Empirische Varianz

Def. 3.1 (Empirische Varianz)

Gegeben sei eine Stichprobe mit n Daten und empirischem Mittelwert $\bar{x}$. Dann ist ihre ***empirische Varianz*** *definiert als*

$$v = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{4}$$

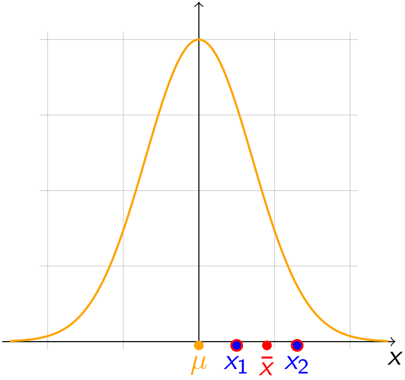
Warum steht hier der Faktor $\frac{1}{n-1}$?

Addon „Erwartungstreue“[3]

Warum steht bei der (Ko-)Varianz der Faktor $\frac{1}{n-1}$ und nicht $\frac{1}{n}$?

Anschauliche Erklärung:

- Schwankende x_i-Werte in der Stichprobe ziehen den Mittelwert $\bar{x}$ ein Stück weit hinter sich her.
- Z. B. $n = 2$: Es kann passieren, dass beide x_i oberhalb des wahren Mittelwertes μ liegen.
- Dann ist $(x_i - \bar{x})^2$ im Mittel kleiner als $(x_i - \mu)^2$ (s. Bild).
- Faktor $\frac{n}{n-1}$ gleicht das aus.



[Maas, 2014, S. 67]

[3] nach: Maas, Christoph, Stochastik für Dummies, Wiley-VCH, 2014.

Addon „Erwartungstreue“ (2)

Warum steht bei der (Ko-)Varianz der Faktor $\frac{1}{n-1}$ und nicht $\frac{1}{n}$?

Rechnerische und exakte Erklärung:

- Zufallsvariable X mit wahrem Mittelwert μ und Varianz σ^2.
- Stichproben der Größe n mit X_i = Zufallsvariable der i-ten Ziehung, $i = 1, ..., n$.
- $\bar{X}$ = Zufallsvariable „Mittelwert der Stichproben“
- Es gilt $E(\bar{X}) = \mu$ und $E((\bar{X} - \mu)^2) = \frac{\sigma^2}{n}$

- (empirische) Varianz *einer* Stichprobe: $v = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$
- V = Zufallsvariable „empirische Varianz der Stichproben“
- wahre Varianz $E((X_i - \mu)^2) = \sigma^2$

[Maas, 2014, S. 183ff]

Addon „Erwartungstreue“ (3)

Warum steht bei der (Ko-)Varianz der Faktor $\frac{1}{n-1}$ und nicht $\frac{1}{n}$?

Rechnerische Erklärung (Forts.): Wir zeigen, dass *nur mit* $\frac{1}{n-1}$ gilt:
Der Erwartungswert der empirischen Varianz ist die wahre Varianz.

$$\begin{aligned}
E(V) &= E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\right) \\
&= \frac{1}{n-1}E\left(\sum_{i=1}^{n}(X_i-\mu+\mu-\bar{X})^2\right) \\
&= \frac{1}{n-1}E\left(\sum_{i=1}^{n}((X_i-\mu)-(\bar{X}-\mu))^2\right) \\
&= \frac{1}{n-1}E\left(\sum_{i=1}^{n}((X_i-\mu)^2-2(X_i-\mu)(\bar{X}-\mu)+(\bar{X}-\mu)^2)\right)
\end{aligned}$$

Addon „Erwartungstreue“ (4)

$$
\begin{aligned}
E(V) &= \frac{1}{n-1} E\left(\sum_{i=1}^{n}(X_i-\mu)^2 - 2\left(\sum_{i=1}^{n}(X_i-\mu)\right)(\bar{X}-\mu) + n\cdot(\bar{X}-\mu)^2\right) \\
&= \frac{1}{n-1} E\left(\sum_{i=1}^{n}(X_i-\mu)^2 - 2\left((\sum_{i=1}^{n}X_i)-n\mu\right)(\bar{X}-\mu) + n\cdot(\bar{X}-\mu)^2\right) \\
&= \frac{1}{n-1} E\left(\sum_{i=1}^{n}(X_i-\mu)^2 - 2n(\bar{X}-\mu)(\bar{X}-\mu) + n\cdot(\bar{X}-\mu)^2\right) \\
&= \frac{1}{n-1} E\left(\sum_{i=1}^{n}(X_i-\mu)^2 - n(\bar{X}-\mu)^2\right) \\
&= \frac{1}{n-1} \left(\sum_{i=1}^{n}E((X_i-\mu)^2) - nE(\bar{X}-\mu)^2\right) \\
&= \frac{1}{n-1} \left(n\sigma^2 - n\frac{\sigma^2}{n}\right) \\
&= \frac{1}{n-1} \left(n\sigma^2 - \sigma^2\right) \\
&= \sigma^2
\end{aligned}
$$

Zurück