

Klausur
Statistik III
Wintersemester 2008/2009

Prof. Dr. Torsten Hothorn
Institut für Statistik

Name: Name, Vorname

Matrikelnummer: 0123456

Wichtig:

- Überprüfen Sie, ob Ihr Klausurexemplar vollständig ist. Die Klausur besteht aus fünf Aufgaben (auf jeweils zwei Seiten, Aufgabe 1 auf drei Seiten), einem Deckblatt und der Standardnormalverteilung im Anhang. Kontrollieren Sie Ihren Namen und die Matrikelnummer.
- Verwenden Sie für Ihre Lösungen ausschließlich den Klausurbogen (Vorder- und Rückseite), Zusatzblätter werden auf Anfrage ausgeteilt.
- Als Hilfsmittel zugelassen sind alle Mitschriften der Vorlesung und der Übung, das ausgedruckte Skript sowie Bücher und ein Taschenrechner.
- Die Bearbeitungszeit beträgt 120 Minuten.

Ich bin damit einverstanden, daß das Klausurergebnis unter Angabe meiner Matrikelnummer auf der Seite

<http://www.stat.uni-muenchen.de/~hothorn>

veröffentlicht wird. (Bei Nichtzutreffen bitten streichen.)

Viel Erfolg!

Punkte:

Note:

Aufgabe 1

Sind die folgenden Aussagen richtig oder falsch (mit kurzer Begründung)?

- a) Eine Polizeistation ist mit 10 Beamten besetzt. Fünf davon sind auf Streife. Zwei haben Innendienst. Drei sind in der Station auf Bereitschaft. Es gibt 2520 verschiedene Aufteilungen der 10 Beamten auf die drei Gruppen. (**1 Pkt.**)

Lösung: richtig. Nach Satz 1.6 (b) gilt $\binom{10}{5,2,3} = \frac{10!}{5!2!3!} = 2520$.

- b) Sei Ω eine (beliebige) nichtleere Menge. Zu einem beliebigen $A \subset \Omega$ bezeichne $\sigma(\{A\})$ die kleinste σ -Algebra, die A enthält. Dann hat $\sigma(\{A\})$ abzählbar viele Elemente. (**1 Pkt.**)

Lösung: richtig, die σ -Algebra besteht aus vier Elementen.

- c) Betrachtet wird der zweidimensionale Einheitskreis $A = \{x \in \mathbb{R}^2 : \|x\|_2 \leq 1\}$. Für Zähl- und Lebesgue-Maß gelten $\mu_Z(A) = 5$ und $\lambda(A) = \pi$. (**1 Pkt.**)

Lösung: falsch. $\mu(A) = \infty$, da A überabzählbar viele Elemente enthält. $\lambda(A) = \text{Vol}(A) = \pi \cdot r(A) = \pi$ ist richtig, vgl. Blatt 5, Aufgabe 1.

- d) Seien $A_1, A_2, \dots \in \mathcal{F}$. Falls $A_1 \subset A_2 \subset \dots$, dann gilt

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n).$$

(**1 Pkt.**)

Lösung: richtig. Dies folgt aus Satz 2.24 (i) (Stetigkeit von Maßen).

- e) Für zwei Ereignisse A und B gelte $\mathbb{P}(A) = \frac{1}{4}$, $\mathbb{P}(B) = \frac{1}{3}$ und $\mathbb{P}(A \cup B) = \frac{1}{2}$. Mit einer Wahrscheinlichkeit von $\frac{7}{12}$ tritt genau eines der beiden Ereignisse A, B ein. (**2 Pkt.**)

Lösung: falsch. Gesucht ist $\mathbb{P}(A \triangle B)$. Wie in Blatt 9, Aufgabe 3.2 gezeigt gilt $\mathbb{P}(A \triangle B) = \mathbb{P}(A) + \mathbb{P}(B) - 2\mathbb{P}(A \cap B)$. Zunächst berechnen wir $\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) = \frac{1}{4} + \frac{1}{3} - \frac{1}{2} = \frac{1}{12}$. Damit erhalten wir $\mathbb{P}(A \triangle B) = \frac{7}{12} - 2\frac{1}{12} = \frac{5}{12} \neq \frac{7}{12}$.

- f) Es seien A und B zwei Ereignisse, für die alle im folgenden betrachteten bedingten Wahrscheinlichkeiten definiert sind. Aus $\mathbb{P}(A|B) > \mathbb{P}(B|A)$ folgt die Abschätzung $\mathbb{P}(A) > \mathbb{P}(B)$. (**2 Pkt.**)

Lösung: richtig. Denn nach der Definition der bedingten Wahrscheinlichkeit (Def. 6.1) gilt

$$\mathbb{P}(A|B) > \mathbb{P}(B|A) \Leftrightarrow \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} > \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \Leftrightarrow \frac{1}{\mathbb{P}(B)} > \frac{1}{\mathbb{P}(A)} \Leftrightarrow \mathbb{P}(A) > \mathbb{P}(B).$$

- g) Sei $X = (X_1, X_2)$ ein Zufallsvektor mit $X \sim N_2(\mu, \Sigma)$, Kovarianzmatrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix},$$

und Korrelationsmatrix

$$R = \begin{pmatrix} 1 & \rho(X_1, X_2) \\ \rho(X_2, X_1) & 1 \end{pmatrix}.$$

Sei $D = \text{diag}(\sigma_1, \sigma_2)$ die Diagonalmatrix der Standardabweichungen. Dann gilt

$$D^{-1}(X - \mu) \sim N_2(0, R).$$

(2 Pkt.)

Lösung: richtig. Betrachte die Zufallsvariable $Y = D^{-1}(X - \mu)$. Nach Definition 8.23 ist $X - \mu \sim N_2(0, \Sigma)$. Nach Satz 7.17 gilt $\mathbb{V}(Y) = D^{-1}\Sigma D^{-1}$. Nach Definition 7.19 gilt $\rho(X_1, X_2) = \frac{\sigma_{12}}{\sigma_1\sigma_2}$. Damit erhalten wir

$$\begin{aligned} D^{-1}\Sigma D^{-1} &= \begin{pmatrix} \sigma_1^{-1} & 0 \\ 0 & \sigma_2^{-1} \end{pmatrix} \begin{pmatrix} 1 & \frac{\sigma_{12}}{\sigma_1\sigma_2} \\ \frac{\sigma_{12}}{\sigma_1\sigma_2} & 1 \end{pmatrix} \begin{pmatrix} \sigma_1^{-1} & 0 \\ 0 & \sigma_2^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1 & \frac{\sigma_{12}}{\sigma_1} \\ \frac{\sigma_{12}}{\sigma_2} & \sigma_2 \end{pmatrix} \begin{pmatrix} \sigma_1^{-1} & 0 \\ 0 & \sigma_2^{-1} \end{pmatrix} \\ &= \begin{pmatrix} 1 & \frac{\sigma_{12}}{\sigma_1\sigma_2} \\ \frac{\sigma_{12}}{\sigma_1\sigma_2} & 1 \end{pmatrix} = R. \end{aligned}$$

Da $D^{-1}(X - \mu)$ zwei Linearkombinationen von X darstellt, handelt es sich nach Satz 8.26 bei der resultierenden Verteilung ebenfalls um eine zweidimensionale Normalverteilung.

h) Ist $X_n \sim \text{Exp}(n)$, dann gilt $X_n \xrightarrow{\mathbb{P}} 0$. (2 Pkt.)

Lösung: richtig. Nach Definition 9.1 (c) und weil exponentialverteilte ZVen einen nichtnegativen Träger haben gilt

$$X_n \xrightarrow{\mathbb{P}} 0 \Leftrightarrow P(|X_n| > \epsilon) \rightarrow 0 \forall \epsilon > 0 \Leftrightarrow P(X_n > \epsilon) \rightarrow 0 \forall \epsilon > 0.$$

Es gilt

$$\begin{aligned} P(X_n > \epsilon) &= 1 - P(X_n \leq \epsilon) \\ &= 1 - n \int_0^\epsilon e^{-nx_n} dx_n \\ &= 1 - n \left[-\frac{1}{n} e^{-nx_n} \right]_0^\epsilon \\ &= 1 - n \left(\frac{1}{n} (1 - e^{-n\epsilon}) \right) \\ &= 1 - (1 - e^{-n\epsilon}) = e^{-n\epsilon} \rightarrow 0 \end{aligned}$$

für $n \rightarrow \infty$.

Aufgabe 2

Sei X geometrisch verteilt mit Parameter $p \in (0, 1)$ und Wahrscheinlichkeitsfunktion

$$\mathbb{P}(X = x) = p(1 - p)^{x-1}, x \in \mathbb{N}.$$

a) Zeigen Sie, daß die momentenerzeugende Funktion

$$M_X(s) = \frac{pe^s}{1 - (1 - p)e^s}$$

lautet. (**3 Pkt.**)

(*Hinweis:* $e^{sx} = e^s e^{s(x-1)}$.)

b) Für welche Werte $s \in \mathbb{R}$ ist $M_X(s)$ definiert? (**2 Pkt.**)

c) Bestimmen Sie mit Hilfe der momentenerzeugenden Funktion $\mathbb{E}(X)$. (**3 Pkt.**)

Lösung

a) Nach Def. 6.10 gilt

$$\begin{aligned} M_X(s) &= \mathbb{E}(e^{sX}) = \sum_{x=1}^{\infty} e^{sx} p (1-p)^{x-1} \\ &\stackrel{e^{sx} = e^{s(x-1)} e^s}{=} p e^s \sum_{x=1}^{\infty} e^{s(x-1)} (1-p)^{x-1} \\ &= p e^s \sum_{x=1}^{\infty} [e^s (1-p)]^{x-1} \\ &\stackrel{\sum_{n=0}^{\infty} |\rho|^n = \frac{1}{1-|\rho|}, |\rho| < 1}{=} p e^s \frac{1}{1 - (1-p)e^s} \\ &\quad \text{(Summenformel für konvergente geom. Reihen)} \end{aligned}$$

b) Wertebereich für s : Für Konvergenz in a) muss gelten

$$\begin{aligned} |e^s (1-p)| &< 1 \\ \Leftrightarrow e^s (1-p) &< 1 \\ \Leftrightarrow s &< \log\left(\frac{1}{1-p}\right) \\ \Leftrightarrow s &< -\log(1-p) \\ \Rightarrow \mathcal{D} &= \{s \in \mathbb{R} \mid s < -\log(1-p)\} \end{aligned}$$

c)

$$\begin{aligned} \mathbb{E}(X) &= \left. \frac{\partial M(s)}{\partial s} \right|_{s=0} \\ \frac{\partial M(s)}{\partial s} &\stackrel{\text{Quotientenregel}}{=} \frac{p e^s (1 - (1-p)e^s) + p e^s (1-p) e^s}{[1 - (1-p)e^s]^2} \\ &= \frac{p e^s}{[1 - (1-p)e^s]^2} \\ \left. \frac{\partial M(s)}{\partial s} \right|_{s=0} &= \frac{p}{[1 - (1-p)]^2} \\ &= \frac{1}{p} \end{aligned}$$

Aufgabe 3

Die Zufallsvariablen X und Y seien unabhängig und stetig verteilt mit den Dichten f_X und f_Y . Zeigen Sie, daß die Dichte der Zufallsvariablen $Z = X \cdot Y$ gegeben ist durch

$$f_{X \cdot Y}(z) = \int_{\mathbb{R}} \frac{1}{|t|} f_X(t) f_Y\left(\frac{z}{t}\right) dt, \quad \forall z \in \mathbb{R}.$$

(8 Pkt.)

Lösung

- Betrachte gemeinsame Verteilung

$$g(X, Y) = (X, X \cdot Y) = (A, B)$$

- Für $h(A, B) = g^{-1}(A, B)$ gilt

$$\begin{aligned} A &= X &\Leftrightarrow X &= A \\ B &= X \cdot Y &\Leftrightarrow Y &= \frac{B}{A} \end{aligned}$$

also $h(A, B) = \left(A, \frac{B}{A}\right)$

- Die Jacobi-Matrix von h lautet

$$H = \begin{bmatrix} \frac{\partial h_1}{\partial A} & \frac{\partial h_1}{\partial B} \\ \frac{\partial h_2}{\partial A} & \frac{\partial h_2}{\partial B} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{B}{A^2} & \frac{1}{A} \end{bmatrix}$$

$$\Rightarrow \det H = \frac{1}{A}$$

- Anwendung des Dichtetransformationssatzes (Satz 7.7) ergibt

$$\begin{aligned} f_{A,B}(a, b) &= f_{X,Y}(h(a, b)) |\det(H)| \\ &\stackrel{X,Y \text{ stu}}{=} f_X(a) f_Y\left(\frac{b}{a}\right) \cdot \frac{1}{|a|} \end{aligned}$$

- Berechnung der Randdichte von B (durch "Herausintegrieren" von A)

$$f_B(b) = \int_{\mathbb{R}} f_X(a) f_Y\left(\frac{b}{a}\right) \cdot \frac{1}{|a|} da$$

w.z.b.w.

Aufgabe 4

Die Rechenzeiten (in Sekunden) von $n = 100$ Programmen auf einem Großrechner seien durch 100 Zufallsvariablen X_1, \dots, X_{100} beschrieben, die stochastisch unabhängig und identisch verteilt sind mit $\mathbb{E}(X_i) = 20$ und $\mathbb{V}(X_i) = 100$. Geben Sie für die Wahrscheinlichkeit, daß die Gesamtrechenzeit der 100 Programme zwischen 1800 und 2200 Sekunden liegt, einen Näherungswert durch Anwendung des zentralen Grenzwertsatzes an. (6 Pkt.)

Lösung

$$\mathbb{E}(X_i) = 20, \quad \mathbb{V}(X_i) = 100, \quad n = 100.$$

Nach ZGWS gilt

$$\begin{aligned} \sqrt{n} \frac{\bar{X} - \mu}{\sigma} &= \frac{\sqrt{n} \frac{1}{n} \sum_{i=1}^n X_i - \sqrt{n} \mu}{\sigma} \\ &= \frac{\frac{1}{\sqrt{n}} (\sum_{i=1}^n X_i - n\mu)}{\sigma} \\ &= \frac{\sum_{i=1}^n X_i - n \mathbb{E}[X_i]}{\sqrt{n} \sqrt{\mathbb{V}(X_i)}} \xrightarrow{d} \mathcal{N}(0, 1) \\ \Rightarrow \quad &\mathbb{P} \left(1800 \leq \sum_{i=1}^{100} X_i \leq 2200 \right) \\ &= \mathbb{P} \left(\frac{1800 - 2000}{100} \leq \frac{\sum_{i=1}^{100} X_i - 2000}{100} \leq \frac{2200 - 2000}{100} \right) \\ &\xrightarrow{d} \Phi(2) - \Phi(-2) \\ \stackrel{\text{Symmetrie NV}}{=} &2 \cdot \Phi(2) - 1 \\ &= 2 \cdot 0.9772 - 1 \\ &= 0.9544 \end{aligned}$$

Aufgabe 5

Es seien X und Y zwei reellwertige Zufallsvariablen über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit der gemeinsamen Dichtefunktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$f(x, y) = 2e^{-x-y}I_{(0,y)}(x)I_{(0,\infty)}(y).$$

- a) Überprüfen Sie, ob $f(x, y)$ eine Dichtefunktion ist. (2 Pkt.)
- b) Berechnen Sie die Randdichte von X . Berücksichtigen Sie hierbei, dass $X < Y$ gelten muss. Was ergibt sich dadurch für den Integrationsbereich? Wie ist X verteilt? (2 Pkt.)
- c) Überprüfen Sie, ob X und Y stochastisch unabhängig sind. (2 Pkt.)

Lösung

a)

$$\begin{aligned}\int_0^\infty \int_0^y 2e^{-x-y} dx dy &= 2 \int_0^\infty e^{-y} \int_0^y e^{-x} dx dy \\ &= 2 \int_0^\infty e^{-y} [-e^{-x}]_0^y dy \\ &= 2 \int_0^\infty e^{-y} (1 - e^{-y}) dy \\ &= 2 \int_0^\infty e^{-y} dy - 2 \int_0^\infty e^{-2y} dy \\ &= 2 [-e^{-y}]_0^\infty - 2 \left[-\frac{1}{2} e^{-y} \right]_0^\infty \\ &= 2 - 2 \cdot \frac{1}{2} = 1.\end{aligned}$$

Da weiterhin $e^{-x-y} > 0$ für alle $x, y \in \mathbb{R}_+$, handelt sich um eine Dichtefunktion.

b) Da $X < Y < \infty$ gelten muss, ergibt sich

$$\begin{aligned}f_X(x) &= \int_{-\infty}^\infty f(x, y) dy \\ &= \int_0^\infty 2e^{-x-y} I_{(0,y)}(x) dy \\ &= 2e^{-x} \int_x^\infty e^{-y} dy I_{(0,\infty)}(x) \\ &= 2e^{-x} [-e^{-y}]_x^\infty I_{(0,\infty)}(x) \\ &= 2e^{-2x} I_{(0,\infty)}(x)\end{aligned}$$

Daher ist $X \sim \text{Exp}(2)$.

c) Nach Satz 7.10 sind zwei Zufallsvariablen unabhängig, wenn ihre gemeinsame Dichte als Produkt der beiden Randdichten dargestellt werden kann. Für die Randdichte $f_Y(y)$ erhalten wir

$$\begin{aligned}f_Y(y) &= 2e^{-y} \int_0^y e^{-x} dx I_{(0,\infty)}(y) \\ &= 2e^{-y} [-e^{-x}]_0^y I_{(0,\infty)}(y) \\ &= 2e^{-y} (1 - e^{-y}) I_{(0,\infty)}(y).\end{aligned}$$

Da $f(x, y) \neq f_X(x)f_Y(y)$ sind X und Y nicht stochastisch unabhängig.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.05	0.5199	1.05	0.8531	2.05	0.9798
0.10	0.5398	1.10	0.8643	2.10	0.9821
0.15	0.5596	1.15	0.8749	2.15	0.9842
0.20	0.5793	1.20	0.8849	2.20	0.9861
0.25	0.5987	1.25	0.8944	2.25	0.9878
0.30	0.6179	1.30	0.9032	2.30	0.9893
0.35	0.6368	1.35	0.9115	2.35	0.9906
0.40	0.6554	1.40	0.9192	2.40	0.9918
0.45	0.6736	1.45	0.9265	2.45	0.9929
0.50	0.6915	1.50	0.9332	2.50	0.9938
0.55	0.7088	1.55	0.9394	2.55	0.9946
0.60	0.7257	1.60	0.9452	2.60	0.9953
0.65	0.7422	1.65	0.9505	2.65	0.9960
0.70	0.7580	1.70	0.9554	2.70	0.9965
0.75	0.7734	1.75	0.9599	2.75	0.9970
0.80	0.7881	1.80	0.9641	2.80	0.9974
0.85	0.8023	1.85	0.9678	2.85	0.9978
0.90	0.8159	1.90	0.9713	2.90	0.9981
0.95	0.8289	1.95	0.9744	2.95	0.9984
1.00	0.8413	2.00	0.9772	3.00	0.9987

Tabelle 1: Verteilungsfunktion der Standardnormalverteilung.