Stochastik II

Sommersemester 2022

Vorlesung: Prof. Dr. Thorsten Schmidt
Übung: Moritz Ritter

Anwesenheitsblatt 2

Abgabe: Freitag, 27.05.2021, um 10:00 Uhr.

Aufgabe 1 (4 Punkte)**.** Geben Sie ein identifizierbares aber nicht reguläres statistisches Modell an. Konstruieren Sie ein statistisches Modell, das regulär aber nicht identifizierbar ist.

Aufgabe 2 (4 Punkte)**.** Beim Roulette-Spiel gibt es 18 rote Zahlen, 18 schwarze Zahlen und die grüne 0. Bei 1000 beobachteten Spielen sei X_1 die Anzahl der dabei aufgetretenen roten Zahlen, X_2 die der schwarzen Zahlen und X_3 die Anzahl der Nullen. Berechnen Sie die bedingten Verteilungen von X_2 gegeben X_1 und von X_3 gegeben $X_1 + X_2$.

Aufgabe 3 (4 Punkte)**.** Zeigen Sie, dass die Statistik $T := \sum_{i=1}^{n} X_i$ im Bernoulli-Verteilungsmodell suffizient ist. Beweisen Sie dies mit und ohne dem Faktorisierungssatz.

Aufgabe 4 (4 Punkte)**.** Seien X_1 und X_2 unabhängige, geometrisch verteilte Zufallsvariablen mit unbekanntem Parameter $\theta \in (0, 1)$, das heißt

$$P_\theta(X_1 = k) = P_\theta(X_2 = k) = \theta(1-\theta)^{k-1}\,, k \in \mathbb{N}.$$

Zeigen Sie nach Definition der Suffizienz, dass die Statistik $T = X_1 + X_2$ suffizient für $\mathscr{P} = \{P_\theta, \theta \in (0, 1)\}$ ist.

Stochastik 2 Blatt 02

Luis Juschke, Lorenz Bung

Aufgabe 01:

- Ein nicht identifizierbares aber reguläres Modell

  $\mathcal{P}_1 = \{ \mathcal{N}(\theta_1 + \theta_2, 1) \mid (\theta_1, \theta_2) \in \mathbb{R}^2 = \Theta \}$

  $\Rightarrow$ für $\theta = (0,0)^T$, $\theta' = (-1,1)^T$ gilt

  $\mathcal{N}(\theta, 1) = \mathcal{N}(0,1) = \mathcal{N}(\theta', 1)$ und $\theta \neq \theta'$

  $\mathcal{P}_1$ ist regulär da die Normalverteilung eine Dichte hat

- Ein nicht reguläres aber identifizierbares Modell

  $\mathcal{P}_2 = \{ U(0,\theta) + \delta_{\frac{1}{2}\theta} \mid \theta \in [1,2] = \Theta \}$

  $\mathcal{P}_2$ ist klarerweise identifizierbar da für $\theta_1 \neq \theta_2$ gilt $U(0,\theta_1) + \delta_{\frac{1}{2}\theta_1} = U(0,\theta_2) + \delta_{\frac{1}{2}\theta_2}$

  Außerdem ist $\mathcal{P}_2$ nicht regulär da weder diskret noch stetig

Aufgabe 02:

Fasst man schwarze Zahlen und die grüne 0 bzw. rote Zahlen und die grüne 0 zusammen so sind X_1, X_2 binomialverteilt, für $k + h \leq 1000$

$$\Rightarrow P(X_2 = k \mid X_1 = h) = \frac{P(X_2 = k, X_1 = h)}{P(X_1 = h)} =$$

$$\frac{\binom{1000}{k+h}\left(\frac{18}{37}\right)^{k+h}\left(\frac{1}{37}\right)^{1000-h-k}}{\binom{1000}{h}\left(\frac{18}{37}\right)^{h}\left(\frac{19}{37}\right)^{1000-h}} =$$

$$\binom{1000-h}{k}\left(\frac{18}{37}\right)^{k}\left(\frac{1}{37}\right)^{1000-h-k}\left(\frac{37}{19}\right)^{1000-h} =$$

$$\binom{1000-h}{k}\left(\frac{18}{37}\right)^{k}\left(\frac{1}{19}\right)^{1000-h}\cdot\left(\frac{1}{37}\right)^{-k} =$$

$$\binom{1000-h}{k}\cdot\frac{18^k}{19^{1000-h}}$$

Aufgabe 03

- ohne Faktorisierungssatz

$X_i \sim Ber(\theta) \quad \forall i = 1, \dots, n \quad , X = (X_1, \dots, X_n)$

$\Rightarrow T = \sum_{i=1}^{n} X_i \sim Bin(n, \theta)$

$$\Rightarrow P(X \mid T(x)) = \frac{P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n, T(x) = t)}{P(T(x) = t)} \overset{*}{=}$$

$$\frac{P(X_1 = x_1, \dots, X_n = x_n)}{P(T(x) = t)} \overset{X_i \text{ unabhängig}}{=} \frac{P(X_1 = x_1) P(X_2 = x_2) \dots P(X_n = x_n)}{P(T(x) = t)} =$$

$$\frac{\theta^{x_1}(1-\theta)^{1-x_1} \cdot \theta^{x_2}(1-\theta)^{1-x_2} \cdot \ldots \cdot \theta^{x_n}(1-\theta)^{1-x_n}}{\binom{n}{t} \theta^t \cdot (1-\theta)^{n-t}} \underset{\sum_{i=1}^{n} x_i = t}{=}$$

$$\frac{\cancel{\theta^t \cdot (1-\theta)^{n-t}}}{\binom{n}{t} \cancel{\theta^t (1-\theta)^{n-t}}} = \frac{1}{\binom{n}{t}} = \binom{n}{t}^{-1}$$

- mit Faktorisierungssatz

wähle $g(T(x), \theta) = P_\theta(T(x) = t) = \binom{n}{t} \theta^t (1-\theta)^{n-t}$

$h(x) = P_\theta(X = x \mid T(x) = t) = \binom{n}{t}^{-1}$

dann gilt

$p(x, \theta) = \theta^{T(x)} (1-\theta)^{n - T(x)} = g(T(x), \theta) \cdot h(x)$

Aufgabe 04

Sei $T = X_1 + X_2$, $X = (X_1, X_2)$

$$P_\theta(X \mid T(x) = t) = \frac{P_\theta(X_1 = x_1, X_2 = x_2, T(x) = t)}{P_\theta(T(x) = t)} =$$

$$\frac{P_\theta(X_1 = x_1, X_2 = x_2, X_1 + X_2 = t)}{P_\theta(X_1 + X_2 = t)} =$$

$t = x_1 + x_2$ sonst im Zähler Wahrscheinlichkeit 0

x_1, x_2 unabhängig

$$\frac{P_\theta(X_1 = x_1, X_2 = x_2)}{P_\theta(X_1 + X_2 = t)} = \frac{P_\theta(X_1 = x_1)\, P_\theta(X_2 = x_2)}{P_\theta(X_1 + X_2 = t)} =$$

$x_1 + x_2 = t$

$$\frac{\theta^2 (1-\theta)^{x_1 + x_2 - 2}}{\binom{t-1}{1} \theta^2 (1-\theta)^{t-2}} = \frac{\theta^2 (1-\theta)^{t-2}}{(t-1)\,\theta^2 (1-\theta)^{t-2}} = \frac{1}{t-1}$$

$\Rightarrow T = X_1 + X_2$ ist eine suffiziente Statistik