**Definition**

- Methode um geeignete [[Schätzer]] zu konstruieren
- $$1.68,\ 1.77,\ 1.70,\ 169.$$

  Die obige Stichprobe erscheint für $\mu = 2.10$ weniger plausibel als für $\mu = 1.70$.

  Gesucht wird jener Schätzer, für den die beobachtete Stichprobe *am plausibelsten* ist.
- gegeben ist [[Dichtefunktionen PDF]] (oder [[Wahrscheinlichkeitsfunktion PMF]])
  - Dichtefunktion für alle ZV gleich
  - $$f_X(x_1, \dots, x_n|\theta) = \prod_{i=1}^{n} f_{X_1}(x_i|\theta).$$
- Likelihood-Funktion
  - Funktion in $\theta$ für feste Parameter
  - $$L(\theta|x) = L(\theta|x_1, \dots, x_n) = f_X(x_1, \dots, x_n|\theta),$$
- maximum likelihood estimator MLE
  - $$\hat{\theta} = \hat{\theta}(X) = \arg\max_{\theta \in \Theta} L(\theta|X).$$
  - mögliche Probleme
    - Existiert überhaupt ein Maximum in $\Theta$?

      Ist das Maximum eindeutig?

      In einigen Fällen ist es schwierig, ein explizites Maximum zu berechnen. Stattdessen können numerische Methoden verwendet werden.

      Für das Auffinden des globalen Maximums müssen auch die Ränder des Parameterraums untersucht werden.

**Beispiel**

- Bernoulli-Verteilung

geg.: $X_1, \dots, X_n \overset{iid}{\sim} \text{Bernoulli}(p) \quad p \in [0,1]$

ges.: MLE $\hat{p}$

$$P(X_1 = k) = p^k (1-p)^{1-k} \qquad k = 0, 1$$

$$L(p \mid x) = \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i}$$

$$x = (x_1, \dots, x_n) \qquad = p^{S_n} (1-p)^{n - S_n} \qquad S_n := \sum_{i=1}^{n} x_i$$

$$\log(L(p \mid x)) = S_n \log(p) + (n - S_n) \log(1-p)$$

Fall: $0 < S_n < n$:

–

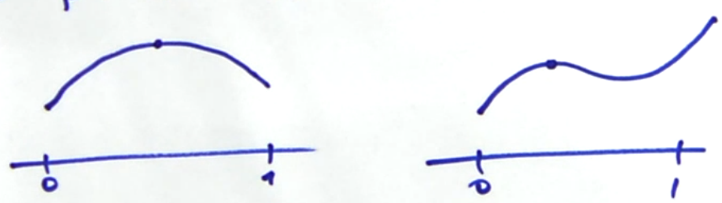
$$\frac{\partial}{\partial p} \log(L(p \mid x)) = \frac{S_n}{p} - \frac{n - S_n}{1-p}$$

$$\frac{S_n}{p} - \frac{n - S_n}{1-p} \overset{!}{=} 0 \iff S_n(1-\cancel{p}) - (n-\cancel{S_n})p \overset{!}{=} 0$$

$$\Rightarrow \quad S_n - np = 0 \quad \Rightarrow \quad \hat{p} = \frac{1}{n} S_n = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{X}.$$

$$\frac{\partial^2}{\partial p^2} \log(L(p \mid x)) = -\frac{S_n}{p^2} - \frac{(n - S_n)}{(1-p)^2} < 0.$$

$\Rightarrow \hat{p} = \bar{X}$ ist lokales Maximum




–

– Alternative: Verhalten am Rand analysieren

$$\log(L(p \mid x)) = S_n \log(p) + (n - S_n) \log(1-p) \underset{\substack{p \to 1 \\ p \to 0}}{\longrightarrow} -\infty$$

*

Fall: $S_n = 0$: $\quad \log(L(p \mid x)) = n \log(1-p)$

$$\hat{p} = 0 = \frac{1}{n} S_n = \bar{X}$$

–

– Fall: $S_n = n$

$\log(L(p|x)) = n \log(p)$

$\Rightarrow \hat{p} = 1 = \frac{1}{n} S_n = \bar{x}$.

In allen Fällen ist $\hat{p} = \bar{x}$ das globale Maximum

$\Rightarrow \hat{p} = \bar{x}$ ist MLE von $p$.

- Normalverteilung 1

  Sei $X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathrm{N}(\mu, 1)$. Die Likelihood–Funktion ist gegeben durch

$$L(\mu|X) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(X_i-\mu)^2}{2}} = (2\pi)^{-n/2} e^{-\frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)^2}.$$

  Der obige Ausdruck ist maximal, wenn $\sum_{i=1}^{n}(X_i-\mu)^2$ minimal ist.

  – Man kann zeigen, dass dies der Fall ist für $\hat{\mu} = \bar{X}$.

- Normalverteilung 2

  – geg.: $x_1, \dots x_n \overset{iid}{\sim} N(\mu, \sigma^2)$ ges.: MLE $\hat{\mu}, \hat{\sigma}^2$

$$L(\mu, \sigma^2|x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\log(L(\mu, \sigma^2|x))$$

$$= \cancel{-\frac{n}{2}\log(2\pi)} - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\ell(\mu, \sigma^2|x) = -n\log(\sigma^2) - \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

  – erste Term vernachlässigbar
    * verändert Position des Maximums nicht

$$\frac{\partial}{\partial \mu} \ell(\mu,\sigma^2|x) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) \overset{!}{=} 0 \quad (\mathrm{I})$$

$$\frac{\partial}{\partial \sigma^2} \ell(\mu,\sigma^2|x) = -\frac{n}{\sigma^2} + \frac{1}{\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2 \overset{!}{=} 0 \quad (\mathrm{II})$$

Aus (I): $\cancel{\frac{1}{\sigma^2}} \sum_{i=1}^{n} (x_i - \mu) \overset{!}{=} 0$

$$\Leftrightarrow \sum_{i=1}^{n} x_i - n\mu \overset{!}{=} 0 \Rightarrow \boxed{\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}.}$$

Setze $\hat{\mu} = \bar{x}$ ein in (II):

$$-\frac{n}{\sigma^2} + \frac{1}{\sigma^4} \sum_{i=1}^{n} (x_i - \bar{x})^2 \overset{!}{=} 0$$

$$-n\sigma^2 + \sum_{i=1}^{n} (x_i - \bar{x})^2 \Rightarrow \boxed{\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

– Nullsetzen der obigen Gleichungen führt zu den Kandidaten

$$\hat{\mu} = \bar{X}, \qquad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 = S_L^2.$$

– Es lässt sich zeigen, dass dies auch ein globales Maximum ist.