Prof. Dr. Z. Kabluchko
Philipp Godland

Wintersemester 2016/17
14. November 2016

Mathematische Statistik

Aufgaben zum Üben

Keine Abgabe

Schätzer

Aufgabe 1

Es seien $X_1, \dots, X_n$ unabhängige und identisch verteilte Zufallsvariablen mit $X_i \sim \mathrm{N}(\mu, 1)$, wobei die Varianz $\sigma^2 = 1$ bekannt sei und über den Parameter μ zusätzlich bekannt sei, dass $\mu \geq 0$. Bestimmen Sie den Maximum–Likelihood–Schätzer für μ.

Hinweis: Eine Fallunterscheidung ist notwendig.

Aufgabe 2

Es seien $X_1, \dots, X_n$ unabhängige und identisch verteilte Zufallsvariablen mit der Dichte

$$h_\theta(x) = \begin{cases} \frac{2x}{\theta^2}, & \text{falls } x \in (0, \theta], \\ 0, & \text{sonst.} \end{cases}$$

Dabei sei $\theta > 0$ ein unbekannter Parameter. Sei $(x_1, \dots, x_n) \in (0, \infty)^n$ eine Realisierung von $(X_1, \dots, X_n)$.

(a) Stellen Sie die Likelihood-Funktion $L(x_1, \dots, x_n; \theta)$ auf.

(b) Schätzen Sie θ mit der Maximum-Likelihood-Methode.
Hinweis: Eine Skizze der Funktion $L(\theta)$ könnte hilfreich sein.

(c) Schätzen Sie θ mit der Momentenmethode.

Aufgabe 3

Es seien $X_1 \sim \mathrm{N}(\mu, \sigma_1^2), \dots, X_n \sim \mathrm{N}(\mu, \sigma_n^2)$ unabhängige Zufallsvariablen, wobei $\sigma_1^2, \dots, \sigma_n^2 > 0$ bekannt seien und μ der unbekannte Parameter sei.

(a) Bestimmen Sie den Maximum–Likelihood–Schätzer für μ.

(b) Zeigen Sie, dass die Statistik $T(x_1, \dots, x_n) := x_1/\sigma_1^2 + \dots + x_n/\sigma_n^2$ suffizient ist.

Aufgabe 4

Seien $X_1, \dots, X_n$ unabhängige, auf dem Intervall $[\theta, 2\theta]$ gleichverteilte Zufallsvariablen. Dabei sei $\theta > 0$ ein unbekannter Parameter.

(a) Bestimmen Sie den Momentenschätzer für θ.

(b) Zeigen Sie, dass $(X_{(1)}, X_{(n)})$ eine suffiziente Statistik ist.

(c) Zeigen Sie, dass $(X_{(1)}, X_{(n)})$ **keine** vollständige Statistik ist.

Aufgabe 5

Seien $X_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, m$, unabhängige Zufallsvariablen mit $X_{i,j} \sim \mathrm{N}(\mu_i, \sigma^2)$. Dabei seien $\mu_1 \in \mathbb{R}, \dots, \mu_n \in \mathbb{R}, \sigma^2 > 0$ unbekannt (alle Varianzen sind gleich). Schätzen Sie diese Parameter mit der Maximum–Likelihood–Methode.

Aufgabe 6

Bei einem gegebenen Parameter $\theta \in (0, 1)$ seien die Zufallsvariablen $X_1, \dots, X_n$ unabhängig und negativ binomialverteilt mit Parametern $r \in \mathbb{N}$ und θ:

$$\mathbb{P}_\theta[X_i = k] = \binom{k+r-1}{k} \theta^r (1-\theta)^k, \quad k = 0, 1, 2, \dots.$$

Für θ werde eine a-priori Betaverteilung mit der Dichte

$$q(\theta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}, \quad \theta \in (0, 1),$$

angenommen. Dabei seien $r \in \mathbb{N}$, $\alpha > 0$, $\beta > 0$ bekannte Parameter. Bestimmen Sie die a-posteriori-Verteilung von θ und den Bayes–Schätzer für θ.

Aufgabe 7

Seien $X_1, \dots, X_n \sim \mathrm{Bern}(\theta)$ unabhängige Zufallsvariablen, wobei über den Parameter θ zusätzlich bekannt sei, dass $\theta \in \Theta \subset (0, 1)$.

(a) Es gelte $|\Theta| \geq n+1$. Zeigen Sie, dass die Statistik $X_1 + \ldots + X_n$ vollständig ist und bestimmen Sie den besten erwartungstreuen Schätzer für θ.

(b) Es gelte $\Theta = \{\frac{1}{m+1} : m \in \mathbb{N}\}$. Untersuchen Sie, ob ein bester erwartungstreuer Schätzer für $\gamma(\theta) = \frac{1}{\theta}$ existiert.

Aufgabe 8

Es seien $X_1, \dots, X_n$ identisch $\mathrm{Poi}(2\lambda)$ und $Y_1, \dots, Y_m$ identisch $\mathrm{Poi}(3\lambda)$-verteilte Zufallsgrößen, die allesamt unabhängig voneinander seien. Der Parameter $\lambda > 0$ sei unbekannt.

(a) Beschreiben Sie die Situation durch ein geeignetes statistisches Experiment und bestimmen Sie den Maximum–Likelihood–Schätzer für λ.

(b) Bestimmen Sie einen besten erwartungstreuen Schätzer für λ. *Hinweis:* Die Vollständigkeit kann mit Hilfe der Definition gezeigt werden.

Aufgabe 9

Die Zufallsvariablen $X_1, \dots, X_n$ seien unabhängig und auf einem Intervall $[\theta_1, \theta_2]$ gleichverteilt, wobei $\theta_1, \theta_2 \in \mathbb{R}$ mit $\theta_1 < \theta_2$ unbekannte Parameter seien. Man beobachtet eine Realisierung $(x_1, \dots, x_n)$ von $(X_1, \dots, X_n)$. Es seien $x_{(1)} \leq \ldots \leq x_{(n)}$ die entsprechenden Ordnungsstatistiken.

(a) Zeigen Sie, dass die Statistik $(x_{(1)}, x_{(n)})$ suffizient ist.

(b) Zeigen Sie, dass die Statistik $(x_{(1)}, x_{(n)})$ vollständig ist.

(c) Finden Sie die besten erwartungstreuen Schätzer für die Länge $l := \theta_2 - \theta_1$ und den Mittelpunkt $m := \frac{\theta_1+\theta_2}{2}$.

Hinweis zu Teil b: Es handelt sich **nicht** um eine Exponentialfamilie. Eine Statistik T ist vollständig, wenn es nicht möglich ist, einen erwartungstreuen Schätzer von 0 der Form $h(T)$ zu konstruieren (es sei denn $h(T) = 0$ $\mathbb{P}_{\theta_1,\theta_2}$-f.s. für alle θ_1, θ_2).

Aufgabe 10

Seien $X_1, X_2 \sim \text{Exp}(\lambda)$ unabhängig. Zeigen Sie, dass $X_1 + X_2$ und X_1/X_2 unabhängig sind.

Aufgabe 11

Die Zufallsvariablen $X_1, \ldots, X_n$ seien unabhängig und identisch verteilt mit $X_i \sim \text{Bin}(m, \theta)$, wobei $m \in \mathbb{N}$ **bekannt** sei und $\theta \in (0, 1)$ der zu schätzende **unbekannte** Parameter ist. Sei $(x_1, \ldots, x_n)$ mit $x_i \in \{0, 1, \ldots, m\}$ eine Realisierung von $(X_1, \ldots, X_n)$.

(a) Zeigen Sie, dass $T(x_1, \ldots, x_n) = x_1 + \ldots + x_n$ eine suffiziente Statistik ist.

(b) Bestimmen Sie die Fisher–Information $I(\theta)$.

(c) Geben Sie einen Cramér–Rao–effizienten Schätzer für θ an.

Hinweis: Ein Cramér–Rao–effizienter Schätzer muss per Definition erwartungstreu sein. Formeln für Erwartungswert und Varianz der Binomialverteilung können als bekannt vorausgesetzt werden.

Aufgabe 12

Seien $X_1, \ldots, X_n$, wobei $n \geq 2$, unabhängige und mit Parameter $\theta \in (0, 1)$ Bernoulli–verteilte Zufallsvariablen. Sei $\bar{X}_n = \frac{X_1+\ldots+X_n}{n}$.

(a) Zeigen Sie, dass es **keinen** erwartungstreuen Schätzer für e^θ gibt (für alle n).

(b) Zeigen Sie, dass $\bar{X}_n^2$ **kein** erwartungstreuer Schätzer für θ^2 ist.

(c) Konstruieren Sie den besten erwartungstreuen Schätzer für θ^2.

Aufgabe 13

Es sei $(x_1, \ldots, x_n)$ eine Realisierung von unabhängigen und identisch verteilten Zufallsvariablen $(X_1, \ldots, X_n)$ mit Zähldichte

$$\mathbb{P}_\theta[X_i = t] = \begin{cases} 1 - 2\theta, & \text{falls } t = 0, \\ \theta, & \text{falls } t = 1 \text{ oder } t = 2, \\ 0, & \text{sonst.} \end{cases}$$

Dabei sei $\theta \in (0, 1/2)$ ein unbekannter Parameter.

(a) Bestimmen Sie $\mathbb{E}_\theta X_1$ und $\text{Var}_\theta X_1$.

(b) Zeigen Sie, dass

$$T = \frac{1}{3} \cdot \frac{X_1 + \ldots + X_n}{n} = \frac{1}{3} \cdot \bar{X}_n$$

ein erwartungstreuer, stark konsistenter und asymptotisch normalverteilter Schätzer für θ ist.

(c) Konstruieren Sie, basierend auf T, ein symmetrisches asymptotisches Konfidenzintervall für θ zum Niveau $1-\alpha$, wobei $\alpha \in (0,1)$ vorgegeben sei. Sie können (müssen aber nicht) den Satz von Slutsky benutzen, um Ihre Berechnungen zu vereinfachen.

Aufgabe 14

Betrachten Sie das lineare Modell $Y_i = \alpha + \beta x_i + \varepsilon_i$, $i = 1, \dots, n$, wobei $\varepsilon_1, \dots, \varepsilon_n \sim \mathrm{N}(0, \sigma^2)$ unabhängig sind. Dabei seien $x_1, \dots, x_n \in \mathbb{R}$ bekannt. Es seien $\hat{\alpha}$ und $\hat{\beta}$ die Kleinste–Quadrate–Schätzer für α und β. Sei außerdem

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{\alpha} - \hat{\beta} x_i)^2.$$

Zeigen Sie, dass die Statistik $(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2)$ suffizient ist.

Aufgabe 15

Betrachten Sie das lineare Modell $Y = X\beta + \varepsilon$, wobei Y ein n-dimensionaler Zufallsvektor ist, X eine bekannte $n \times m$–Matrix, und $(\varepsilon_1, \dots, \varepsilon_n) \sim N(0, \sigma^2)$ unabhängige Zufallsvariablen. Unbekannte Parameter sind in diesem Modell $\beta \in \mathbb{R}^m$ und $\sigma^2 > 0$. Das Modell sei von vollem Rang, d.h. $\mathrm{Ker} X = \{0\}$.

(a) Stellen Sie die Likelihood–Funktion $L(\beta, \sigma^2; y)$ (also die Dichte von Y unter $\mathbb{P}_{\beta,\sigma^2}$) auf und bestimmen Sie den Maximum–Likelihood Schätzer $(\hat{\beta}, \hat{\sigma}^2)$.

(b) Zeigen Sie, dass dieses Modell eine $(m+1)$-parametrige Exponentialfamilie bildet, d.h. es gibt eine Darstellung

$$L(\beta, \sigma^2; y) = a(\beta, \sigma^2) b(y) e^{c_1(\beta,\sigma^2) d_1(y) + \dots + c_{m+1}(\beta,\sigma^2) d_{m+1}(y)}.$$

(c) Zeigen Sie, dass die Statistik $T(y) = (X^\top y, |y|^2)$ suffizient ist.

(d) Zeigen Sie, dass die Statistik $(\hat{\beta}, \hat{\sigma}^2)$ suffizient ist.

Tests und Konfidenzintervalle

Aufgabe 16

Ein Spielwürfel wurde 180 Mal geworfen und zeigte dabei 15 Mal eine 1.

Testen Sie die Nullhypothese H_0: “Die Wahrscheinlichkeit von 1 ist 1/6” gegen die Alternativhypothese H_1: “Die Wahrscheinlichkeit von 1 ist ungleich 1/6”. Als Niveau sei $\alpha = 0.05$ vorgegeben.

Hinweis: Sie können einen asymptotischen Test benutzen. Sie können (müssen aber nicht) den Satz von Slutsky benutzen, um Ihre Berechnungen zu vereinfachen. Quantile der Standardnormalverteilung: $z_{0.925} = 1.4395$, $z_{0.95} = 1.6449$, $z_{0.975} = 1.960$.

Aufgabe 17

Die Zufallsvariablen $X_1, \dots, X_n \sim \mathrm{N}(0, \sigma^2)$ seien unabhängig und normalverteilt mit Erwartungswert 0 und Varianz σ^2. Dabei sei $\sigma^2 > 0$ ein unbekannter Parameter.

(a) Zeigen Sie, dass

$$T := \frac{X_1^2 + \ldots + X_n^2}{n}$$

ein erwartungstreuer Schätzer für σ^2 ist.

(b) Bestimmen Sie die Verteilung von

$$\frac{nT}{\sigma^2}.$$

(c) Konstruieren Sie ein *einseitiges* Konfidenzintervall der Form $[0, \bar{\theta}]$ für σ^2 zum Niveau $1-\alpha$, wobei $\alpha \in (0,1)$ vorgegeben ist. In dieser Teilaufgabe ist somit die Größe $\bar{\theta}$ zu bestimmen.

Aufgabe 18

Seien $X_1, \ldots, X_n \sim \mathrm{N}(\mu, \sigma^2)$ unabhängig, wobei sowohl μ als auch σ^2 unbekannt seien.

(a) Konstruieren Sie ein einseitiges Konfidenzintervall der Form $(-\infty, A)$ zum Niveau $1-\alpha$ für μ.

(b) Konstruieren Sie ein einseitiges Konfidenzintervall der Form $[0, B)$ zum Niveau $1-\alpha$ für σ^2.

Aufgabe 19

Es seien $X_1, \ldots, X_n$ unabhängige Zufallsvariablen mit Dichten der Form

$$h_\lambda(t) = \frac{\lambda}{2} e^{-\lambda|t|}, \quad \lambda > 0, t \in \mathbb{R}.$$

(a) Bestimmen Sie für $\lambda_0 > 0$ einen gleichmäßig besten Test zum Niveau $\alpha \in (0,1)$ für $H_0 : \lambda \leq \lambda_0$ gegen $H_1 : \lambda > \lambda_0$.

(b) Zeigen Sie, dass

$$2\lambda \sum_{i=1}^{n} |X_i| \sim \chi^2_{2n}.$$

(c) Bei einer Studie wird bei $n = 50$ Beobachtungen ein Wert von $\frac{1}{n}\sum_{i=1}^{n} |x_i| = 87.5$ gemessen. Lässt sich die Nullhypothese $\lambda \leq 0.01$ zum Niveau $\alpha = 0.1$ verwerfen?

Hinweis: Verwenden Sie die angegebene Vertafelung von Quantilen der χ^2_{100}-Verteilung.

γ	0,900	0,750	0,500	0,250	0,100
$\chi^2_{100,\gamma}$-Quantil	118,5	109,1	99,33	90,13	82,36

Aufgabe 20

Seien $X_1, \ldots, X_n \sim \mathrm{N}(\mu, 1)$ unabhängig. Bestimmen Sie den gleichmäßig besten Test zum Niveau α für $H_0 : \mu = 0$ gegen $H_1 : \mu = 1$.

Aufgabe 21

Eine Brauerei besitzt eine Abfüllanlage für Halbliterflaschen. Die in eine Flasche abgefüllte Menge in Liter sei dabei $\mathrm{N}(\mu, \sigma^2)$-verteilt mit unbekanntem μ und $\sigma^2 = 0.0009$.

(a) Die Brauerei möchte aus naheliegenden Gründen nicht zuviel Bier in die Flaschen abfüllen. Eine Testreihe, bei der die Füllmengen von 200 Flaschen untersucht wurden, ergab

$$\bar{x}_{200} = 0.5051$$

Geben Sie einen gleichmäßig besten Test zum Niveau $\alpha = 0.01$ für die Hypothese $H_0 : \mu \leq 0.5$ an.

(b) Eine Verbraucherorganisation führt alternativ einen Test über die Füllmengen der Halbliterflaschen der Brauerei durch. Sie untersuchte ebenfalls 200 Flaschen und erhielt

$$\bar{x}_{200} = 0.49486$$

Zu große Füllmengen stören die Verbraucherorganisation natürlich nicht. Testen Sie (mit einem gleichmäßig besten Test) zum Niveau $\alpha = 0.01$ die Hypothese $H_0 : \mu \geq 0.5$.

Hinweis: Verwenden Sie die folgende Tabelle der Verteilungsfunktion $\Phi(x)$ der Standard–Normalverteilung. Tabelliert sind die Werte $\Phi(x + \Delta)$.

.	Δ									
x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500	0.504	0.508	0.512	0.516	0.520	0.524	0.528	0.532	0.536
0.1	0.540	0.544	0.548	0.552	0.556	0.560	0.564	0.567	0.571	0.575
0.2	0.579	0.583	0.587	0.591	0.595	0.599	0.603	0.606	0.610	0.614
0.3	0.618	0.622	0.626	0.629	0.633	0.637	0.641	0.644	0.648	0.652
0.4	0.655	0.659	0.663	0.666	0.670	0.674	0.677	0.681	0.684	0.688
0.5	0.691	0.695	0.698	0.702	0.705	0.709	0.712	0.716	0.719	0.722
0.6	0.726	0.729	0.732	0.736	0.739	0.742	0.745	0.749	0.752	0.755
0.7	0.758	0.761	0.764	0.767	0.770	0.773	0.776	0.779	0.782	0.785
0.8	0.788	0.791	0.794	0.797	0.800	0.802	0.805	0.808	0.811	0.813
0.9	0.816	0.819	0.821	0.824	0.826	0.829	0.831	0.834	0.836	0.839
1.0	0.841	0.844	0.846	0.848	0.851	0.853	0.855	0.858	0.860	0.862
1.1	0.864	0.867	0.869	0.871	0.873	0.875	0.877	0.879	0.881	0.883
1.2	0.885	0.887	0.889	0.891	0.893	0.894	0.896	0.898	0.900	0.901
1.3	0.903	0.905	0.907	0.908	0.910	0.911	0.913	0.915	0.916	0.918
1.4	0.919	0.921	0.922	0.924	0.925	0.926	0.928	0.929	0.931	0.932
1.5	0.933	0.934	0.936	0.937	0.938	0.939	0.941	0.942	0.943	0.944
1.6	0.945	0.946	0.947	0.948	0.949	0.951	0.952	0.953	0.954	0.954
1.7	0.955	0.956	0.957	0.958	0.959	0.960	0.961	0.962	0.962	0.963
1.8	0.964	0.965	0.966	0.966	0.967	0.968	0.969	0.969	0.970	0.971
1.9	0.971	0.972	0.973	0.973	0.974	0.974	0.975	0.976	0.976	0.977
2.0	0.977	0.978	0.978	0.979	0.979	0.980	0.980	0.981	0.981	0.982
2.1	0.982	0.983	0.983	0.983	0.984	0.984	0.985	0.985	0.985	0.986
2.2	0.986	0.986	0.987	0.987	0.987	0.988	0.988	0.988	0.989	0.989
2.3	0.989	0.990	0.990	0.990	0.990	0.991	0.991	0.991	0.991	0.992
2.4	0.992	0.992	0.992	0.992	0.993	0.993	0.993	0.993	0.993	0.994
2.5	0.994	0.994	0.994	0.994	0.994	0.995	0.995	0.995	0.995	0.995
2.6	0.995	0.995	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996
2.7	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997
2.8	0.997	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998
2.9	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.999	0.999	0.999

Aufgabe 22

Betrachten Sie ein statistisches Modell $(\mathfrak{X}, \mathcal{A}, \{\mathbb{P}_0, \mathbb{P}_1\})$, wobei $\mathbb{P}_0$ und $\mathbb{P}_1$ die Dichten h_0 und h_1 bzgl. eines σ-endlichen Maßes λ auf $\mathfrak{X}$ besitzen. Ein Test φ für $H_0 = \{0\}$ gegen $H_1 = \{1\}$ heißt ein Minimax–Test, wenn das Maximum der Irrtumswahrscheinlichkeiten erster und zweiter Art minimal unter allen Tests ist. Zeigen Sie:

a) Es gibt einen Likelihood–Quotienten Test φ mit $\mathbb{E}_0[\varphi] = \mathbb{E}_1[1 - \varphi]$.

b) Der Test aus (a) ist ein Minimax–Test.

Aufgabe 23

Bei einer radioaktiven Substanz handelt sich entweder um Stoff A oder um Stoff B. Um zu bestimmen, um welchen Stoff es sich handelt, hat man die Anzahl der Zerfälle X innerhalb von 10 Minuten gemessen. Bei Stoff A geht man davon aus, dass $X \sim \text{Poi}(1)$, bei Stoff B geht man davon aus, dass $X \sim \text{Poi}(10)$.

(a) Bestimmen Sie den besten randomisierten Test zum Niveau $\alpha = 0.05$ für die Nullhypothese H_0 : "Es handelt sich um Stoff A" gegen H_1 : "Es handelt sich um Stoff B".

(b) Wie entscheiden Sie sich, wenn Sie $n = 4$ Zerfälle messen?

Hinweis: Sie können verwenden, dass $0.36 < 1/e < 0.37$.

Aufgabe 24

Seien X_1, X_2, X_3 drei unabhängige und auf dem Intervall $[0, \theta]$ gleichverteilte Zufallsvariablen. Bestimmen Sie einen besten (randomisierten!) Test zum Niveau $\alpha = 0.05$ für die Hypothese $H_0 : \theta = 2$ gegen die Hypothese $H_1 : \theta = 1$.