{E(X)} = \frac{1{,}049}{3} = 0{,}350$$

„Mean Absolute Deviation"

$$MAD = \sum_{i=1}^{K}|x_i-E(X)|\cdot P(X=x_i) = |1-3|\cdot 0{,}1+|2-3|\cdot 0{,}2\ldots|5-3|\cdot 0{,}05 = 0{,}8$$

④ Maschine 1 (x) 83,2 83,7 84,1 84,8 85,2 86,1 87,1

Maschine 2 (y) 81,4 82,3 83,7 84,1 86,5 87,4 89,8

$\bar{x} = 84{,}89$ $\bar{y} = 85{,}029$

$s_x = 1{,}375$ $(n-1)$* $s_y = 2{,}994$ $(n-1)$

$\sigma_x = 1{,}273$ (n) $\sigma_y = 2{,}772$ (n)

* n-1 : hier wurde die Standardabweichung mit n-1 gerechnet (für Stichproben)

n : hier wurde die Standardabweichung mit n gerechnet (für Population) ✓

wir haben alle Werte gegeben (Grundgesamtheit), nicht nur ein paar Werte wie bei der Stichprobe

Werte eingeben unter $\boxed{\text{data}}$

Auswertung über 2nd $\boxed{\text{data}}$ (stat)

MDA-2023-A-2

$\bar{x} = \frac{1}{5} \cdot (3{,}145 + 3{,}332 \ldots + 4{,}035) = 3{,}629$ [kg]

$\bar{d} = \frac{1}{5} \left(|3{,}145 - 3{,}629| + |3{,}332 - 3{,}629| \ldots + |4{,}035 - 3{,}629| \right) = 0{,}3124$ [kg]

TR: Mittelwert unter „x" abspeichern: 3,629 $\boxed{\text{STO}\rightarrow}$ (store) $\boxed{x^{\,y z}}$ $\boxed{\text{enter}}$

dann Rechnung eingeben: $\frac{1}{5}\left(|3{,}145 - x| + \ldots \right.$

$Var = \frac{1}{5} \cdot \left((3{,}145 - 3{,}629)^2 + (3{,}332 - 3{,}629)^2 \ldots + (4{,}035 - 3{,}629)^2 \right) = 0{,}1137$ [kg²]

oder

$Var = \frac{1}{5} \cdot (3{,}145^2 + 3{,}332^2 \ldots + 4{,}035^2) - 3{,}629^2 = 0{,}1137$ [kg²]

$\sigma = \sqrt{0{,}1137} = 0{,}3373$ [kg]

$v = \frac{0{,}3373}{3{,}629} = 0{,}093$ (= 9,3%)

MDA-2023-B-2

$\bar{x} = \frac{1}{5} \cdot (360 + 520 + \ldots 2230) = 1110$ [€]

$\bar{d} = \frac{1}{5} \left(|360 - 1110| + |520 - 1110| \ldots + |2230 - 1110| \right) = 588$ [€]

$Var = \frac{1}{5} (360^2 + 520^2 + \ldots 2230^2) - 1110^2 = 460.880$ [€²]

$\sigma = \sqrt{Var} = 678{,}88$ [€]

$v = \frac{678{,}88}{1110} = 0{,}612$ (= 61,2%)

Mobil & whatsapp 0173-3047 206 • Internet franko-mathe.de/SRH • E-Mail info@franko-mathe.de • Dozent Rawe Franko • Bereich SRH Fachschulen Heidelberg

STREUMAßE – FORMELN

bei Population

arithmetisches Mittel

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

mittlere absolute Abweichung

$$\bar{d} = \frac{1}{n}\cdot\sum_{i=1}^{n} |x_i - \bar{x}|$$

(durchschnittliche Abweichung vom Mittelwert)

Varianz

oder „Var"

$$Var(X) = \frac{1}{n}\cdot\sum_{i=1}^{n} (x_i - \bar{x})^2$$

(logischere Formel)

oder

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

(schnellere Formel)

neu

um einen realistischen Wert für die Population zu schätzen, wird hier bei Stichproben mit $n-1$ gerechnet

Standardabweichung

$$\sigma = \sqrt{Var(X)}$$

Variationskoeffizient

$$V = \frac{\sigma}{\bar{x}}$$

oder CV

bei Wahrscheinlichkeiten

Erwartungswert

$$E(X) = \sum_{i=1}^{k} x_i \cdot P(X=x_i)$$

oder μ

Mittlere absolute Abweichung („Mean Absolute Deviation")

$$MAD = \sum_{i=1}^{k} |x_i - E(X)| \cdot P(X=x_i)$$

Varianz

$$Var(X) = \sum_{i=1}^{k} (x_i - E(X))^2 \cdot P(X=x_i)$$

Standardabweichung

$$\sigma(X) = \sqrt{Var(X)}$$

Variationskoeffizient

$$V = \frac{\sigma}{E(X)}$$

(oder international CV für Coefficient of Variation)

Mobil & whatsapp 0173-3047 206 • Internet franko-mathe.de/SRH • E-Mail info@franko-mathe.de • Dozent Rawe Franko • Bereich SRH Fachschulen Heidelberg