

Erwartungstreue

Die Varianz der m Werte x_1, x_2, \dots, x_m wird mit

$$\sigma^2 = \frac{1}{m} [(x_1 - \bar{x})^2 + \dots + (x_m - \bar{x})^2]$$

bestimmt.

Um diese Varianz mit einer (kleinen) Stichprobe vom Umfang n zu schätzen, wird die sogenannte empirische Varianz (Stichprobenvarianz)

$$s^2 = \frac{1}{n-1} [(x_1 - \bar{x}^*)^2 + \dots + (x_n - \bar{x}^*)^2]$$

verwendet. \bar{x}^* ist der Mittelwert der Stichprobe.

Es wird also $n - 1$ statt n verwendet.

Dies soll im Folgenden erhellt werden.

Hierzu holen wir etwas weiter aus und betrachten das Schätzen des Mittelwerts von $\{1, 2, \dots, 6\}$ mit einer Stichprobe vom Umfang 2.

Zerlegen wir zunächst die Zahlenmenge in Teilmengen, z. B.

$$\{1, 5\}, \{2, 6\}, \{3, 4\}.$$

Wenn wir nun jeweils die Mittelwerte jeder Teilmenge

$$\bar{x}_1 = \frac{a_1+a_2}{2}, \quad \bar{x}_2 = \frac{a_3+a_4}{2}, \quad \bar{x}_3 = \frac{a_5+a_6}{2},$$

bestimmen und den Mittelwert von $\bar{x}_1, \bar{x}_2, \bar{x}_3$ bilden, erhalten wir:

$$\frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3} = \dots = \frac{a_1 + a_2 + \dots + a_6}{6} = \bar{x} \quad (= 3,5)$$

Der Mittelwert einer Teilmenge ist daher geeignet, den Mittelwert aller Werte zu schätzen.

Beim Ziehen ohne Zurücklegen gibt es $\binom{6}{2}$ Zahlenpaare und daher doppelt so viele Zahlen, sie sind alle mit der gleichen Häufigkeit $k = 2 \cdot \binom{6}{2} / 6$ vorhanden. Dies führt ebenfalls zu:

$$\frac{\frac{1}{2}(ka_1 + ka_2 + \dots + ka_6)}{\binom{6}{2}} = \bar{x}$$

Ähnlich kann beim Ziehen mit Zurücklegen (36 Zahlenpaare) argumentiert werden.

Die Mittelwertbildungen sind, wie man sagt, erwartungstreu.

Berechnen wir nun die Varianzen $\sigma_i^2 = \frac{1}{2} [(a_i - \bar{x}^*)^2 + (a_{i+1} - \bar{x}^*)^2]$ der obigen Teilmengen:

$$\sigma_1^2 = 4, \quad \sigma_2^2 = 4, \quad \sigma_3^2 = 0,25, \quad \text{gemittelt: } \bar{\sigma}^2 = 2,75$$

$$\bar{\sigma}^2 \text{ ist kleiner als } \sigma^2 = 2,92.$$

Das ist nicht verwunderlich, da \bar{x}^* jeweils σ_i^2 minimiert, (siehe: 14.) *Verschiedenes, Mittelwert, ...*) wohingegen \bar{x} in der Berechnung von σ^2 konstant bleibt.

Es kann gezeigt werden, dass für Stichproben mit Zurücklegen vom Umfang n gilt: $\frac{n}{n-1} \bar{\sigma}^2 = \sigma^2$ (siehe: 9.) *Stochastik, Schätzgrößen, ...*).

Berechne für $\{1, 2, 3, 4\}$ und 2-Stichproben mit Zurücklegen σ^2 und $\bar{\sigma}^2$.

Begründe nun die Schätzgröße s^2 .

Erwartungstreue Fortsetzung

Berechne für $\{1, 2, 3, 4\}$ und 2-Stichproben mit Zurücklegen σ^2 und $\bar{\sigma}^2$.

Lösung:

$$\sigma^2 = 1,25$$

Stichprobe	Mittelwert	Varianz
(1 1)	1	0
(1 2)	1,5	0,25
(1 3)	2	1
(1 4)	2,5	2,25
(2 1)	1,5	0,25
(2 2)	2	0
(2 3)	2,5	0,25
(2 4)	3	1
(3 1)	2	1
(3 2)	2,5	0,25
(3 3)	3	0
(3 4)	3,5	0,25
(4 1)	2,5	2,25
(4 2)	3	1
(4 3)	3,5	0,25
(4 4)	4	0
	2,5	$\bar{\sigma}^2 = 0,625$

σ^2 ist tatsächlich doppelt so groß wie $\bar{\sigma}^2$.

Die Division durch $n - 1$ statt n (hier 2) ergibt eine bessere, eine erwartungstreue Schätzung.