Vorlesung am 02.02.2009

Sind X_1, X_2, . . . unabhängige und identisch verteilte reelle Zufallsvariablen, so gilt nach dem **zentralen Grenzwertsatz**, dass für “große” n

$$\frac{\sqrt{n}}{\sqrt{V(X_1)}} \cdot \left(\frac{1}{n} \sum_{i=1}^{n} X_i - \mathbf{E}X_1 \right)$$

annähernd $N(0,1)$-verteilt ist.

Also ist z.B. für “große” n

$$\mathbf{P}\left[\left| \frac{\sqrt{n}}{\sqrt{V(X_1)}} \cdot \left(\frac{1}{n} \sum_{i=1}^{n} X_i - \mathbf{E}X_1 \right) \right| \leq 1.96 \right] \approx 0.95$$

Da

$$\left| \frac{\sqrt{n}}{\sqrt{V(X_1)}} \cdot \left(\frac{1}{n} \sum_{i=1}^{n} X_i - \mathbf{E} X_1 \right) \right| \leq 1.96$$

äquivalent ist zu

$$\frac{\sqrt{n}}{\sqrt{V(X_1)}} \cdot \left(\frac{1}{n} \sum_{i=1}^{n} X_i - \mathbf{E} X_1 \right) \leq 1.96$$

und

$$\frac{\sqrt{n}}{\sqrt{V(X_1)}} \cdot \left(\frac{1}{n} \sum_{i=1}^{n} X_i - \mathbf{E} X_1 \right) \geq -1.96,$$

was wiederum das gleiche bedeutet wie

$$\mathbf{E}X_1 \geq \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{\sqrt{V(X_1)}}{\sqrt{n}} \cdot 1.96$$

und

$$\mathbf{E}X_1 \leq \frac{1}{n}\sum_{i=1}^{n} X_i + \frac{\sqrt{V(X_1)}}{\sqrt{n}} \cdot 1.96,$$

folgt daraus:

$$\mathbf{P}\left[\mathbf{E}X_1 \in \left[\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{\sqrt{V(X_1)}}{\sqrt{n}} \cdot 1.96, \frac{1}{n}\sum_{i=1}^{n} X_i + \frac{\sqrt{V(X_1)}}{\sqrt{n}} \cdot 1.96\right]\right] \approx 0.95$$

Wir schätzen nun $V(X_1)$ durch die empirische Varianz

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2 = \frac{n}{n-1}\cdot\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2 - (\frac{1}{n}\sum_{j=1}^{n}X_j)^2\right)$$

und erhalten daraus als **Konfidenzintervall zum Konfidenzniveau 0.95:**

$$C(X_1,\ldots,X_n) = \left[\frac{1}{n}\sum_{j=1}^{n}X_j - \frac{\sqrt{S_n^2}}{\sqrt{n}}\cdot 1.96, \frac{1}{n}\sum_{j=1}^{n}X_j + \frac{\sqrt{S_n^2}}{\sqrt{n}}\cdot 1.96\right]$$