Vorlesung *Mathematatische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

1. Übungsblatt

1. Beweisen Sie den *Korrespondenzsatz* zwischen Tests und Konfidenzbereichen: Sind für jeden Parameter $\vartheta_0 \in \Theta$ Tests φ_{ϑ_0} auf die Hypothese $H_0 : \vartheta = \vartheta_0$ (gegen irgendeine Alternative) vom Niveau $\alpha \in (0,1)$ gegeben, so bildet

$$C := \left\{ \vartheta_0 \in \Theta \,\middle|\, \varphi_{\vartheta_0} \text{ akzeptiert } H_0 : \vartheta = \vartheta_0 \right\}$$

eine $(1-\alpha)$-Konfidenzmenge. Andersherum kann aus einer $(1-\alpha)$-Konfidenzmenge eine Familie solcher Tests (φ_{ϑ_0}) konstruiert werden (wie?).

2. Es seien X und Y rellwertige Zufallsvariablen mit $\mathbb{E}[|Y|] < \infty$ und gemeinsamer Lebesgue-Dichte $f^{X,Y}$. Als *bedingte Dichte von Y gegeben $X = x$* definiert man

$$f^{Y|X=x}(y) := \frac{f^{X,Y}(x,y)}{\int_{\mathbb{R}} f^{X,Y}(x,\eta)\, d\eta}.$$

   (a) Zeigen Sie, dass $f^{Y|X=x}$ für P^X-fast alle x wohl definiert ist.

   (b) Betrachten Sie die Funktion $g(x) = \int_{\mathbb{R}} y f^{Y|X=x}(y)\, dy$, falls die rechte Seite wohldefiniert ist, und $g(x) = 0$ andernfalls. Weisen Sie nach, dass $g(X)$ eine Version der bedingten Erwartung $\mathbb{E}[Y|X]$ ist.

   (c) X und Y mögen gemeinsam normalverteilt sein mit Erwartungswert $\mu \in \mathbb{R}^2$ und Kovarianzmatrix $\Sigma \in \mathbb{R}^{2\times 2}$. Bestimmen Sie $\mathbb{E}[Y|X]$.

3. Für X und Y wie in Aufgabe 2 ist $m(x)$ ein *bedingter Median von Y gegeben $X = x$*, sofern

$$\int_{-\infty}^{m(x)} f^{Y|X=x}(y)\, dy = 1/2 \text{ (falls wohldefiniert)}$$

gilt. Zeigen Sie, dass m die Minimalitätseigenschaft

$$\mathbb{E}[|Y - m(X)|] = \inf_h \mathbb{E}[|Y - h(X)|]$$

besitzt, wobei sich das Infimum über alle Borel-messbaren $h : \mathbb{R} \to \mathbb{R}$ erstreckt.

Welches Kriterium minimieren die entsprechend definierten bedingten α-Quantile, $\alpha \in (0,1)$?

4. Bei acht Absolventen werden anhand einer Befragung die Studiendauer und das Einstiegsgehalt (in 1000€) ermittelt:

Studiendauer x_i	10	9	11	9	11	12	10	11
Einstiegsgehalt Y_i	35	35	34	36	41	39	40	38

   (a) Modellieren Sie dies als ein lineares Modell und bestimmen Sie die Regressionsgerade. Zeichnen Sie Daten und Regressionsgerade in ein geeignetes Koordinatensystem ein.

   (b) Es stellt sich heraus, dass die ersten Vier ein anderes Fach studiert haben als die anderen Vier. Bestimmen und zeichnen Sie die Regressionsgeraden für beide Studienfächer getrennt.

   (c) Wie erklären Sie die unterschiedlichen Ergebnisse in (a) und (b)?

Abgabe vor der Vorlesung am Freitag, dem 24.4.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

2. Übungsblatt

1. Es sei $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$ ein statistisches Experiment sowie π eine a-priori-Verteilung auf $(\Theta, \mathcal{F}_\Theta)$, so dass $\mathbb{P}_\vartheta \ll \mu$ für alle $\vartheta \in \Theta$ sowie $\pi \ll \nu$ gilt mit σ-endlichen Maßen μ und ν und Dichten $f_{X|\boldsymbol{\Theta}=\vartheta}$ bzw. $f_{\boldsymbol{\Theta}}$. Zeigen Sie für $(\mathcal{F}\otimes\mathcal{F}_\Theta)$-messbare Funktionen $(x,\vartheta)\mapsto f_{X|\boldsymbol{\Theta}=\vartheta}(x)\in[0,\infty)$:

   (a) Für $f_X(x) := \int_\Theta f_{X|\boldsymbol{\Theta}=\vartheta}(x) f_{\boldsymbol{\Theta}}(\vartheta)\nu(\mathrm{d}\vartheta)$ in $(0,\infty)$ definiere $f_{\boldsymbol{\Theta}|X=x}(\vartheta)$ durch
   $$f_{\boldsymbol{\Theta}|X=x}(\vartheta) := \frac{f_{X|\boldsymbol{\Theta}=\vartheta}(x) f_{\boldsymbol{\Theta}}(\vartheta)}{f_X(x)}, \qquad \vartheta\in\Theta,$$
   und sonst durch $f_{\boldsymbol{\Theta}|X=x}(\vartheta) := f_{\boldsymbol{\Theta}}(\vartheta)$, $\vartheta\in\Theta$, dann ist $f_{\boldsymbol{\Theta}|X=x}$ eine ν-Wahrscheinlichkeitsdichte für alle $x\in\mathcal{X}$.

   (b) Es seien X und $\boldsymbol{\Theta}$ Zufallsvariablen mit der gemeinsamen Verteilung $\tilde{\mathbb{P}}(\mathrm{d}x,\mathrm{d}\vartheta) = \mathbb{P}_\vartheta(\mathrm{d}x)\pi(\mathrm{d}\vartheta)$. Die Funktion $g:\mathcal{X}\times\Theta\to\mathbb{R}$ sei $(\mathcal{F}\otimes\mathcal{F}_\Theta)$-messbar. Ist g nichtnegativ oder ist $\mathbb{E}_{\tilde{\mathbb{P}}}[|g(X,\boldsymbol{\Theta})|]<\infty$, dann gilt
   $$\mathbb{E}_{\tilde{\mathbb{P}}}[g(X,\boldsymbol{\Theta})|X=x] = \int_\Theta g(x,\vartheta) f_{\boldsymbol{\Theta}|X=x}(\vartheta)\nu(\mathrm{d}\vartheta)$$
   und speziell $\int_A f_{\boldsymbol{\Theta}|X=x}(\vartheta)\nu(\mathrm{d}\vartheta) = \mathbb{E}_{\tilde{\mathbb{P}}}[\mathbf{1}_{\{\boldsymbol{\Theta}\in A\}}|X=x] =: \tilde{\mathbb{P}}(\boldsymbol{\Theta}\in A|X=x)$.

   (c) Bestimmen Sie $f_{X|\boldsymbol{\Theta}=\bullet}$ sowie $\mathbb{E}_{\tilde{\mathbb{P}}}[(X-\boldsymbol{\Theta})^2\,|\,X=x]$ im Fall $\Theta=\mathbb{R}$, $P_\vartheta = N(\vartheta,1)$, $\pi = N(0,\delta^2)$ für festes $\delta>0$. Was ergibt sich für $\delta\to\infty$ und $\delta\to 0$?

2. Wenn man in die Bayesformel statt einer Dichte $f_{\boldsymbol{\Theta}}$ eine nichtnegative, messbare Funktion $f_{\boldsymbol{\Theta}}$ einsetzt und $f_{\boldsymbol{\Theta}|X=x}(\vartheta)$ weiterhin wohldefiniert ist, so ergibt sich aus der a-posteriori-Verteilung ein *verallgemeinerter Bayesschätzer*. Es sei nun $X_1,\dots,X_n$ eine $N(\mu, E_d)$-verteilte mathematische Stichprobe mit $\mu\in\mathbb{R}^d$ unbekannt.

   (a) Beweisen Sie: $\bar{X} := \frac{1}{n}\sum_{i=1}^n X_i$ ist ein verallgemeinerter Bayesschätzer von μ zum quadratischen Risiko bzgl. dem Lebesguemaß als verallgemeinerter a-priori-Verteilung.

   (b) Berechnen Sie den verallgemeinerten Bayesschätzer $\hat{\mu}_{a,b}$ zum quadratischen Risiko für $d=1$ und $f_{\boldsymbol{\Theta}}(\vartheta) = \mathbf{1}_{(a,b)}(\vartheta)$ mit $a,b\in\mathbb{R}\cup\{-\infty,\infty\}$. Zeichnen Sie $\hat{\mu}_{0,1}$ für $n=1$ als Funktion von $\bar{X}$.

3. Gegeben sei das gewöhnliche lineare Modell $Y = X\beta + \varepsilon$. In der *ridge regression* verwendet man den Schätzer $\hat{\beta}_a = (X^\top X + a^2 E_p)^{-1} X^\top Y$. Weisen Sie nach: Der ridge-regression-Schätzer $\hat{\beta}_a$ ist Bayes-optimaler Schätzer für quadratisches Risiko bei a-priori-Verteilung $\beta \sim N(0, \eta^2 E_p)$ mit $\eta = \eta(a, \sigma) \geqslant 0$ geeignet.

4. Beweisen Sie für Entscheidungsregeln ρ basierend auf einem statistischen Experiment $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta \in \Theta})$ mit Verlustfunktion l:

   (a) Ist ρ minimax und eindeutig in dem Sinn, dass jede andere Minimax-Regel die gleiche Risikofunktion besitzt, so ist ρ zulässig.

   (b) Ist ρ zulässig mit konstanter Risikofunktion, so ist ρ minimax.

   (c) Ist ρ eine Bayesregel (bzgl. π) und eindeutig in dem Sinn, dass jede andere Bayesregel (bzgl. π) die gleiche Risikofunktion besitzt, so ist ρ zulässig.

   (d) Die Parametermenge Θ bilde einen metrischen Raum mit Borel-σ-Algebra $\mathcal{F}_\Theta$. Ist ρ eine Bayesregel (bzgl. π), so ist ρ zulässig, falls (i) $R_\pi(\rho) < \infty$; (ii) für jede nichtleere offene Menge U in Θ gilt $\pi(U) > 0$; (iii) für jede Regel ρ' mit $R_\pi(\rho') \leq R_\pi(\rho)$ ist $\vartheta \mapsto R(\vartheta, \rho')$ stetig.

5. Eine Krankheit kommt bei ca. $0,1\%$ der Bevölkerung vor. Ein Test zur Erkennung der Krankheit führt bei 97% der Kranken, aber auch bei 2% der Gesunden zu einer Reaktion. Auf Grund des Tests wird eine Person als krank bzw. gesund klassifiziert. Mit $\ell_0 \geqslant 0$ (bzw. $\ell_1 \geqslant 0$) werde der Verlust bei der Klassifizierung *krank* (bzw. *gesund*) eines gesunden (bzw. kranken) Patienten bewertet. Formulieren Sie dies als Bayessches Entscheidungsproblem und geben Sie eine Bayes-optimale Entscheidungsregel in Abhängigkeit von ℓ_0, ℓ_1 an.

6. Die Beta-Verteilung $B(a, b)$ auf $[0, 1]$ ist gegeben durch die Dichte

$$f_{a,b}(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \quad x \in (0, 1),$$

   wobei $a, b > 0$ und Γ die Gamma-Funktion bezeichnet. $B(a, b)$ hat Erwartungswert $\mu_{a,b} = \frac{a}{a+b}$ und Varianz $\sigma^2_{a,b} = \frac{ab}{(a+b)^2(a+b+1)}$.

   (a) Skizzieren Sie $f_{a,b}$ für $(a, b) \in \{0.5; 1; 10\}^2$ (Computereinsatz gestattet).

   (b) Es sei eine $\mathrm{Bin}(n, p)$-verteilte math. Stichprobe X gegeben, wobei $n \geqslant 1$ bekannt ist sowie p gemäß $B(a, b)$ a priori verteilt ist. Zeigen Sie, dass die bedingte Dichte von p gegeben $X = x$ zur Beta-Verteilung $B(a + x, b + n - x)$ gehört.

   (c) Schließen Sie, dass der Bayesschätzer unter quadratischem Risiko gegeben ist durch $\hat{p}_{a,b} = \frac{a+X}{a+b+n}$. Bestimmen Sie sein quadratisches Risiko als Funktion von p und sein zugehöriges Bayesrisiko.

Abgabe vor der Vorlesung am Freitag, dem 8.5.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

3. Übungsblatt

1. $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$, $\Theta \subseteq \mathbb{R}^d$, sei ein statistischen Experiment mit a-priori-Verteilung π und ρ sei eine Bayesregel (bzgl. π) zum quadratischen Risiko (d.h. $l(\vartheta, a) = |a-\vartheta|^2$). Zeigen Sie: ρ kann nur dann erwartungstreu sein, wenn $R_\pi(\rho) = 0$.

2. Es sei $X_1, \ldots, X_n$ eine $N(\mu, E_d)$-verteilte mathematische Stichprobe. Der James-Stein-Schätzer mit positivem Gewicht ist definiert als $\hat{\mu}_{JS+} = \big(1 - \frac{d-2}{n|\overline{X}|^2}\big)_+ \overline{X}$. Beweisen Sie für alle $d \geqslant 3$ und $\mu \in \mathbb{R}^d$ schrittweise folgenden Risikovergleich mit dem klassischen James-Stein-Schätzer:
$$\mathbb{E}_\mu[|\hat{\mu}_{JS+} - \mu|^2] < \mathbb{E}_\mu[|\hat{\mu}_{JS} - \mu|^2].$$
   (a) Die Abschätzung ist korrekt für $\mu = 0$.
   (b) Die Abschätzung folgt aus der Ungleichung $\mathbb{E}_\mu[\mu_i \overline{X}_i |G| \mathbf{1}_{\{G\leqslant 0\}}] > 0$ für $G = 1 - \frac{d-2}{n|\overline{X}|^2}$ und alle $i = 1, \ldots, d$ mit $\mu_i \neq 0$.
   (c) Für $a > 0$ und $\mu_i \neq 0$ gilt $\mathbb{E}_\mu[\mu_i \overline{X}_i \,|\, (\overline{X}_i)^2 = a^2] = a\mu_i \tanh(na\mu_i) > 0$. Dies ergibt die Ungleichung in (b) durch Einfügen einer auf $((\overline{X}_1)^2, \ldots, (\overline{X}_d)^2)$ bedingten Erwartung.

3. Gegeben sei $X \sim N(\mu, \sigma^2 E_d)$ mit $\sigma > 0$ bekannt und $\mu \in \mathbb{R}^d$ unbekannt.
   (a) Zeigen Sie: Soll in einem statistischen Experiment $g(\vartheta) \in \mathbb{R}^d$ durch $\hat{g}$ geschätzt werden, so gilt die *Bias-Varianz-Zerlegung*:
$$\mathbb{E}_\vartheta[|\hat{g} - g(\vartheta)|^2] = |\,\mathbb{E}_\vartheta[\hat{g}] - g(\vartheta)|^2 + \mathbb{E}_\vartheta[|\hat{g} - \mathbb{E}_\vartheta[\hat{g}]|^2]$$
   (b) Bestimmen Sie die Bias-Varianz-Zerlegung für $\hat{\mu}_\alpha = \alpha X$, $\alpha \in \mathbb{R}$, und zeigen Sie, dass $\alpha_{\text{Orakel}} := 1 - \frac{\sigma^2 d}{|\mu|^2 + \sigma^2 d}$ das quadratische Risiko minimiert, falls μ der wahre Parameter ist. Erklären Sie das Verhalten für $d \to \infty$ bei konstantem $|\mu|$.
   (c) Weisen Sie nach, dass $|X|^2$ ein erwartungstreuer Schätzer von $|\mu|^2 + \sigma^2 d$ ist und setze $\hat{\alpha} := 1 - \frac{\sigma^2 d}{|X|^2}$. Schließen Sie durch Berechnen von $\operatorname{Var}(|X|^2)$, dass $\forall \epsilon, R > 0\ \exists K > 0:$
$$\mathbb{P}_\mu\left(\left|\frac{|X|^2}{\sigma^2 d} - \frac{|\mu|^2 + \sigma^2 d}{\sigma^2 d}\right| \geqslant \frac{K}{\sqrt{d}}\right) \leqslant \epsilon, \quad \forall d \geqslant 1\ \forall \mu \in \mathbb{R}^d \text{ mit } |\mu| \leqslant R.$$
   Folgere, dass $|\hat{\alpha} - \alpha_{\text{Orakel}}| = O_P(d^{-1/2})$ für $d \to \infty$ und $|\mu| \leqslant R$ gilt, d.h. $\forall \epsilon, R > 0\ \exists K' > 0:$
$$\mathbb{P}_\mu(|\hat{\alpha} - \alpha_{\text{Orakel}}| > K' d^{-1/2}) \leqslant \epsilon, \quad \forall d \geqslant 1\ \forall \mu \in \mathbb{R}^d \text{ mit } |\mu| \leqslant R.$$

4. Beweisen Sie das Stein- bzw. Chen-Stein-Lemma:

   (a) Eine reellwertige Zufallsvariable X ist $N(\mu, \sigma^2)$-verteilt genau dann, wenn $\mathbb{E}[(X-\mu)f(X)] = \sigma^2\,\mathbb{E}[f'(X)]$ für alle $f \in C^1(\mathbb{R})$ mit $\mathbb{E}[|f'(X)|] < \infty$ gilt.

   (b) Eine Zufallsvariable N mit Werten in $\mathbb{N}_0$ ist Poiss(λ)-verteilt genau dann, wenn $\mathbb{E}[Nf(N)] = \lambda\,\mathbb{E}[f(N+1)]$ für jedes beschränkte $f : \mathbb{N}_0 \to \mathbb{R}$ gilt.

Abgabe vor der Vorlesung am Freitag, dem 15.5.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

4. Übungsblatt

1. Es sei $X_1, \dots, X_n$ eine $N(\mu, 1)$-verteilte mathematische Stichprobe mit $\mu \in \mathbb{R}$ unbekannt.

   (a) Geben Sie das zugehörige statistische Experiment auf $\mathcal{X} = \mathbb{R}^n$ an und zeigen Sie, dass es vom Produktmaß $N(0,1)^{\otimes n}$ dominiert wird.

   (b) Bestimmen Sie die Likelihoodfunktion für das dominierende Maß in (a). Welcher Wert $\mu \in \mathbb{R}$ maximiert die Likelihoodfunktion zu gegebenem $x \in \mathbb{R}^n$ (dies ist der Maximum-Likelihood-Schätzer bei Beobachtung $X = x$)?

2. Beweisen oder widerlegen Sie die Aussage, dass folgende Verteilungen Exponentialfamilien bilden. Bestimmen Sie gegebenenfalls den natürlichen Parameterraum.

   (a) Multinomialverteilung $(M(p_0, \dots, p_s; n))_{0<p_i<1, \sum p_i = 1}$;

   (b) Poissonverteilung $(\mathrm{Poiss}(\lambda))_{\lambda>0}$;

   (c) Gleichmäßige Verteilung $(U([0,\vartheta]))_{\vartheta>0}$;

   (d) Gammaverteilung $(\Gamma(a,b))_{a,b>0}$.

3. Ein Physiker untersucht die Radioaktivität bei zwei verschiedenen Präparaten. Die unabhängig gemessene Zahl der Zerfälle in einer Zeiteinheit bei Präparat 1 sei $X_1, \dots, X_{m_1}$ (m_1 Messungen), bei Präparat 2 $Y_1, \dots, Y_{m_2}$ (m_2 Messungen). Geben Sie eine vernünftige Regel an, um zu entscheiden, welches Präparat stärker radioaktiv ist. Begründen Sie dazu, weshalb die Annahme einer Poissonverteilung gerechtfertigt ist, und geben Sie ein Suffizienzargument.

4. Beweisen Sie: Es sei $(\mathbb{P}_\vartheta)_{\vartheta \in \mathcal{Z}}$ eine Exponentialfamilie mit natürlichem Parameterraum $\mathcal{Z} \subseteq \mathbb{R}^k$ und Darstellung

$$\frac{d\,\mathbb{P}_\vartheta}{d\mu}(x) = C(\vartheta) h(x) \exp(\langle \vartheta, T(x) \rangle) = h(x) \exp(\langle \vartheta, T(x) \rangle - A(\vartheta)),$$

   wobei $A(\vartheta) = \log\big(\int h(x) \exp(\langle \vartheta, T(x) \rangle) \mu(dx)\big)$. Ist $\bar\vartheta$ ein innerer Punkt von $\mathcal{Z}$, so ist die *erzeugende Funktion* von T $\psi_{\bar\vartheta}(s) = \mathbb{E}_{\bar\vartheta}[e^{\langle T, s \rangle}]$, $s \in \mathbb{R}^k$, in einer Umgebung der Null wohldefiniert und beliebig oft differenzierbar. Es gilt $\psi_{\bar\vartheta}(s) = \exp(A(\bar\vartheta + s) - A(\bar\vartheta))$ für alle s mit $\bar\vartheta + s \in \mathcal{Z}$. Für $i, j = 1, \dots, k$ folgt $\mathbb{E}_{\bar\vartheta}[T_i] = \frac{dA}{d\vartheta_i}(\bar\vartheta)$ und $\mathrm{Cov}_{\bar\vartheta}(T_i, T_j) = \frac{d^2 A}{d\vartheta_i d\vartheta_j}(\bar\vartheta)$.

Abgabe vor der Vorlesung am Freitag, dem 22.5.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

5. Übungsblatt

1. Eine suffiziente Statistik T^* heißt *minimalsuffizient*, wenn es zu jeder suffizienten Statistik T eine messbare Funktion h gibt, so dass $T^* = h(T)$ $\mathbb{P}_\vartheta$-f.s. für alle $\vartheta \in \Theta$ gilt. Beweisen Sie, dass jede $\mathbb{R}^d$-wertige, suffiziente und vollständige Statistik minimalsuffizient ist, sofern eine minimalsuffiziente Statistik überhaupt existiert. Gilt die Umkehrung für $\mathbb{R}^d$-wertige Statistiken?
   Hinweis: Man kann zeigen, dass minimalsuffiziente Statistiken für dominierte Experimente auf separablen Messräumen (wie $(\mathbb{R}^d, \mathscr{B}_{\mathbb{R}^d})$) stets existieren.

2. Es sei $(B_t, t \geqslant 0)$ eine Brownsche Bewegung. Es wird $X_t := \sigma B_t + at$ mit $\sigma > 0$ unbekannt und $a \in \mathbb{R}$ unbekannt zu den n Zeitpunkten $h, 2h, \ldots, T := nh$ mit $h > 0$ beobachtet.

   (a) Bestimmen Sie die gemeinsame Verteilung der $\Delta X_k := X_{kh} - X_{(k-1)h}$, $k \in \{1, \ldots, n\}$.

   (b) $\mathbb{P}_{a,\sigma^2}$ bezeichne die Verteilung von $(\Delta X_1, \Delta X_2, \ldots, \Delta X_n)$ mit $X_t := \sigma B_t + at$. Bestimmen Sie die Likelihoodfunktion bezüglich $\mathbb{P}_{0,1}$ und weisen Sie nach, dass $(X_T, \sum_{k=1}^n (\Delta X_k)^2)$ eine suffiziente Statistik ist.

   (c) Berechnen Sie das quadratische Risiko von $\hat{a} = X_T/T$ und $\hat{\sigma}^2 = \sum_{k=1}^n (\Delta X_k)^2/T$ und diskutieren Sie jeweils das Verhalten für $T \to \infty$ bei festem h und für $h \to 0$ bei festem T.

   (d) Simulieren Sie 1000 Realisierungen von $X_t = B_t$ sowie $X_t = 0.5B_t + 4t$ auf dem Intervall $[0, 1]$ und berechnen Sie $\hat{\sigma}^2$ jeweils für $h \in \{0.1, 0.01, 10^{-4}\}$ anhand der Beobachtungen $X_h, X_{2h}, \ldots, X_1$. Stellen Sie in jedem der sechs Fälle die Verteilung des Schätzfehlers $\hat{\sigma} - \sigma$ in einem Histogramm dar. Formulieren Sie eine Vermutung, gegen welche Verteilung $\hat{\sigma} - \sigma$ bei richtiger Skalierung für $h \to 0$ konvergiert.

   Hinweis: Eine *Brownsche Bewegung* $(B_t, t \geqslant 0)$ ist durch folgende Eigenschaften charakterisiert:

   (i) es gilt $B_0 = 0$ und $B_t \sim N(0, t)$, $t > 0$;

   (ii) die Inkremente sind stationär und unabhängig: für $0 \leqslant t_0 < t_1 < \cdots < t_m$ gilt $(B_{t_1} - B_{t_0}, \ldots, B_{t_m} - B_{t_{m-1}}) \sim N(0, \mathrm{diag}(t_1 - t_0, \ldots, t_m - t_{m-1}))$;

   (iii) B hat stetige Pfade.

Abgabe vor der Vorlesung am Freitag, dem 29.5.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

6. Übungsblatt

1. Für zwei Wahrscheinlichkeitsmaße $\mathbb{P}, \mathbb{Q}$ auf demselben Raum $(\mathcal{X}, \mathscr{F})$ bezeichnet

$$H^2(\mathbb{P}, \mathbb{Q}) := \int_{\mathcal{X}} \Big(\sqrt{p(x)} - \sqrt{q(x)}\Big)^2 \mu(dx)$$

den quadrierten *Hellinger-Abstand*, wobei μ ein Maß ist, das $\mathbb{P}$ und $\mathbb{Q}$ dominiert (z.B. $\mu = \mathbb{P} + \mathbb{Q}$), und p, q die entsprechenden μ-Dichten bezeichnet. Zeigen Sie:

(a) Der Hellingerabstand H ist unabhängig vom dominierenden Maß μ und definiert eine Metrik.

(b) Es gilt $H^2(\mathbb{P}, \mathbb{Q}) = 2 - 2\int_{\mathcal{X}} \sqrt{p(x)q(x)}\, \mu(dx) \in [0, 2]$.

(c) Für Produktmaße $\mathbb{P} = \otimes_{i=1}^n \mathbb{P}_i$, $\mathbb{Q} = \otimes_{i=1}^n \mathbb{Q}_i$ gilt

$$H^2(\mathbb{P}, \mathbb{Q}) = 2\Big(1 - \prod_{i=1}^n \Big(1 - \frac{H^2(\mathbb{P}_i, \mathbb{Q}_i)}{2}\Big)\Big) \leqslant \sum_{i=1}^n H^2(\mathbb{P}_i, \mathbb{Q}_i).$$

2. Für zwei Wahrscheinlichkeitsmaße $\mathbb{P}, \mathbb{Q}$ auf $(\mathcal{X}, \mathscr{F})$ bezeichnet

$$\|\mathbb{P} - \mathbb{Q}\|_{TV} := \sup_{A \in \mathscr{F}} |\mathbb{P}(A) - \mathbb{Q}(A)|$$

den *Totalvariationsabstand*. Weisen Sie die Ungleichung von Le Cam nach:

$$\tfrac{1}{2} H^2(\mathbb{P}, \mathbb{Q}) \leqslant \|\mathbb{P} - \mathbb{Q}\|_{TV} \leqslant H(\mathbb{P}, \mathbb{Q}) \sqrt{1 - \tfrac{1}{4} H^2(\mathbb{P}, \mathbb{Q})}.$$

Insbesondere konvergiert also $\mathbb{P}_n \to \mathbb{P}$ im Totalvariationsabstand genau dann, wenn Konvergenz im Hellingerabstand vorliegt.
Anleitung:
Weise mit μ-Dichten p, q nach: $\|\mathbb{P} - \mathbb{Q}\|_{TV} = \frac{1}{2}\int |p - q| d\mu = 1 - \int \min(p, q) d\mu$ sowie $(\int \sqrt{pq} d\mu)^2 \leqslant 2 \int \min(p, q) d\mu$.

3. Betrachten Sie für eine Lebesguedichte $f : \mathbb{R} \to [0, \infty)$ das *Lokationsmodell* einer mathematischen Stichprobe $X_1, \dots, X_n$, die gemäß $f(\bullet - \vartheta)$ mit $\vartheta \in \mathbb{R}$ unbekannt verteilt ist. Zeigen Sie, dass diese Verteilungsfamilie L^2-differenzierbar ist, wenn $f \in C^1(\mathbb{R})$, $f > 0$ und $\int f'(x)^2 f(x)^{-1} dx < \infty$ gilt, und bestimmen Sie die Fisher-Information $I_n(\vartheta)$.
   Können Sie die Bedingungen so abschwächen, dass auch die L^2-Differenzierbarkeit im Fall der Laplace-Verteilung $f(x) = \frac{1}{2} e^{-|x|/2}$ folgt?

4. Bestimmen Sie die Fisher-Informationsmatrix für eine $N(\mu, \sigma^2)$-verteilte mathematische Stichprobe $X_1, \dots X_N$ mit unbekannten Werten $\mu \in \mathbb{R}$ und $\sigma^2 > 0$ sowie für $X \sim \mathrm{Bin}(n, p)$ mit $p \in (0, 1)$ unbekannt und n bekannt. Finden Sie jeweils einen erwartungstreuen Schätzer für μ und für p, der die Cramér-Rao-Schranke erreicht. Bestimmen Sie einen erwartungstreuen Schätzer für σ^2 bei $N \geqslant 2$ Beobachtungen, der zumindest asymptotisch für $N \to \infty$ die Cramér-Rao-Schranke erreicht (bei Reskalierung mit N).

Abgabe vor der Vorlesung am Freitag, dem 5.6.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

7. Übungsblatt

1. (a) Es sei $X_1, \dots, X_n$ eine $U([0, \vartheta])$-verteilte mathematische Stichprobe mit $\vartheta > 0$ unbekannt. Weisen Sie nach, dass $\hat{\vartheta}_n := \max_i X_i$ ein Maximum-Likelihood-Schätzer für ϑ ist und $n(\vartheta - \hat{\vartheta}_n) \xrightarrow{d} \mathrm{Exp}(1/\vartheta)$ für $n \to \infty$ gilt.

   (b) Zeigen Sie für eine mathematische Stichprobe $X_1, \dots, X_n$ bei zugrundeliegender Lebesguedichte $f_m(x) = \frac{1}{2} e^{-|x-m|}$, $x \in \mathbb{R}$, mit $m \in \mathbb{R}$ unbekannt, dass der Stichproben-Median ein Maximum-Likelihood-Schätzer von m ist.

2. Es seien $X_1, \dots, X_n$ unabhängig und gemäß der stetigen Lebesgue-Dichte $f : \mathbb{R} \to \mathbb{R}^+$ verteilt. Es gebe ein $m \in \mathbb{R}$ (*Median*) mit $\int_{-\infty}^m f = \int_m^\infty f = 1/2$ und es gelte $f(m) > 0$. Dann gilt für jeden Stichprobenmedian $\hat{m}_n$ die asymptotische Normalität
$$\sqrt{n}(\hat{m}_n - m) \xrightarrow{d} N\Big(0, \frac{1}{4f(m)^2}\Big).$$
   Zeigen Sie dies gemäß einem der folgenden Ansätze:

   (a) Für eine $U([0, 1])$-verteilte mathematische Stichprobe $U_1, \dots, U_n$ ist die k-te Ordnungsstatistik $X_{(k)}$ Beta(k,n+1-k)-verteilt auf $[0, 1]$. Setzt amn $k = n/2$ bzw. $k = (n \pm 1)/2$, so folgt, dass ihr Stichprobenmedian $\hat{m}_n^U$ obige asymptotische Normalität erfüllt. Benutze die Quantilstransformation $X_i = F^{-1}(U_i)$ für eine geeignete Inverse F^{-1} der Verteilungsfunktion von X_i, um die Asymptotik von $\hat{m}_n$ herzuleiten.

   (b) Der Stichprobenmedian kann als Funktional der empirischen Verteilungsfunktion dargestellt werden und eine funktionale Δ-Methode kann angewendet werden, wobei der Satz von Donsker zur Asymptotik des empirischen Prozesses als bekannt vorausgesetzt sei. Argumentieren Sie zunächst heuristisch und erklären Sie dann kurz die notwendigen formalen Schritte z.B. aus A. van der Vaart, *Asymptotic Statistics*.

3. Für zwei Wahrscheinlichkeitsmaße $\mathbb{P}$ und $\mathbb{Q}$ auf demselben Messraum $(\mathcal{X}, \mathscr{F})$ heißt die Funktion

$$\mathrm{KL}(\mathbb{P} \mid \mathbb{Q}) = \begin{cases} \int_{\mathcal{X}} \log\left(\frac{d\,\mathbb{P}}{d\,\mathbb{Q}}(x)\right) \mathbb{P}(dx), & \text{falls } \mathbb{P} \ll \mathbb{Q}, \\ +\infty, & \text{sonst} \end{cases}$$

Kullback-Leibler-Divergenz.

   (a) Zeigen Sie, dass $f : (0,\infty) \to \mathbb{R}, x \mapsto x\log(x)$ konvex ist und somit (benutze $\mathrm{d}\,\mathbb{P} = \frac{\mathrm{d}\,\mathbb{P}}{\mathrm{d}\,\mathbb{Q}}\,\mathrm{d}\,\mathbb{Q}$)

   $$\mathrm{KL}(\mathbb{P} \mid \mathbb{Q}) \geqslant 0 \quad \text{und} \quad \mathrm{KL}(\mathbb{P} \mid \mathbb{Q}) = 0 \Longleftrightarrow \mathbb{P} = \mathbb{Q}\,.$$

   Finden Sie zwei äquivalente Wahrscheinlichkeitsmaße $\mathbb{P}$ und $\mathbb{Q}$ mit

   $$\mathrm{KL}(\mathbb{P} \mid \mathbb{Q}) \neq \mathrm{KL}(\mathbb{Q} \mid \mathbb{P}).$$

   (b) Beweisen Sie für Produktmaße:

   $$\mathrm{KL}\left(\bigotimes_{i=1}^{n} \mathbb{P}_i \mid \bigotimes_{i=1}^{n} \mathbb{Q}_i\right) = \sum_{i=1}^{n} \mathrm{KL}(\mathbb{P}_i \mid \mathbb{Q}_i).$$

4. (a) Beweisen Sie: Bildet $(\mathbb{P}_\vartheta)_{\vartheta\in\Theta}$ eine natürliche Exponentialfamilien und ist ϑ_0 innerer Punkt von $\Theta \subseteq \mathbb{R}^k$, so gilt $\mathrm{KL}(\mathbb{P}_{\vartheta_0} \mid \mathbb{P}_\vartheta) = A(\vartheta) - A(\vartheta_0) + \langle \dot{A}(\vartheta_0), \vartheta_0 - \vartheta\rangle$. Folgern Sie

   $$\ddot{\mathrm{KL}}(\mathbb{P}_{\vartheta_0} \mid \mathbb{P}_\vartheta)|_{\vartheta=\vartheta_0} = I(\vartheta_0).$$

   (b) Leiten Sie eine entsprechende Formel für die Hesse-Matrix bei ϑ_0 des quadrierten Hellingerabstandes $\vartheta \mapsto H^2(\mathbb{P}_{\vartheta_0} \mid \mathbb{P}_\vartheta)$ her.

Abgabe vor der Vorlesung am Freitag, dem 12.6.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

8. Übungsblatt

1. Ein Teich enthält eine unbekannte Anzahl ϑ von Karpfen. Zur Schätzung von ϑ werden zunächst w Fische gefangen, markiert und wieder freigelassen. Wenn sich die markierten Fische wieder gut verteilt haben, werden n Fische gefangen, von denen x markiert sind.
   Modellieren Sie die Schätzung des Fischbestandes ϑ durch ein statistisches Experiment mit hypergeometrischen Verteilungen und bestimmen Sie einen MLE, diskutieren Sie dabei den Fall $x = 0$ separat.

2. Sei $X_1, \dots, X_n$ eine mathematische Stichprobe bezüglich der Lebesguedichte

$$f_\vartheta(x) = \frac{1-\vartheta}{\varphi(\vartheta)} \left(1 - \frac{|x-\vartheta|}{\varphi(\vartheta)}\right)_+ + \frac{\vartheta}{2} \mathbf{1}_{[-1,1]}(x),$$

   wobei $\vartheta \in [0,1)$ und $\varphi : [0,1] \to [0,1]$ eine stetige, fallende Funktion mit $\varphi(0) = 1$ und $0 < \varphi(\vartheta) \leqslant 1 - \vartheta$ für $\vartheta \in (0,1)$ ist. Setze weiterhin $f_1(x) = \frac{1}{2}\mathbf{1}_{[-1,1]}(x)$. Ziel ist es für geeignetes φ zu sehen, dass für alle $\vartheta \in [0,1]$ jeder MLE fast sicher gegen Eins konvergiert und insbesondere inkonsistent ist. Zeigen Sie:

   (a) Es existiert ein Maximum-Likelihood-Schätzer $\hat{\vartheta}_n$.

   (b) Für $\vartheta < 1$ ist $f_\vartheta(x) < 1/\varphi(\vartheta) + 1/2$ und daraus folgt, dass für die Loglikelihoodfunktion ℓ_n bei n Beobachtungen und für jedes $\alpha < 1$

$$\max_{0 \leqslant \vartheta \leqslant \alpha} \frac{\ell_n(\vartheta)}{n} \leqslant \log\left(\frac{1}{\varphi(\alpha)} + \frac{1}{2}\right) < \infty$$

   gilt. Um zu beweisen, dass $\lim_{n\to\infty} \hat{\vartheta}_n = 1$ f.s. für alle $\vartheta \in [0,1]$, reicht es $\max_{0\leqslant\vartheta\leqslant 1} \ell_n(\vartheta)/n \to \infty$ f.s. zu zeigen.

   (c) Mit $X_{(n)} = \max\{X_1, \dots, X_n\}$ gilt f.s.

$$\max_{0 \leqslant \vartheta \leqslant 1} \frac{\ell_n(\vartheta)}{n} \geqslant \frac{n-1}{n} \log\left(\frac{X_{(n)}}{2}\right) + \frac{1}{n} \log\left(\frac{1 - X_{(n)}}{\varphi(X_{(n)})}\right).$$

   (d) Aus dem Lemma von Borel-Cantelli folgt $n^{1/4}(1 - X_{(n)}) \to 0$ f.s. für $\vartheta = 0$ und auch für alle $\vartheta \in [0,1]$. Mit $\varphi(\vartheta) := (1-\vartheta)\exp(-(1-\vartheta)^{-4} + 1)$ folgt $\liminf_{n\to\infty}(1/n)\log((1 - X_{(n)})/\varphi(X_{(n)})) = \infty$ f.s. und damit die gewünschte Aussage.

3. Es sei $\Theta \subseteq \mathbb{R}^k$ kompakt sowie $(X_n(\vartheta),\, \vartheta \in \Theta)_{n\geqslant 1}$ und $(X(\vartheta),\, \vartheta \in \Theta)$ stetige Prozesse mit $X_n(\vartheta) \xrightarrow{\mathbb{P}} X(\vartheta)$ für $n \to \infty$ für alle $\vartheta \in \Theta$. Beweisen Sie:

   (a) $\forall \delta > 0\, \exists U_\delta \subseteq \Theta$ endlich: $\sup_{\vartheta\in\Theta} \inf_{\vartheta' \in U_\delta} |\vartheta - \vartheta'| \leqslant \delta$.

   (b) Mit dem Stetigkeitsmodul $\omega_\delta(f) := \sup_{|\vartheta_1 - \vartheta_2| \leqslant \delta} |f(\vartheta_1) - f(\vartheta_2)|$ gilt

$$\sup_{\vartheta\in\Theta} |X_n(\vartheta) - X(\vartheta)| \leqslant \omega_\delta(X_n) + \omega_\delta(X) + \max_{\vartheta \in U_\delta} |X_n(\vartheta) - X(\vartheta)|.$$

   (c) Es gilt $\max_{\vartheta\in U_\delta} |X_n(\vartheta) - X(\vartheta)| \xrightarrow{\mathbb{P}} 0$ für $n \to \infty$ sowie $\omega_\delta(X) \to 0$ fast sicher für $\delta \to 0$.

   Schließen Sie daraus, dass aus der Straffheitsbedingung

$$\forall \varepsilon, \eta > 0\, \exists \delta > 0: \limsup_{n\to\infty} \mathbb{P}(\omega_\delta(X_n) \geqslant \varepsilon) \leqslant \eta$$

   die gleichmäßige Konvergenz $\sup_{\vartheta\in\Theta} |X_n(\vartheta) - X(\vartheta)| \xrightarrow{\mathbb{P}} 0$ für $n \to \infty$ folgt.

4. Im nichtlinearen Regressionsmodell der Beobachtungen

$$Y_i = f_\vartheta(i/n) + \varepsilon_i, \quad i = 1, \ldots, n,\ f_\vartheta \in C([0,1]),\ (\varepsilon_i)_{1\leqslant i\leqslant n} \text{ iid},$$

   mit $\mathbb{E}[\varepsilon_i] = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, $\mathbb{E}[\varepsilon_i^4] < \infty$, $\sigma > 0$ betrachte den Kleinste-Quadrate-Schätzer $\hat\vartheta_n = \mathrm{argmin}_{\vartheta\in\Theta} \sum_{i=1}^n (Y_i - f_\vartheta(i/n))^2$. Geben Sie Voraussetzungen für die Parametrisierung $\vartheta \mapsto f_\vartheta$ an, um auf die asymptotische Normalität von $\hat\vartheta_n$ für $n \to \infty$ zu schließen und bestimmen Sie die asymptotische Varianz. Berechnen Sie $\hat\vartheta_n$ und die asymptotische Varianz konkret für das Modell $f_\vartheta(x) = \sin(2\pi(x + \vartheta))$, $\vartheta \in [0,1)$.
   Hinweis: Schlagen Sie ggf. den ZGWS unter Lindeberg- oder Lyapunov-Bedingung nach.

Abgabe vor der Übung am Freitag, den 19.6.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

9. Übungsblatt

1. Im Beobachtungsmodell $Y_i = \gamma + \varepsilon_i$, $i = 1, \dots, n$, mit (ε_i) i.i.d. und symmetrisch um Null verteilt betrachte für $\kappa > 0$ und $\Gamma \subseteq \mathbb{R}$ den Huber-Schätzer

$$\hat{\gamma}_n = \operatorname{argmin}_{\gamma \in \Gamma} \frac{1}{n} \sum_{i=1}^{n} \rho(Y_i - \gamma), \quad \rho(x) = \begin{cases} \frac{1}{2}x^2, & \text{falls } |x| \leqslant \kappa, \\ \kappa|x| - \frac{\kappa^2}{2}, & \text{falls } |x| > \kappa. \end{cases}$$

   (a) Im Fall $\Gamma = [a, b]$ mit $-\infty < a < b < \infty$ weise die Konsistenz von $\hat{\gamma}_n$, z.B. gemäß Bedingungen A1-A3, nach.

   (b) Für $\gamma_0 \in (a, b)$, $\Gamma = [a, b]$ und $\mathbb{P}(\varepsilon_i = x) = 0$ für alle $x \in \mathbb{R}$ beweise rigoros die asymptotische Normalität

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} N\Big(0, \frac{\mathbb{E}[\varepsilon_i^2 \wedge \kappa^2]}{\mathbb{P}(|\varepsilon_i| \leqslant \kappa)^2}\Big).$$

   Tipp: Der Mittelwertsatz gilt für $\dot{K}_n$ bis auf einen maximalen Fehler $\frac{1}{n}$.

   (c*) Untersuche den Fall $\Gamma = \mathbb{R}$.

2. Vergleichen Sie im Modell aus Aufgabe 1 die asymptotische Varianz von Stichprobenmittel, Stichprobenmedian, Huber-Schätzer (für geeignete Wahlen von κ, ggf. Computereinsatz) im Fall folgender Verteilungen:

   (a) $\varepsilon_i \sim N(0, \sigma^2)$, (b) $f_\varepsilon(x) = \frac{s}{2} e^{-s|x|}$, (c) $\varepsilon_i \sim t(k)$, $k \in \{1, 2, 4, 10\}$.

   Welche Gesichtspunkte sollten daher bei der Auswahl des Schätzers Beachtung finden?

3. Es sei (Y, Z) gemäß der Dichte $f(y, z, \vartheta)$, $\vartheta \in \Theta$, bezüglich $\mu \otimes \nu$ verteilt, wobei μ und ν σ-endliche Maße seien. Nur $Y = y$ wird beobachtet. Der EM-Algorithmus zur Berechnung eines MLE besteht aus der Wahl eines Startwertes ϑ_0 mit $L(\vartheta_0) = f_Y(y, \vartheta_0) > 0$ und aus der Wiederholung für $j = 0, 1, \dots$ der Schritte (1) und (2):

   (1) Berechne

$$J(\vartheta, \vartheta_j) = \mathbb{E}_{\vartheta_j}\left[\log\left(\frac{f(Y, Z, \vartheta)}{f(Y, Z, \vartheta_j)}\right) \Big| Y = y\right].$$

   (2) Setze $\vartheta_{j+1} = \operatorname{argmax}_\vartheta J(\vartheta, \vartheta_j)$.

   Zeigen Sie die Gleichung

$$J(\vartheta_{j+1}, \vartheta_j) = \log\left(\frac{f_Y(y, \vartheta_{j+1})}{f_Y(y, \vartheta_j)}\right) + \int \log\left(\frac{f_{Z|Y=y}(z, \vartheta_{j+1})}{f_{Z|Y=y}(z, \vartheta_j)}\right) f_{Z|Y=y}(z, \vartheta_j) \nu(\mathrm{d}z)$$

   und folgern Sie, dass im EM-Algorithmus $L(\vartheta_{j+1}) \geqslant L(\vartheta_j)$ gilt.

4. Betrachten Sie eine mathematische Stichprobe $Y_1, \dots, Y_n$, die gemäß einer Mischung zweier Normalverteilungen verteilt ist: Gegeben $Z_i = 0$ ist $Y_i \sim N(a,1)$ und gegeben $Z_i = 1$ ist $Y_i \sim N(b,1)$, wobei $\mathbb{P}_{a,b}(Z_i = 0) = \mathbb{P}_{a,b}(Z_i = 1) = 1/2$ und $(Y_1, Z_1), \dots, (Y_n, Z_n)$ unabhängig. μ sei das Lebesguemaß auf $\mathbb{R}^n$ und ν sei gegeben durch $\nu(\{z\}) = 1/2^n$ für alle $z \in \{0,1\}^n$.

   (a) Bestimmen Sie $f_Y(y,a,b)$ für $n = 1$ und zeigen Sie für beliebige n

   $$f(y,z,a,b) = \prod_{i=1}^n \varphi(y_i - a)^{1-z_i} \varphi(y_i - b)^{z_i}, \quad \text{mit } \varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

   (b) Zeigen Sie: Es gilt im EM-Algorithmus

   $$a_{j+1} = \frac{\sum_{i=1}^n (1-\tau_i) y_i}{\sum_{i=1}^n (1-\tau_i)} \quad \text{und} \quad b_{j+1} = \frac{\sum_{i=1}^n \tau_i y_i}{\sum_{i=1}^n \tau_i},$$

   wobei $\tau_i := \varphi(y_i - b_j)/(\varphi(y_i - a_j) + \varphi(y_i - b_j))$.

   (c*) Simulieren Sie einen numerischen MLE und den EM-Algorithmus für $a = 1$, $b = 2$, $n = 100$ und für verschiedene Werte von j . Konvergiert ϑ_j für $j \to \infty$ gegen den numerischen MLE?

Abgabe vor der Übung am Freitag, dem 26.6.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

10. Übungsblatt

1. Betrachten Sie eine mathematische Stichprobe $X_1, \dots, X_n$ mit Lebesguedichte $f_{\mu,\sigma}(x) = \sigma^{-1} f((x-\mu)/\sigma)$, $f \in C^2(\mathbb{R})$, mit $(\mu, \sigma) \in \Theta \subseteq \mathbb{R} \times \mathbb{R}^+$ unbekannt (*Lokations-Skalen-Familie*). Welche Annahmen an f garantieren, dass der MLE asymptotisch normalverteilt ist? Geben Sie die asymptotische Kovarianzmatrix an.

2. Betrachten Sie für ein binäres statistisches Modell $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$ mit $\Theta = \{0, 1\}$ das Testproblem $H_0 : \vartheta = 0$ gegen $H_1 : \vartheta = 1$. Beweisen Sie:

   (a) Im Neyman-Pearson-Lemma gilt auch die Umkehrung: Jeder gleichmäßig beste Test φ für $H_0 : \vartheta = 0$ gegen $H_1 : \vartheta = 1$ zum Niveau $\mathbb{E}_0[\varphi] \in (0, 1)$ besitzt fast sicher die Form eines Neyman-Pearson-Tests.

   (b) Bei einem Minimax-Test bezüglich 0-1-Verlust sind Fehler 1. und 2. Art gleich: $\mathbb{E}_0[\varphi] = 1 - \mathbb{E}_1[\varphi]$.

   (c) Für einen Test φ mit $\mathbb{E}_0[\varphi] \in (0, 1)$ sind äquivalent:

   - φ ist ein Minimax-Test bezüglich 0-1-Verlust.
   - φ besitz fast sicher die Form eines Neyman-Pearson-Tests und es gilt $\mathbb{E}_0[\varphi] = 1 - \mathbb{E}_1[\varphi]$.

   *H*inweis: (b) kann indirekt durch betrachten von Tests der Form $\tilde{\varphi} := \chi\varphi$ bzw. $\tilde{\varphi} := \chi\varphi + (1-\chi)$ mit $\chi \in (0, 1)$ bewiesen werden.

3. Es sei $X_1, \dots, X_n$ eine $\mathrm{Exp}(\vartheta)$-verteilte mathematische Stichprobe mit $\vartheta > 0$ unbekannt. Konstruieren Sie einen gleichmäßig besten Test zum Niveau $\alpha \in (0, 1)$ für das Testproblem $H_0 : \vartheta \leqslant 1$ gegen $H_1 : \vartheta > 1$. Geben Sie für $n = 1$ den kritischen Wert k explizit an und zeichnen Sie die Gütefunktion für $\alpha = 0,05$.

4. Es sei $X_1, \dots, X_n$ eine $U([0, \vartheta])$-verteilte mathematische Stichprobe mit $\vartheta > 0$ unbekannt. Setze $X = (X_1, \dots, X_n)$. Zeigen Sie: Für das Testproblem $H_0 : \vartheta \le \vartheta_0$ gegen $H_1 : \vartheta > \vartheta_0$ mit festem ϑ_0 ist jeder Test φ mit $\mathbb{E}_{\vartheta_0}[\varphi(X)] = \alpha$, $\mathbb{E}_\vartheta[\varphi(X)] \le \alpha$ für alle $\vartheta \le \vartheta_0$ und $\varphi(X) = 1$ für $\max(X_1, \dots, X_n) > \vartheta_0$ ein gleichmäßig bester Test zum Niveau $\alpha \in (0, 1)$.

Abgabe vor der Übung am Freitag, dem 3.7.15.

Vorlesung *Mathematische Statistik*
Sommersemester 2015
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

11. Übungsblatt

1. Beweisen Sie das verallgemeinerte Neyman-Pearson-Lemma aus der Vorlesung. *Hinweis:* Orientieren Sie sich am Beweis des Neyman-Pearson-Lemmas.

2. Es seien $X_1, \ldots, X_n$ eine $\mathrm{Exp}(\vartheta)$-verteilte mathematische Stichprobe mit $\vartheta > 0$ unbekannt sowie $\vartheta_0 > 0$. Konstruieren Sie einen gleichmäßig besten unverfälschten Test zum Niveau $\alpha \in (0, 1)$ für das zweiseitige Testproblem $H_0 : \vartheta = \vartheta_0$ gegen $H_1 : \vartheta \neq \vartheta_0$. Berechnen Sie (zumindest approximativ) die kritischen Werte im Fall $\vartheta_0 = 1$, $n = 5$, $\alpha = 0{,}05$.

3. Es werden die zwei mathematischen Stichproben $X_1, \ldots, X_n \sim \mathrm{Bin}(1, p_1)$ und $Y_1, \ldots, Y_n \sim \mathrm{Bin}(1, p_2)$ mit $p_1, p_2 \in (0, 1)$ unabhängig beobachtet. Konstruieren Sie einen gleichmäßig besten unverfälschten Test vom Niveau $\alpha \in (0, 1)$ für das Testproblem $H_0 : p_1 = p_2$ gegen $H_1 : p_1 \neq p_2$.

Abgabe vor der Übung am Freitag, dem 10.7.15.