Standardisierung einer Zufallsvariablen

Jan Sellner

15. Juni 2014

1 Grundsätzliche Definition

Das Ziel der Standardisierung ist es, aus einer vorhandenen Zufallsvariablen X eine neue Zufallsvariable $\tilde{X}$ zu erzeuge. $\tilde{X}$ hat dabei die besonderen Eigenschaften

$$E(\tilde{X}) = 0 \qquad \text{und} \qquad Var(\tilde{X}) = 1.$$

Um das zu erreichen, geht man folgendermaßen vor

1. X wird in Richtung des Ursprungs verschoben, so dass sich ein Erwartungswert von 0 ergibt. Dadurch entsteht die neue Zufallsvariable
$$\hat{X} = X - E(X).$$
2. Diese Zufallsvariable wird nun erneut verändert, so dass die Varianz 1 beträgt. Auch hier entsteht wieder eine neue Zufallsvariable
$$\tilde{X} = \frac{\hat{X}}{\sigma_X}.$$

Im Gesamten hat mal also für die Standardisierung den Schritt

$$\boxed{\tilde{X} = \frac{X - E(X)}{\sigma_X}}$$

vollzogen.

2 Herleitung

Der 1. Schritt ist durch die Rechenregeln des Erwartungswertes definiert: $E(\hat{X}) = E(X - E(X)) = E(X) - E(X) = 0$. Für den 2. Schritt muss man bereits etwas tiefer in die Trickkiste greifen. Zuerst einmal ist die Feststellung

$$\sigma_X^2 = Var(X) = Var(\hat{X}) = E((\hat{X} - E(\hat{X}))^2) = E(\hat{X}^2) \tag{1}$$

wichtig[1]. Das nun auch der 2. Schritt das gewünschte Ergebnis liefert, kann folgendermaßen gesehen werden

$$\begin{aligned} Var(\tilde{X}) &= E((\tilde{X} - E(\tilde{X}))^2) = E(\tilde{X}^2) = \sum_{i=1}^{n} \tilde{x}_i^2 p_i = \sum_{i=1}^{n} \left(\frac{\hat{x}_i}{\sigma_X}\right)^2 p_i \\ &= \sum_{i=1}^{n} \frac{\hat{x}_i^2}{\sigma_X^2} p_i = \frac{1}{\sigma_X^2} \sum_{i=1}^{n} \hat{x}_i^2 p_i = \frac{1}{\sigma_X^2} E(\hat{X}^2) = \frac{\sigma_X^2}{\sigma_X^2} = 1. \end{aligned}$$

[1] Man beachte, dass sich bei einer gleichmäßigen Verschiebung der Zufallsvariablen um einen bestimmten Wert (hier: $E(X)$) keinerlei Veränderungen an der Varianz ergeben (die Streuung bleibt ja gleich).

Zu Beginn wurde verwendet, dass auch $E(\tilde{X}) = 0$ gilt, da beim Schritt von $\hat{X}$ nach $\tilde{X}$ die Zufallswerte nur mit einem Skalierungsfaktor versehen wurden ($\frac{1}{\sigma_X}$) und es gilt

$$E(\tilde{X}) = E(\hat{X} \cdot \frac{1}{\sigma_X}) = \frac{1}{\sigma_X} \cdot E(\hat{X}) = \frac{1}{\sigma_X} \cdot 0 = 0.$$

Gegen Ende wurde dann noch das Ergebnis aus Gleichung 1 verwendet.

3 Grafische Veranschaulichung

Der Vorgang der Standardisierung sei nun am Beispiel von $X =$ *Augenzahl beim Wurf von zwei Würfeln* veranschaulicht. Den 1. Schritt der Verschiebung kann man dazu in Abbildung 1 sehen.

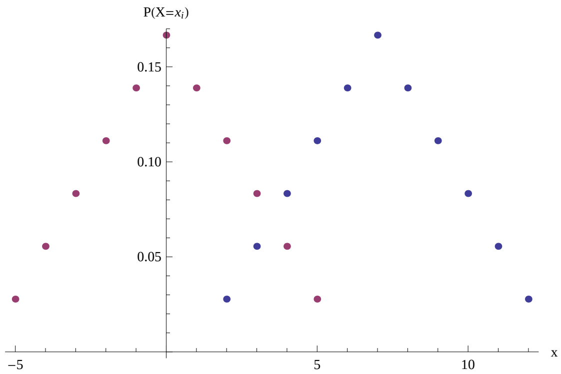

Abbildung 1: Verschieben der Zufallsvariablen $\hat{X} = X - E(X)$ in Richtung Ursprung.

Der 2. Schritt, wo die Skalierung vorgenommen wird, kann anhand von Abbildung 2 veranschaulicht werden.

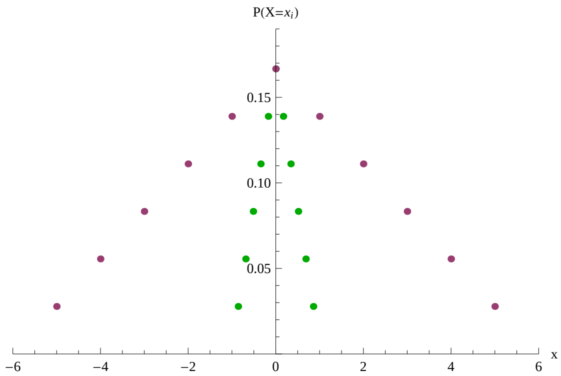

Abbildung 2: Skalieren der Zufallsvariablen $\tilde{X} = \frac{\hat{X}}{\sigma_X}$ am Ursprung.

Wie man sieht, werden hierbei die Zufallswerte nach innen gestaucht. Der Vorteil dabei ist, dass so eine gute Vergleichbarkeit verschiedener Zufallsvariablen möglich ist. Zudem ist die Standardisierung auch bei der Normalverteilung von großer Bedeutung. Hier wird sie verwendet, um die entsprechenden Wahrscheinlichkeiten der Verteilung überhaupt erst vernünftig berechnen zu können.