Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

1. Übungsblatt

1. Ist $A \in \mathbb{R}^{n \times k}$, dann nennt man $B \in \mathbb{R}^{k \times n}$ eine *Moore-Penrose-Inverse* von A, wenn gilt:

   - $ABA = A$ und $BAB = B$,
   - AB und BA sind symmetrisch.

   (a) Sei $k \leqslant n$ und $A \in \mathbb{R}^{n \times k}$ eine Matrix mit vollem Rang k. Zeige, dass $A^\top A$ invertierbar ist und dass $(A^\top A)^{-1} A^\top$ eine Moore-Penrose-Inverse von A ist.

   (b) Sei $A \in \mathbb{R}^{n \times k}$ und $b \in \mathbb{R}^n$. Sei A^+ eine Moore-Penrose-Inverse von A. Zeige: Wenn das Gleichungssystem $Ax = b$ lösbar ist, dann ist $A^+ b$ eine Lösung und hat unter allen Lösungen die kleinste euklidische Norm.

2. Formuliere und beweise den Satz von Gauß-Markov für das lineare Modell mit allgemeiner Kovarianzmatrix $\Sigma > 0$.

3. Die F-Verteilung mit (m, n) Freiheitsgraden auf $(\mathbb{R}, \mathfrak{B}_\mathbb{R})$ ist gegeben durch die Lebesguedichte

$$f_{m,n}(x) = \frac{m^{m/2} n^{n/2}}{B(m/2, n/2)} \frac{x^{m/2-1}}{(mx+n)^{(m+n)/2}} \mathbf{1}_{\mathbb{R}^+}(x), \qquad x \in \mathbb{R},$$

   wobei $B(z, w) := \Gamma(z)\Gamma(w)/\Gamma(z+w)$ für $z, w > 0$ die *Betafunktion* ist. Es seien U und V unabhängige Zufallsvariablen mit den Lebesguedichten f_U bzw. f_V. Bezeichne $f_{|U|}$ die Lebesguedichte von $|U|$. Es seien $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ unabhängige $N(0,1)$-verteilte Zufallsvariablen. Zeige folgende Aussagen:

   (a) Die Zufallsvariable U^2 hat die Lebesguedichte $f_{U^2}(x) = f_{|U|}(\sqrt{x})/2\sqrt{x}$ für $x > 0$ und $f_{U^2}(x) = 0$ sonst.

   (b) Benutze $\Gamma(1/2) = \sqrt{\pi}$ und $B(z, w) = \int_0^1 t^{z-1}(1-t)^{w-1} \mathrm{d}t$, um zu zeigen, dass $X = \sum_{i=1}^m X_i^2$ die Lebesguedichte

$$f_X(x) := \frac{x^{m/2-1}}{\Gamma(m/2) 2^{m/2}} e^{-x/2} \mathbf{1}_{\mathbb{R}^+}(x), \qquad x \in \mathbb{R},$$

   der $\chi^2(m)$-Verteilung hat.

   (c) Die Zufallsvariable $W := U/V$ ist fast sicher definiert und hat die Lebesguedichte $f_W(x) = \int_{-\infty}^{\infty} f_U(xy) f_V(y) |y| \mathrm{d}y$.

(d) Es ist

$$F_{m,n} := \frac{\frac{1}{m}\sum_{i=1}^{m} X_i^2}{\frac{1}{n}\sum_{j=1}^{n} Y_j^2}$$

gemäß einer F-Verteilung mit (m,n) Freiheitsgraden verteilt.

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) Modelliere dies als ein lineares Modell und bestimme die Regressionsgerade. Zeichne Messwerte 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. Bestimme die Regressionsgraden für beide Studienfächer getrennt und zeichne sie ein.

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

Abgabe vor der Vorlesung am Freitag, den 23.04.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

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 $f_{X|T=\vartheta}(\bullet)$ bzw. $f_T(\bullet)$. Zeige für $(\mathcal{F} \otimes \mathcal{F}_\Theta)$-messbare Funktionen $f_{X|T} : \mathcal{X} \times \Theta \to [0,\infty), (x,\vartheta) \to f_{X|T=\vartheta}(x)$:

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

   (b) Es seien X und T 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 $\mathbb{E}_{\tilde{\mathbb{P}}}[|g(X,T)|] < \infty$, dann gilt
   $$\mathbb{E}_{\tilde{\mathbb{P}}}[g(X,T)|X=x] = \int_\Theta g(x,\vartheta) f_{T|X=x}(\vartheta)\nu(\mathrm{d}\vartheta)$$
   und speziell $\int_A f_{T|X=x}(\vartheta)\nu(\mathrm{d}\vartheta) = \mathbb{E}_{\tilde{\mathbb{P}}}[\mathbf{1}_{\{T\in A\}}|X=x] =: \tilde{\mathbb{P}}(T \in A|X=x)$.

2. Beweise 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 $R_\pi(\rho') \le R_\pi(\rho)$ ist $\vartheta \mapsto R(\vartheta, \rho')$ stetig.

3. 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. Formuliere dies als Bayessches Entscheidungsproblem und gib eine Bayes-optimale Entscheidungsregel in Abhängigkeit von ℓ_0, ℓ_1 an.

4. 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) Skizziere $f_{a,b}$ 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. Zeige, dass die bedingte Dichte von p gegeben $X = x$ zur Beta-Verteilung $B(a + x, b + n - x)$ gehört.

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

Abgabe vor der Vorlesung am Freitag, den 30.04.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

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. $l(\vartheta, a) = |a-\vartheta|^2$). Zeige: ρ kann nur dann erwartungstreu sein, wenn $R_\pi(\rho) = 0$.

2. Gegeben sei das gewöhnlich lineare Modell $Y = X\beta + \epsilon$ mit der Kovarianzmatrix $\Sigma = \sigma^2 E_n$. In der *ridge regression* verwendet man den Schätzer $\hat\beta_a = (X^\top X + a^2 E_k)^{-1} X^\top Y$. Die a-priori-Verteilung π von β sei eine zentrierte Normalverteilung mit Varianz $\eta^2 E_k$. Zeige: Für quadratisches Risiko ist der Bayes-optimale Schätzer $\hat\beta_\pi$ gleich dem ridge-regression-Schätzer $\hat\beta_{\frac{\sigma}{\eta}}$.

3. Wenn man in die Bayesformel statt einer Dichte $f_T(\vartheta)$ eine nichtnegative, messbare Funktion $f_T(\vartheta)$ einsetzt und $f_{T|X=x}(\vartheta)$ 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) Zeige: $\bar X := \frac{1}{n}\sum_{i=1}^n X_i$ ist ein verallgemeinerter Bayesschätzer von μ zum quadratischen Risiko bzgl. dem Lebesguemaß als verallgemeinerter a-priori-Verteilung.

   (b) Berechne den verallgemeinerten Bayesschätzer $\hat\mu_{a,b}$ zum quadratischen Risiko für $d = 1$ und $f_T(\vartheta) = \mathbf{1}_{(a,b)}(\vartheta)$ mit $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$. Zeichne $\hat\mu_{0,1}$ für $n = 1$ als Funktion von $\bar X$.

4. Es sei $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$ ein statistischen Experiment und μ ein σ-endliches Maß auf $(\mathcal{X}, \mathcal{F})$, so dass $L^1(\mathcal{X}, \mathcal{F}, \mu)$ separabel ist und $\mathbb{P}_\vartheta \ll \mu$ für alle $\vartheta \in \Theta$. $\Phi = \{\rho : \mathcal{X} \to [0,1] |\ \rho \text{ messbar}\}$ sei die Menge aller Tests.

   (a) Schließe mit dem Satz von Alaoglu aus der Funktionalanalysis, dass Φ *schwach folgenkompakt* ist, d. h. es gibt zu jeder Folge $(\rho_n)_n \subseteq \Phi$ eine Teilfolge $(\rho_{n_k})_k$ und ein $\rho' \in \Phi$ mit

   $$\int_{\mathcal{X}} \rho_{n_k}(x) f(x) \mu(\,\mathrm{d}x) \to \int_{\mathcal{X}} \rho'(x) f(x) \mu(\,\mathrm{d}x) \qquad \forall f \in L^1(\mathcal{X}, \mathcal{F}, \mu).$$

   (b) Für ein $\vartheta_1 \in \Theta$ definiere $\Theta_0 := \Theta\backslash\{\vartheta_1\}$. Es gelte $\mathbb{P}_{\vartheta_1} \notin \{\mathbb{P}_\vartheta\,|\vartheta \in \Theta_0\}$. Sei Φ_α die Menge der α-Niveau-Tests für $\mathrm{H}_0 : \vartheta \in \Theta_0$ gegen $\mathrm{H}_1 : \vartheta = \vartheta_1$. Dann existiert ein bester α-Niveau Test $\tilde\rho$, d. h. $\tilde\rho \in \Phi_\alpha$ und $\mathbb{E}_{\vartheta_1}[\tilde\rho] = \sup_{\rho\in\Phi_\alpha} \mathbb{E}_{\vartheta_1}[\rho]$.

Hinweis: Ist X ein normierter Raum, so ist die Einheitskugel $B_{X'}$ des Dualraums X' kompakt bezüglich der schwach-*-Topologie (Satz von Alaoglu). Ist X ein separabler normierter Raum, so ist die schwach-*-Topologie auf $B_{X'}$ metrisierbar.

Abgabe vor der Vorlesung am Freitag, den 07.05.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

4. Übungsblatt

1. Es sei $X_1, \dots, X_n$ eine $N(\mu, E_d)$-verteilte mathematische Stichprobe. Der James-Stein-Schätzer mit positivem Gewicht ist definiert als $\hat{\mu}_{JS+} = \big(1 - \frac{d-2}{n|\overline{X}|^2}\big)_+ \overline{X}$. Beweise 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 $i = 1, \dots, d$ 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 $((\overline{X}_1)^2, \dots, (\overline{X}_d)^2)$ bedingten Erwartung.

2. Gegeben sei $X \sim N(\mu, \sigma^2 E_d)$ mit $\sigma > 0$ bekannt und $\mu \in \mathbb{R}^d$ unbekannt.
   (a) Zeige: 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) Berechne die Bias-Varianz-Zerlegung für $\hat{\mu}_\alpha = \alpha X$, $\alpha \in \mathbb{R}$, und zeige, dass $\alpha_{\text{Orakel}} := 1 - \frac{\sigma^2 d}{|\mu|^2 + \sigma^2 d}$ das quadratische Risiko minimiert, falls μ der wahre Parameter ist.
   (c) Wähle $R > 0$. Weise nach, dass $|X|^2$ ein erwartungstreuer Schätzer von $|\mu|^2 + \sigma^2 d$ ist und setze $\hat{\alpha} := 1 - \frac{\sigma^2 d}{|X|^2}$. Schließe durch Berechnen von $\operatorname{Var}(|X|^2)$, dass $\forall \epsilon > 0\ \exists K > 0$:
$$\mathbb{P}_\mu\left(\left|\frac{|X|^2}{\sigma^2 d} - \frac{|\mu|^2 + \sigma^2 d}{\sigma^2 d}\right| \geqslant \frac{K}{\sqrt{d}}\right) \leqslant \epsilon, \quad \forall d \geqslant 1\ \forall \mu \in \mathbb{R}^d \text{ mit } |\mu| \leqslant R.$$
   Weise nach, dass $\forall \epsilon > 0\ \exists K' > 0$:
$$\mathbb{P}_\mu(|\hat{\alpha} - \alpha_{\text{Orakel}}| > K'\, d^{-1/2}) \leqslant \epsilon, \quad \forall d \geqslant 1\ \forall \mu \in \mathbb{R}^d \text{ mit } |\mu| \leqslant R.$$
   Folgere, dass insbesondere für $|\mu| \leqslant R$ die Normen $|\hat{\alpha} - \alpha_{\text{Orakel}}|$ für $d \to \infty$ stochastisch gegen 0 konvergieren.

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

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

   (b) Bestimme 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 $X = x$)?

4. Beweise oder widerlege die Aussage, dass folgende Verteilungen Exponentialfamilien bilden. Bestimme gegebenenfalls den natürlichen Parameterraum.

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

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

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

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

Abgabe vor der Vorlesung am Freitag, den 14.05.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

5. Übungsblatt

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

2. Beweise: 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 $i, j = 1, \ldots, k$ folgt $\mathbb{E}_{\bar{\vartheta}}[T_i] = \frac{dA}{d\vartheta_i}(\bar{\vartheta})$ und $\operatorname{Cov}_{\bar{\vartheta}}(T_i, T_j) = \frac{d^2 A}{d\vartheta_i d\vartheta_j}(\bar{\vartheta})$.

3. 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. Beweise, 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.

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

   (a) Bestimme 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 $X_t := \sigma B_t + at$. Bestimme die Likelihoodfunktion bezüglich $\mathbb{P}_{0,1}$ und weise nach, dass $(X_T, \sum_{k=1}^n (\Delta X_k)^2)$ eine suffiziente Statistik ist.

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

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

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 der Aufgaben 1-3 vor der Vorlesung am Freitag, den 21.05.10,
Abgabe von Aufgabe 4 am Freitag, den 28.05.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

6. Übungsblatt

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

2. Weise die Bedingungen der Cramér-Rao-Ungleichung für eine mathematische Stichprobe $X_1, \dots, X_n$ der Doppelexponentialverteilung mit Lebesguedichte $\frac{1}{2}e^{-|x-\mu|}$, $\mu \in \mathbb{R}$ unbekannt, und den Schätzer $\hat{\mu}_1 := \frac{1}{n}\sum_{k=1}^{n} X_k$ von μ nach. Bestimme die Fisher-Information und überprüfe, ob Gleichheit in der Cramér-Rao-Schranke für $\hat{\mu}_1$ gilt.
   Zeige, dass der Stichprobenmedian $\hat{\mu}_2$ den Wert μ ebenfalls erwartungstreu schätzt. Simuliere $\hat{\mu}_1$ und $\hat{\mu}_2$ in 1000 Monte-Carlo-Iterationen für $n = 100$ Beobachtungen zum Wert $\mu = 0$ und vergleiche die Monte-Carlo-Varianzen von $\hat{\mu}_1$ und $\hat{\mu}_2$.

Abgabe vor der Übung am Freitag, den 28.05.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

7. Übungsblatt

1. Es sei $X_1, \dots, X_n$ eine mathematische Stichprobe mit Werten in $\mathbb{R}$. Es seien $\vartheta_0 \in \Theta$, $g : \Theta \to \mathbb{R}$ und $k \in \mathbb{N}$, so dass $\operatorname{Var}_{\vartheta_0}(X_1^k)$ endlich ist und $\varphi(\vartheta) := E_\vartheta[X_1^k]$ für alle $\vartheta \in \Theta$ endlich ist. Es gebe eine Borel-messbare Funktion $G : \varphi(\Theta) \to g(\Theta)$ mit $G \circ \varphi = g$. Zeige für einen inneren Punkt $\varphi(\vartheta_0)$ von $\varphi(\Theta)$:

   (a) $\frac{1}{n}\sum_{i=1}^n X_i^k$ liegt bezüglich $\mathbb{P}_{\vartheta_0}$ für $n \to \infty$ mit gegen Eins gehender Wahrscheinlichkeit in $\varphi(\Theta)$. Ist G stetig in $\varphi(\vartheta_0)$, dann konvergiert der *Momenentenschätzer* $\hat{g}_n = G(\frac{1}{n}\sum_{i=1}^n X_i^k)$ $\mathbb{P}_{\vartheta_0}$-f.s. gegen $g(\vartheta_0)$.

   (b) Ist G in $\varphi(\vartheta_0)$ differenzierbar mit $\sigma^2 := \operatorname{Var}_{\vartheta_0}(X_1^k)(G'(\varphi(\vartheta_0)))^2 > 0$, dann ist $\hat{g}_n$ unter $\mathbb{P}_{\vartheta_0}$ asymptotisch normalverteilt mit asymptotischer Varianz σ^2:
   $$\sqrt{n}(\hat{g}_n - g(\vartheta_0)) \xrightarrow{d} N(0, \sigma^2).$$

2. Es seien $X_1, \dots, X_n \sim \operatorname{Exp}(\lambda)$ eine mathematische Stichprobe mit $\lambda > 0$ unbekannt. Zeige:

   (a) Für $k \in \mathbb{N}$ ist der Momentenschätzer von λ gegeben durch
   $$\hat{\lambda}_{k,n} = \left(\frac{k!}{\frac{1}{n}\sum_{i=1}^n X_i^k} \right)^{1/k}.$$

   (b) Die Momentenschätzer $\hat{\lambda}_{k,n}$ sind asymptotisch normalverteilt mit asymptotischer Varianz $\sigma_k^2 = \lambda_0^2 k^{-2}((2k)!/(k!)^2 - 1)$. Für welches $k \in \mathbb{N}$ ist die asymptotische Varianz minimal?

Abgabe vor der Vorlesung am Freitag, den 04.06.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

8. Übungsblatt

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

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

Finde 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) Beweise für Produktmaße:

$$\mathrm{KL}(\mathbb{P}_1 \otimes \mathbb{P}_2 \mid \mathbb{Q}_1 \otimes \mathbb{Q}_2) = \mathrm{KL}(\mathbb{P}_1 \mid \mathbb{Q}_1) + \mathrm{KL}(\mathbb{P}_2 \mid \mathbb{Q}_2).$$

2. (a) Zeige: Bildet $(\mathbb{P}_\vartheta)_{\vartheta\in\Theta}$ eine natürliche Exponentialfamilien und ist ϑ_0 innerer Punkt von Θ, so gilt $\mathrm{KL}(\mathbb{P}_{\vartheta_0} \mid \mathbb{P}_\vartheta) = A(\vartheta) - A(\vartheta_0) + \langle \dot{A}(\vartheta_0), \vartheta_0 - \vartheta\rangle$. Folgere

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

(b) Finde allgemeine Voraussetzungen, so dass folgende Gleichungen gelten:

$$\dot{\mathrm{KL}}(\mathbb{P}_{\vartheta_0} \mid \mathbb{P}_\vartheta)\Big|_{\vartheta=\vartheta_0} = 0, \quad \ddot{\mathrm{KL}}(\mathbb{P}_{\vartheta_0} \mid \mathbb{P}_\vartheta)\Big|_{\vartheta=\vartheta_0} = -\int \ddot{\ell}(\vartheta_0)\,\mathrm{d}\,\mathbb{P}_{\vartheta_0}\,.$$

3. (a) Zeige für eine mathematische Stichprobe $X_1,\ldots,X_n$ 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.

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

4. Betrachte eine mathematische Stichprobe $X_1,\ldots,X_n$ mit Lebesguedichte $f_{\mu,\sigma}(x) = \sigma^{-1}f((x-\mu)/\sigma)$, $f \in C^1(\mathbb{R})$, mit $(\mu,\sigma) \in \Theta \subseteq \mathbb{R}\times\mathbb{R}^+$ (*Lokations-Skalen-Familie*). Bestimme die Fisher-Information für die Fälle, dass (a) f bekannt und μ, σ unbekannt sowie (b) f, σ bekannt und μ unbekannt sind. Welche Annahmen garantieren, dass die jeweiligen MLE asymptotisch normalverteilt sind?

Abgabe vor der Vorlesung am Freitag, den 11.06.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

9. 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.
   Modelliere die Schätzung des Fischbestandes ϑ durch ein statistisches Experiment mit hypergeometrischen Verteilungen und bestimme einen MLE, diskutiere 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. 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. Zeige:

   (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

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

   (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 $\vartheta \in \Theta$. Beweise:

   (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 $\delta \to 0$.

   Schließe 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

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

   mit $\mathbb{E}[\varepsilon_i] = 0$, $\operatorname{Var}(\varepsilon_i) = \sigma^2$, $\mathbb{E}[\varepsilon_i^4] < \infty$, $\sigma > 0$ betrachte den Kleinste-Quadrate-Schätzer $\hat\vartheta_n = \operatorname{argmin}_{\vartheta\in\Theta} \sum_{i=1}^n (Y_i - g_\vartheta(i/n))^2$. Gib Voraussetzungen für die Parametrisierung $\vartheta \mapsto g_\vartheta$ an, um auf die asymptotische Normalität von $\hat\vartheta_n$ für $n \to \infty$ zu schließen und bestimme die asymptotische Varianz.

Abgabe vor der Vorlesung am Freitag, den 18.06.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

10. Übungsblatt

1. Im linearen Regressionsmodell der Beobachtungen

$$Y_i = g_\vartheta(i/n) + \varepsilon_i, \quad i = 1, \dots, n, \ (\varepsilon_i)_{1\leqslant i\leqslant n} \text{ iid},$$

mit $g_\vartheta(x) = \sum_{l=1}^k \vartheta_l \, g_l(x)$, $g_l \in C([0,1])$, $k < n$, $\mathbb{E}[\varepsilon_i] = 0$, $\operatorname{Var}(\varepsilon_i) = \sigma^2$, $\sigma > 0$ wird $\vartheta \in \Theta = \mathbb{R}^k$ geschätzt.

   (a) Schreibe dies unter einer Rangbedingung als ein gewöhnliches lineares Modell und bezeichne mit $\hat\vartheta$ den Kleinste-Quadrate-Schätzer.

   (b) Unter dem zugrundeliegenden Wahrscheinlichkeitsmaß $\mathbb{P}$ gilt allerdings $Y_i = g(i/n) + \varepsilon_i$ mit einer Funktion $g \in C([0,1])$. Bestimme $g_{\hat\vartheta}(i/n)$ und das quadratische Risiko $\mathbb{E}_\mathbb{P}[\|g_{\hat\vartheta} - g\|_n^2]$.

2.* In Aufgabe 1 setze (i) $g_\vartheta(x) = \vartheta_0 + \vartheta_1 x$ bzw. (ii) $g_\vartheta(x) = \vartheta_0 + \vartheta_1 x + \vartheta_2 x^2$ und $n = 10, 100, 1.000, 10.000$. Die Beobachtungen seien wie in Aufgabe 1(b) mit $g(x) = \sin(\pi x/2)$, $\varepsilon_i \sim N(0,1)$ iid, verteilt.

   (a) Zeichne jeweils ein Diagramm mit einer Realisierung von Y, $g_{\hat\vartheta}$ und g.

   (b) Bestimme jeweils das quadratische Monte-Carlo-Risiko, aufgeteilt nach Bias- und Varianzanteil, bei 10.000 Iterationen.

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

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

4. Betrachte 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) Bestimme $f_Y(y, a, b)$ für $n = 1$ und zeige 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) Zeige: 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*) Simuliere 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 Vorlesung am Freitag, den 25.06.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

11. Übungsblatt

1. Seien $\Theta \subseteq \mathbb{R}^k$ beschränkt und $\vartheta_0 \in \Theta$. Definiere $\varrho := \sup_{\vartheta \in \Theta} |\vartheta - \vartheta_0|$. Θ' heißt ε-Netz von Θ, wenn $\Theta' \subseteq \Theta$ und $\sup_{\vartheta \in \Theta} \inf_{\vartheta' \in \Theta'} |\vartheta - \vartheta'| \leqslant \varepsilon$ gilt. Zeige: Für alle $j \in \mathbb{N}_0$ gibt es $2^{-j}\varrho$-Netze Θ_j mit $\Theta_0 = \{\vartheta_0\}$, $\Theta_j \subseteq \Theta_{j+1}$ und $|\Theta_j| \leqslant C_1 2^{jk}$, wobei $C_1 > 0$ unabhängig von j und ϱ ist.

2. Es sei $(A_\vartheta, \vartheta \in \Theta)$ mit $\Theta \subseteq \mathbb{R}^k$ beschränkt ein stetiger Prozess und es möge die Ungleichung
$$\forall \vartheta, \vartheta' \in \Theta : \mathbb{P}(|A_\vartheta - A_{\vartheta'}| \geqslant r|\vartheta - \vartheta'|) \leqslant C\, r^{-p}, \quad r \in (0, \infty),$$
mit einer Potenz $p > k$ und einer Konstanten $C > 0$ gelten. Beweise, dass es für $\vartheta_0 \in \Theta$ eine Konstante $\tilde{C}$ gibt, so dass mit $\rho := \sup_{\vartheta \in \Theta} |\vartheta - \vartheta_0|$ gilt:
$$\mathbb{P}(\sup_{\vartheta \in \Theta} |A_\vartheta - A_{\vartheta_0}| \geqslant r\rho) \leqslant \tilde{C}\, r^{-p}, \quad r \in (0, \infty).$$

3. Im Gaußschen linearen Regressionsmodell der Beobachtungen
$$Y_i = g_\vartheta(i/n) + \varepsilon_i, \qquad i = 1, \dots, n, \quad \varepsilon \sim N(0, \sigma^2 E_n),$$
mit $g_\vartheta(x) = \sum_{l=1}^k \vartheta_l\, g_l(x)$, $g_l \in C([0,1])$, $n > k$, $\sigma > 0$ wird $\vartheta \in \Theta = \mathbb{R}^k$ geschätzt. Die (g_l) seien orthonormal bezüglich $\langle f, f' \rangle_n := (1/n) \sum_{i=1}^n f(i/n) f'(i/n)$.

   (a) Zeige, dass im falsch spezifizierten Modell die Darstellung $\|g_{\hat{\vartheta}} - g\|_n^2 = \inf_{\vartheta \in \Theta} \|g_\vartheta - g\|_n^2 + \frac{Z}{n}$ mit $Z := n \sum_{l=1}^k \langle \varepsilon, g_l \rangle_n^2$ gilt und dass $(\langle \varepsilon, g_l \rangle_n)_{l=1,\dots,k}$ unabhängige $N(0, \sigma^2/n)$-verteilten Zufallsvariablen sind.

   (b) Schließe, dass $U := (1/\sigma^2) Z$ eine $\chi^2(k)$-verteilt Zufallsvariable ist, und berechne $\mathbb{E}[e^{\alpha U}]$ für $\alpha < 1/2$.

   (c) Setze $\delta = 1/2 - \alpha$ und zeige für $\delta \in (0, 1/2)$ mit der verallgemeinerten Tschebyschew-Ungleichung $\mathbb{P}(Z \geqslant \kappa) \leqslant (2\delta)^{-k/2} \exp(-(1/2 - \delta)\kappa/\sigma^2)$, $\kappa > 0$.

   (d*) Bestimme numerisch oder analytisch den Wert von $\mathbb{P}(Z \geqslant \kappa)$ und vergleiche mit der Schranke.

4. Es soll getestet werden, ob ein Wahrsager verdeckt liegende Karten erkennen kann. Dem Wahrsager wird $n_1 = 20$ Mal rein zufällig entweder die Kreuz-Dame oder der Kreuz-König vorgelegt und er soll jeweils nur bei einer Kreuz-Dame die Karte aufdecken.

   (a) Gib ein statistisches Modell mittels einer Binomialverteilung $\mathrm{Bin}(n_1, p)$ an und zeige, dass $\varphi_1 = \mathbf{1}_{\{c_1,\dots,n_1\}}$ mit $c_1 = 15$ ein nichtrandomisierter Test von $H_0 : p = 1/2$ gegen $H_1 : p > 1/2$ zum Niveau $\alpha = 0,05$ ist.

   (b) Zeige, dass die erste Vermutung, bei $n_2 = 40$ müsse der Wahrsager entsprechend mindestens 30 Karten korrekt erkennen, falsch ist. Bestimme dazu einen Test $\varphi_2 = \mathbf{1}_{\{c_2,\dots,n_2\}}$ zum Niveau $\alpha = 0,05$ mit minimalem c_2.

   (c) Skizziere die *Gütefunktionen* $G_{\varphi_j}(p) = \mathbb{E}_p[\varphi_j]$, $p \in [1/2, 1]$, $j \in \{1, 2\}$, der beiden Tests.

Abgabe vor der Vorlesung am Freitag, den 02.07.10.

Markus Reiß, Jakob Söhl

Vorlesung *Mathematische Statistik*
Sommersemester 2010
Humboldt-Universität zu Berlin

12. Übungsblatt

1. Ist $(\mathcal{X}, \mathscr{F}, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$ ein statistisches Experiment, so ist ein Konfidenzbereich S zum Niveau $1-\alpha$ eine Abbildung $S : \mathcal{X} \to \mathscr{P}(\Theta)$, so dass die Ereignisse $\{x \in \mathcal{X} \,|\, \vartheta \in S(x)\}$ messbar sind und $\mathbb{P}_\vartheta(\vartheta \in S) \geqslant 1-\alpha$ für alle $\vartheta \in \Theta$ gilt. Zeige: jeder Konfidenzbereich S zum Niveau $1-\alpha$ generiert durch $\varphi_{\vartheta_0}(x) = 1 - \mathbf{1}_{S(x)}(\vartheta_0)$ einen nichtrandomisierten Test auf $H_0 : \vartheta = \vartheta_0$ zum Niveau α, und jede Familie $(\varphi_{\vartheta_0})_{\vartheta_0\in\Theta}$ von nichtrandomisierten Tests auf $H_0 : \vartheta = \vartheta_0$ zum Niveau α generiert einen Konfidenzbereich zum Niveau $1-\alpha$.

2. Betrachte 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 $H_1 : \vartheta = 1$. Zeige:

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

4. Beweise das verallgemeinerte Neyman-Pearson-Lemma aus der Vorlesung.

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

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

7.* 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. Konstruiere 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 Vorlesung am Freitag, den 09.07.10. Die mit * gekennzeichneten Aufgaben werden nur zum Erreichen des 50% Kriteriums korrigiert und können bis Mittwoch, den 14.07.10 abgegen werden.