Aufgabe 5

(Hausaufgabe, Abgabe freiwillig)

Konstruieren Sie ein Beispiel, so dass $E_\theta[X_i^{2k}] < \infty$ und $h(\theta) = g(m_1(\theta), ..., m_k(\theta))$ für alle $\theta \in \Theta$, aber $\hat{h(\theta)} = g(\hat{m}_1, ..., \hat{m}_k)$ nicht konsistent ist für $h(\theta)$ ist.

Theorem

Sei $(F_{\theta, X_1, ..., X_n})_{\theta \in \Theta}$ ein statistisches Modell, so dass $(X_n)_{n \in \mathbb{N}}$ u.i.v. und für alle $\theta \in \Theta$, $j \leq k$ gilt: $E_\theta[X_i^{2j}] < \infty$. Sei $g : \mathbb{R}^k \to \mathbb{R}$ stetig, dann konvergiert $g(\hat{m}_1, ..., \hat{m}_k)$ in Wahrscheinlichkeit gegen $g(m_1(\theta), ..., m_k(\theta))$.

$X_1 \sim \text{Bernoulli}(\Theta)$, $\Theta \in [0,1] = \Theta$

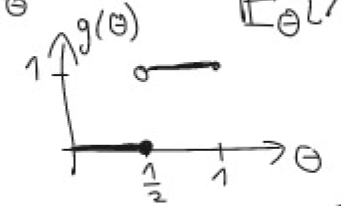
$\mathbb{P}_\Theta(X_1 = 1) = p_{\Theta, X_1} = \Theta = 1 - \mathbb{P}_\Theta(X_1 = 0)$

$k=1$: $m_1(\Theta) = \mathbb{E}_\Theta[X_1] = 1 \cdot \mathbb{P}_\Theta(X_1 = 1) + 0 \cdot \mathbb{P}(X_1 = 0) = \Theta$

$\mathbb{E}_\Theta[X_1^{2k}] = \mathbb{E}_\Theta[X_1^2] = 1 \cdot \mathbb{P}_\Theta(X_1^2 = 1) + 0 \cdot \mathbb{P}_\Theta(X_1^2 = 0) = 1 \cdot \Theta = \Theta < \infty$

$g : [0,1] \to \mathbb{R}$

$\Theta \mapsto \begin{cases} 0 & \Theta \leq \frac{1}{2} \\ 1 & \Theta > \frac{1}{2} \end{cases}$



$g(m_1(\Theta)) = g(\Theta) = h(\Theta)$ ✓

$h = g$

Noch zu zeigen: $\hat{h(\Theta)} = g(\hat{m}_1) = g\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) = \begin{cases} 1 & \text{falls } \frac{1}{n}\sum_{i=1}^n X_i > \frac{1}{2} \\ 0 & \text{falls } \frac{1}{n}\sum_{i=1}^n X_i \leq \frac{1}{2} \end{cases}$

ist nicht konsistent für $h(\Theta)$, d.h.

$\exists \Theta \in [0,1]$ und ein $\varepsilon > 0$ sodass $\lim_{n \to \infty} \mathbb{P}_\Theta(|\hat{h(\Theta)} - h(\Theta)| > \varepsilon) \neq 0$

Wähle $\Theta := \frac{1}{2}$ Wähle $\varepsilon := 0{,}8 = \frac{8}{10} = \frac{4}{5}$

$\mathbb{P}_{\frac{1}{2}}\left(|\hat{h(\Theta)} - \underbrace{h(\tfrac{1}{2})}_{=0}| > \frac{4}{5}\right) = \mathbb{P}_{\frac{1}{2}}\left(g\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) > \frac{4}{5}\right)$

$= \mathbb{P}_{\frac{1}{2}}\left(g\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) = 1\right) =$

$= \mathbb{P}_{\frac{1}{2}}\left(\frac{1}{n}\sum_{i=1}^n X_i > \frac{1}{2}\right)$

$= \mathbb{P}_{\frac{1}{2}}\left(\sum_{i=1}^n X_i > \frac{n}{2}\right) \stackrel{\downarrow}{=} \frac{1}{2} \not\xrightarrow{n \to \infty} 0$

n ungerade $\Rightarrow \mathbb{P}_{\frac{1}{2}}\left(\sum_{i=1}^n X_i = \frac{n}{2}\right) = 0$

$\sum_{i=1}^n X_i \sim \text{Binom}(n, \frac{1}{2})$

$\Rightarrow \hat{h}(\Theta)$ nicht konsistent für $h(\Theta)$

Definition

Sei für alle $n \in \mathbb{N}$ ein statistisches Modell $(F_{\theta, X_1, ..., X_n})_{\theta \in \Theta}$ gegeben. Ein Schätzer $\hat{h(\theta)}$ für $h(\theta)$ heißt konsistent, wenn er in Wahrscheinlichkeit gegen $h(\theta)$ kovergiert, d.h. für alle $\varepsilon > 0$ gilt

$$P\left(\left|\hat{h(\theta)} - h(\theta)\right| > \varepsilon\right) \xrightarrow{n \to \infty} 0.$$

$\hat{h(\Theta)}$ konsistent für $h(\Theta)$:$\Leftrightarrow$

$\forall \Theta \in \Theta \ \forall \varepsilon > 0$: $\lim_{n \to \infty} \mathbb{P}(|\hat{h(\Theta)} - h(\Theta)| > \varepsilon) = 0$

$\neg(\hat{h(\Theta)}$ konsistent für $h(\Theta)) \equiv \exists \Theta \in \Theta \ \exists \varepsilon > 0$:

$\lim_{n \to \infty} \mathbb{P}(|\hat{h(\Theta)} - h(\Theta)| > \varepsilon) \neq 0$

$\boxed{\neg(\forall x : A(x)) \equiv \exists x : \neg A(x)}$

$\boxed{\neg(\exists x : A(x)) \equiv \forall x : \neg A(x)}$

i) X_1 = Anzahl Misserfolge bis zum ersten Erfolg

$$\mathbb{P}_p(X_1 = \underset{k\in\mathbb{N}_0}{k}) = (1-p)^k \cdot p \qquad p\in[0,1]=\Theta$$

ii)
$$L(p,k_1,\dots,k_n) = \prod_{i=1}^{n} \mathbb{P}_p(X_i = k_i) = \prod_{i=1}^{n}\left((1-p)^{k_i}\cdot p\right) = (1-p)^{\sum_{i=1}^{n} k_i} \cdot p^n$$

$$\hat{p} = \underset{p\in[0,1]}{\arg\max}\, L(p,k_1,\dots,k_n) = \underset{p\in[0,1]}{\arg\max}\, \log\left(L(p,k_1,\dots,k_n)\right)$$

$$\frac{\partial L}{\partial p} = -\sum_{i=1}^{n} k_i \cdot (1-p)^{\sum_{i=1}^{n} k_i - 1} \cdot p^n + (1-p)^{\sum_{i=1}^{n} k_i} \cdot n \cdot p^{n-1}$$

$$= (1-p)^{\sum_{i=1}^{n} k_i - 1} \cdot p^{n-1} \underbrace{\left(-\sum_{i=1}^{n} k_i \cdot p + n\cdot(1-p)\right)}_{\overset{!}{=}0} \overset{!}{=} 0$$

$$\Rightarrow$$

$$-\sum_{i=1}^{n} k_i \cdot p + n - np = 0$$

$$p\cdot\left(-\sum_{i=1}^{n} k_i - n\right) = -n$$

$$p = \frac{n}{\sum_{i=1}^{n} k_i + n}$$

$$\Rightarrow \hat{p} = \frac{n}{\sum_{i=1}^{n} X_i + n} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} X_i + 1}$$ ist Maximum-Likelihood-Schätzer

Aufgabe 3

$$\frac{1}{\frac{1}{n}\sum_{i=1}^{n} X_i + 1}$$

Aufgabe 2
(gemeinsame Bearbeitung während des Tutoriums)

Ein Experiment mit zwei möglichen Ausgängen (Erfolg und Misserfolg) wird solange wiederholt, bis der erste Erfolg auftritt und die Anzahl der Misserfolge wird notiert.

i) Geben Sie ein geeignetes statistisches Modell an.

ii) Finden Sie einen Schätzer für die Erfolgswahrscheinlichkeit mit einer Methode Ihrer Wahl.

$|q|<1$ geometrische Summenformel

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$

$$\mathbb{P}_p(X_1 = k) = (1-p)^k \cdot p, \quad k \in \mathbb{N}_0$$

$$m_1(p) = \mathbb{E}_p[X_1] = \sum_{k=0}^{\infty} k \cdot (1-p)^k \cdot p = p \cdot (1-p) \cdot \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} = p \cdot (1-p) \cdot (-1) \sum_{k=1}^{\infty} \frac{d}{dp}\left((1-p)^k\right)$$

$$= p \cdot (1-p) \cdot (-1) \cdot \frac{d}{dp} \sum_{k=1}^{\infty} (1-p)^k = \sum_{k=0}^{\infty}(1-p)^k - 1$$

$$= p \cdot (1-p) \cdot (-1) \cdot \frac{d}{dp}\left(\frac{1}{1-(1-p)} - 1\right)$$

$$= p \cdot (1-p) \cdot (-1) \cdot \frac{d}{dp} \frac{1}{p} = p \cdot (1-p) \cdot (-1) \cdot (-1) \frac{1}{p^2}$$

$$= \frac{1-p}{p}$$

$$\Rightarrow m_1(p) = \frac{1-p}{p} \Rightarrow$$

$$m_1(p) = \frac{1-p}{p}$$

$$p \cdot m_1(p) = 1-p$$

$$p = \frac{1}{1+m_1(p)}$$

$$p = g(m_1(p))$$

$$\Rightarrow g(x) = \frac{1}{1+x}$$

$\Rightarrow$ Momentenmethodenschätzer ist $g(\hat{m}_1) = \dfrac{1}{1+\hat{m}_1} = \dfrac{1}{1+\frac{1}{n}\sum_{i=1}^{n} X_i}$