**Aufgabe 5**

(Hausaufgabe, Abgabe freiwillig) Es seien $X_1, ..., X_n$ unabhängig und diskret gleichverteilt auf der Menge $\{1, 2, ..., \theta\}$, wobei $\theta \in \Theta = \mathbb{N}$ unbekannt ist. Finden Sie eine suffiziente Statistik der Dimension 1 und nutzen Sie dies, um einen unverzerrten Schätzer geringer Varianz für $\theta$ zu finden.

$$\forall k \in \{1,2,..,\Theta\}: \mathbb{P}_\Theta(X_1 = k) = \frac{1}{\Theta}\cdot \mathbb{1}_{\{1,2,..,\Theta\}}(k) = p_{\Theta,X_1}(k)$$

**Lemma (Faktorisierungslemma)**

*Sei $(X_1, ..., X_n)^t$ ein Zufallsvektor und $(F_{\theta,X_1,...,X_n})_{\theta\in\Theta}$ ein reguläres statistisches Modell. Falls für eine Statistik $T$ gilt:*

$$p_{\theta,X_1,...,X_n}(x_1,...,x_n) = g_\theta(T(x_1,...,x_n))h(x_1,...,x_n) \quad \text{(diskreter Fall)},$$
$$\text{bzw. } f_{\theta,X_1,...,X_n}(x_1,...,x_n) = g_\theta(T(x_1,...,x_n))h(x_1,...,x_n) \quad \text{(stetiger Fall)},$$

*dann ist $T$ suffizient.*

$$p_{\Theta,X_1,...,X_n}(x_1,...,x_n) = \prod_{i=1}^{n} p_{\Theta,X_i}(x_i) = \frac{1}{\Theta^n}\cdot \prod_{i=1}^{n} \mathbb{1}_{\{1,2,..,\Theta\}}(x_i) =$$

$$\prod_{i=1}^{n} \mathbb{1}_{\{1,2,..,\Theta\}}(x_i) = \mathbb{1}_{\{1,..,\Theta\}}(x_1)\cdot \mathbb{1}_{\{1,..,\Theta\}}(x_2)\cdot \ldots$$

$$= g_\Theta(T(x_1,...,x_n))\cdot h(x_1,...,x_n)$$

$$h(x_1,..,x_n) = \mathbb{1}_{\mathbb{N}}(x_1)\cdot \ldots \cdot \mathbb{1}_{\mathbb{N}}(x_n)$$

$$g_\Theta(x) = \frac{1}{\Theta^n}\cdot \mathbb{1}_{\{1,..,\Theta\}}(x)$$

$$T(x_1,..,x_n) = \max\{x_1,..,x_n\}$$

$$= \underbrace{\frac{1}{\Theta^n}\cdot \mathbb{1}_{\{1,..,\Theta\}}(\underbrace{\max\{x_1,..,x_n\}}_{T(x_1,..,x_n)})}_{g_\Theta(T(x_1,..,x_n))}\cdot \underbrace{\mathbb{1}_{\mathbb{N}}(x_1)\cdot \ldots \cdot \mathbb{1}_{\mathbb{N}}(x_n)}_{=h(x_1,..,x_n)}$$

$$\xRightarrow{\text{Faktorisierungslemma}} T(x_1,..,x_n) = \max\{x_1,..,x_n\}$$

suffizienter Schätzer

**Theorem (Rao-Blackwell)**

*Sei $h(\hat\theta)$ ein erwartungsteuer Schätzer für $h(\theta)$ und $T$ eine suffiziente Statistik. Dann hängt $E[h(\hat\theta)|T]$ nicht von $\theta$ ab und dies ist ebenfalls ein erwartungsteuer Schätzer. Außerdem gilt für alle $\theta \in \Theta$*

$$\mathrm{Var}_\theta\left[E[h(\hat\theta)|T]\right] \le \mathrm{Var}_\theta\left[h(\hat\theta)\right]$$

$$\hat\Theta = \hat\Theta(X_1,\dots,X_n)$$

$\hat\Theta$ erwartungstreu heißt $\forall \Theta \in \Theta$:

$$\mathbb{E}_\Theta[\hat\Theta] = \Theta$$

$X_1 \in \{1,\dots,\Theta\}$

$$\mathbb{E}_\Theta[X_1] = 1\cdot \mathbb{P}_\Theta(X_1=1) + 2\cdot\mathbb{P}_\Theta(X_1=2) + \dots + \Theta\cdot\mathbb{P}_\Theta(X_1=\Theta)$$

$$= 1\cdot\frac{1}{\Theta} + 2\cdot\frac{1}{\Theta} + \dots + \Theta\cdot\frac{1}{\Theta}$$

$$= \frac{1}{\Theta}\cdot(1+2+\dots+\Theta) = \frac{1}{\cancel{\Theta}}\cdot\frac{\cancel{\Theta}(\Theta+1)}{2} = \frac{\Theta+1}{2} \neq \Theta \Rightarrow X_1 \text{ nicht erwartungstreu}$$

$$\boxed{\sum_{i=1}^{n} i = 1+2+\dots+n = \frac{n\cdot(n+1)}{2}}$$

$$\mathbb{E}_\Theta[X_1+\dots+X_n] = \mathbb{E}_\Theta[X_1]+\dots+\mathbb{E}_\Theta[X_n] = \frac{\Theta+1}{2}+\dots+\frac{\Theta+1}{2}$$

$$= n\cdot\frac{\Theta+1}{2} \neq \Theta \Rightarrow X_1+\dots+X_n \text{ nicht erwartungstreu}$$

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$\hat\Theta = 2\cdot\frac{1}{n}\cdot(X_1+\dots+X_n) - 1$$

$$\mathbb{E}[\hat\Theta] = 2\cdot\frac{1}{n}\cdot\mathbb{E}[X_1+\dots+X_n] - 1 = \cancel{2\cdot\frac{1}{n}\cdot n}\cdot\frac{\Theta+1}{\cancel{2}} - 1$$

$$= \Theta + 1 - 1 = \Theta$$

$$\Rightarrow \hat\Theta = \frac{2}{n}\cdot(X_1+\dots+X_n) - 1$$

ist erwartungstreuer Schätzer für $\Theta$

Nach Rao-Blackwell ist $\mathbb{E}[\hat\Theta \mid T(X_1,\dots,X_n)] = \mathbb{E}_\Theta\left[\frac{2}{n}(X_1+\dots+X_n) - 1 \mid \max\{X_1,\dots,X_n\}\right] = \frac{2}{n}\mathbb{E}_\Theta[X_1+\dots+X_n \mid \max\{X_1,\dots,X_n\}] - 1$

auch ein erwartungstreuer Schätzer mit besserer Varianz

**Aufgabe 4**
(Hausaufgabe, Abgabe freiwillig) Es seien $X_1, ..., X_n$ u.i.v. $\exp(\lambda)$-verteilt. Zeigen Sie, dass $T = \sum_{i=1}^{n} X_i$ eine suffiziente Statistik ist. Nutzen Sie dies, um einen unverzerrten Schätzer geringer Varianz für $\lambda$ zu finden.

$$f_{\lambda, X_1}(x_1) = \lambda \cdot e^{-\lambda x_1} \cdot \mathbb{1}_{[0,\infty[}(x_1)$$

$$f_{\lambda,(X_1,\ldots,X_n)}(x_1,\ldots,x_n) = f_{\lambda,X_1}(x_1) \cdot \ldots \cdot f_{\lambda,X_n}(x_n) =$$

$$= \lambda \cdot e^{-\lambda x_1} \cdot \mathbb{1}_{[0,\infty[}(x_1) \cdot \ldots \cdot \lambda \cdot e^{-\lambda x_2} \cdot \mathbb{1}_{[0,\infty[}(x_n)$$

$$\underbrace{\lambda \cdot \lambda \cdot \ldots \cdot \lambda}_{n\text{-mal}} =$$

$$= \underbrace{\lambda^n \cdot e^{-\lambda \underbrace{\sum_{i=1}^{n} x_i}_{T(x_1,\ldots,x_n)}}}_{g_\lambda(T(x_1,\ldots,x_n))} \cdot \underbrace{\mathbb{1}_{[0,\infty[}(x_1) \cdot \ldots \cdot \mathbb{1}_{[0,\infty[}(x_n)}_{= h(x_1,\ldots,x_n)}$$

$$e^{-\lambda x_1} \cdot e^{-\lambda x_2} = e^{-\lambda x_1 - \lambda x_2}$$

$$\boxed{a^n \cdot a^m = a^{n+m}}$$

$$\boxed{g_\lambda(x) = \lambda^n \cdot e^{-\lambda x}}$$

$$\Longrightarrow T(x_1,\ldots,x_n) = x_1 + \ldots + x_n$$ ist suffiziente Statistik

Faktorisierungs-lemma

**Theorem (Rao-Blackwell)**

Sei $\hat{h(\theta)}$ ein erwartungstreuer Schätzer für $h(\theta)$ und $T$ eine suffiziente Statistik. Dann hängt $E[\hat{h(\theta)}|T]$ nicht von $\theta$ ab und dies ist ebenfalls ein erwartungstreuer Schätzer. Außerdem gilt für alle $\theta \in \Theta$

$$\mathrm{Var}_\theta\left[E[\hat{h(\theta)}|T]\right] \leq \mathrm{Var}_\theta\left[\hat{h(\theta)}\right]$$

$$\mathbb{E}_\delta[X_1] = \frac{1}{\delta} \neq \delta \Rightarrow X_1$$ kein erwartungstreuer Schätzer für $\delta$

Aber: $X_1$ ist erwartungstreuer Schätzer für $\frac{1}{\delta} =: h(\delta)$

$$\min\{X_1, \dots, X_n\} \sim \mathrm{Exp}(n \cdot \delta)$$

$$\mathbb{E}_\delta[\min\{X_1, \dots, X_n\}] = \frac{1}{n \cdot \delta}$$

**Weitere Eigenschaften** [ Bearbeiten | Quelltext bearbeiten ]

Sind $X_1 \sim \mathrm{Exp}(\lambda_1), \dots, X_n \sim \mathrm{Exp}(\lambda_n)$ stochastisch unabhängig, so ist $\min(X_1, \dots, X_n) \sim \mathrm{Exp}(\lambda_1 + \dots + \lambda_n)$

$$\widehat{\left(\frac{1}{\delta}\right)} := n \cdot \min\{X_1, \dots, X_n\}$$

$$\Rightarrow \mathbb{E}_\delta\left[\widehat{\left(\frac{1}{\delta}\right)}\right] = \frac{1}{\delta} \quad \forall \delta > 0 \Rightarrow \widehat{\left(\frac{1}{\delta}\right)}$$ ist erwartungstreue Schätzer für $\frac{1}{\delta}$

$\xRightarrow{\text{Rao-Blackwell}}$

$$\mathbb{E}\left[\widehat{\left(\frac{1}{\delta}\right)} \,\middle|\, T(X_1, \dots, X_n)\right] =$$

$$= \mathbb{E}[n \cdot \min\{X_1, \dots, X_n\} | X_1 + \dots + X_n] = n \cdot \mathbb{E}[\min\{X_1, \dots, X_n\} | X_1 + \dots + X_n]$$

ist auch erwartungstreue Schätzer für $\frac{1}{\delta}$

mit Varianz $\leq \mathrm{Var}_\delta\left(\widehat{\left(\frac{1}{\delta}\right)}\right) \; \forall \delta > 0$

miro