SCHÄTZEN VON PARAMETERN

„Realität"
Messgröße folgt einer Verteilung

→ **Modellierung**
$\Theta \subseteq \mathbb{R}^k$ Parameterraum. Verteilung entstammt einer Familie $P = \{P_\theta : \theta \in \Theta\}$, hängt von Parameter θ ab und hat Dichte $f_\theta(x)$.

→ **Schätzen**
Sammle Beobachtungen $X_1, \dots, X_n \overset{iid}{\sim} P_{\theta_0}$ der Messgröße und ermittle den „wahren Parameter" θ_0 mit einem Schätzer θ_n

z.B. Messgröße = Mehltütengewicht

z.B. $P = \{N(\mu, (2g)^2) : \mu \in \mathbb{R}\}$ Gewicht normalverteilt

z.B. Messe einige Mehltütengewichte $X_1, \dots, X_n \overset{iid}{\sim} N(\mu_0, (2g)^2)$

Fisher-Informations-Matrix: $I(\theta) = \mathbb{E}_\theta\left[\nabla \log f_\theta(X_1) \cdot (\nabla \log f_\theta(X_1))^T\right]$

Schätzer θ_n von θ
Sind $X_1, \dots, X_n \overset{iid}{\sim} P_\theta$, dann ist eine Funktion $\theta_n = \theta_n(X_1, \dots, X_n)$ Schätzer von θ.

(Schwache) Konsistenz
$\theta_n \xrightarrow{P} \theta$

Asymptotische Normalität
$\sqrt{n} \cdot (\theta_n - \theta) \xrightarrow{D} \mathcal{N}(0, \Sigma)$

Risiko / Mittlerer quadr. Fehler → min.
$MSE_\theta(\theta_n) = \mathbb{E}_\theta[(\theta_n - \theta)^2]$
$= Var_\theta(\theta_n) + (BIAS_\theta(\theta_n))^2$

Cramer-Rao-Schranke:
$\Sigma \geq I(\theta)^{-1}$
$Var_\theta(\theta_n) \geq \dfrac{\left(1 + \frac{\partial}{\partial\theta} BIAS_\theta(\theta_n)\right)^2}{n \cdot I(\theta)}$

Erwartungstreue
$\mathbb{E}_\theta[\theta_n] = \theta$
$\iff BIAS_\theta(\theta_n) = \mathbb{E}_\theta[\theta_n - \theta] = 0$

(Asymptotische) Effizienz
$\Sigma = I(\theta)^{-1}$
$Var_\theta(\theta_n) = \dfrac{\left(1 + \frac{\partial}{\partial\theta} BIAS_\theta(\theta_n)\right)^2}{n \cdot I(\theta)}$

$P = (P_\theta)$ Exponentialfamilie + Voraus. 15.6.

$P = (P_\theta)$ Exponentialfamilie + Voraus. 15.9.

Maximum-Likelihood-Schätzer
$\theta_n = \underset{\theta \in \Theta}{\arg\max}\, f_\theta^{(n)}(x) \overset{iid}{=} \underset{\theta \in \Theta}{\arg\max} \prod_{i=1}^n f_\theta(X_i)$
bzw.
$\theta_n = \underset{\theta \in \Theta}{\arg\min}\, L_n(\theta), \quad L_n(\theta) \overset{iid}{=} -\frac{1}{n} \sum_{i=1}^n \log f_\theta(X_i)$

$P = \{P_\theta : \theta \in \Theta\}$ ist Exponentialfamilie, falls Dichte
$f_\theta(x) = C(\theta) \exp\left(\sum_{j=1}^s \eta_j(\theta) T_j(x)\right) \cdot h(x)$
mit $h(x) \geq 0 \ \forall x \in \mathbb{R}$,
$C(\theta) > 0 \ \forall \theta \in \Theta$,
$\eta_j(\theta), C(\theta)$ stetig auf $\mathbb{R}$