Schätzen

Ein Prozess oder Experiment wird durch eine Verteilung $f(x, \theta)$ beschrieben, die von einem (oder mehreren) Parameter(n) θ abhängt.

Beispiel: Poissonverteilung mit Zerfallsrate λ beim radioaktiven Zerfall
Temperatur T bei der Maxwellverteilung

Ziel: Schätzen des Parameters θ anhand einer Stichprobe $\{x_1, \ldots, x_n\}$ mittels einer geeigneten Schätzfunktion $\hat{\theta}(x_1, \ldots, x_n)$

Ein guter **Schätzer** $\hat{\theta}$ ist eine von den $\{x_i\}$ abhängige ZV, die folgende Eigenschaften haben sollte:
1. Erwartungstreue: $E(\hat{\theta}) = \theta$
2. Konsistenz: wird mit wachsender Stichprobengröße besser
3. Effizienz (kleine Varianz)

Ausgangssituation:

X Zufallsvariable mit Erwartungswert μ und Varianz σ^2:

Stichprobe: N Messungen (Zufallsvariablen) $X_1, X_2, \ldots, X_N$

Für jede Einzelmessung (Zufallsvariable) X_i gilt:

$$E(X_i) = \mu \quad \text{und} \quad V(X_i) = \sigma^2$$

Schätzen des Erwartungswertes μ der Verteilung durch den Mittelwert der Stichprobe:

$$\boxed{\langle X \rangle = \bar{X} := \frac{1}{N} \sum_{i=1}^{N} X_i}$$

Wie man leicht nachrechnet, gilt für den Erwartungswert dieser Schätzung $E(\bar{X}) = \mu$, d.h. die Schätzung ist **erwartungstreu**.
Für die Varianz dieser Schätzung findet man:

$$V(\bar{X}) = V\left(\frac{1}{N} \sum_{i=1}^{N} X_i\right) = \frac{1}{N^2} \sum_{i=1}^{N} V(X_i) = \frac{1}{N^2} \sum_{i=1}^{N} \sigma^2 = \frac{\sigma^2}{N}$$

d.h. die Schätzung wird mit wachsendem N besser!
Dies ist die Grundlage für die Bestimmung des **Fehlers des Mittelwerts**!

Schätzen der Varianz σ^2 aus der Stichprobe $\{X_1, \ldots, X_N\}$:

Da das wahre μ in der Praxis fast nie bekannt ist, wird für die Schätzung von V(X) der Schätzwert $\bar{X}$ benutzt:

also naheliegend: Schätzen der Varianz der Verteilung σ^2 durch:

$$s_1^2 = \frac{1}{N} \sum_{i=1}^{N} (X_i - \bar{X})^2$$

Für den Erwartungswert dieser Schätzung findet man:

$$E(s_1^2) = \cdots = \frac{N-1}{N} \sigma^2 \ ,$$

d.h. sie ist **nicht erwartungstreu**.

Offenbar kann man aber mit dem Faktor $\frac{N}{N-1}$ einen erwartungstreuen Schätzer für die Varianz konstruieren (**Bessel Korrektur**).

Die Stichprobenvarianz $$\boxed{s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2}$$

liefert eine erwartungstreue Schätzung von σ^2 bei unbekanntem μ.

Maximum-Likelihood Prinzip (advanced topic)

Gegeben: vom Parameter θ abhängige Verteilung $f(x, \theta)$ für die ZV x

Die **Likelihood** $\boxed{L(\theta, x) := f(x_1, \theta) \cdot f(x_2, \theta) \cdots f(x_n, \theta)}$

ist die Wahrscheinlichkeit die Werte $\{x_1, \ldots, x_n\}$ zu beobachten, falls θ vorliegt.

Das Maximum Likelihood Prinzip besagt nun, dass $L(\theta, x)$ für das "wahrscheinlich richtige" θ ein Maximum hat.
Der ML-Schätzer $\hat{\theta}$ wird daher aus der Bedingung

$$\boxed{\frac{dL}{d\theta} = 0} \quad \text{oder} \quad \frac{d(\ln L)}{d\theta} = 0$$

bestimmt. Zumeist hängt die Verteilung von mehreren Parametern ab, d.h. θ ist oft ein Vektor $\theta = (\theta_1, \ldots, \theta_m)$.

Gewichtetes Mittel

Bisher: Messung *derselben Größe* mit *demselben Instrument*

Oft: Dieselbe Größe wird mit verschiedenen Instrumenten und verschiedenen Methoden gemessen, dh. die einzelnen Messungen sind unterschiedlich genau (die Varianzen σ_i^2 der zugrundeliegenden Einzelverteilungen sind verschieden):

$$x_1 \pm \sigma_1, \ldots, x_n \pm \sigma_n$$

Die Maximum-Likelihood liefert dann den folgenden Schätzer für den Erwartungswert der gesuchten Größe:

$$\boxed{\hat{\mu} = \frac{\displaystyle\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}}{\displaystyle\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}} \qquad \text{mit} \qquad \boxed{\sigma^2(\hat{\mu}) = \frac{1}{\displaystyle\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}}$$

Falls alle σ_i gleich sind, kommen die alten Formeln heraus!