# Übungsblatt 10
# Stochastik 2

Abgabe von: **Linus Mußmächer**

23. Juni 2023

## 10.1 Zentralübung

Die Parameter der geometrischen Verteilung sind $\theta \in \Theta = (0,1)$, der Stichprobenraum ist $\mathcal{X} = \mathbb{N}^n$. Aus der Definition der geometrischen Verteilung bestimmen wir die Likelihood-Funktion und direkt daraus folgend die Log-Likelihood-Funktion:

$$L : \Theta \times \mathcal{X} \to [0,1],\ (\theta, x) \mapsto \prod_{k=1}^{n} (1-\theta)^{x_k - 1}\theta$$

$$\mathcal{L} : \Theta \times \mathcal{X} \to [0,1],\ (\theta, x) \mapsto \log\left(\prod_{k=1}^{n} (1-\theta)^{x_k - 1}\theta\right) = n \cdot \log\theta + \log(1-\theta) \cdot \sum_{k=1}^{n} (x_k - 1)$$

Wir bestimmen ihr Maximum durch die Nullstellen der Ableitung:

$$\frac{\partial}{\partial\theta}\mathcal{L}(\theta, x) = \frac{n}{\theta} - \frac{1}{1-\theta}\sum_{k=1}^{n}(x_k - 1) = \frac{n}{\theta} - \frac{1}{1-\theta}\left(-n + \sum_{k=1}^{n} x_k\right) = 0$$

$$\Leftrightarrow n \cdot (1-\theta) = \theta \sum_{k=1}^{n} x_k - n\theta \Leftrightarrow n = \theta \sum_{k=1}^{n} x_k \Leftrightarrow \theta = \frac{n}{\sum_{k=1}^{n} x_k} = \frac{1}{\overline{x}}$$

wobei $\overline{x}$ den Durschnitt von $x_1, \ldots, x_n$ bezeichne. Der (in diesem Fall aufgrund des an einer eindeutigen Stelle angenommen Maximums eindeutige) ML-Schätzer ist daher

$$\hat{\theta}^{ML} : \mathcal{X} \to \overline{\Theta}, x \mapsto \frac{1}{\overline{x}}.$$

## 10.2

(i) Es ist

$$\mathbb{E}_\theta[X] = \frac{2}{3}\theta \cdot 0 + \frac{1}{3}\theta \cdot 1 + \frac{2}{3}(1-\theta) \cdot 2 + \frac{1}{3}(1-\theta) \cdot 3 = 0 + \frac{1}{3}\theta + \frac{4}{3} - \frac{4}{3}\theta + 1 - \theta = \frac{7}{3} - 2\theta.$$

(ii) Uns ist der Erwartungswert (in Abhängigkeit von $\theta$) bereits bekannt. Wir können nun für eine Menge an Zufallsvariablen $X_1, \ldots, X_n$ die unabhängig identisch wie $X$ verteilt sind, den Momentenschätzer dieser Variablen unter Rückgriff auf den Durchschnitt $\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_n$ berechnen als

$$\overline{X} = \mathbb{E}_\theta[X] = \frac{7}{3} - 2\theta \Leftrightarrow \theta = \frac{7 - 3\overline{X}}{6}.$$

Insbesondere fällt auf, dass dieser für extreme Durchschnitte (z.B. $\overline{X} = 3$) auch Werte außerhalb von $\Theta = [0,1]$ annehmen kann.

(iii) Der Durchschnitt der Beobachtungen ist $\overline{X} = \frac{1}{10} \cdot 15 = \frac{3}{2}$. Damit ist der Wert des Momentenschätzers gleich $\frac{7-\frac{9}{2}}{6} = \frac{5}{12}$.

Um den Wert des ML-Schätzers für diese Beobachtungen zu bestimmen, müssen wir nicht die gesamte Funktion $\hat{\theta}^{ML}$ bestimmen, sondern lediglich die Wahrscheinlichkeit genau dieses Ereignisses bestimmen:

$$\mathbb{P}(3,0,2,1,3,2,1,0,2,1) = \left(\frac{2}{3}\theta\right)^2 \cdot \left(\frac{1}{3}\theta\right)^3 \cdot \left(\frac{2}{3}(1-\theta)\right)^3 \cdot \left(\frac{1}{3}(1-\theta)\right)^2 = \frac{32\theta^5(1-\theta^5)}{59049}.$$

Diese Funktion wird am Maximum von $\theta^5(1-\theta)^5$ maximiert, also bei $\theta = \frac{1}{2}$. Damit ist dies der Wert des Maximum-Likelihood-Schätzers.

Dies ist auch insofern plausibel, dass jedes Teilereignis $0, 1$ genauso oft auftritt wie das 'symmetrische' Ereignis $2, 3$ und insofern die Wahrscheinlichkeiten (vermutlich) gleich hoch sein sollten. Der Momentenschätzer, der nur den Durchschnitt 'sieht', kann dies allerdings nicht erkennen und tendiert daher zu niedrigerem $\theta$.

## 10.3

Wir bestimmen zuerst die Likelihood-Funktion der gemeinsamen Verteilung, indem wir die einzelnen Dichten als Funktionen in $y_k, \mu$ betrachten und multiplizieren:

$$L((y_1,\dots,y_n),\mu) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left(-\sum_{k=1}^n \frac{(\log(y_k)-\mu)^2}{2\sigma^2}\right) \prod_{k=1}^n \frac{1}{y_k}$$

Um unsere Rechnungen zu vereinfachen, betrachten wir die Log-Likelihood-Funktion

$$\mathcal{L}((y_1,\dots,y_n),\mu) = -\frac{n}{2} \cdot \log\bigl(2\pi\sigma^2\bigr) - \sum_{k=1}^n \frac{(\log(y_k)-\mu)^2}{2\sigma^2} - \sum_{k=1}^n \log(y_k)$$

und leiten für festes, aber beliebiges $y$ nach $\mu$ ab

$$\frac{\partial}{\partial\mu}\mathcal{L}((y_1,\dots,y_n),\mu) = \frac{1}{\sigma^2}\sum_{k=1}^n (\mu - \log y_k).$$

Damit folgt

$$\frac{\partial}{\partial\mu}\mathcal{L}((y_1,\dots,y_n),\mu) = 0 \Leftrightarrow n\mu = \sum_{k=1}^n \log y_k \Leftrightarrow \mu = \frac{1}{n}\sum_{k=1}^n \log y_k$$

und dies ist somit das $\mu$, das für festes $y$ die Log-Likelihood-Funktion und damit auch die Likelihood-Funktion maximiert. Der ML-Schätzer ist somit

$$\hat{\theta}^{ML} : \mathbb{R}_+^n \to \mathbb{R}, (y_1,\dots,y_n) \mapsto \frac{1}{n}\sum_{k=1}^n \log y_k$$

## 10.4

(i) Wir wollen das Maximum von

$$f(x,\nu) = \nu x_0^\nu x^{-(\nu+1)} \mathbb{1}_{x \geq x_0} = \nu(x_0 x^{-1})^\nu x^{-1} \mathbb{1}_{x \geq x_0}$$

für festes $x$, aber variables $\nu$ bestimmen. Sei dazu $x \geq x_0$ beliebig, aber fest. Wir leiten nach $\nu$ ab:

$$\frac{\partial}{\partial \nu} f(x,\nu) = 1 \cdot (x_0 x^{-1})^{\nu} x^{-1} + \nu \ln(x x_0^{-1})(x_0 x^{-1})^{\nu} x^{-1}$$

und damit

$$\begin{aligned}
&\frac{\partial}{\partial \nu} f(x,\nu) = 0 \\
\Leftrightarrow\ & (x_0 x^{-1})^{\nu} x^{-1} = -\nu ln(x_0 x^{-1})(x_0 x^{-1})^{\nu} x^{-1} \\
\Leftrightarrow\ & 1 = -\nu \ln(x_0 x^{-1}) \\
\Leftrightarrow\ & \nu = \frac{1}{\ln(x/x_0)}
\end{aligned}$$

und damit, da $\nu \in (1,\infty)$:

$$\hat{\theta}^{ML}(x) : \mathcal{X} \to [1,\infty), x \mapsto \max\left(1, \frac{1}{\ln(x) - \ln(x_0)}\right)$$

(ii) Für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ und $\nu \in \mathbb{R}$ ist die Likelihood-Funktion

$$f(x,\nu) = \begin{cases} 1 & \forall_{1 \leq k \leq n}\ x_k \in [\nu, \nu+1] \\ 0 & \exists_{1 \leq k \leq n}\ x_k \notin [\nu, \nu+1] \end{cases}$$

Setze $x_{(0)} = \min\{x_k \mid 1 \leq k \leq n\}$ und $x_{(n)} = \max\{x_k \mid 1 \leq k \leq n\}$.

Falls $x_{(n)} - x_{(0)} > 1$, so existiert kein $\nu$ für das $f(x,\nu) = 1$ ist. Da wir aber annehmen, dass unsere Datenpunkte wie angegeben verteilt sind, kann dies nicht auftreten.

Betrachten wir nun die Funktion

$$\hat{\theta}^{ML} : \{(x_1, \ldots, x_k) \in \mathbb{R}^n \mid x_{(n)} - x_{(0)} \leq 1\} \to \mathbb{R}, (x_1, \ldots, x_k) \mapsto x_{(0)}.$$

Sicher gilt $x_{(0)} \in [x_{(0)}, x_{(0)} + 1]$ und ebenso $x_{(n)} \in [x_{(0)}, x_{(0)} + 1]$, da $\max\{x_k\} - x_{(0)} \leq 1$. Damit gilt $x_k \in [x_{(0)}, x_{(0)} + 1]$ für alle $k$, also ist $f(x, \hat{\theta}^M L(x)) = 1$, was dem Supremum von $f$ entspricht (da $f$ ja nur zwei Werte annimmt).

Betrachten wir aber analog $\hat{\Theta}^{ML_2}(x) = x_{(n)} - 1$, so gilt $x_{(0)} \in [x_{(n)} - 1, x_{(n)}]$ aufgrund der Abstandbedingung und $x_{(n)} \in [x_{(n)} - 1, x_{(n)}]$. Wieder folgt daher $f(x, \hat{\Theta}^{ML_2}(x)) = 1$ für alle $x \in \mathcal{X}$ und auch $\hat{\Theta}^{ML_2}(x)$ ist ein ML-Schätzer.

Da insbesondere $x$ mit $x_{(0)} - x_{(n)} < 1$ existieren, sind diese beiden Schätzer außerdem unterschiedlich und "der" ML-Schätzer somit nicht eindeutig festgelegt. Man sollte also stets von "einem", nicht "dem", ML-Schätzer sprechen.