DAV

DEUTSCHE
AKTUARVEREINIGUNG e.V.

Zulassungsprüfung Stochastik

der Deutschen Aktuarvereinigung e. V.

am 12. Oktober 2018

Hinweise:

- Als Hilfsmittel ist ein Taschenrechner zugelassen.
- Die Gesamtpunktzahl beträgt 120 Punkte. Die Klausur ist bestanden, wenn mindestens 60 Punkte erreicht werden.
- Bitte prüfen Sie die Ihnen vorliegende Prüfungsklausur auf Vollständigkeit. Die Klausur besteht aus 4 Seiten und einem Beiblatt.
- Alle Antworten sind zu begründen und bei Rechenaufgaben muss der Lösungsweg ersichtlich sein.

Wir gehen stets von einem Maßraum $(\Omega, \mathcal{A}, \mu)$ bzw. einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ aus. Die Borel σ-Algebra auf $\mathbb{R}^n$ wird mit $\mathcal{B}^n$ bezeichnet, das Lebesgue Maß auf $\mathbb{R}^n$ wird mit λ^n bezeichnet.

Sollten Ihnen in Teilaufgaben Ergebnisse fehlen, dann treffen Sie eine plausible Annahme dafür.

Aufgabe 1. *[24 Punkte]*

(a) *[12 Punkte]* Sei $f : \mathbb{R} \longrightarrow [0, \infty)$ eine stetige, monoton fallende Funktion, $(g_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge messbarer Funktionen $g_n : \Omega \longrightarrow \mathbb{R}$ mit Grenzwert $g := \lim_{n\to\infty} g_n \geq 0$. Beweisen Sie, dass

$$\lim_{n\to\infty} \int_\Omega f(g_n)\, d\mu = \int_\Omega f(g)\, d\mu$$

gilt. Zeigen Sie hierzu, dass $(f(g_n))_{n\in\mathbb{N}}$ eine monoton wachsende Folge messbarer Funktionen ist.

(b) *[12 Punkte]* Sei $h : \Omega \longrightarrow [0, \infty)$ integrierbar, $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ eine Folge von Teilmengen von Ω mit $A_{n+1} \subset A_n$ für alle $n \in \mathbb{N}$ und $A := \bigcap_{n\in\mathbb{N}} A_n$. Beweisen Sie mit dem Satz von der dominierten Konvergenz

$$\lim_{n\to\infty} \int_{A_n} h\, d\mu = \int_A h\, d\mu.$$

Aufgabe 2. *[18 Punkte]*

Gegeben seien Zufallsvariablen $X, Y : \Omega \longrightarrow \mathbb{R}$ mit $E(X^2) < \infty$ und eine messbare Funktion $g : \mathbb{R} \longrightarrow \mathbb{R}$ mit $E(g(Y)^2) < \infty$. Beweisen Sie

(a) *[6 Punkte]* $E(X + g(Y)|Y) = E(X|Y) + g(Y),$

(b) *[6 Punkte]* $\mathrm{Var}(X + g(Y)|Y) = \mathrm{Var}(X|Y),$

(c) *[6 Punkte]* $\mathrm{Cov}(X, g(Y)|Y) = 0.$

Aufgabe 3. *[27 Punkte]*

Sei $T \sim \mathrm{Exp}(\lambda)$, $\lambda > 0$, die Lebensdauer eines Geräts in Jahren. Ein neues Produktionsverfahren führt zu einer längeren Lebensdauer $T_\alpha := \dfrac{T}{\alpha}$ mit $\alpha \in (0, 1)$.

(a) *[3 Punkte]* Beweisen Sie, dass $T_\alpha \sim \mathrm{Exp}(\alpha\lambda)$ gilt.

(b) *[12 Punkte]* Beweisen Sie, dass

$$E(\min(T,1)) = \frac{1-e^{-\lambda}}{\lambda} \text{ und } E\left(1_{[0,1]}(T)\right) = 1 - e^{-\lambda}$$

gilt.

(c) *[4 Punkte]* Besitzen die Zufallsvariablen $\min(T,1)$ und $1_{[0,1]}(T)$ Lebesgue-Dichten? Begründen Sie Ihre Antwort.

(d) *[8 Punkte]* Nun werden n Geräte betrachtet mit Lebensdauern $T_i \overset{\text{iid}}{\sim} \text{Exp}(\lambda)$, $i = 1, \dots, n$.

(i) *[3 Punkte]* Sei $N = \sum_{i=1}^{n} 1_{\{T_i \le 1\}}$. Bestimmen Sie $E(N)$. Wie kann man N interpretieren?

(ii) *[2 Punkte]* Sei $T := \sum_{i=1}^{n} T_i$. Bestimmen Sie $\text{Var}(T)$.

(iii) *[3 Punkte]* Sei $\tilde{T} := \sum_{i=1}^{n} \min(T_i, 1)$. Bestimmen Sie $E(\tilde{T})$. Wie kann man $\tilde{T}$ interpretieren?

Aufgabe 4. *[21 Punkte]*

(a) *[5 Punkte]* Seien $T_1, \dots, T_n \overset{\text{iid}}{\sim} \text{Exp}(\lambda)$, $\lambda > 0$ und $X := \min(T_1, \dots, T_n)$. Beweisen Sie, dass X die Dichte

$$f(x) = \begin{cases} \lambda n e^{-\lambda n x}, & x \ge 0, \\ 0, & x < 0, \end{cases}$$

besitzt.

(b) *[12 Punkte]* Die Lebensdauer eines Objekts sei exponential verteilt $\text{Exp}(\lambda)$ mit $\lambda > 0$. Um λ zu schätzen wird wie folgt vorgegangen: Es werden k Gruppen von jeweils n_i, $i = 1, \dots, k$, Objekten gebildet und in jeder Gruppe die Ausfallzeit des ersten ausgefallenen Objekts beobachtet. Bestimmen Sie einen ML-Schätzer für λ. Gehen Sie hierzu von $T_{ij} \overset{\text{iid}}{\sim} \text{Exp}(\lambda)$, $i = 1, \dots, k, j = 1, \dots, n_i$ und $X_i := \min(T_{i1}, \dots, T_{in_i})$, $i = 1, \dots, k$, aus.

(c) *[4 Punkte]* Für die Situation in (b) seien die folgenden Beobachtungen gegeben:

i	1	2	3	4	5
n_i	10	20	30	30	10
x_i	2,3	0,9	0,4	0,2	0,1

Bestimmen Sie den sich ergebenden Schätzwert $\hat{\lambda}$.

Aufgabe 5 *[30 Punkte]* Es wird der maximale Puls Y im Alter x untersucht. Für die Zufallsvariable Y wird für festes $x \in [20, 70]$ das Modell

$$Y = a + bx + \varepsilon$$

angenommen, wobei $\varepsilon \sim (0, \sigma^2)$-verteilt ist.

Die zur Untersuchung vorliegenden Daten bestehen aus den Altern x_i und den Realisierungen y_i der Zufallsvariablen $Y_i = a + bx_i + \varepsilon_i$ mit $\varepsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2)$, $i = 1, \ldots, n$. Die Größe der Stichprobe sei $n = 21$. Es werden die folgenden Kennzahlen berechnet:

$$\overline{x} = 42,9, \qquad \overline{y} = 176,8, \qquad \sum_{i=1}^{21}(x_i - \overline{x})^2 = 2708,2,$$

$$\hat{b} = -0,58, \qquad \hat{a} = 201,6, \qquad \hat{\sigma}^2 = 43,7,$$

$$\text{se}(\hat{b}) = 0,13, \qquad \text{se}(\hat{a}) = 5,6.$$

Eine Graphik finden Sie auf dem Beiblatt (s. Teil (b)).

(a) *[6 Punkte]* Ist der Schätzer $\hat{b} = \dfrac{\sum_{i=1}^{n}(Y_i - \overline{Y})(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$ normalverteilt? Begründen Sie Ihre Antwort. Sie können ohne Beweis verwenden:

$$\sum_{i=1}^{n}(Y_i - \overline{Y})(x_i - \overline{x}) = \sum_{i=1}^{n} Y_i(x_i - \overline{x}).$$

(b) *[6 Punkte]* Zeichnen Sie die Ausgleichsgerade in das Beiblatt ein (***bitte mit abgeben!***). Wie beurteilen Sie die Anpassung?

(c) *[6 Punkte]* Gibt es einen Einfluss des Alters auf den maximalen Puls? Begründen Sie Ihre Einschätzung zu einem Niveau von 5 %.

(d) *[12 Punkte]*

(i) *[6 Punkte]* Geben Sie ein 95 % Schätzintervall für den erwarteten maximalen Pulsschlag für das Alter 62 an.

(ii) *[6 Punkte]* Eine Statistik-Software gibt ein Schätzintervall (157,6, 173,7) mit dem Niveau $1 - \alpha$ für den erwarteten maximalen Pulsschlag für das Alter 62 an. Das Niveau ist aber nicht leserlich. Gilt $1 - \frac{\alpha}{2} < 0,975$ oder $1 - \frac{\alpha}{2} > 0,975$? Begründen Sie Ihre Antwort.

Lösungsvorschläge

Aufgabe 1 [24 Punkte]

Zu (a) *[12 Punkte]*

Sei $n \in \mathbb{N}$, $h_n := f(g_n)$, $\omega \in \Omega$. Es gilt $g_{n+1}(\omega) \leq g_n(\omega)$ und somit wegen der Monotonie von f auch $f(g_{n+1}(\omega)) \geq f(g_n(\omega))$ also $h_{n+1} \geq h_n$. Somit ist $(h_n)_{n\in\mathbb{N}}$ eine monoton wachsende Folge messbarer Funktionen. Wegen $f \geq 0$ gilt auch $h_n \geq 0$ für alle $n \in \mathbb{N}$. Wegen der Stetigkeit von f gilt

$$\lim_{n\to\infty} h_n = \lim_{n\to\infty} f(g_n) = f(g).$$

Da f stetig ist, ist f messbar uns somit auch h_n als Verkettung messbarer Funktionen. Mit dem Satz von der monotonen Konvergenz folgt

$$\lim_{n\to\infty} \int_\Omega f(g_n)\, d\mu = \lim_{n\to\infty} \int_\Omega h_n\, d\mu = \lim_{n\to\infty} \int_\Omega f(g)\, d\mu.$$

Zu (b) *[12 Punkte]*

Es gilt $\lim_{n\to\infty} 1_{A_n} = 1_A$:

1. Fall: $\omega \in A = \bigcap_{n=1}^\infty A_n$. Dann gilt $\omega \in A_n$ für alle $n \in \mathbb{N}$, also $1_{A_n}(\omega) = 1$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} 1_{A_n}(\omega) = 1 = 1_A(\omega)$.

2. Fall: $\omega \notin A$. Es gibt $n_0 \in \mathbb{N}$ so, dass $\omega \notin A_{n_0}$. Damit gilt auch $\omega \notin A_n$ für alle $n \geq n_0$ und $\lim_{n\to\infty} 1_{A_n}(\omega) = 0 = 1_A(\omega)$.

Somit gilt für $h_n := 1_{A_n} h$

$$\begin{aligned} \lim_{n\to\infty} h_n &= \lim_{n\to\infty} 1_{A_n} h = 1_A h \\ h_n &= h 1_{A_n} \leq h \text{ für alle } n \in \mathbb{N}. \end{aligned}$$

Es folgt mit dem Satz von der dominierten Konvergenz

$$\lim_{n\to\infty} \int_{A_n} h d\mu = \lim_{n\to\infty} \int_\Omega h_n\, d\mu = \int_\Omega 1_A h\, d\mu = \int_A h\, d\mu.$$

Aufgabe 2. *[je 6 Punkte]*

Zu (a) *[6 Punkte]*

Der bedingte Erwartungswert ist linear und $g(Y)$ ist $\sigma(Y)$-messbar. Somit gilt:

$$E(X + g(Y)|Y) = E(X|Y) + E(g(Y)|Y) = E(X|Y) + g(Y)E(1|Y) = E(X|Y) + g(Y).$$

Zu (b) *[6 Punkte]*

Mit (a) und der Definition der bedingten Erwartung gilt

$$\begin{aligned}\operatorname{Var}(X+g(Y)|Y) &= E\left((X+g(Y)-E(X+g(Y)|Y))^2\,|Y\right)\\ &= E\left((X+g(Y)-E(X|Y)-g(Y))^2\,|Y\right)\\ &= E\left((X-E(X|Y))^2\,|Y\right) = \operatorname{Var}(X|Y).\end{aligned}$$

Zu (c) *[6 Punkte]*

Wiederum mit der Definiton der bedingten Kovarianz und der $\sigma(Y)$-Messbarkeit von $g(Y)$, gilt

$$\begin{aligned}\operatorname{Cov}(X,g(Y)|Y) &= E\left((X-E(X|Y))\,(g(Y)-E(g(Y)|Y))\right)\\ &= E\left((X-E(X|Y))\,(g(Y)-g(Y))\right) = 0.\end{aligned}$$

Aufgabe 3

Zu (a) *[3 Punkte]*

Sei $t \geq 0$. Es gilt wegen $T \sim \operatorname{Exp}(\lambda)$

$$P\left(\frac{T}{\alpha} \leq t\right) = P\left(T \leq t\alpha\right) = 1 - e^{-\lambda(\alpha t)} = 1 - e^{-(\alpha\lambda)t}.$$

Zu (b) *[12 Punkte]*

Es gilt

$$E\left(\min(T,1)\right) = \int_0^\infty \min(t,1)\lambda e^{-\lambda t}\,dt = \int_0^1 t\lambda e^{-\lambda t}\,dt + \int_1^\infty \lambda e^{-\lambda t}\,dt.$$

Für den ersten Summanden erhalten wir mit partieller Integration

$$\int_0^1 t\lambda e^{-\lambda t}\,dt = -te^{-\lambda t}\Big|_0^1 + \int_0^1 e^{-\lambda t} = -e^{-\lambda} - \frac{1}{\lambda}\left(e^{-\lambda}-1\right).$$

Es ergibt sich weiter

$$E\left(\min(T,1)\right) = \left[-e^{-\lambda} - \frac{1}{\lambda}\left(e^{-\lambda}-1\right)\right] + e^{-\lambda} = \frac{1-e^{-\lambda}}{\lambda}.$$

In analoger Weise gilt

$$E\left(1_{[0,1]}(T)\right) = \int_0^\infty 1_{[0,1]}(t)\lambda e^{-\lambda t}\,dt = \int_0^1 \lambda e^{-\lambda t}\,dt = 1 - e^{-\lambda}.$$

Zu (c) *[4 Punkte]*

Beide Zufallsvariablen besitzen keine Dichten, da

$$P(\min(T,1)=1)=P(T\geq 1)=e^{-\lambda}>0$$
$$P\left(1_{[0,1]}(T)=0\right)=P(T\geq 1)=e^{-\lambda}>0$$

gilt. Beide Terme müssten im Falle vorliegender Dichten den Wert 0 annehmen.

Zu (d) *[8 Punkte]*

(i) *[3 Punkte]* Wegen $1_{\{T_i\leq 1\}} = 1_{[0,1]}(T_i)$ gilt mit (b)

$$E\left(\sum_{i=1}^{n}1_{\{T_i\leq 1\}}\right)=\sum_{i=1}^{n}E\left(1_{\{T_i\leq 1\}}\right)=\sum_{i=1}^{n}(1-e^{-\lambda})=n(1-e^{-\lambda}).$$

N ist die Anzahl der Geräte, die innerhalb eines Jahres defekt werden.

(ii) *[2 Punkte]* Da die T_i unabhängig sind, gilt

$$\operatorname{Var}\left(\sum_{i=1}^{n}T_i\right)=\sum_{i=1}^{n}\operatorname{Var}(T_i)=\sum_{i=1}^{n}\frac{1}{\lambda^2}=\frac{n}{\lambda^2}.$$

(iii) *[3 Punkte]* Mit (b) folgt wegen der Linearität des Erwartungswerts

$$E(\tilde{T})=\frac{n}{\lambda}\left(1-e^{-\lambda}\right).$$

$\tilde{T}$ kann als die Anzahl Betriebsjahre interpretiert werden, die die Geräte im ersten Betriebsjahr funktionieren.

Aufgabe 4[21]

Zu (a) *[5 Punkte]*

Sei $x\geq 0$. Sei F die Verteilungsfunktion von $X=\min(T_1,\dots,T_n)$. Dann gilt wegen der Unabhängigkeit der T_i

$$\begin{aligned}F(x)&=P\left(\min(T_1,\dots,T_n)\leq x\right)=1-P\left(\min(T_1,\dots,T_n)>x\right)=1-P\left(T_1>x,\dots,T_n>x\right)\\&=1-\prod_{i=1}^{n}P(T_i>x)=1-\prod_{i=1}^{n}(1-P(T_i\leq x))=1-\prod_{i=1}^{n}\left(1-(1-e^{-\lambda x})\right)\\&=1-e^{-n\lambda x}.\end{aligned}$$

Für $x<0$ gilt $F(x)=0$. Differenzieren nach x führt zu

$$\frac{d}{dx}F(x)=\begin{cases}n\lambda e^{-n\lambda x} & x>0\\ 0 & x<0.\end{cases}$$

Zu (b)*[12 Punkte]*

Die gemeinsame Dichte der $X_1, \ldots, X_n$ ist wegen der Unabhängigkeit gegeben durch

$$f(x_1, \ldots, x_k) = \prod_{i=1}^{k} n_i \lambda e^{-n_i \lambda x_i}$$

Geht man zur Log-Likelihood-Funktion über folgt

$$\begin{aligned}
\ell(\lambda) &= \sum_{i=1}^{k} \ln(n_i) + k \ln \lambda - \sum_{i=1}^{k} n_i \lambda x_i, \\
\frac{d}{d\lambda}\ell(\lambda) &= \frac{k}{\lambda} - \sum_{i=1}^{k} n_i x_i, \\
\frac{d^2}{d\lambda^2}\ell(\lambda) &= -\frac{k}{\lambda^2} < 0.
\end{aligned}$$

Damit ergibt sich das Maximum $\hat{\lambda}$

$$\begin{aligned}
\frac{k}{\hat{\lambda}} &= \sum_{i=1}^{k} n_i x_i \\
\hat{\lambda} &= \frac{k}{\sum_{i=1}^{k} n_i x_i}.
\end{aligned}$$

Somit ergibt sich der Schätzer $\hat{\lambda} = \dfrac{k}{\sum_{i=1}^{k} n_i X_i}$.

Zu (c) *[4 Punkte]*

Mit den Daten ergibt sich

$$\hat{\lambda} = \frac{5}{23 + 18 + 12 + 6 + 1} = \frac{1}{12}.$$

Aufgabe 5*[30 Punkte]*

Zu (a) *[6 Punkte]*

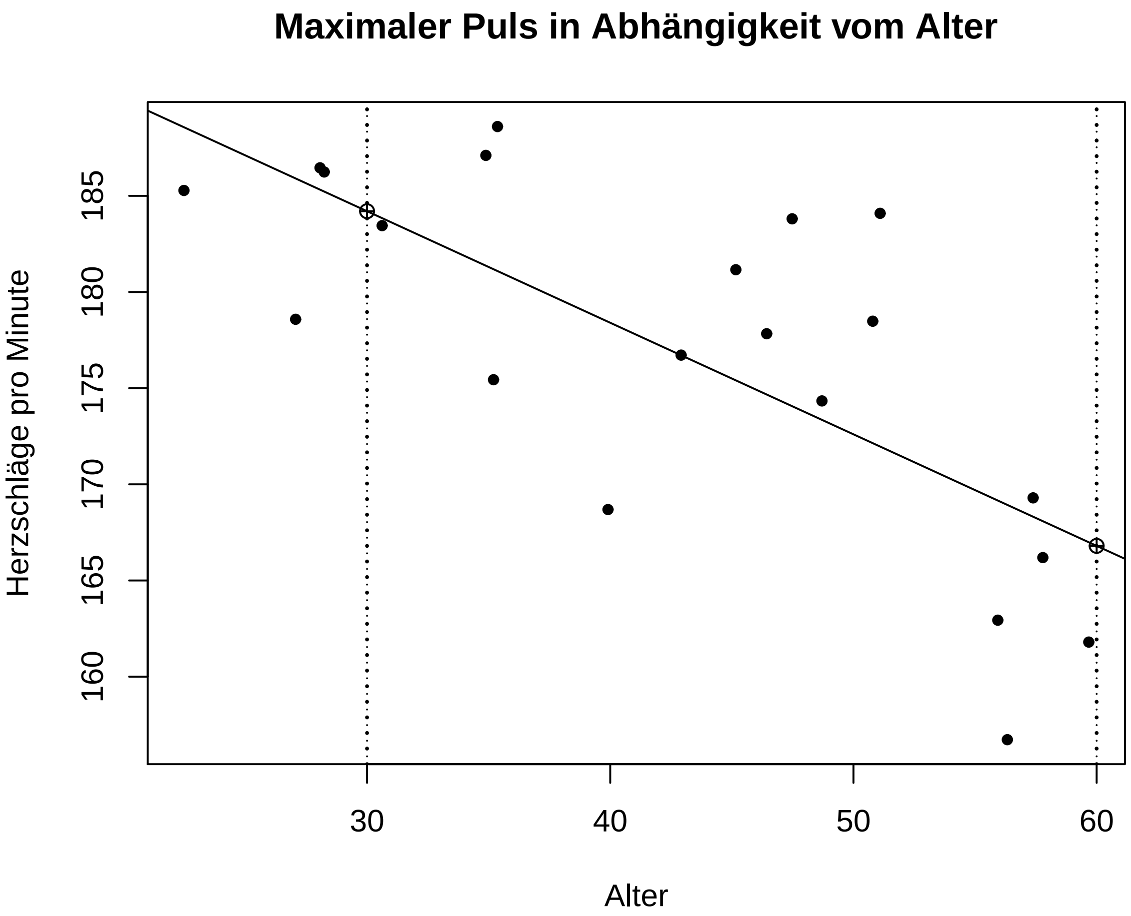
Aus $Y_i = a + bx_i + \varepsilon_i$ mit $\varepsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2)$ folgt, dass $Y_i \sim N(a + bx_i, \sigma^2)$ verteilt ist und dass die Y_i unabhängig sind. Mit der Abkürzung $sxx := \sum_{i=1}^{n}(x_i - \overline{x})$ ist

$$\hat{b} = \sum_{i=1}^{n} Y_i \frac{x_i - \overline{x}}{sxx}$$

eine Linearkombination unabhängiger normalverteilter Zufallsvariablen, also normalverteilt.

Zu (b) *[6 Punkte]*

Für die Alter 30 und 60 ergibt sich der Puls $184,2$ und $166,8$. Die Anpassung ist aktzeptabel, ein negativer Trend entspricht der Datenlage.



Zu (c) *[6 Punkte]*

Es wird die Nullhypothese $H_0 : b = 0$ gegen $H_1 : b \neq 0$ geprüft. Wegen

$$t = \frac{\hat{b}}{\mathrm{se}(\hat{b})} = \frac{-0,58}{0,13} = -4,46$$

$$t_{19,0.975} = 2,093 < |t| = 4,46$$

wird H_0 verworfen. Das Alter hat einen Einfluss auf den maximalen Puls.

Zu (d) *[12 Punkte]*

(i) *[6 Punkte]* Zu bestimmen ist zunächst

$$\mathrm{se}(\hat{a}+\hat{b}x) = \sqrt{\sigma^2\left(\frac{1}{n}+\frac{(x-\overline{x})^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right)} = \sqrt{43,7\left(\frac{1}{21}+\cdot\frac{(62-42,9)^2}{2708,2}\right)} = 2,82.$$

Der erwartete maximale Puls ergibt sich zu $201,6 - 0,58 \cdot 62 \approx 165,6$, das Schätzintervall besitzt die Breite $2 \cdot 2,093 \cdot 2,82 \approx 11,8$:

$$(159,7,\ 171,5).$$

(ii) *[6 Punkte]* Es gilt für die oberen Grenzen

$$180,6 = 165,6 + \mathrm{se}(\hat{a}+\hat{b}62)\cdot t_{19,1-\frac{\alpha}{2}} > 165,6 + \mathrm{se}(\hat{a}+\hat{b}62)t_{19,0,975}$$

also durch auflösen $t_{19,1-\frac{\alpha}{2}} > t_{19,0,975}$. Daraus ergibt sich $1-\frac{\alpha}{2} > 0,975$.