LÖSUNG

Übungen zu Diskrete Wahrscheinlichkeitstheorie

Dieses Blatt ist am 20. Juli bis 10:15 abzugeben und wird am 20./21. Juli besprochen.

Beachten Sie: Alle Ergebnisse sind zu begründen!

Aufgabe 10.1 — 1P+1P+2P

Ziel dieser Aufgabe ist es, den ML-Schätzer für die Erfolgsw'keit $p \in (0,1)$ einer negativ binomial-verteilten ZV zu bestimmen.

Hierfür seien $X_1, \ldots, X_n$ unabhängige ZV, wobei jede ZV X_i negativ binomial-verteilt mit Erfolgsw'keit p bei m zu erzielenden Erfolgen sei, d.h., jeder der ZV hat die Dichte

$$f(k) = \binom{k-1}{m-1} p^m (1-p)^{k-m} \text{ (mit } k \geq m).$$

Der Parameter m sei bekannt, zu schätzen ist nur p.

(a) Es sei $\vec{k} = (k_1, \ldots, k_n)$ ein Stichprobenvektor (mit $k_i \geq m$). Stellen Sie die Likelihood-Funktion $L(\vec{k}; p)$ auf.

(b) Maximimieren Sie $L(\vec{k}; p)$ und bestimmen Sie hiermit den ML-Schätzer für p.

(c) Zeigen Sie, dass der ML-Schätzer aus (b) i.A. nicht erwartungstreu ist.

Hinweis: Betrachten Sie den Fall $n = m = 1$ und verwenden Sie:

$$\ln(1+x) = \sum_{i=1}^{\infty} (-1)^{i+1} \frac{x^i}{i} \text{ für } x \in (-1, 1].$$

Lösungsvorschlag:

(a)

$$\begin{aligned}
L(\vec{k}; p) &= \prod_{i=1}^{n} f_p(k_i) \\
&= \prod_{i=1}^{n} \left(\binom{k_i - 1}{m-1} \left(\frac{p}{1-p} \right)^m (1-p)^{k_i} \right) \\
&= \left(\frac{p}{1-p} \right)^{mn} \cdot (1-p)^{\overbrace{\sum_{i=1}^{n} k_i}^{=:K}} \cdot \underbrace{\prod_{i=1}^{n} \binom{k_i - 1}{m-1}}_{=:C} \\
&= \left(\frac{p}{1-p} \right)^{mn} \cdot (1-p)^K \cdot C.
\end{aligned}$$

$$\begin{aligned}
\ln L(\vec{k}; p) &= mn(\ln p - \ln(1-p)) + K \ln(1-p) + \ln C \\
&= mn \ln p + (K - mn) \ln(1-p) + \ln C.
\end{aligned}$$

(b) Da ln streng monoton wachsend ist auf $(0,\infty)$, nehmen L und $\ln L$ im selben Punkt ihr Maximum an. Daher:

$$\frac{d}{dp}\ln L = mn\frac{1}{p} - (K-mn)\frac{1}{1-p} \stackrel{!}{=} 0 \Leftrightarrow mn(1-p) = (K-mn)p \Leftrightarrow p = \frac{mn}{K}.$$

$$\frac{d^2}{dp^2}\ln L = -mn\frac{1}{p^2} - (K-mn)\frac{1}{(1-p)^2} \stackrel{p=\frac{mn}{K}}{=} \frac{-K}{mn} - \frac{K^2}{K-mn} \stackrel{K\geq mn}{<} 0.$$

$$\lim_{p\searrow 0} L = 0 \text{ und } \lim_{p\nearrow 0} L = 0\,(\text{da } K \geq mn).$$

Da $K = \sum_{i=1}^n k_i$ und k_i ein Ausgang des Experiments X_i ist, erhält man den ML-Schätzer:

$$U = \frac{mn}{\sum_{i=1}^n X_i}.$$

(c) U ist erwartungstreu, falls $\mathbb{E}[U] = p$ gilt. Für $m = n = 1$ gilt $U = 1/X_1$ und $X_1 \sim \mathrm{Geo}(p)$:

$$\mathbb{E}[U] = \mathbb{E}[\frac{1}{X_1}] = \sum_{i=1}^{\infty}\frac{1}{i}p\cdot(1-p)^{i-1} = \frac{p}{1-p}\cdot\sum_{i=1}^{\infty}\frac{(1-p)^i}{i}.$$

Mit der Taylorreihe des ln folgt:

$$-\ln p = -\ln(1+(p-1)) = -\sum_{i=1}^{\infty}(-1)^{i+1}\frac{(p-1)^i}{i} = \sum_{i=1}^{\infty}\frac{(1-p)^i}{i}.$$

Also für $m = n = 1$:

$$\mathbb{E}[U] = p\cdot\frac{-\ln p}{1-p}$$

Für $p \in [0,1)$ gilt

$$\frac{-\ln p}{1-p} > 1\,.$$

Also ist U i.A. nicht erwartungstreu.

Aufgabe 10.2 **2P+1P+1P+2P**

Es seien $X_1, X_2, \ldots, X_n$ unabhängige und identisch verteilte ZV mit der gemeinsamen Dichte ($\gamma > 0$):

$$f(x) := \begin{cases} \frac{1}{2\gamma^3}x^2e^{-x/\gamma} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

(a) Überprüfen Sie, dass es sich bei $f(x)$ tatsächlich um eine Dichte handelt. Bestimmen Sie auch $\mathbb{E}[X_i]$.

(b) Es sei $\vec{k} = (k_1,\ldots,k_n)$ ein Stichprobenvektor. Stellen Sie die Likelihood-Funktion $L(\vec{k};\gamma)$ für den Parameter $\gamma > 0$ auf.

(c) Bestimmen Sie den ML-Schätzer für γ durch Maximieren von $L(\vec{k};\gamma)$.

(d) Ist der ML-Schätzer aus (c) erwartungstreu und konsistent im quadratischen Mittel?

Hinweis: Sie dürfen verwenden, dass $\mathrm{Var}[X_i] < \infty$ gilt.

Lösungsvorschlag:

(a) Zu zeigen ist, dass $\int_0^\infty f(x)dx = 1$.

$$\begin{aligned}\int_0^\infty x^2e^{-x/\gamma}dx &= \left[x^2\cdot(-\gamma e^{-x/\gamma})\right]_0^\infty + 2\gamma\int_0^\infty xe^{-x/\gamma}dx \\ &= 2\gamma\left(\left[x\cdot(-\gamma e^{-x/\gamma})\right]_0^\infty + \gamma\int_0^\infty e^{-x/\gamma}dx\right) \\ &= 2\gamma^2\cdot\left[-\gamma e^{-x/\gamma}\right]|_0^\infty \\ &= 2\gamma^3.\end{aligned}$$

$$\begin{aligned}\mathbb{E}[X_i] &= \frac{1}{2\gamma^3}\int_0^\infty x^3 e^{-x/\gamma}dx \\ &= \frac{1}{2\gamma^3}\left(\left[x^3\cdot(-\gamma e^{-x/\gamma})\right]_0^\infty + 3\gamma\int_0^\infty x^2 e^{-x/\gamma}dx\right) \\ &= 3\gamma\cdot\int_0^\infty f(x)dx \\ &= 3\gamma.\end{aligned}$$

(b)

$$\begin{aligned}L(\vec{k};\gamma) &= \prod_{i=1}^n\left(\frac{1}{2\gamma^3}k_i^2 e^{-k_i/\gamma}\right) \\ &= 2^{-n}\gamma^{-3n}\left(\prod_{i=1}^n k_i^2\right)e^{-\frac{1}{\gamma}\sum_{i=1}^n k_i} \\ &(C := 2^{-n}\textstyle\prod_{i=1}^n k_i^2,\ K := \sum_{i=1}^n k_i) \\ &= C\gamma^{-3n}e^{-\frac{K}{\gamma}}\end{aligned}$$

(c)

$$\begin{aligned}\frac{d}{d\gamma}\ln L(\vec{k};\gamma) &= \frac{d}{d\gamma}\left(\ln C - 3n\ln\gamma - \frac{K}{\gamma}\right) \\ &= \frac{-3n}{\gamma} + \frac{K}{\gamma^2} \\ &\overset{!}{=} 0.\end{aligned}$$

Somit $\gamma = \frac{K}{3n}$.

Wegen

$$\begin{aligned}\frac{d^2}{dp^2}\ln L(\vec{k};\gamma) &= \frac{d}{dp}\left(\frac{-3n}{\gamma} + \frac{K}{\gamma^2}\right) \\ &= \frac{3n}{\gamma^2} - 2\frac{K}{\gamma^3} \\ &\overset{\gamma=\frac{K}{3n}}{=} \frac{3n}{K^2/(3n)^2} - 2\frac{K}{K^3/(3n)^3} = -\frac{(3n)^3}{K^2} < 0.\end{aligned}$$

handelt es sich hierbei zumindest um ein lokales Maximum. Offensichtlich gilt auch $L \to 0$ sowohl für $\gamma \to \infty$ als auch für $\gamma \to 0$. Somit ist L in $\gamma = \frac{K}{3n}$ global maximal, und der ML-Schätzer ergibt sich zu

$$U = \frac{\sum_{i=1}^n X_i}{3n}.$$

(d)

$$\mathbb{E}[U] = \frac{1}{3}\mathbb{E}[X_i] = \frac{1}{3}3\gamma = \gamma.$$

Der ML-Schätzer ist also erwartungstreu.

$$\mathbb{E}[(U-\gamma)^2] = \mathbb{E}[(U-\mathbb{E}[U])] = \mathrm{Var}[U] = \frac{1}{9n^2}n\mathrm{Var}[X_i] = \frac{\mathrm{Var}[X_i]}{9n} \xrightarrow{n\to\infty} 0.$$

Damit strebt der MSE mit $n \to \infty$ gegen 0. Der Schätzer ist also auch konsistent im quadratischen Mittel.

Aufgabe 10.3 **1P+1P+1P+1P+1P**

Sie wissen von Blatt 1, dass Benjamins drei Geschwister ihn beschuldigen, dass er ungeschickter beim Abspülen sei als sie. Die Eltern wollen nun die Anschulding Benjamins durch die Geschwister statistisch testen. Dazu warten sie die nächsten 4 Abspültage, an denen Geschirr zerbrochen wird, ab. Sie stellen fest, dass Benjamin an 3 von diesen 4 Tagen die Scherben verursacht hat.

Sei für $i \in \{1,2,3,4\}$

$$X_i = \begin{cases} 1 & \text{wenn Benjamin am } i\text{-ten Abspültag mit Scherben schuld ist} \\ 0 & \text{wenn ein anderes Kind schuld ist}\,. \end{cases}$$

Es gelte $X_i \sim \mathrm{Bin}(1;p)$. Es wird ein Test für den Parameter p entworfen.

(a) Der Fehlerwahrscheinlichkeit 1. Art soll auf 0.05 begrenzt werden, und Benjamin soll nicht zu Unrecht des Ungeschicks bezichtigt werden. Wie muss die Nullhypothese gewählt werden? Begründung! Geben Sie H_0 als Teilmenge von $[0,1]$ an.

(b) Wählen Sie eine geeignete Testgröße T.
Hinweis: T soll von X_1, X_2, X_3, X_4 abhängen.

(c) Wählen Sie den Ablehnungsbereich so, dass H_0 abgelehnt wird und die Fehlerw'keit 1. Art minimiert wird.

(d) Berechnen Sie für den in (c) gewählten Ablehnungsbereich die Fehlerw'keit 1. Art.

(e) Was folgt daraus für die Frage, wie geschickt/ungeschickt Benjamin ist?

Lösungsvorschlag:

(a) Die Fehlerw'keit 1. Art wird begrenzt, das ist die W'keit, dass die Nullhypothese H_0 abgelehnt wird, obwohl sie gilt. Gleichzeitig ist die W'keit, dass Benjamin zu Unrecht beschuldigt wird, zu begrenzen. Folglich ist H_0, dass Benjamin höchstens so ungeschickt ist wie die Geschwister, also $H_0 = \{p \in [0,1] \mid p \le \frac{1}{4}\} = [0, \frac{1}{4}]$.

(b) $T = X_1 + X_2 + X_3 + X_4$

(c) Als Ablehnungsbereich K ist – bei Wahl der Testgröße wie in (b) – nur eine Menge der Form $\{a, a+1, \ldots, 4\}$ für ein $a \in \{1,2,3,4\}$ sinnvoll. Damit H_0 abgelehnt wird, muss $3 \in K$ gelten, folglich $a \in \{1,2,3\}$. Damit die Fehlerw'keit 1. Art minimiert wird, muss K möglichst klein sein, folglich wähle $a = 3$, also $K = \{3,4\}$.

(d) Für die Fehlerw'keit 1. Art gilt

$$\sup_{p \in H_0} \Pr_p[T \in K] = \sup_{0 \le p \le 1/4} \Pr_p[T \ge 3]\,.$$

Da $T \sim \mathrm{Bin}(4;p)$, wird das Supremum von $\Pr_p[T \ge 3]$ für $p = 1/4$ angenommen. Folglich ist die Fehlerw'keit 1. Art

$$\Pr[T \ge 3] \ \text{ wobei } T \sim \mathrm{Bin}(4; 1/4).$$

Diese W'keit wurde in Aufgabe 1.2 (b) und (c) berechnet zu $13/256 \approx 0.051$.

(e) Die Nullhypothese H_0 kann mit einer Signifikanz von ≈ 0.949 abgelehnt werden. Dieses Signifikanzniveau wird normalerweise als nicht ausreichend angesehen. Die Beweise für eine Verurteilung Benjamins reichen nicht aus!

Aufgabe 10.4 **2.5P+2.5P**

$X_1, \ldots, X_n$ seien unabhängige und identisch verteilt mit $X_i \sim \mathcal{N}(\mu, 1)$. Wir betrachten das Hypothesenpaar

$$H_0 \,:\, \mu \ge \mu_0 \text{ vs. } H_1 \,:\, \mu < \mu_1,$$

wobei $\mu_0 \ge \mu_1$ gelte.

Testgröße T sei das Stichprobenmittel, d.h.,

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

Da $\mathbb{E}[T] = \mu$ gilt, wählen wir als Ablehnungsbereich für H_0 das Intervall $K = (-\infty, a)$ mit $a \in \mathbb{R}$.

(a) Bestimmen Sie a so, dass die Fehlerw'keit 1. Art gerade α beträgt, d.h., wählen Sie a so, dass

$$\sup_{\mu \ge \mu_0} \Pr[T \in K] = \alpha.$$

Welchen Wert nimmt in diesem Fall die Fehlerw'keit 2. Art an? Betrachten Sie insbesondere den Fall $\mu_0 = \mu_1$.

(b) Bestimmen Sie a so, dass die Fehlerw'keit 2. Art gerade α beträgt, d.h., wählen Sie a so, dass

$$\sup_{\mu \le \mu_1} \Pr[T \notin K]\,.$$

Welchen Wert nimmt in diesem Fall die Fehlerw'keit 1. Art an? Betrachten Sie insbesondere den Fall $\mu_0 = \mu_1$.

Lösungsvorschlag:

(a)

$$\begin{aligned}
\sup_{\mu\geq\mu_0} \Pr[T\in K] &= \sup_{\mu\geq\mu_0} \Pr[T<a] \\
&= \sup_{\mu\geq\mu_0} \Pr\left[\frac{T-\mathbb{E}[T]}{\sqrt{\mathrm{Var}[T]}} < \sqrt{n}(a-\mu)\right] \\
&= \sup_{\mu\geq\mu_0} \Phi(\sqrt{n}(a-\mu))
\end{aligned}$$

$$(\Phi(\sqrt{n}(a-\mu)) \text{ fällt für wachsendes } \mu)$$

$$= \Phi(\sqrt{n}(a-\mu_0))$$

Fehler'wkeit 1. Art gleich α:

$$\Phi(\sqrt{n}(a-\mu_0)) = \alpha \Leftrightarrow \sqrt{n}(a-\mu_0) = z_\alpha \Leftrightarrow a = \mu_0 + \frac{z_\alpha}{\sqrt{n}}.$$

Dann ergibt sich für die Fehlerw'keit 2. Art (siehe (b)):

$$\Phi(\sqrt{n}(\mu_1 - (\mu_0 + \frac{z_\alpha}{\sqrt{n}}))) = \Phi(-z_\alpha - \sqrt{n}(\mu_0-\mu_1)) \leq \Phi(-z_\alpha) = 1-\alpha.$$

(b)

$$\begin{aligned}
\sup_{\mu\leq\mu_1} \Pr[T\notin K] &= \sup_{\mu\leq\mu_1} \Pr[T\geq a] \\
&= \sup_{\mu\leq\mu_1} \Pr\left[\frac{T-\mathbb{E}[T]}{\sqrt{\mathrm{Var}[T]}} \geq \sqrt{n}(a-\mu)\right] \\
&= \sup_{\mu\leq\mu_1} (1-\Phi(\sqrt{n}(a-\mu)))
\end{aligned}$$

$$(\Phi(\sqrt{n}(a-\mu)) \text{ fällt für wachsendes } \mu)$$

$$\begin{aligned}
&= 1-\Phi(\sqrt{n}(a-\mu_1)) \\
&= \Phi(\sqrt{n}(\mu_1-a))
\end{aligned}$$

Fehlerw'keit 2. Art gleich α:

$$\Phi(\sqrt{n}(\mu_1-a)) = \alpha \Leftrightarrow \sqrt{n}(\mu_1-a) = z_\alpha \Leftrightarrow a = \mu_1 - \frac{z_\alpha}{\sqrt{n}} = \mu_1 + \frac{z_{1-\alpha}}{\sqrt{n}}.$$

In diesem Fall beträgt die Fehlerw'keit 1. Art:

$$\Phi(\sqrt{n}(a-\mu_0)) = \Phi(\sqrt{n}((\mu_1 + \frac{z_{1-\alpha}}{\sqrt{n}}) - \mu_0)) = \Phi(z_{1-\alpha} - \sqrt{n}(\mu_0-\mu_1)) \leq \Phi(z_{1-\alpha}) = 1-\alpha.$$