Prof. Dr. E. Spodarev
D. Meschenmoser

Wintersemester 2009 / 2010
19. November 2009

Statistik II – Übungsblatt 6

Abgabe: 26. November 2009, vor den Übungen

Aufgabe 1
In dieser Aufgabe soll der Datensatz `staub.txt` untersucht werden. Er enthält Daten aus einer Studie zum Zusammenhang zwischen Staubbelastung am Arbeitsplatz und Rauchgewohnheiten. Die Daten sind wie folgt bezeichnet:

staub:	Staubbelastung am Arbeitsplatz (in mg/m^3)
raucher:	Ist die Person Raucher? (1 = Ja, 0 = Nein)
dauer:	Dauer der Staubbelastung in Jahren

Es soll die Vermutung überprüft werden, ob die Staubbelastung am Arbeitsplatz eher zum Rauchen verleitet. Teste deshalb, ob die Raucher die gleiche Verteilung der Belastung haben wie die Nichtraucher gegen die Alternative, dass die Raucher eine andere Verteilung der Staubbelatsung haben als die Nichtraucher. Es sollen die Datensätze analysiert werden, die eine Belastung von mehr als 45 Jahren haben. Verwende dafür den Iterationstest von Wald-Wolfowitz. Implementiere den Test selbst. Gehe dabei wie folgt vor:

a) Implementiere die Verteilungsfunktion der Testgröße beim Iterationstest in R. (2)

b) Berechne mit R den Wert der Teststatistik für die gegebenen Daten. (2)

c) Führe in R einen exakten und einen asymptotischen Test zum Niveau $\alpha = 0.05$ durch. (2)

Hinweise:

- Zur Modulo-Rechnung kann in R der Befehl `%%` verwendet werden.
- Die Daten in dem Datensatz sind nach der Staubbelastung geordnet. Auch wenn die Zahlen auf zwei Nachkommastellen gleich sind, kannst Du davon ausgehen, dass sie sich in Wahrheit doch unterscheiden.

Aufgaben 2 (3)
Weise nach, dass die multivariate Normalverteilung faltungsstabil ist, d.h. zeige, dass für zwei unabhängige n-dimensionale Zufallsvektoren $X \sim \mathcal{N}(\mu_X, K_X)$ und $Y \sim \mathcal{N}(\mu_Y, K_Y)$ gilt: $X + Y \sim \mathcal{N}(\mu_X + \mu_Y, K_X + K_Y)$.

Aufgabe 3
Sei $X = (X_1, \dots, X_n)^\top \sim \mathcal{N}(\mu, K)$ mit $\mu \in \mathbb{R}^n$ und $K = (k_{ij})_{i,j=1,\dots,n} \in \mathbb{R}^{n \times n}$ symmetrisch und nichtnegativ definit.

a) Berechne den Erwartungswert und die Kovarianzmatrix von X. (3)

b) Zeige, dass ein beliebiger Teilvektor von X auch normalverteilt ist, wobei der Erwartungswertvektor der entsprechende Teilvektor von μ ist und die Kovarianzmatrix die entsprechende Teilmatrix von K. (2)

c) Zeige, dass für $i \neq j$ die Zufallsvariablen X_i und X_j genau dann unabhängig sind, wenn $k_{ij} = 0$ ist. (2)

d) Sei $\mu = 0$. Zeige, dass für beliebige $i, j, l, m \in \{1, \dots, n\}$ gilt: $\mathbb{E}(X_i X_j X_l) = 0$ und $\mathbb{E}(X_i X_j X_l X_m) = k_{ij}k_{lm} + k_{il}k_{jm} + k_{im}k_{jl}$. (3)

Hinweis: Für $j_1, \dots, j_n \in \mathbb{N}_0$ mit $r = j_1 + \dots + j_n$ gilt: $\mathbb{E}\left(X_1^{j_1} \cdots X_n^{j_n}\right) = i^{-r} \frac{\partial^r \varphi(t)}{\partial t_1^{j_1} \cdots \partial t_n^{j_n}}\Big|_{t=0}$ wobei φ die charakteristische Funktion von X bezeichnet.

Aufgabe 4 (2)
Gib ein Beispiel für zwei (reellwertige) normalverteilte Zufallsvariablen X und Y, so dass der Vektor $(X, Y)^\top$ nicht multivariat normalverteilt ist (auch nicht singulär).

Hinweis: Diese Aufgabe kann (muss aber nicht) mit Hilfe von *Copulas* gelöst werden.

Aufgabe 5
Sei $X = (X_1, X_2, X_3)^\top$ multivariat normalverteilt mit Erwartungswertvektor $\mu = (1, 2, 3)^\top$ und Kovarianzmatrix $K = \sigma^2 \begin{pmatrix} 2 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 2 \end{pmatrix}$.

a) Bestimme die Verteilungen von X_2 und (X_1, X_3) (2)

b) Für welchen Wert von ρ sind die Zufallsvariablen $X_1 + X_2 + X_3$ und $X_1 - X_2 - X_3$ unabhängig? (3)