Vorlesung 6b+

Zum Schätzen von Mittelwerten

Eine große Population von Werten

$$w_1, \ldots, w_g$$

ist auf der Zahlengeraden verteilt.

Man interessiert sich für den Populationsmittelwert:

$$\mu := \frac{1}{g} \sum_{j=1}^{g} w_j.$$

Zur Verfügung stehen die Werte einer

rein zufällig aus der Population gezogene Stichprobe

$$x_1, \ldots, x_n$$

$$m := \frac{1}{n}(x_1 + \ldots + x_n)$$

ist ein *Schätzwert* für μ.

Wie zuverlässig ist diese Schätzung?

Das hängt ab

– von der “Streuung” der Werte w_j in der Population

und

– von der Größe n der Stichprobe.

Goldene Idee der Statistik:

Man fasst $x_1, \ldots, x_n$ auf als Ergebnis eines (rein) zufälligen Ziehens aus der Population:

$$X_1 := w_{J_1}, X_2 := w_{J_2}, \ldots$$

mit $J_1, J_2, \ldots$ rein zufällige Wahl aus $\{1, \ldots, g\}$.

Ein Maß für die Variabilität in der Population ist

$$\sigma^2 := \frac{1}{g} \sum_{j=1}^{g} (w_j - \mu)^2$$

$$X_1, X_2, \ldots$$

unabhängig und identisch verteilt

mit Erwartungswert μ und Varianz σ^2

$$m = \frac{1}{n}(x_1 + \ldots + x_n)$$

fasst man auf

als eine Realisierung (einen Ausgang)

der Zufallsvariable

$$\bar{X} := \frac{1}{n}(X_1 + \ldots + X_n)$$

Wie ist $\bar{X}$ verteilt?

Wie ist $\bar{X}$ verteilt?

Der Zentrale Grenzwertsatz gibt eine Antwort:

$\bar{X}$ ist approximativ N$(\mu, \frac{\sigma^2}{n})$-verteilt.

Ein Problem in der Praxis: Man kennt σ^2 nicht.

Auch σ^2 muss man schätzen.

Ein Vorschlag:

$$s^2 := \frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2$$

Warum im Nenner $n-1$ und nicht n?

Eine Antwort dazu gibt's auf dem nächsten Übungsblatt.

Eine anschauliche Präsentation der Thematik

“Wie genau ist eine Schätzung des Mittelwertes?”

vom Standpunkt der Anwendung aus

finden Sie auf den Folien der Vorlesung

“Statistik für Biologen”

http://ismi.math.uni-frankfurt.de/wakolbinger/teaching/statbio10/statbio.html

Folien 2, “Der Standardfehler”