Vorlesung *Mathematische Statistik*
Sommersemester 2018
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

1. Übungsblatt

1. 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.

2. Es sei $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$ ein statistisches Modell sowie π eine a-priori-Verteilung auf $(\Theta, \mathcal{F}_\Theta)$, so dass $\mathbb{P}_\vartheta$ die μ-Dichte $f_{X|\boldsymbol{\Theta}=\vartheta}$ für alle $\vartheta \in \Theta$ und π die ν-Dichte $f_{\boldsymbol{\Theta}}$ besitzen bezüglich σ-endlichen Maßen μ und ν. Zeigen Sie für $(\mathcal{F} \otimes \mathcal{F}_\Theta)$-messbare Dichten $(x, \vartheta) \mapsto f_{X|\boldsymbol{\Theta}=\vartheta}(x) \in [0, \infty)$:

   (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 $\mathbb{E}_{\tilde{\mathbb{P}}}[|g(X, \boldsymbol{\Theta})|] < \infty$, so ist $\int_\Theta g(x, \vartheta) f_{\boldsymbol{\Theta}|X=x}(\vartheta)\nu(\mathrm{d}\vartheta)$ eine Version der bedingten Erwartung $\mathbb{E}_{\tilde{\mathbb{P}}}[g(X, \boldsymbol{\Theta})|X = x]$. Speziell ergibt sich für $A \in \mathcal{F}_\Theta$ die *a-posteriori*-Wahrscheinlichkeit:
   $$\tilde{\mathbb{P}}(\boldsymbol{\Theta} \in A | X = x) := \mathbb{E}_{\tilde{\mathbb{P}}}[\mathbf{1}_{\{\boldsymbol{\Theta}\in A\}} | X = x] = \int_A f_{\boldsymbol{\Theta}|X=x}(\vartheta)\nu(\mathrm{d}\vartheta).$$

3. 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.

4. 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.

Einzel-Abgabe vor der Vorlesung am **Montag, dem 30.4.18**.

Vorlesung *Mathematische Statistik*
Sommersemester 2018
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

2. Übungsblatt

1. *Truncated normal means.* Es sei $X_1, \dots, X_n$ eine mathematische $N(\vartheta, 1)$-verteilte Stichprobe sowie $\Theta_1 = [\vartheta_0, \infty)$ und $\Theta_2 = [-m, m]$ Parametermengen für $\vartheta_0 \in \mathbb{R}$, $m > 0$. $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ bezeichne das Stichprobenmittel. Lesen Sie Example 5.2.9 auf Seite 327 in Lehmann/Casella `https://link.springer.com/content/pdf/10.1007%2Fb98854.pdf` und beweisen Sie vollständig:

   (a) In beiden Fällen $\vartheta \in \Theta_1$ und $\vartheta \in \Theta_2$ ist $\bar{X}$ nicht zulässig.

   (b) Im Fall $\vartheta \in \Theta_1$ ist $\bar{X}$ minimax, im Fall $\vartheta \in \Theta_2$ ist $\bar{X}$ nicht minimax.

   (c*) *freiwillig:* Der Schätzer (2.11) auf Seite 328 mit $m < n^{-1/2}$ ist minimax und besser als $\bar{X}$ für $\vartheta \in \Theta_2$.

2. Beweisen Sie das Lemma von Stein bzw. Stein-Chen:

   (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.

   Hinweis: Dies ist Ausgangspunkt der mächtigen *Steinschen Methode*.

Abgabe vor der Vorlesung am Dienstag, dem 8.5.18.

Vorlesung *Mathematische Statistik*
Sommersemester 2018
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

3. Übungsblatt

1. 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 setzen Sie $\hat{\alpha} := 1 - \frac{\sigma^2 d}{|X|^2}$. Schließen Sie durch Berechnen von $\operatorname{Var}(|X|^2)$, dass $\forall \varepsilon, R > 0\ \exists K > 0$:

   $$\mathbb{P}_\mu\left(\left|\frac{|X|^2}{\sigma^2 d} - \frac{|\mu|^2 + \sigma^2 d}{\sigma^2 d}\right| \geqslant \frac{K}{\sqrt{d}}\right) \leqslant \varepsilon, \quad \forall d \geqslant 1\ \forall \mu \in \mathbb{R}^d \text{ mit } |\mu| \leqslant R.$$

   Folgern Sie, dass $|\hat{\alpha} - \alpha_{\text{Orakel}}| = O_P(d^{-1/2})$ für $d \to \infty$ und $|\mu| \leqslant R$ gilt, d.h. $\forall \varepsilon, R > 0\ \exists K' > 0$:

   $$\mathbb{P}_\mu(|\hat{\alpha} - \alpha_{\text{Orakel}}| > K'\, d^{-1/2}) \leqslant \varepsilon, \quad \forall d \geqslant 1\ \forall \mu \in \mathbb{R}^d \text{ mit } |\mu| \leqslant R.$$

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. Lesen Sie Abschnitt 2.4 im Skript von 2015 und beweisen Sie Lemma 2.33 und Lemma 2.34. Machen Sie sich dazu mit dem Konzept des randomisierten Tests vertraut.

4. Es sei $X_1, \ldots, X_n$ 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 $X = x$)?

Abgabe vor der Übung am Dienstag, dem 15.5.18.
Korrektor: Marvin van Stegen, erreichbar unter `marvin@vanstegen.eu`

4. Übungsblatt

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

2. Es sei $(X_1,\dots,X_n)$ eine $\mathrm{Poiss}(\lambda)$-verteilte mathematische Stichprobe mit $\lambda>0$ unbekannt.

   (a) Zeigen Sie, dass das Stichprobenmittel $\bar{X}$ eine suffiziente Statistik ist.

   (b) Finden Sie mittels des Satzes von Rao-Blackwell einen erwartungstreuen Schätzer von λ, der kleinere Varianz als $\hat{\lambda}_1 = X_1$ besitzt.

   (c) Finden Sie mittels des Satzes von Rao-Blackwell einen erwartungstreuen Schätzer von $\vartheta = e^{-\lambda}$, der kleinere Varianz als $\hat{\vartheta}_1 = \mathbf{1}(X_1 = 0)$ besitzt.

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. 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 $h, 2h, \dots, T := nh$ mit $h > 0$ beobachtet.

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

   (b) $\mathbb{P}_{a,\sigma^2}$ bezeichne die Verteilung von $(\Delta X_1, \Delta X_2, \dots, \Delta X_n)$ 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 $(B_{t_1} - B_{t_0}, \ldots, B_{t_m} - B_{t_{m-1}}) \sim N(0, \operatorname{diag}(t_1 - t_0, \ldots, t_m - t_{m-1}))$;

(iii) B hat stetige Pfade.

Abgabe vor der Übung am Dienstag, dem 22.5.18.

Vorlesung *Mathematische Statistik*
Sommersemester 2018
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. Beim linearen Modell unter Normalverteilung wird $Y = X\beta + \sigma\varepsilon$ beobachtet mit $\varepsilon \sim N(0, E_n)$, $X \in \mathbb{R}^{n\times k}$ von vollem Rang $k < n$ und mit $\beta \in \mathbb{R}^k$, $\sigma > 0$ unbekannt.

   (a) Bestimmen Sie die Likelihoodfunktion $L(\beta, \sigma)$ bezüglich dem $(k+1)$-dimensionalen Lebesguemaß. Folgern Sie, dass eine Exponentialfamilie in $T(Y) = (X^\top Y, |Y|^2)^\top \in \mathbb{R}^k \times \mathbb{R}^+$ und $\eta(\beta,\sigma) = \sigma^{-2}(\beta, -1/2)^\top \in \mathbb{R}^k \times \mathbb{R}^-$ vorliegt.

   (b) Begründe Sie, weshalb T eine suffiziente und vollständige Statistik ist. Schließen Sie mittels bijektiver Abbildungen, dass das auch für $\tilde{T} = (\hat{\beta}, \hat{\sigma}^2)$ gilt mit dem Kleinste-Quadrate-Schätzer $\hat{\beta} = (X^\top X)^{-1} X^\top Y$ und $\hat{\sigma}^2 = \frac{|Y - X\hat{\beta}|^2}{n-k}$. Weisen Sie damit nach, dass $\hat{\beta}$ und $\hat{\sigma}^2$ jeweils erwartungstreue Schätzer von β und σ^2 von minimaler Varianz in der Klasse aller erwartungstreuer Schätzer sind.
   Hinweis: Sie dürfen für die Erwartungstreue von $\hat{\beta}$ und $\hat{\sigma}^2$ auch auf eine Literaturstelle verweisen, sofern dort ein Beweis gegeben wird.

   (c*) *Freiwillig:* Zeigen Sie anhand eines Beispiels für die Verteilung von $\varepsilon_1, \ldots, \varepsilon_n$ i.i.d. mit $\mathbb{E}[\varepsilon_i] = 0$, $\mathrm{Var}(\varepsilon_i) = 1$ und für die Designmatrix X, dass es im Allgemeinen erwartungstreue Schätzer von kleinerer Varianz geben kann als $\hat{\beta}$. Wie ist die Beziehung zum *Satz von Gauß-Markov*?

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

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

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

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

Abgabe vor der Übung am Dienstag, dem 29.5.18.

Vorlesung *Mathematische Statistik*
Sommersemester 2018
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

6. Übungsblatt

1. Zeigen Sie im Detail, dass die *score*-Funktion und die Fisher-Information bei Hellinger-differenzierbaren Modellen unabhängig vom dominierenden Maß μ sind.

2. 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?

3. 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).

4. Es sei $X_1, \dots, X_n$ eine $U([0, \vartheta])$-verteilte mathematische Stichprobe mit $\vartheta > 0$ unbekannt.

   (a) Zeigen Sie anhand der Definition, dass das zugrundeliegende statistische Modell nicht Hellinger-differenzierbar ist.

   (b) 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.

   Bemerkung: Sowohl die Konvergenzrate n^{-1} als auch die Grenzverteilung werden sich später als irreguläre Ausnahme für MLE herausstellen.

Abgabe vor der Übung am Dienstag, dem 5.6.18.

Vorlesung *Mathematische Statistik*
Sommersemester 2018
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

7. Übungsblatt

1. Es sei $(\mathcal{X}, \mathscr{F}, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$ ein statistisches Modell mit $\Theta \subseteq \mathbb{R}^k$, $k \geqslant 2$, das die Voraussetzungen an die Cramér-Rao-Ungleichung erfüllt mit Fisher-Informationsmatrix $I(\vartheta) > 0$.

   (a) Zeigen Sie, dass stets $(I(\vartheta))_{11}(I(\vartheta)^{-1})_{11} \geqslant 1$ gilt.
   Tipp: Schreiben Sie den ersten Einheitsvektor $e_1 = \sum_i \langle e_1, v_i\rangle v_i$ in einer Orthonormalbasis $(v_1, \dots, v_k)$ von Eigenvektoren von $I(\vartheta)$.

   (b) Schließen Sie, dass die Cramér-Rao-Schranke für $g(\vartheta) = g(\vartheta_1, \dots, \vartheta_n) = \vartheta_1$ stets mindestens $1/(I(\vartheta))_{11}$ beträgt. In welchen Fällen gilt Gleichheit?

   (c) Interpretieren Sie dieses Ergebnis statistisch, indem Sie mit dem Modell vergleichen, wo nur ϑ_1 unbekannt ist, während $\vartheta_2, \dots, \vartheta_k$ bekannt sind.

2. Es sei $(X, Y)^\top$ bivariat normalverteilt mit $\mathbb{E}[X] = \mathbb{E}[Y] = 0$, $\mathrm{Var}(X) = \mathrm{Var}(Y) = \sigma^2$ und $\mathrm{Cov}(X, Y) = \sigma^2\rho$, wobei $\vartheta = (\sigma^2, \rho) \in \mathbb{R}^+ \times (-1, 1)$ der unbekannte Parameter sei.

   (a) Weisen Sie nach, dass für $\hat\sigma^2 = \frac{1}{2}(X^2 + Y^2)$ gilt $\mathbb{E}[\hat\sigma^2] = \sigma^2$ und $\mathrm{Var}(\hat\sigma^2) = (1 + \rho^2)\sigma^4$.

   (b) Nehmen Sie nun an, dass ρ bekannt ist, und konstruieren Sie einen erwartungstreuen Schätzer $\tilde\sigma^2_\rho$ mit $\mathrm{Var}(\tilde\sigma^2_\rho) = \sigma^4 < \mathrm{Var}(\hat\sigma^2)$ im Fall $\rho \neq 0$.
   Tipps: Diagonalisieren Sie die Kovarianzmatrix Σ von $(X, Y)^\top$ und betrachten Sie die Beobachtung in der Eigenbasis von Σ ('whitening', 'decorrelation'). Alternativ bestimmen Sie direkt den MLE.

   (c) Weisen Sie nach, dass die Varianz von $\tilde\sigma^2_\rho$ die Cramér-Rao-Schranke bei bekanntem ρ erreicht.

   (d*) Bestimmen Sie die Cramér-Rao-Schranke für σ^2 im Fall von unbekanntem ρ (Einsatz von Mathematik-Software bei Angabe von Zwischenergebnissen gestattet).

3. Δ-Methode. Für Zufallsvariablen (X_n) und $x \in \mathbb{R}$, $\sigma > 0$ gelte

$$\sqrt{n}(X_n - x) \xrightarrow{d} N(0, \sigma^2).$$

   Zeigen Sie für eine bei x differenzierbare reellwertige Funktion f:

   (a) Es gilt $X_n \to x$ und $|f(X_n) - f(x) - f'(x)(X_n - x)|/|X_n - x| \to 0$ jeweils in stochastischer Konvergenz.

   (b) Schließen Sie, z.B. mittels dem Slutsky-Lemma aus der Literatur, dass $\sqrt{n}|f(X_n) - f(x) - f'(x)(X_n - x)| \to 0$ stochastisch.

   (c) Wiederum durch Anwendung des Slutsky-Lemmas folgt

$$\sqrt{n}\big(f(X_n) - f(x)\big) \xrightarrow{d} N(0, \sigma^2 f'(x)^2).$$

4. Momentenschätzer. Es sei $X_1, \ldots, X_n$ eine $\mathrm{Exp}(\lambda)$-verteilte mathematische Stichprobe mit $\lambda > 0$ unbekannt.

   (a) Zeigen Sie $m_k(\lambda) := \mathbb{E}[X_1^k] = \lambda^{-k} k!$, $k \in \mathbb{N}$, und schließen Sie aus dem zentralen Grenzwertsatz $\sqrt{n}(M_{k,n} - m_k(\lambda)) \xrightarrow{d} N(0, \sigma_k^2)$ mit $M_{k,n} := \frac{1}{n}\sum_{i=1}^n X_i^k$, $\sigma_k^2 = ((2k)! - (k!)^2)\lambda^{-2k}$.

   (b) Der *k-te Momentenschätzer* von λ ist definiert durch

$$\hat{\lambda}_{k,n} := m_k^{-1}(M_{k,n}).$$

   Beweisen Sie mit der Δ-Methode, dass für $n \to \infty$ gilt

$$\sqrt{n}(\hat{\lambda}_{k,n} - \lambda) \xrightarrow{d} N\Big(0, \lambda^2 k^{-2}\big((2k)!/(k!)^2 - 1\big)\Big).$$

   Für welches Moment k ist die asymptotische Varianz am Kleinsten?

Abgabe vor der Übung am Dienstag, dem 12.6.18.

Vorlesung *Mathematische Statistik*
Sommersemester 2018
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 $X_1, \ldots, X_n$ 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 $\vartheta \in \Theta = [0,1]$ 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 $X_{(n)} = \max\{X_1, \ldots, X_n\}$ 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 $(\mathcal{X}, \mathscr{F}, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$ ein statistisches Modell mit Likelihoodfunktion $L(\vartheta)$ und MLE $\hat\vartheta$ sowie $g : \Theta \to \Theta'$ eine injektive Funktion.

   (a) Zeigen Sie, dass $\hat\vartheta' := g(\hat\vartheta)$ MLE ist für das reparametrisierte Modell mit Verteilungen $(\mathbb{P}_{g^{-1}(\vartheta')})_{\vartheta'\in g(\Theta)}$ (*Plug-in-Prinzip*)

   (b) Betrachten Sie nun $(\mathbb{R}^n, \mathfrak{B}_{\mathbb{R}^n}, (N(\mu, 1)^{\otimes n})_{\mu\in\mathbb{R}})$ mit MLE $\hat\mu_n = \bar X$. Zeigen Sie für $g(\mu) = \mu^{2k+1}$, $k \in \mathbb{N}$, dass $\widehat{\mu^{2k+1}} = (\hat\mu_n)^{2k+1}$ MLE, aber nicht erwartungstreu ist. Finden Sie durch Biaskorrektur von $\widehat{\mu^3}$ einen erwartungstreuen Schätzer $\widetilde{\mu^3}$ von μ^3, der einen kleineren quadratischen Fehler als $\widehat{\mu^3}$ besitzt und UMVU ist. Vergleichen Sie asymptotisch für $n \to \infty$ die quadratischen Fehler von $\widehat{\mu^3}$, $\widetilde{\mu^3}$ und die entsprechende Cramér-Rao-Schranke.

4. Betrachten Sie für $\vartheta \in \Theta := C([0, 1])$ das Regressionsmodell

$$Y_i = \vartheta(i/n) + \varepsilon_i, \quad i = 1, \ldots, n,$$

   mit ε_i i.i.d., $\mathbb{E}[\varepsilon_i] = 0$, $\mathbb{E}[\varepsilon_i^4] < \infty$. Es bezeichne $\Gamma = \{\sum_{k=0}^d c_k x^k\} \subseteq \Theta$ die Menge aller Polynome vom Grad kleiner oder gleich d sowie $g(\vartheta) = \Pi_\Gamma\vartheta$ die Orthogonalprojektion von Θ auf Γ in $L^2([0, 1])$.

   (a) Zeigen Sie, dass $K_n(\gamma) := \frac{1}{n}\sum_{i=1}^n (Y_i - \gamma(i/n))^2$, $\gamma \in \Gamma$, ein Kontrastprozess ist zur Kontrastfunktion

$$K(\vartheta_0, \gamma) = \|\vartheta_0 - \Pi_\Gamma\vartheta_0\|^2_{L^2([0,1])} + \|\Pi_\Gamma\vartheta_0 - \gamma\|^2_{L^2([0,1])} + \mathbb{E}[\varepsilon_i^2].$$

   (b) Schließen Sie, dass der Kleinste-Quadrate-Schätzer der Polynomregression $\hat\gamma_n = \operatorname{argmin}_{\gamma\in\Gamma} \sum_{i=1}^n (Y_i - \gamma(i/n))^2$ Minimum-Kontrast-Schätzer ist und gegen die Orthogonalprojektion $g(\vartheta_0)$ $\mathbb{P}^n_{\vartheta_0}$-stochastisch konvergiert. *Hinweis:* Betrachten Sie zunächst eine kompakte Umgebung $\Gamma_0 \subseteq \Gamma$ von $g(\vartheta_0)$ und zeigen Sie dann $\mathbb{P}^n_{\vartheta_0}(\hat\gamma_n \notin \Gamma_0) \to 0$.

   (c*) Weisen Sie $\sqrt{n}(\hat\gamma_n - g(\vartheta_0)) \xrightarrow{d} N(0, V^{-1}UV^{-1})$ unter $\mathbb{P}^n_{\vartheta_0}$ nach mit $U = 2\,\mathbb{E}[\varepsilon_i^2]V$, $V = 2(\int_0^1 x^k x^l dx)_{0\leqslant k,l\leqslant d}$ (in der Basis der Monome).

Abgabe vor der Übung am Dienstag, dem 19.6.18.

Vorlesung *Mathematische Statistik*
Sommersemester 2018
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

9. 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. Es sei $v : (-1,1) \to H$ eine Hilbertraum-wertige Abbildung mit $\|v(t)\| = 1$ für alle $t \in (-1,1)$, die bei $t = 0$ differenzierbar sei mit Ableitung $\dot{v}(0)$. Zeigen Sie:

   (a) Es gilt $0 = 2\langle v(0), v(t) - v(0)\rangle + \|v(t) - v(0)\|^2$ für alle $t \in (-1,1)$, und dies impliziert $\langle \dot{v}(0), v(0)\rangle = 0$ sowie
   $$\langle v(t), v(0)\rangle = 1 - \tfrac{1}{2}\|\dot{v}(0)\|^2 t^2 + o(t^2) \text{ für } t \to 0.$$
   Insbesondere ist $f(t) = \langle v(t), v(0)\rangle$ zweimal differenzierbar bei $t = 0$ mit $f''(0) = -\|\dot{v}(0)\|^2$.

   (b) Ist $L(\vartheta)$ die Likelihoodfunktion in einem bei $\vartheta_0 \in \operatorname{int}(\Theta) \subseteq \mathbb{R}$ Hellinger-differenzierbaren Modell $(\mathcal{X}, \mathscr{F}, (\mathbb{P}_\vartheta)_{\vartheta \in \Theta})$, so folgt mit $v(t) = \sqrt{L(\vartheta_0 + t)}$ und $\|\dot{v}(0)\|^2 = \frac{1}{4} I(\vartheta_0)$ für den quadrierten Hellingerabstand:
   $$\tfrac{d}{d\vartheta} H^2(\mathbb{P}_{\vartheta_0}, \mathbb{P}_\vartheta)|_{\vartheta=\vartheta_0} = 0, \quad \tfrac{d^2}{d\vartheta^2} H^2(\mathbb{P}_{\vartheta_0}, \mathbb{P}_\vartheta)|_{\vartheta=\vartheta_0} = \tfrac{1}{2} I(\vartheta_0).$$

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

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

   Zeigen Sie die Gleichung
   $$J(\vartheta_{j+1}, \vartheta_j) = \log\left(\frac{f_Y(y, \vartheta_{j+1})}{f_Y(y, \vartheta_j)}\right) + \int \log\left(\frac{f_{Z|Y=y}(z, \vartheta_{j+1})}{f_{Z|Y=y}(z, \vartheta_j)}\right) f_{Z|Y=y}(z, \vartheta_j) \nu(\,dz)$$
   und folgern Sie, dass im EM-Algorithmus $L(\vartheta_{j+1}) \geqslant L(\vartheta_j)$ gilt, d.h. die Likelihood in jedem Schritt größer wird.

4. Betrachten Sie eine mathematische Stichprobe $Y_1, \ldots, Y_n$, 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 $(Y_1, Z_1), \ldots, (Y_n, Z_n)$ unabhängig. μ sei das Lebesguemaß auf $\mathbb{R}^n$ und ν sei die Gleichverteilung auf $\{0, 1\}^n$.

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

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

   (b) Zeigen Sie, dass im EM-Algorithmus für $\vartheta_j = (a_j, b_j)$ gilt

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

   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 Dienstag, dem 26.6.18.

Vorlesung *Mathematische Statistik*
Sommersemester 2018
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

10. Übungsblatt

1. Ein Zufallszahlengenerator erzeugt bei 1 000 Aufrufen die Ziffern $0, 1, \dots, 9$ mit folgenden Häufigkeiten:

Ziffer	0	1	2	3	4	5	6	7	8	9
Anzahl	123	91	105	83	89	110	108	85	90	116

Testen Sie mittels χ^2-Test die Nullhypothese, dass eine Multinomialverteilung mit Zifferwahrscheinlichkeiten $1/10$ vorliegt, zum asymptotischen Niveau $\alpha = 0.05$.

2. *Supereffizienz.* Für eine $N(\mu, 1)$-verteilte Stichprobe $X_1, \dots, X_n$ sei

$$\hat{\mu}_n := \begin{cases} \bar{X}, & \text{falls } |\bar{X}| > n^{-1/4} \\ a\bar{X}, & \text{falls } |\bar{X}| \leqslant n^{-1/4} \end{cases}$$

der sogenannte *Hodges-Schätzer* mit $|a| < 1$. Zeigen Sie

$$\sqrt{n}(\hat{\mu}_n - \mu) \xrightarrow{d} N(0, v(\mu))$$

mit $v(\mu) = 1$ für $\mu \neq 0$ und $v(0) = a^2 < 1$, während die Fisher-Information bei n Beobachtungen $I_n(\mu) = n$ für alle $\mu \in \mathbb{R}$ beträgt.

3. Zeigen Sie, dass für $\mathbb{Q}_n = N(\mu_n, 1)$ und $\mathbb{P}_n = N(0, 1)$ Kontiguität $\mathbb{Q}_n \lhd \mathbb{P}_n$ vorliegt, falls $(\mu_n) \subseteq \mathbb{R}$ eine beschränke Folge ist. Gibt es auch eine unbeschränkte Folge (μ_n), so dass $\mathbb{Q}_n \lhd \mathbb{P}_n$ gilt?

4. Es sei $(\mathcal{X}^n, \mathcal{F}^{\otimes n}, (P_\vartheta^{\otimes n})_{\vartheta \in \Theta})$ ein bei $\vartheta_0 \in \operatorname{int}(\Theta)$ Hellinger-differenzierbares statistisches Modell mit score-Funktion $\dot{\ell}(\vartheta_0)$ und Fisher-Information $I(\vartheta_0)$, so dass der Maximum-Likelihood-Schätzer der Entwicklung

$$\sqrt{n}(\hat{\vartheta}_n(x) - \vartheta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} I(\vartheta_0)^{-1} \dot{\ell}(\vartheta_0, x_i) + o_{P_{\vartheta_0}^{\otimes n}}(1)$$

genügt. Zeigen Sie, dass unter $P_{\vartheta_0^n}^{\otimes n}$ mit $\vartheta_0^n := \vartheta_0 + n^{-1/2}h,\ h \in \mathbb{R}^k$, folgende Konvergenz gilt:

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_0^n) \to \mathcal{N}(0, I(\vartheta_0)^{-1}).$$

Abgabe vor der Übung am Dienstag, dem 3.7.18.

Vorlesung *Mathematische Statistik*
Sommersemester 2018
Humboldt-Universität zu Berlin

Prof. Dr. Markus Reiß

11. Übungsblatt

1. Es sei $X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ eine mathematische Stichprobe und $\hat{\sigma}_\alpha^2 := \alpha \sum_{i=1}^n (X_i - \bar{X})^2$ mit $\alpha > 0$ Schätzer von σ^2. Zeigen Sie, dass $\hat{\sigma}_\alpha^2$ für $\alpha^* = \frac{1}{n+1}$ minimales Risiko unter quadratischem Verlust besitzt.

2. Weisen Sie nach, dass die Familie $(\mathrm{Beta}(\alpha, \beta))_{\alpha,\beta>0}$ der Beta-Verteilungen eine Exponentialfamilie bildet, und finden Sie eine suffiziente und vollständige Statistik für $\vartheta = (\alpha, \beta)^\top$.

3. Prüfen Sie, ob folgende Modelle der Form $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), (P_\vartheta)_{\vartheta>0})$ Hellinger-differenzierbar sind und geben Sie gegebenfalls die score-Funktion $\dot{\ell}(\vartheta)$ sowie die Fisher-Information $I(\vartheta)$ an.

   (a) $P_\vartheta \sim \delta_\vartheta$, wobei δ_x das Dirac-Maß in x bezeichnet.

   (b) $P_\vartheta \sim \mathrm{Exp}(\vartheta)$.

4. Begründen Sie, warum ein Bayestest zum Niveau $\alpha \in (0, 1)$ minimalen Fehler 2. Art in der Klasse aller Tests zum Niveau α besitzt. Sie dürfen sich auf nichtrandomisierte Tests beschränken.

5. Sei $X_1, \dots, X_n \sim \mathrm{Poisson}(\lambda)$ eine mathematische Stichprobe. Konstruieren Sie auf Basis des MLE $\hat{\lambda}_n$ von λ einen konsistenten Schätzer $\hat{\vartheta}_n$ von $\vartheta = \sqrt{\lambda}$, so dass die asymptotische (reskalierte) Varianz von $\hat{\vartheta}_n$ nicht von λ abhängt.

6. Zeigen Sie: Genügt für $\vartheta_0, h \in \mathbb{R}^k$ die Zufallsgröße $\log(dP^{\otimes n}_{\vartheta_0+n^{-1/2}h}/dP^{\otimes n}_{\vartheta_0})$ unter $P^{\otimes n}_{\vartheta_0}$ einer LAN-Entwicklung, so gilt $(P^{\otimes n}_{\vartheta_0+n^{-1/2}h})_{n\geqslant 1} \triangleleft (P^{\otimes n}_{\vartheta_0})_{n\geqslant 1}$.

Abgabe vor der Übung am Dienstag, dem 10.7.18.