28
Varianz. Standardabweichung

Jörn Loviscach

Versionsstand: 21. September 2013, 15:55

Die nummerierten Felder sind absichtlich leer, zum Ausfüllen beim Ansehen der Videos: http://www.j3L7h.de/videos.html

Bitte hier notieren, was beim Bearbeiten unklar geblieben ist

1 Varianz und Standardabweichung

Der Erwartungswert $E[X]$ einer Zufallsvariablen X sagt nur etwas über den Schwerpunkt der Verteilung aus, aber nichts über ihre Breite:

1

Schön wäre, auch eine Größe zu haben, welche die Breite der Verteilung beschreibt, also die mittlere Abweichung vom Erwartungswert. Eine erste Idee wäre $E[X-E[X]]$, aber das klappt leider nicht:

2

Man könnte $E[|X-E[X]|]$ nehmen, aber das ist wegen der Betragsfunktion umständlich handzuhaben. Üblich ist deshalb die Varianz [variance] σ^2:

3

Randbemerkung: Es gibt pathologische Fälle, in denen dieser Erwartungswert unendlich wird und damit die Varianz nicht existiert.

Die Varianz ist die mittlere *quadratische* Abweichung vom Erwartungswert. Wenn X in Metern gemessen wird, hat die Varianz σ^2 also die Einheit Quadratmeter. Das ist ungeschickt, wenn man sich die Breite der Verteilung vorstellen will. Also nimmt man dazu die Wurzel der Varianz, die Standardabweichung [standard deviation] σ:

4

Daher kommt die Bezeichnung σ^2 für die Varianz.

Zur Unterscheidung spricht man hier gerne von der „Varianz der Grundgesamtheit“. Später kommt noch eine „Varianz der Stichprobe“.

2 Berechnung der Varianz

Durch Ausmultiplizieren lässt sich die Varianz noch etwas geschickter berechnen:

5

Beispiel 1: Varianz eines idealen Würfels:

6

Beispiel 2: Varianz einer gleichförmig zwischen 1 und 6 verteilten Zufallszahl:

7

3 Zentraler Grenzwertsatz, Normalverteilung

Selbst wenn der Erwartungswert $\mu = E[X]$ und die Varianz $\sigma^2 = E[X^2] - (E[X])^2$ einer Zufallsvariablen X bekannt sind, gibt es noch immer wilde Möglichkeiten für die Verteilung von X:

8

Es gibt aber eine Situation, in der man mehr weiß. Der zentrale Grenzwertsatz [central limit theorem] besagt anschaulich: Ist die Zufallsvariable X ein fester Wert plus eine Summe vieler unkorrelierter Fehler, die den Erwartungswert 0 und eine kleine Varianz haben, dann ist die Verteilung von X ungefähr die Normalverteilung [normal distribution] zu diesem μ und diesem σ. Die hat folgende Wahrscheinlichkeitsdichte:

9

Damit die Gesamtwahrscheinlichkeit, also die Fläche unter dieser Kurve, gleich 1 wird, wählt man die Konstante C vor der Exponentialfunktion gleich $\frac{1}{\sigma\sqrt{2\pi}}$. Bei der Normalverteilung liegen etwa 68 Prozent der Werte innerhalb einer Standardabweichung vom Erwartungswert; etwa 3 Promille der Werte liegen außerhalb von drei Standardabweichungen vom Erwartungswert:

10