

Die Verteilung von Stichprobenmittelwerten einer normalverteilten Variablen

GRUNDLAGEN

Die Normalverteilung ist eine reproduzierende Verteilung, d.h., dass die Summe mehrerer voneinander unabhängiger, normalverteilter Variablen wieder normalverteilt ist. Der Mittelwert der Summe ist die Summe der einzelnen Mittelwerte, analog die Varianz.

Formal:

X_1, X_2, \dots, X_n seien voneinander unabhängige, normalverteilte Zufallsvariablen mit den Mittelwerten $\mu_1, \mu_2, \dots, \mu_n$ und den Varianzen $\sigma_1, \sigma_2, \dots, \sigma_n$. Die Zufallsvariable $Y = \sum_i X_i$ ist dann ebenfalls normalverteilt, wobei $E(Y) = \sum_i \mu_i$ und $\sigma_Y^2 = \sum_i \sigma_i^2$ gilt.

DIE VERTEILUNG VON \bar{X}

Die Zufallsvariable X habe in der (unendlich großen) Gesamtpopulation die Verteilung $N(\mu_X, \sigma_X)$. \bar{X} sei jene Zufallsvariable, die man erhält, indem man wiederholt Stichproben der Größe n zieht und deren Mittelwerte \bar{x} berechnet.

Man kann sie anschreiben als:

$$\bar{X} = \frac{1}{n} \cdot (X_1 + X_2 + \dots + X_n)$$

X_1 ist hier die Zufallsvariable für die jeweils erste Zahl jeder Stichprobe, X_2 für die jeweils zweite gezogene Zahl usw. All diese X_i sind - wie X - zufällig aus der Gesamtpopulation gezogene Werte und haben somit - wie X - die Verteilung $N(\mu_X, \sigma_X)$. Da die Ziehungen zufällig erfolgen, ist auch die Voraussetzung der Unabhängigkeit der X_i erfüllt. Also ist ihre Summe $(X_1 + X_2 + \dots + X_n)$ ebenfalls normalverteilt:

$$X_1 + X_2 + \dots + X_n \sim N(n \cdot \mu_X, \sqrt{n} \cdot \sigma_X)$$

Die Division durch n stellt eine einfache Lineartransformation dar. Diese wirkt sich auf die *Parameter* der Verteilung, nicht aber auf den Verteilungstyp aus:

$$\frac{1}{n} \cdot (X_1 + X_2 + \dots + X_n) \sim N\left(\frac{n \cdot \mu_X}{n}, \frac{\sqrt{n} \cdot \sigma_X}{n}\right)$$

$$X \sim N\left(\mu_X, \frac{\sigma_X}{\sqrt{n}}\right)$$

In Worten

Ist die Zufallsvariable normalverteilt, so sind die Stichprobenmittelwerte ebenfalls normalverteilt mit demselben Mittelwert und der um den Faktor $\frac{1}{\sqrt{\text{Stichprobenumfang}}}$ verringerten Standardabweichung.

Hinweis des Autors:

Dieses Dokument wurde am 23.04.2018 überarbeitet. Die Änderungen betreffen im Wesentlichen die Voraussetzung der Unabhängigkeit der X_i , die im ursprünglichen Dokument aus Versehen nicht erwähnt worden war. Vielen Dank an Dr. Michael Grosser für den wichtigen Hinweis.