Übungen zur Angewandten Stochastik

Blatt 8

1. Aufgabe Es wird vermutet, dass bei einem Pferderennen auf einem ovalen Parcours die Startposition einen Einfluß auf die Gewinnchance hat. In 144 Rennen hatten die Sieger die Startpositionen $1, 2, \ldots, 8$ mit den folgenden Häufigkeiten: $29, 19, 18, 25, 17, 10, 15, 11$. Testen Sie die Hypothese, dass die Startposition keinen Einfluß auf die Siegchance hat, zum Niveau 5%.

2. Aufgabe (χ^2-Anpassungstest als (asymptotischer) Likelihood-Quotiententest) Sei $s \in \mathbb{N}$, unter $P_\vartheta, \vartheta \in \Theta := (0,1)^s$ sei $h_n = (h_n(1), \ldots, h_n(s))$ Multinomial$(n, \vartheta_1, \ldots, \vartheta_s)$-verteilt. Um die Nullhypothese $\vartheta = \rho$ gegen die Alternative $\vartheta \in \Theta \setminus \{\rho\}$ für ein festes $\rho \in \Theta$ zu testen, haben wir in der Vorlesung den χ^2-Anpassungstest betrachtet, der auf der Statistik

$$D_n := \sum_{i=1}^{s} \frac{(h_n(i) - n\rho(i))^2}{n\rho(i)}$$

beruht (d.h. wir lehnen die Nullhypothese ab, wenn D_n einen gewissen Schwellwert c_n überschreitet). Ein (zunächst) anderer Ansatz für einen Test fußt auf dem Likelihood-Quotienten

$$R_n := \frac{\sup_{\vartheta \in \Theta} \prod_{i=1}^{s} \vartheta(i)^{h_n(i)}}{\prod_{i=1}^{s} \rho(i)^{h_n(i)}}$$

und lehnt die Nullhypothese ab, wenn R_n einen gewissen Schwellwert c'_n überschreitet. Was haben diese beiden Tests miteinander zu tun? Zeigen Sie dazu: $\log R_n = nH(\widetilde{h}_n|\rho)$, wo $\widetilde{h}_n(i) = h_n(i)/n$ die empirische Häufigkeitsverteilung ist, und

$$2nH(\widetilde{h}_n|\rho) - D_n \underset{n\to\infty}{\longrightarrow} 0 \quad \text{in Verteilung unter } P_\rho.$$

(Hinweis: Taylorentwicklung für $\psi(x) = 1 - x + x\log x$ um $x = 1$.)

3. Aufgabe Jede nichtnegativ definite $(n \times n)$-Matrix $C = (c_{ij})$ besitzt eine Darstellung der Form $C = ODO^t$ mit einer orthogonalen $(n \times n)$-Matrix O und einer $(n \times n)$-Diagonalmatrix $D = (d_{ij})$ mit nichtnegativen Einträgen (Hauptachsentransformation). Für eine ein solches C und $\mu \in \mathbb{R}^n$ sei $\mathcal{N}(\mu, C)$, die n-dimensionale Normalverteilung mit Mittelwertvektor μ und Kovarianzmatrix C, definiert als die Verteilung von

$$X := \mu + O\widetilde{Z},$$

wo $\widetilde{Z} = (\sqrt{d_{11}}Z_1, \ldots, \sqrt{d_{nn}}Z_n)^t$ mit $Z_1, \ldots, Z_n$ u.a., 1-dim. standard-normalverteilt.
a) Zeigen Sie: $\mathcal{N}(\mu, C)$ ist wohldefiniert, d.h. die Verteilung von X hängt nicht von der Wahl von O und D in der Zerlegung $C = ODO^t$ ab, und es gilt tatsächlich $\mathbb{E}\, X_i = \mu_i$, $\mathbb{E}\big[(X_i - \mu_i)(X_j - \mu_j)\big] = c_{ij}$. Wenn C invertierbar ist, so hat X die Dichte

$$\big((2\pi)^n \det(C)\big)^{-1/2} \exp\Big(-\frac{1}{2}(x-\mu)^t C^{-1}(x-\mu)\Big),$$

wenn Rang$(C) = m < n$, so ist die Verteilung von X auf einem m-dimensionalen affinen Teilraum des $\mathbb{R}^n$ konzentriert.
b) Sei $A = (a_{ij})$ irgendeine $(n \times n)$-Matrix, X wie oben, dann hat $Y := AX$ die Verteilung $\mathcal{N}(A\mu, ACA^t)$.

4. Aufgabe Kuss et al. ("The fouled player should not take the penalty himself": An empirical investigation of an old German football myth, *J. Sports Sciences* **25**, no. 9, 963–967) berichten über die Strafelfmeter in der 1. Fußballbundesliga (der Herren) von August 1993 bis Februar 1995:

	verwandelt	nicht verwandelt
der Gefoulte schießt selbst	74	28
anderer Spieler schießt	547	186

Stützen diese Daten die These, dass der Gefoulte den Elfmeter nicht selbst schießen sollte?