Dr. Ingo Steinke

Universität Mannheim
Lehrstuhl für Statistik

Klausur zur Veranstaltung Grundlagen der Statistik

FSS 2014, 1. Termin, 05. Juni 2014
Version: A

Die Klausur besteht aus einem losen Lösungsbogen, einem gehefteten Teil und einer Formelsammlung. Vergleichen Sie die Version von Lösungsbogen und geheftetem Teil.
Der geheftete Teil mit 10 Seiten besteht aus diesem Deckblatt und Aufgabenzetteln mit **60** Aufgaben. Überprüfen Sie bitte Ihr Exemplar auf Vollständigkeit!

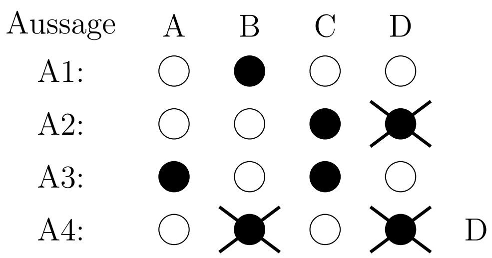
Zu jeder Aufgabe sind zwei bis vier Antworten angegeben. Von den Antwortmöglichkeiten (A) bis (D) ist genau eine richtig. Ermitteln Sie dazu die Werte, auf deren Grundlage Sie Ihre Entscheidung treffen, möglichst genau. Markieren Sie Ihre Antwort bitte auf dem Lösungsbogen. Am Ende der Klausur ist nur der **unterschriebene Lösungsbogen** abzugeben.

Unbearbeitete Aussagen, d.h. wenn keine Option markiert wurde, oder undeutliche oder unmögliche Angaben zu einer Aufgabe werden mit 0 Punkten bewertet.
Korrekt bearbeitete Aufgaben werden mit 1 Punkt, nicht korrekt bearbeitete mit 0 Punkten bewertet. Die Gesamtpunktzahl ergibt sich durch Addition aller Punkte.

Für Ihre Auswahl müssen Sie die passenden Kreise auf dem Lösungsbogen ausmalen. Bei einer ungewollten Auswahl streichen Sie den nicht korrekt markierten Kreis mit einem X durch.

Illustrierende Beispiele:

Lösungsbogen



Sie denken Aussage ...

- A1: B ist richtig.
- A2: C ist richtig, zuerst wurde D markiert.
- A3: A *oder* C ist richtig. Kein Punkt.
- A4: D ist richtig. Zuerst wurde D, dann aber B markiert.

In der Klausur werden 60 Punkte vergeben. Ab einer Punktzahl von 30 Punkten gilt die Klausur auf jeden Fall als bestanden. Die Bestehensquote wird auf minimal 60% begrenzt, d.h die Durchfallquote beträgt maximal 40%.

Die Bearbeitungszeit beträgt **180** Minuten.

Als Hilfsmittel ist nur ein nicht-programmierbarer Taschenrechner zugelassen.

Viel Erfolg!

Aufgabe
Beim zweimaligen Würfeln sei A das Ereignis, beim ersten Wurf eine ‚6' zu werfen, und B das Ereignis, beim zweiten Wurf eine ‚6' zu werfen.
Das Ereignis, höchstens eine ‚6' zu werfen, wird dann ausgedrückt durch

A1: (A) $\overline{A} \cup \overline{B}$. (B) $A \cap B$. (C) $A \cup B$. (D) $\overline{A \cup B}$.

Aufgabe
Für die Endrunde bei den Weltmeisterschaften im „Schnick-Schnack-Schnuck" haben sich 16 Teilnehmer qualifiziert. In der ersten Tournierphase werden aus den 16 Teilnehmern 8 Paarungen zufällig zusammengestellt, die dann gegeneinander antreten müssen.
Die Anzahl der verschiedenen Zusammenstellungen der Paarungen liegt in (Angaben in 1000)

A2: (A) $[0, 1000)$. (B) $[1000, 1500)$. (C) $[1500, 2000)$. (D) $[2000, \infty)$.

Aufgabe
In einem Schuhschrank befinden sich 7 verschiedene Paare von Schuhen. Es werden zufällig zwei Schuhe (ohne Zurücklegen) dem Schuhschrank entnommen.
Die Wahrscheinlichkeit, dass beide Schuhe ein Paar bilden, liegt in

A3: (A) $[0, 0.07)$. (B) $[0.07, 0.08)$. (C) $[0.08, 0.09)$. (D) $[0.09, 1]$.

Aufgabe
Von vier Fußballspielern sei bekannt, dass Sie Elfmeter unabhängig voneinander mit Wahrscheinlichkeit 0.9, 0.8, 0.7 und 0.6 verwandeln. Jeder der vier Spieler schieße genau ein Mal.
Die Wahrscheinlichkeit, dass mindestens zwei der vier geschossenen Elfmeter nicht ins Tor gehen, liegt in

A4: (A) $[0, 0.21)$. (B) $[0.21, 0.24)$. (C) $[0.24, 0.27)$. (D) $[0.27, 1]$.

Aufgabe
A, B, C und D seien beliebige Ereignisse.
(a) Die Ungleichung $P(A \cup B) > P(A) + P(B)$ ist dann

A5: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

(b) A, B, C, D seien unabhängig. Folgende Ereignisse sind dann *nicht* stets unabhängig:

A6: (A) $A \cap B$ und $B \cap C$. (B) A und $B \cap C$. (C) $A \setminus D$ und $\bar{B}$.

Aufgabe
In einem Schuhlager sind 45% der Schuhe schwarz, 35% der Schuhe braun und die restlichen Schuhe rot. Der Anteil der Damenschuhe an den schwarzen, braunen bzw. roten Schuhen betrage 40%, 30% bzw. 90%.
Der Anteil der Damenschuhe an allen Schuhen im Lager liegt dann in (Angaben in %)

A7: (A) $[0, 42)$. (B) $[42, 45)$. (C) $[45, 48)$. (D) $[48, 100]$.

Der Anteil der Damenschuhe, die eine schwarze Farbe besitzen, an allen Damenschuhen ist dann in (Angaben in %)

A8: (A) $[0, 40)$. (B) $[40, 43)$. (C) $[43, 46)$. (D) $[46, 100]$.

Aufgabe
X sei eine diskrete Zufallsgröße mit folgender Wahrscheinlichkeitstabelle

a_j	0	1	2	3	4
$P(X = a_j)$	0.2	0.1	0.4	0	0.3.

Der Träger zu X ist dann

A9: (A) $\{0, 1, 2, 3, 4\}$. (B) $\{0, 1, 2, 4\}$. (C) $\mathbb{R}$. (D) (A)-(C) sind falsch.

$P(|X - 1| \leq 1)$ befindet sich im Intervall

A10: (A) $[0, 0.6)$. (B) $[0.6, 0.7)$. (C) $[0.7, 0.8)$. (D) $[0.8, 1]$.

$E(X^2)$ ist in

A11: (A) $(-\infty, 7.0)$. (B) $[7.0, 8.0)$. (C) $[8.0, 9.0)$. (D) $[9.0, \infty)$.

Aufgabe
Aus einer Urne mit drei Kugeln, die mit den Zahlen 1, 2 und 4 beschriftet sind, werden nacheinander zwei Kugeln ohne Zurücklegen gezogen. Sei X_1 die erste gezogene Zahl und X_2 die zweite gezogene Zahl. Weiter sei $D = \max(X_1, X_2) - \min(X_1, X_2)$, also die Differenz aus der größeren und der kleineren der beiden Zahlen X_1 und X_2.
Dann sind X_1 und X_2

A12: (A) unabhängig. (B) nicht unabhängig.

$E(D)$ liegt in

A13: (A) $(-\infty, 1.2)$. (B) $[1.2, 1.5)$. (C) $[1.5, 1.8)$. (D) $[1.8, \infty)$.

Aufgabe
X sei eine stetige Zufallsvariable mit Träger (0,1) und Verteilungsfunktion F.
Die Aussage „$F(x) < F(y)$ für alle reellen $x < y$“ ist dann

A14: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

Aufgabe
X sei stetig verteilt mit Dichtefunktion $f(x) = (\frac{1}{4} + \frac{1}{4}\,x)\, I_{[0,2]}(x)$.
$P(1.5 < X < 2.5)$ liegt in

A15: (A) $[0.20, 0.24)$. (B) $[0.24, 0.28)$. (C) $[0.28, 0.32)$. (D) $[0.32, 0.36)$.

$E(X)$ ist in

A16: (A) $[1.0, 1.1)$. (B) $[1.1, 1.2)$. (C) $[1.2, 1.3)$. (D) $[1.3, 1.4)$.

Das 0.8-Quantil zu X ist in

A17: (A) $[1.70, 1.76)$. (B) $[1.76, 1.82)$. (C) $[1.82, 1.88)$. (D) $[1.88, 1.94)$.

Aufgabe
X und Y seien unabhängige Zufallsvariablen mit E(X)=1, E(Y)=2, Var(X)=3 und Var(Y)=1. Var$(2X - Y)$ ist in

A18: (A) $(-\infty, 10.5)$. (B) $[10.5, 11.5)$. (C) $[11.5, 12.5)$. (D) $[12.5, \infty)$.

E$(Y \cdot X^2)$ ist in

A19: (A) $(-\infty, 8)$. (B) $[8, 9)$. (C) $[9, 10)$. (D) $[10, \infty)$.

Aufgabe
In einer Bevölkerung seien 20% gegen eine bestimmte Steuerreform. Eine Lokalzeitung befragt 10 zufällig ausgewählte Passanten unabhängig voneinander nach ihrer Meinung.
Die Wahrscheinlichkeit, dass höchstens ein Passant gegen die Steuerreform ist, liegt in

A20: (A) $[0, 0.36)$. (B) $[0.36, 0.39)$. (C) $[0.39, 0.42)$. (D) $[0.42, 1]$.

Aufgabe
Die Zeitdauer X zwischen dem Auftauchen zweier Kunden an einem Kundenschalter sei exponentialverteilt und betrage im Mittel 4 Minuten.
Die Wahrscheinlichkeit, dass nach dem Eintreffen eines Kunden mindestens 5 Minuten lang kein Kunde kommt, ist in

A21: (A) $[0, 0.21)$. (B) $[0.21, 0.24)$. (C) $[0.24, 0.27)$. (D) $[0.27, 1]$.

Aufgabe
Es sei $X \sim N(1, 4)$.
Das 0.975-Quantil zu X liegt in

A22: (A) $(-\infty, 4.7)$. (B) $[4.7, 5.0)$. (C) $[5.0, 5.3)$. (D) $[5.3, \infty)$.

$P(X^2 \leq 3)$ liegt in

A23: (A) $[0, 0.57)$. (B) $[0.57, 0.60)$. (C) $[0.60, 0.63)$. (D) $[0.63, 1]$.

Die Verteilung von $Y = 3 + 2 \cdot X$ ist dann

A24: (A) $N(5, 8)$. (B) $N(5, 16)$. (C) $N(4, 4)$. (D) (A)-(C) sind falsch.

Aufgabe
Die Zufallsvariable U sei auf $[0, 2]$ stetig gleichverteilt und $V = 0.5\,U^3$.
Der Wert der Verteilungsfunktion von U an der Stelle 1.5 ist dann in

A25: (A) $[0, 0.5)$. (B) $[0.5, 0.6)$. (C) $[0.6, 0.7)$. (D) $[0.7, \infty)$.

$E(V)$ liegt in

A26: (A) $(-\infty, 1.05)$. (B) $[1.05, 1.15)$. (C) $[1.15, 1.25)$. (D) $[1.25, \infty)$.

Aufgabe
Für den Zusammenhang der beiden Merkmale X = Anzahl von Personen im Haushalt und Y = Anzahl der vorhandenen Fernseher ergebe sich folgende Kontingenztabelle:

X|Y	0	1	2	3
1	0.05	0.15	0	0
2	0	0.25	0.15	0
3	0.05	0.20	0.10	0.05

Die im Mittel zu erwartende Anzahl von Fernsehern in einem Haushalt ist in

A27: (A) $[0, 1.3)$. (B) $[1.3, 1.4)$. (C) $[1.4, 1.5)$. (D) $[1.5, \infty)$.

Die im Mittel zu erwartende Anzahl von Fernsehern in einem Zweipersonenhaushalt ist in

A28: (A) $[0, 1.3)$. (B) $[1.3, 1.4)$. (C) $[1.4, 1.5)$. (D) $[1.5, \infty)$.

$\text{Cov}(X, Y)$ liegt in

A29: (A) $(-\infty, 0.0)$. (B) $[0.0, 0.1)$. (C) $[0.1, 0.2)$. (D) $[0.2, \infty)$.

Aufgabe
$X_1, \ldots, X_n \sim U(0, 1)$ seien unabhängig.
Dann strebt $P(\bar{X}_n > 0.3)$ für $n \to \infty$ gegen einen Wert aus

A30: (A) $[0, 0.25)$. (B) $[0.25, 0.5)$. (C) $[0.5, 0.75)$. (D) $[0.75, 1.0]$.

Aufgabe
U und V seien unkorrelierte Zufallsvariablen mit $E(U) = E(V) = 1$, $Var(U) = 2$ und $Var(V) = 3$. Es seien $X = 2U - V$ und $Y = U + V$.
Die Aussage, dass U und X unabhängig sind, ist dann

A31: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

$\text{Cov}(X, Y)$ liegt in

A32: (A) $(-\infty, -0.5)$. (B) $[-0.5, 0.5)$. (C) $[0.5, 1.5)$. (D) $[1.5, \infty)$.

Aufgabe
Ein Cateringservice soll eine Hochzeitsfeier mit 100 Gästen beliefern. Es werden 85 alkoholhaltige Sektgläser für den Sektempfang angerichtet. Erfahrungsgemäß geht man davon aus, dass ein (zufällig ausgewählter) Gast mit Wahrscheinlichkeit 0.8 Sekt trinken möchte.
Die Wahrscheinlichkeit dafür, dass genügend alkoholhaltige Sektgläser zur Verfügung gestellt werden, ist dann in

A33: (A) $[0, 0.865)$. (B) $[0.865, 0.885)$. (C) $[0.885, 0.905)$. (D) $[0.905, 1]$.

Aufgabe
$X_1, \ldots, X_n \sim B(1, \pi)$ seien u.i.v. und

$$S = \frac{2X_1 + X_n}{2}, \quad T = 2X_2 - 1, \quad U_n = \bar{X}_n(1 - \bar{X}_n), \quad V_n = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X}_n)^2.$$

Für π sind folgende Schätzer erwartungstreu:

A34: (A) S, aber nicht T. (B) T, aber nicht S. (C) weder S noch T. (D) S und T.

Für $\theta = \pi(1-\pi)$ sind folgende Schätzer konsistent:

A35: (A) U_n, aber nicht V_n. (B) V_n, aber nicht U_n. (C) weder U_n noch V_n. (D) U_n und V_n.

Aufgabe
Die Zufallsvariablen $X_1, \ldots, X_n$ seien diskret u.i.v. gemäß einer diskreten Verteilungsfamilie mit unbekanntem Parameter θ, die mit folgender Wahrscheinlichkeitstabelle beschrieben wird:

a_i	θ	$\theta+1$	$\theta+2$	$\theta+3$
$P(X=a_i)$	0.1	0.3	0.5	0.1

Der Momentenschätzer für θ ist

A36: (A) $\bar{X}_n - 1.6$. (B) $\bar{X}_n$. (C) $\bar{X}_n - 1.2$. (D) (A)-(C) sind falsch.

Aufgabe
Die Befragung von 5 zufällig ausgewählten Passanten in der Mannheimer Innenstadt nach den Ausgaben für Geburtstagsgeschenke für den Ehepartner (in Euro) ergab folgende Werte:

50, 200, 1200, 500, 120.

Zur Vereinfachung nehme man an, dass die Ausgaben unabhängig und normalverteilt sind mit Erwartungswert μ und Varianz σ^2.
Die Stichprobenvarianz liegt in (Angaben in 1000)

A37: (A) $[0, 190)$. (B) $[190, 210)$. (C) $[210, 230)$. (D) $[230, \infty)$.

Die obere Grenze des 0.95-Konfidenzintervalls für μ liegt in

A38: (A) $[0, 930)$. (B) $[930, 970)$. (C) $[970, 1020)$. (D) $[1020, \infty)$.

Aufgabe
Eine Firma mit 4000 Angestellten möchte mit Hilfe einer Zufallsstichprobe vom Umfang 35 die mittlere Lohnsteuer μ (in Euro) schätzen, die pro Angestellten abgeführt wird. Die Erhebung ergab als arithmetisches Mittel $\bar{x} = 640$ Euro und als Stichprobenvarianz $s^2 = 400$ Euro^2.
Die Länge des approximativen Konfidenzintervalls für μ zum Niveau 0.95 liegt in

A39: (A) $[0, 13.0)$. (B) $[13.0, 14.0)$. (C) $[14.0, 15.0)$. (D) $[15.0, \infty)$.

Aufgabe
Ein Meinungsforschungsinstitut gebe als (approximatives) 95%-Konfidenzintervall für den Stimmenanteil einer Partei [0.2269, 0.2731] an.
Die Anzahl der Personen, die befragt wurden, liegt in

A40: (A) $[0, 1270)$. (B) $[1270, 1330)$. (C) $[1330, 1390)$. (D) $[1390, \infty)$.

Aufgabe
Für einen Parameter θ wurde aus einer Stichprobe das realisierte 0.90-Konfidenzintervall KI=[0.372,0.656] bestimmt.
Die Wahrscheinlichkeit, dass θ in KI liegt, ist dann in

A41: (A) $[0.82, 0.87)$. (B) $[0.87, 0.92)$. (C) $[0.92, 0.97)$. (D) (A)-(C) sind falsch.

Aufgabe
Eine Maschine verpacke Kartoffeln in 3.5kg-Netze. Das abgepackte Gewicht werde als normalverteilt angenommen mit Standardabweichung $\sigma = 150$ g. In der Qualitätskontrolle beim Kunden erfolgt automatisch eine Reklamation, wenn systematisch geringere Gewichte als das Sollgewicht abgepackt werden. Routinemäßig werden bei jeder Groß-Lieferung 10 Netze entnommen und deren Gesamtgewicht gemessen. Der Toleranzbereich sei so festgelegt, dass die Wahrscheinlichkeit einer irrtümlichen Reklamation 1% beträgt.
Das Gesamtgewicht c (in kg), ab dem die Lieferung abgelehnt werden darf, liegt in

A42: (A) $(-\infty, 33.7)$. (B) $[33.7, 34.2)$. (C) $[34.2, 34.7)$. (D) $[34.7, \infty)$.

Aufgabe
Die Beobachtungen $x_1, ..., x_{20}$ können aufgefasst werden als Realisierungen von unabhängigen, normalverteilten Zufallsvariablen mit unbekanntem μ und bekanntem $\sigma^2 = 1$. Als arithmetisches Mittel der Beobachtungen wurde 2.34 ermittelt. Es ist ein Test zu $H_0 : \mu \leq 2$ gegen $H_1 : \mu > 2$ zum Signifikanzniveau $\alpha = 0.05$ durchzuführen.
Der p-Wert des Tests ist in

A43: (A) $[0, 0.051)$. (B) $[0.051, 0.056)$. (C) $[0.056, 0.061)$. (D) $[0.061, 1]$.

Die Nullhypothese ist

A44: (A) nicht abzulehnen. (B) abzulehnen.

Aufgabe
Für einen Test für das Testproblem $H_0 : \theta \leq 0$ gegen $H_1 : \theta > 0$ wurde die Gütefunktion an der Stelle $\theta = 1$ berechnet und es ergab sich ein Wert von 0.4.
Dann ist 0.6 in diesem Zusammenhang die Wahrscheinlichkeit für

A45: (A) einen Fehler 1.Art. (B) einen Fehler 2.Art. (C) eine richtige Entscheidung.

Aufgabe
Für $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ u.i.v. werde für das Testproblem $H_0 : \mu = 0$ gegen $H_1 : \mu \neq 0$ ein Gauß-Test durchgeführt. Dabei seien $n = 16$ und $\sigma^2 = 1$.
Die Verteilung der Teststatistik für $\mu = 1$ ist dann:

A46: (A) $N(0, 1)$. (B) $N(4, 1)$. (C) $N(1, 1)$. (D) (A)-(C) sind falsch.

Aufgabe
Ein Fahrzeughersteller will den Abgasausstoß eines Motorentyps überprüfen. Testfahrten ergaben folgende Emissionswerte (in mg Kohlenwasserstoffe pro 100km)

i	1	2	3	4	5	6	7	8	9	10
x_i	78	81	112	103	83	97	83	90	64	70

Es soll mit einem statistischen Test zum Signifikanzniveau $\alpha = 0.01$ gezeigt werden, dass die Abgasnorm von im Mittel weniger als 100 mg pro 100km eingehalten wird. Gehen Sie von unabhängigen Messungen und normalverteilten Emissionswerten aus.
Der Wert der Teststatistik liegt in

A47: (A) $(-\infty, -2.7)$. (B) $[-2.7, -2.4)$. (C) $[-2.4, -2.1)$. (D) $[-2.1, \infty)$.

Der kritische Wert des Tests ist in

A48: (A) $(-\infty, -2.95)$. (B) $[-2.95, -2.55)$. (C) $[-2.55, -2.15)$. (D) $[-2.15, \infty)$.

Aufgabe
Bei einer Autobahnkontrolle wurden 200 PKW hinsichtlich korrekter Bereifung kontrolliert. Dabei ergaben sich 40 Beanstandungen. Es soll geprüft werden, ob der Anteil π zu beanstandender Fahrzeuge insgesamt bei einem Niveau von 5% signifikant höher als 15% ist.
Die Nullhypothese lautet H_0 :

A49: (A) $\pi \leq 0.15$. (B) $\pi \geq 0.15$. (C) $\pi = 0.15$. (D) $\pi \neq 0.15$.

Der Wert der Teststatistik ist im Intervall

A50: (A) $(-\infty, 2.05)$. (B) $[2.05, 2.15)$. (C) $[2.15, 2.25)$. (D) $[2.25, \infty)$.

Aufgabe
Bei der Befragung von Studenten hinsichtlich der durchschnittlichen Schlafdauer ergab sich für 119 Männer (x_i) ein Mittelwert von $\bar{x} = 7.4$ Stunden bei einer geschätzten Standardabweichung von 1.3 Stunden. Für die 47 Frauen (y_j) ergab sich ein Mittelwert von $\bar{y} = 7.6$ Stunden bei einer geschätzten Standardabweichung von 0.9 Stunden. Es soll geprüft werden, ob die mittleren Schlafdauern von Männern und Frauen zum Niveau $\alpha = 0.05$ signifikant verschieden sind.
Folgender Test ist durchzuführen (in einer Zweistichprobenversion):

A51: (A) ein Gauß-Test. (B) ein approximativer Gauß-Test. (C) ein t-Test.

Der Wert der Teststatistik liegt in

A52: (A) $(-\infty, -1.2)$. (B) $[-1.2, -1.1)$. (C) $[-1.1, -1.0)$. (D) $[-1.0, \infty)$.

Die Nullhypothese wird

A53: (A) abgelehnt. (B) nicht abgelehnt.

Aufgabe
Es soll untersucht werden, ob in unterschiedlichen Landesteilen der Anteil der Arbeitnehmer, die ihr Mittagessen zu Hause einnehmen, unterschiedlich hoch ausfällt. Dazu wurden sowohl im Norden als auch im Süden des Landes jeweils 500 Personen zufällig ausgewählt und befragt. 330 Personen im Norden und 280 im Süden gaben an, das Mittag zu Hause zu essen.
Man überprüfe mit einem Test zum Niveau 0.05, ob es regionale Unterschiede bzgl. der Mittagseinnahme gibt.
Der Betrag des Wertes der Teststatistik liegt in

A54: (A) $[0, 3.10)$. (B) $[3.10, 3.20)$. (C) $[3.20, 3.30)$. (D) $[3.30, \infty)$.

Der Betrag der kritischen Werte des Tests liegt in

A55: (A) $[0, 1.72)$. (B) $[1.72, 1.82)$. (C) $[1.82, 1.92)$. (D) $[1.92, \infty)$.

Aufgabe
Im Regressionsmodell $Y_i = \alpha + \beta x_i + \varepsilon_i$ seien die $\varepsilon_i \sim N(0, \sigma^2)$ u.i.v. Für n=50 Beobachtungen (x_i, y_i) erhielt man folgende Daten

$$\bar{x} = 1.57, \quad \bar{y} = 10.96, \quad s_x^2 = 0.269, \quad s_y^2 = 25.94, \quad \frac{1}{n}\sum_{i=1}^{n} x_i y_i = 19.07.$$

Der Schätzwert für β ist in

A56: (A) $(-\infty, 6.8)$. (B) $[6.8, 7.0)$. (C) $[7.0, 7.2)$. (D) $[7.2, \infty)$.

Der Schätzwert für $\sigma_{\hat{\beta}}$ ist in

A57: (A) $[0, 1.05)$. (B) $[1.05, 1.20)$. (C) $[1.20, 1.35)$. (D) $[1.35, \infty)$.

Aufgabe
Um die Zuverlässigkeit von Maschinen vom Typ A und Typ B zu untersuchen, wurde von verschiedenen Maschinen die Ausfallzeit für Wartung (y_i, zeit, in Stunden pro Monat) in Abhängigkeit vom Alter (a_i, alter, in Jahren) und vom Typ (t_i, typ, $t_i = 1$ für Typ A und $t_i = 2$ für Typ B) mit einer multiplen linearen Regression

$$Y_i = \beta_0 + \beta_1 a_i + \beta_2 t_i + \varepsilon_i, \quad \varepsilon \sim N(0, \sigma^2) \text{ u.i.v.}$$

modelliert. Die nachfolgende STATA-Ausgabe gibt das Ergebnis der Analyse an:

```
    Source |       SS       df       MS              Number of obs =      20
-----------+------------------------------           F(  2,     17) =  109.08
     Model |  2342.41577     2  1171.20789           Prob > F      =  0.0000
  Residual |  182.534226    17  10.7373074           R-squared     =  0.9277
-----------+------------------------------           Adj R-squared =  0.9192
     Total |     2524.95    19  132.892105           Root MSE      =  3.2768

-----------------------------------------------------------------------------
      zeit |      Coef.   Std. Err.       t    P>|t|     [95% Conf. Interval]
-----------+-----------------------------------------------------------------
     alter |   6.710326   .4544386    14.77   0.000     5.751545    7.669108
       typ |  -12.58514   1.712522    -7.35   0.000    -16.19824    -8.97203
     _cons |   19.54877   2.324515     8.41   0.000     14.64447    24.45306
-----------------------------------------------------------------------------
```

Der Schätzwert für die im Mittel zu erwartende höhere Ausfallzeit (in Stunden pro Monat) pro zusätzlichem Einsatzjahr einer Maschine liegt in

A58: (A) $(-\infty, 5.0)$. (B) $[5.0, 5.5)$. (C) $[5.5, 6.0)$. (D) $[6.5, \infty)$.

Die Teststatistik für das Testproblem $H_0 : \beta_2 \geq -5$ gegen $H_1 : \beta_2 < -5$ liegt in

A59: (A) $(-\infty, -4.7)$. (B) $[-4.7, -4.5)$. (C) $[-4.5, -4.3)$. (D) $[-4.3, \infty)$.

Welcher Maschinentyp ist statistisch signifikant zuverlässiger ($\alpha = 0.05$), d.h. welcher Maschinentyp hat im Mittel kleinere Ausfallzeiten?

A60: (A) Typ B. (B) Beide Typen sind vergleichbar zuverlässig. (C) Typ A.

Dr. Ingo Steinke

Universität Mannheim
Lehrstuhl für Statistik

Klausur zur Veranstaltung Grundlagen der Statistik

FSS 2014, 1. Termin, 05. Juni 2014
Version: B

Die Klausur besteht aus einem losen Lösungsbogen, einem gehefteten Teil und einer Formelsammlung. Vergleichen Sie die Version von Lösungsbogen und geheftetem Teil.
Der geheftete Teil mit 10 Seiten besteht aus diesem Deckblatt und Aufgabenzetteln mit **60** Aufgaben. Überprüfen Sie bitte Ihr Exemplar auf Vollständigkeit!

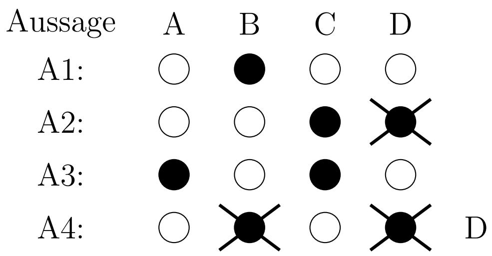
Zu jeder Aufgabe sind zwei bis vier Antworten angegeben. Von den Antwortmöglichkeiten (A) bis (D) ist genau eine richtig. Ermitteln Sie dazu die Werte, auf deren Grundlage Sie Ihre Entscheidung treffen, möglichst genau. Markieren Sie Ihre Antwort bitte auf dem Lösungsbogen. Am Ende der Klausur ist nur der **unterschriebene Lösungsbogen** abzugeben.

Unbearbeitete Aussagen, d.h. wenn keine Option markiert wurde, oder undeutliche oder unmögliche Angaben zu einer Aufgabe werden mit 0 Punkten bewertet.
Korrekt bearbeitete Aufgaben werden mit 1 Punkt, nicht korrekt bearbeitete mit 0 Punkten bewertet. Die Gesamtpunktzahl ergibt sich durch Addition aller Punkte.

Für Ihre Auswahl müssen Sie die passenden Kreise auf dem Lösungsbogen ausmalen. Bei einer ungewollten Auswahl streichen Sie den nicht korrekt markierten Kreis mit einem X durch.

Illustrierende Beispiele:

Lösungsbogen



Sie denken Aussage ...

- A1: B ist richtig.
- A2: C ist richtig, zuerst wurde D markiert.
- A3: A *oder* C ist richtig. Kein Punkt.
- A4: D ist richtig. Zuerst wurde D, dann aber B markiert.

In der Klausur werden 60 Punkte vergeben. Ab einer Punktzahl von 30 Punkten gilt die Klausur auf jeden Fall als bestanden. Die Bestehensquote wird auf minimal 60% begrenzt, d.h die Durchfallquote beträgt maximal 40%.

Die Bearbeitungszeit beträgt **180** Minuten.

Als Hilfsmittel ist nur ein nicht-programmierbarer Taschenrechner zugelassen.

Viel Erfolg!

Aufgabe
Beim zweimaligen Würfeln sei A das Ereignis, beim ersten Wurf eine ‚6' zu werfen, und B das Ereignis, beim zweiten Wurf eine ‚6' zu werfen.
Das Ereignis, höchstens eine ‚6' zu werfen, wird dann ausgedrückt durch

A1: (A) $A \cap B$. (B) $\overline{A} \cup \overline{B}$. (C) $\overline{A \cup B}$. (D) $A \cup B$.

Aufgabe
Für die Endrunde bei den Weltmeisterschaften im „Schnick-Schnack-Schnuck" haben sich 14 Teilnehmer qualifiziert. In der ersten Tournierphase werden aus den 14 Teilnehmern 7 Paarungen zufällig zusammengestellt, die dann gegeneinander antreten müssen.
Die Anzahl der verschiedenen Zusammenstellungen der Paarungen liegt in (Angaben in 1000)

A2: (A) $[0, 150)$. (B) $[150, 300)$. (C) $[300, 600)$. (D) $[600, \infty)$.

Aufgabe
In einem Schuhschrank befinden sich 6 verschiedene Paare von Schuhen. Es werden zufällig zwei Schuhe (ohne Zurücklegen) dem Schuhschrank entnommen.
Die Wahrscheinlichkeit, dass beide Schuhe ein Paar bilden, liegt in

A3: (A) $[0, 0.06)$. (B) $[0.06, 0.08)$. (C) $[0.08, 0.10)$. (D) $[0.10, 1]$.

Aufgabe
Von vier Fußballspielern sei bekannt, dass Sie Elfmeter unabhängig voneinander mit Wahrscheinlichkeit 0.8, 0.7, 0.6 und 0.5 verwandeln. Jeder der vier Spieler schieße genau ein Mal.
Die Wahrscheinlichkeit, dass mindestens zwei der vier geschossenen Elfmeter nicht ins Tor gehen, liegt in

A4: (A) $[0, 0.33)$. (B) $[0.33, 0.37)$. (C) $[0.37, 0.41)$. (D) $[0.41, 1]$.

Aufgabe
A, B, C und D seien beliebige Ereignisse.
(a) Die Ungleichung $P(A \cup B) \leq P(A) + P(B)$ ist dann

A5: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

(b) A, B, C, D seien unabhängig. Folgende Ereignisse sind dann *nicht* stets unabhängig:

A6: (A) $B \cap C$ und $C \cap D$. (B) A und $B \cap D$. (C) $A \setminus C$ und $\bar{B}$.

Aufgabe
In einem Schuhlager sind 55% der Schuhe schwarz, 25% der Schuhe braun und die restlichen Schuhe rot. Der Anteil der Damenschuhe an den schwarzen, braunen bzw. roten Schuhen betrage 50%, 40% bzw. 80%.
Der Anteil der Damenschuhe an allen Schuhen im Lager liegt dann in (Angaben in %)

A7: (A) $[0, 46)$. (B) $[46, 49)$. (C) $[49, 52)$. (D) $[52, 100]$.

Der Anteil der Damenschuhe, die eine braune Farbe besitzen, an allen Damenschuhen ist dann in (Angaben in %)

A8: (A) $[0, 20)$. (B) $[20, 22)$. (C) $[22, 24)$. (D) $[24, 100]$.

Aufgabe
X sei eine diskrete Zufallsgröße mit folgender Wahrscheinlichkeitstabelle

a_j	0	1	2	3	4
$P(X = a_j)$	0.2	0.1	0	0.4	0.3.

Der Träger zu X ist dann

A9: (A) $\mathbb{R}$. (B) $\{0, 1, 2, 3, 4\}$. (C) $\{0, 1, 3, 4\}$. (D) (A)-(C) sind falsch.

$P(|X - 0.5| \leq 0.5)$ befindet sich im Intervall

A10: (A) $[0, 0.2)$. (B) $[0.2, 0.3)$. (C) $[0.3, 0.4)$. (D) $[0.4, 1]$.

$E(X^2)$ ist in

A11: (A) $(-\infty, 7.0)$. (B) $[7.0, 8.0)$. (C) $[8.0, 9.0)$. (D) $[9.0, \infty)$.

Aufgabe
Aus einer Urne mit drei Kugeln, die mit den Zahlen 1, 2 und 3 beschriftet sind, werden nacheinander zwei Kugeln ohne Zurücklegen gezogen. Sei X_1 die erste gezogene Zahl und X_2 die zweite gezogene Zahl. Weiter sei $D = \max(X_1, X_2) - \min(X_1, X_2)$, also die Differenz aus der größeren und der kleineren der beiden Zahlen X_1 und X_2.
Dann sind X_1 und X_2

A12: (A) nicht unabhängig. (B) unabhängig.

$E(D)$ liegt in

A13: (A) $(-\infty, 1.4)$. (B) $[1.4, 1.6)$. (C) $[1.6, 1.8)$. (D) $[1.8, \infty)$.

Aufgabe
X sei eine stetige Zufallsvariable mit Verteilungsfunktion F.
Die Aussage „$F(x) < F(y)$ für alle reellen $x < y$“ ist dann

A14: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

Aufgabe
X sei stetig verteilt mit Dichtefunktion $f(x) = (\frac{1}{6} + \frac{1}{9}\,x)\, I_{[0,3]}(x)$.
$P(2 < X < 3.5)$ liegt in

A15: (A) $[0.38, 0.42)$. (B) $[0.42, 0.46)$. (C) $[0.46, 0.50)$. (D) $[0.50, 0.54)$.

$E(X)$ ist in

A16: (A) $[1.4, 1.5)$. (B) $[1.5, 1.6)$. (C) $[1.6, 1.7)$. (D) $[1.7, 1.8)$.

Das 0.8-Quantil zu X ist in

A17: (A) $[2.43, 2.49)$. (B) $[2.49, 2.55)$. (C) $[2.55, 2.61)$. (D) $[2.61, 2.67)$.

Aufgabe
X und Y seien unabhängige Zufallsvariablen mit E(X)=2, E(Y)=2, Var(X)=2 und Var(Y)=1. Var$(2X - Y)$ ist in

A18: (A) $(-\infty, 6.5)$. (B) $[6.5, 7.5)$. (C) $[7.5, 8.5)$. (D) $[8.5, \infty)$.

E$(Y \cdot X^2)$ ist in

A19: (A) $(-\infty, 10)$. (B) $[10, 11)$. (C) $[11, 12)$. (D) $[12, \infty)$.

Aufgabe
In einer Bevölkerung seien 30% gegen eine bestimmte Steuerreform. Eine Lokalzeitung befragt 8 zufällig ausgewählte Passanten unabhängig voneinander nach ihrer Meinung.
Die Wahrscheinlichkeit, dass höchstens ein Passant gegen die Steuerreform ist, liegt in

A20: (A) $[0, 0.21)$. (B) $[0.21, 0.24)$. (C) $[0.24, 0.27)$. (D) $[0.27, 1]$.

Aufgabe
Die Zeitdauer X zwischen dem Auftauchen zweier Kunden an einem Kundenschalter sei exponentialverteilt und betrage im Mittel 5 Minuten.
Die Wahrscheinlichkeit, dass nach dem Eintreffen eines Kunden mindestens 6 Minuten lang kein Kunde kommt, ist in

A21: (A) $[0, 0.29)$. (B) $[0.29, 0.32)$. (C) $[0.32, 0.35)$. (D) $[0.35, 1]$.

Aufgabe
Es sei $X \sim N(1, 9)$.
Das 0.975-Quantil zu X liegt in

A22: (A) $(-\infty, 7.0)$. (B) $[7.0, 7.3)$. (C) $[7.3, 7.6)$. (D) $[7.6, \infty)$.

$P(X^2 \leq 3)$ liegt in

A23: (A) $[0, 0.34)$. (B) $[0.34, 0.37)$. (C) $[0.37, 0.40)$. (D) $[0.40, 1]$.

Die Verteilung von $Y = 3 + 2 \cdot X$ ist dann

A24: (A) $N(5, 9)$. (B) $N(5, 36)$. (C) $N(5, 18)$. (D) (A)-(C) sind falsch.

Aufgabe
Die Zufallsvariable U sei auf $[0, 3]$ stetig gleichverteilt und $V = 0.5\,U^3$.
Der Wert der Verteilungsfunktion von U an der Stelle 2 ist dann in

A25: (A) $[0, 0.45)$. (B) $[0.45, 0.55)$. (C) $[0.55, 0.65)$. (D) $[0.65, \infty)$.

$E(V)$ liegt in

A26: (A) $(-\infty, 3.1)$. (B) $[3.1, 3.3)$. (C) $[3.3, 3.5)$. (D) $[3.5, \infty)$.

Aufgabe
Für den Zusammenhang der beiden Merkmale X = Anzahl von Personen im Haushalt und Y = Anzahl der vorhandenen Fernseher ergebe sich folgende Kontingenztabelle:

X|Y	0	1	2	3
1	0.05	0.15	0	0
2	0	0.20	0.20	0
3	0.05	0.20	0.10	0.05

Die im Mittel zu erwartende Anzahl von Fernsehern in einem Haushalt ist in

A27: (A) $[0, 1.35)$. (B) $[1.35, 1.45)$. (C) $[1.45, 1.55)$. (D) $[1.55, \infty)$.

Die im Mittel zu erwartende Anzahl von Fernsehern in einem Zweipersonenhaushalt ist in

A28: (A) $[0, 1.55)$. (B) $[1.55, 1.65)$. (C) $[1.65, 1.75)$. (D) $[1.75, \infty)$.

$\mathrm{Cov}(X, Y)$ liegt in

A29: (A) $(-\infty, 0.10)$. (B) $[0.10, 0.15)$. (C) $[0.15, 0.20)$. (D) $[0.20, \infty)$.

Aufgabe
$X_1, \ldots, X_n \sim U(0, 1)$ seien unabhängig.
Dann strebt $P(\bar{X}_n \leq 0.6)$ für $n \to \infty$ gegen einen Wert aus

A30: (A) $[0, 0.25)$. (B) $[0.25, 0.5)$. (C) $[0.5, 0.75)$. (D) $[0.75, 1.0]$.

Aufgabe
U und V seien unabhängige Zufallsvariablen mit $E(U) = E(V) = 1$, $Var(U) = 3$ und $Var(V) = 4$. Es seien $X = 2U - V$ und $Y = U + V$.
Die Aussage, dass U und V unkorreliert sind, ist dann

A31: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

$\mathrm{Cov}(X, Y)$ liegt in

A32: (A) $(-\infty, 0.5)$. (B) $[0.5, 1.5)$. (C) $[1.5, 2.5)$. (D) $[2.5, \infty)$.

Aufgabe
Ein Cateringservice soll eine Hochzeitsfeier mit 120 Gästen beliefern. Es werden 100 alkoholhaltige Sektgläser für den Sektempfang angerichtet. Erfahrungsgemäß geht man davon aus, dass ein (zufällig ausgewählter) Gast mit Wahrscheinlichkeit 0.8 Sekt trinken möchte.
Die Wahrscheinlichkeit dafür, dass genügend alkoholhaltige Sektgläser zur Verfügung gestellt werden, ist dann in

A33: (A) $[0, 0.835)$. (B) $[0.835, 0.855)$. (C) $[0.855, 0.875)$. (D) $[0.875, 1]$.

Aufgabe
$X_1, \ldots, X_n \sim B(1, \pi)$ seien u.i.v. und

$$S = \frac{2X_1 + X_n}{2}, \quad T = 2X_2 - 1, \quad U_n = (1 - \bar{X}_n), \quad V_n = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X}_n)^2.$$

Für π sind folgende Schätzer erwartungstreu:

A34: (A) S, aber nicht T. (B) T, aber nicht S. (C) weder S noch T. (D) S und T.

Für $\theta = \pi(1-\pi)$ sind folgende Schätzer konsistent:

A35: (A) U_n, aber nicht V_n. (B) weder U_n noch V_n. (C) V_n, aber nicht U_n. (D) U_n und V_n.

Aufgabe
Die Zufallsvariablen $X_1, \ldots, X_n$ seien diskret u.i.v. gemäß einer diskreten Verteilungsfamilie mit unbekanntem Parameter θ, die mit folgender Wahrscheinlichkeitstabelle beschrieben wird:

a_i	θ	$\theta+1$	$\theta+2$	$\theta+3$
$P(X=a_i)$	0.1	0.3	0.4	0.2

Der Momentenschätzer für θ ist

A36: (A) $\bar{X}_n - 1.7$. (B) $\bar{X}_n$. (C) $\bar{X}_n - 1.5$. (D) (A)-(C) sind falsch.

Aufgabe
Die Befragung von 5 zufällig ausgewählten Passanten in der Mannheimer Innenstadt nach den Ausgaben für Geburtstagsgeschenke für den Ehepartner (in Euro) ergab folgende Werte:

50, 200, 900, 500, 120.

Zur Vereinfachung nehme man an, dass die Ausgaben unabhängig und normalverteilt sind mit Erwartungswert μ und Varianz σ^2.
Die Stichprobenvarianz liegt in (Angaben in 1000)

A37: (A) $[0, 110)$. (B) $[110, 130)$. (C) $[130, 150)$. (D) $[150, \infty)$.

Die obere Grenze des 0.95-Konfidenzintervalls für μ liegt in

A38: (A) $[0, 680)$. (B) $[680, 720)$. (C) $[720, 760)$. (D) $[760, \infty)$.

Aufgabe
Eine Firma mit 4000 Angestellten möchte mit Hilfe einer Zufallsstichprobe vom Umfang 35 die mittlere Lohnsteuer μ (in Euro) schätzen, die pro Angestellten abgeführt wird. Die Erhebung ergab als arithmetisches Mittel $\bar{x} = 640$ Euro und als Stichprobenvarianz $s^2 = 256$ Euro^2.
Die Länge des approximativen Konfidenzintervalls für μ zum Niveau 0.95 liegt in

A39: (A) $[0, 10.8)$. (B) $[10.8, 11.1)$. (C) $[11.1, 11.4)$. (D) $[11.4, \infty)$.

Aufgabe
Ein Meinungsforschungsinstitut gebe als (approximatives) 95%-Konfidenzintervall für den Stimmenanteil einer Partei [0.3179, 0.3821] an.
Die Anzahl der Personen, die befragt wurden, liegt in

A40: (A) $[0, 780)$. (B) $[780, 820)$. (C) $[820, 860)$. (D) $[860, \infty)$.

Aufgabe
Für einen Parameter θ wurde aus einer Stichprobe das realisierte 0.90-Konfidenzintervall KI=[0.572,0.856] bestimmt.
Die Wahrscheinlichkeit, dass θ in KI liegt, ist dann in

A41: (A) $[0.82, 0.87)$. (B) $[0.87, 0.92)$. (C) $[0.92, 0.97)$. (D) (A)-(C) sind falsch.

Aufgabe
Eine Maschine verpacke Kartoffeln in 3.5kg-Netze. Das abgepackte Gewicht werde als normalverteilt angenommen mit Standardabweichung $\sigma = 150$ g. In der Qualitätskontrolle beim Kunden erfolgt automatisch eine Reklamation, wenn systematisch geringere Gewichte als das Sollgewicht abgepackt werden. Routinemäßig werden bei jeder Groß-Lieferung 20 Netze entnommen und deren Gesamtgewicht gemessen. Der Toleranzbereich sei so festgelegt, dass die Wahrscheinlichkeit einer irrtümlichen Reklamation 2% beträgt.
Das Gesamtgewicht c (in kg), ab dem die Lieferung abgelehnt werden darf, liegt in

A42: (A) $(-\infty, 68.3)$. (B) $[68.3, 68.8)$. (C) $[68.8, 69.3)$. (D) $[69.3, \infty)$.

Aufgabe
Die Beobachtungen $x_1, ..., x_{25}$ können aufgefasst werden als Realisierungen von unabhängigen, normalverteilten Zufallsvariablen mit unbekanntem μ und bekanntem $\sigma^2 = 1$. Als arithmetisches Mittel der Beobachtungen wurde 2.34 ermittelt. Es ist ein Test zu $H_0 : \mu \leq 2$ gegen $H_1 : \mu > 2$ zum Signifikanzniveau $\alpha = 0.05$ durchzuführen.
Der p-Wert des Tests ist in

A43: (A) $[0, 0.037)$. (B) $[0.037, 0.042)$. (C) $[0.042, 0.047)$. (D) $[0.047, 1]$.

Die Nullhypothese ist

A44: (A) abzulehnen. (B) nicht abzulehnen.

Aufgabe
Für einen Test für das Testproblem $H_0 : \theta \leq 0$ gegen $H_1 : \theta > 0$ wurde die Gütefunktion an der Stelle $\theta = -1$ berechnet und es ergab sich ein Wert von 0.1.
Dann ist 0.1 in diesem Zusammenhang die Wahrscheinlichkeit für

A45: (A) einen Fehler 1.Art. (B) einen Fehler 2.Art. (C) eine richtige Entscheidung.

Aufgabe
Für $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ u.i.v. werde für das Testproblem $H_0 : \mu = 0$ gegen $H_1 : \mu \neq 0$ ein Gauß-Test durchgeführt. Dabei seien $n = 25$ und $\sigma^2 = 1$.
Die Verteilung der Teststatistik für $\mu = 1$ ist dann:

A46: (A) $N(0, 1)$. (B) $N(2.5, 1)$. (C) $N(1, 1)$. (D) (A)-(C) sind falsch.

Aufgabe
Ein Fahrzeughersteller will den Abgasausstoß eines Motorentyps überprüfen. Testfahrten ergaben folgende Emissionswerte (in mg Kohlenwasserstoffe pro 100km)

i	1	2	3	4	5	6	7	8	9
x_i	78	81	112	103	83	97	83	90	64

Es soll mit einem statistischen Test zum Signifikanzniveau $\alpha = 0.01$ gezeigt werden, dass die Abgasnorm von im Mittel weniger als 100 mg pro 100km eingehalten wird. Gehen Sie von unabhängigen Messungen und normalverteilten Emissionswerten aus.
Der Wert der Teststatistik liegt in

A47: (A) $(-\infty, -2.4)$. (B) $[-2.4, -2.2)$. (C) $[-2.2, -2.0)$. (D) $[-2.0, \infty)$.

Der kritische Wert des Tests ist in

A48: (A) $(-\infty, -2.95)$. (B) $[-2.95, -2.55)$. (C) $[-2.55, -2.15)$. (D) $[-2.15, \infty)$.

Aufgabe
Bei einer Autobahnkontrolle wurden 300 PKW hinsichtlich korrekter Bereifung kontrolliert. Dabei ergaben sich 60 Beanstandungen. Es soll geprüft werden, ob der Anteil π an Beanstandungen bei einem Niveau von 5% signifikant höher als 15% ist.
Die Nullhypothese lautet H_0 :

A49: (A) $\pi \geq 0.15$. (B) $\pi = 0.15$. (C) $\pi \neq 0.15$. (D) $\pi \leq 0.15$.

Der Wert der Teststatistik ist im Intervall

A50: (A) $(-\infty, 2.4)$. (B) $[2.4, 2.5)$. (C) $[2.5, 2.6)$. (D) $[2.6, \infty)$.

Aufgabe
Bei der Befragung von Studenten hinsichtlich der durchschnittlichen Schlafdauer ergab sich für 119 Männer (x_i) ein Mittelwert von $\bar{x} = 7.4$ Stunden bei einer geschätzten Standardabweichung von 1.3 Stunden. Für die 87 Frauen (y_j) ergab sich ein Mittelwert von $\bar{y} = 7.6$ Stunden bei einer geschätzten Standardabweichung von 0.9 Stunden. Es soll geprüft werden, ob die mittleren Schlafdauern von Männern und Frauen zum Niveau $\alpha = 0.05$ signifikant verschieden sind.
Folgender Test ist durchzuführen (in einer Zweistichprobenversion):

A51: (A) ein Gauß-Test. (B) ein approximativer Gauß-Test. (C) ein t-Test.

Der Wert der Teststatistik liegt in

A52: (A) $(-\infty, -1.45)$. (B) $[-1.45, -1.35)$. (C) $[-1.35, -1.25)$. (D) $[-1.25, \infty)$.

Die Nullhypothese wird

A53: (A) nicht abgelehnt. (B) abgelehnt.

Aufgabe
Es soll untersucht werden, ob in unterschiedlichen Landesteilen der Anteil der Arbeitnehmer, die ihr Mittagessen zu Hause einnehmen, unterschiedlich hoch ausfällt. Dazu wurden sowohl im Norden als auch im Süden des Landes jeweils 500 Personen zufällig ausgewählt und befragt. 330 Personen im Norden und 280 im Süden gaben an, das Mittag zu Hause zu essen.
Man überprüfe mit einem Test zum Niveau 0.05, ob es regionale Unterschiede bzgl. der Mittagseinnahme gibt.
Der Betrag des Wertes der Teststatistik liegt in

A54: (A) $[0, 3.00)$. (B) $[3.00, 3.10)$. (C) $[3.10, 3.20)$. (D) $[3.20, \infty)$.

Der Betrag der kritischen Werte des Tests liegt in

A55: (A) $[0, 1.92)$. (B) $[1.92, 2.02)$. (C) $[2.02, 2.12)$. (D) $[2.12, \infty)$.

Aufgabe
Im Regressionsmodell $Y_i = \alpha + \beta x_i + \varepsilon_i$ seien die $\varepsilon_i \sim N(0, \sigma^2)$ u.i.v. Für n=60 Beobachtungen (x_i, y_i) erhielt man folgende Daten

$$\bar{x} = 1.57, \quad \bar{y} = 10.96, \quad s_x^2 = 0.269, \quad s_y^2 = 25.94, \quad \frac{1}{n}\sum_{i=1}^{n} x_i y_i = 19.07.$$

Der Schätzwert für β ist in

A56: (A) $(-\infty, 6.9)$. (B) $[6.9, 7.1)$. (C) $[7.1, 7.3)$. (D) $[7.3, \infty)$.

Der Schätzwert für $\sigma_{\hat{\beta}}$ ist in

A57: (A) $[0, 0.65)$. (B) $[0.65, 0.80)$. (C) $[0.80, 0.95)$. (D) $[0.95, \infty)$.

Aufgabe
Um die Zuverlässigkeit von Maschinen vom Typ A und Typ B zu untersuchen, wurde von verschiedenen Maschinen die Ausfallzeit für Wartung (y_i, zeit, in Stunden pro Monat) in Abhängigkeit vom Alter (a_i, alter, in Jahren) und vom Typ (t_i, typ, $t_i = 1$ für Typ A und $t_i = 2$ für Typ B) mit einer multiplen linearen Regression

$$Y_i = \beta_0 + \beta_1 a_i + \beta_2 t_i + \varepsilon_i, \quad \varepsilon \sim N(0, \sigma^2) \text{ u.i.v.}$$

modelliert. Die nachfolgende STATA-Ausgabe gibt das Ergebnis der Analyse an:

```
   Source |       SS       df       MS              Number of obs =      18
----------+------------------------------           F(  2,    15) =   77.11
    Model |  1791.73108     2  895.865539           Prob > F      =  0.0000
 Residual |  174.268921    15  11.6179281           R-squared     =  0.9114
----------+------------------------------           Adj R-squared =  0.8995
    Total |        1966    17  115.647059           Root MSE      =  3.4085

------------------------------------------------------------------------------
     zeit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------+-------------------------------------------------------------------
    alter |   6.322355   .5092307    12.42   0.000     5.236956    7.407755
      typ |  -12.55151   1.917668    -6.55   0.000    -16.63892   -8.464094
    _cons |   20.19006   2.550613     7.92   0.000     14.75356    25.62657
------------------------------------------------------------------------------
```

Der Schätzwert für die im Mittel zu erwartende höhere Ausfallzeit (in Stunden pro Monat) pro zusätzlichem Einsatzjahr einer Maschine liegt in

A58: (A) $(-\infty, 6.0)$. (B) $[6.0, 6.5)$. (C) $[6.5, 7.0)$. (D) $[7.0, \infty)$.

Die Teststatistik für das Testproblem $H_0 : \beta_2 \geq -7$ gegen $H_1 : \beta_2 < -7$ liegt in

A59: (A) $(-\infty, -3.4)$. (B) $[-3.4, -3.2)$. (C) $[-3.2, -3.0)$. (D) $[-3.0, \infty)$.

Welcher Maschinentyp ist statistisch signifikant zuverlässiger ($\alpha = 0.05$), d.h. welcher Maschinentyp hat im Mittel kleinere Ausfallzeiten?

A60: (A) Typ A. (B) Typ B. (C) Beide Typen sind vergleichbar zuverlässig.

Dr. Ingo Steinke

Universität Mannheim
Lehrstuhl für Statistik

Klausur zur Veranstaltung Grundlagen der Statistik

FSS 2014, 1. Termin, 05. Juni 2014
Version: C

Die Klausur besteht aus einem losen Lösungsbogen, einem gehefteten Teil und einer Formelsammlung. Vergleichen Sie die Version von Lösungsbogen und geheftetem Teil.
Der geheftete Teil mit 10 Seiten besteht aus diesem Deckblatt und Aufgabenzetteln mit **60** Aufgaben. Überprüfen Sie bitte Ihr Exemplar auf Vollständigkeit!

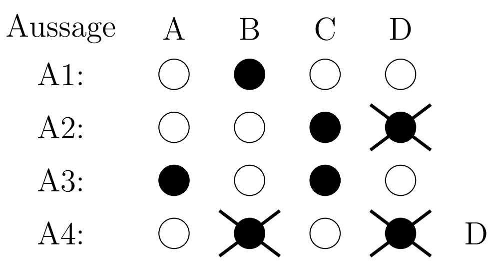
Zu jeder Aufgabe sind zwei bis vier Antworten angegeben. Von den Antwortmöglichkeiten (A) bis (D) ist genau eine richtig. Ermitteln Sie dazu die Werte, auf deren Grundlage Sie Ihre Entscheidung treffen, möglichst genau. Markieren Sie Ihre Antwort bitte auf dem Lösungsbogen. Am Ende der Klausur ist nur der **unterschriebene Lösungsbogen** abzugeben.

Unbearbeitete Aussagen, d.h. wenn keine Option markiert wurde, oder undeutliche oder unmögliche Angaben zu einer Aufgabe werden mit 0 Punkten bewertet.
Korrekt bearbeitete Aufgaben werden mit 1 Punkt, nicht korrekt bearbeitete mit 0 Punkten bewertet. Die Gesamtpunktzahl ergibt sich durch Addition aller Punkte.

Für Ihre Auswahl müssen Sie die passenden Kreise auf dem Lösungsbogen ausmalen. Bei einer ungewollten Auswahl streichen Sie den nicht korrekt markierten Kreis mit einem X durch.

Illustrierende Beispiele:

Lösungsbogen



Sie denken Aussage ...

- A1: B ist richtig.
- A2: C ist richtig, zuerst wurde D markiert.
- A3: A *oder* C ist richtig. Kein Punkt.
- A4: D ist richtig. Zuerst wurde D, dann aber B markiert.

In der Klausur werden 60 Punkte vergeben. Ab einer Punktzahl von 30 Punkten gilt die Klausur auf jeden Fall als bestanden. Die Bestehensquote wird auf minimal 60% begrenzt, d.h die Durchfallquote beträgt maximal 40%.

Die Bearbeitungszeit beträgt **180** Minuten.

Als Hilfsmittel ist nur ein nicht-programmierbarer Taschenrechner zugelassen.

Viel Erfolg!

Aufgabe
Beim zweimaligen Würfeln sei A das Ereignis, beim ersten Wurf eine ‚6' zu werfen, und B das Ereignis, beim zweiten Wurf eine ‚6' zu werfen.
Das Ereignis, höchstens eine ‚6' zu werfen, wird dann ausgedrückt durch

A1: (A) $\overline{A \cup B}$. (B) $\overline{A} \cup \overline{B}$. (C) $A \cap B$. (D) $A \cup B$.

Aufgabe
Für die Endrunde bei den Weltmeisterschaften im „Schnick-Schnack-Schnuck" haben sich 18 Teilnehmer qualifiziert. In der ersten Tournierphase werden aus den 18 Teilnehmern 9 Paarungen zufällig zusammengestellt, die dann gegeneinander antreten müssen.
Die Anzahl der verschiedenen Zusammenstellungen der Paarungen liegt in (Angaben in 10^6)

A2: (A) $[0, 40)$. (B) $[40, 80)$. (C) $[80, 150)$. (D) $[150, \infty)$.

Aufgabe
In einem Schuhschrank befinden sich 5 verschiedene Paare von Schuhen. Es werden zufällig zwei Schuhe (ohne Zurücklegen) dem Schuhschrank entnommen.
Die Wahrscheinlichkeit, dass beide Schuhe ein Paar bilden, liegt in

A3: (A) $[0, 0.08)$. (B) $[0.08, 0.10)$. (C) $[0.10, 0.12)$. (D) $[0.12, 1]$.

Aufgabe
Von vier Fußballspielern sei bekannt, dass Sie Elfmeter unabhängig voneinander mit Wahrscheinlichkeit 0.9, 0.8, 0.7 und 0.5 verwandeln. Jeder der vier Spieler schieße genau ein Mal.
Die Wahrscheinlichkeit, dass mindestens zwei der vier geschossenen Elfmeter nicht ins Tor gehen, liegt in

A4: (A) $[0, 0.28)$. (B) $[0.28, 0.31)$. (C) $[0.31, 0.34)$. (D) $[0.34, 1]$.

Aufgabe
A, B, C und D seien beliebige Ereignisse.
(a) Die Ungleichung $P(A \cup B) < P(A) + P(B)$ ist dann

A5: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

(b) A, B, C, D seien unabhängig. Folgende Ereignisse sind dann *nicht* stets unabhängig:

A6: (A) D und $A \cap C$. (B) $A \cap B$ und $B \cap C$. (C) $A \setminus D$ und $\bar{B} \cup C$.

Aufgabe
In einem Schuhlager sind 45% der Schuhe schwarz, 25% der Schuhe braun und die restlichen Schuhe rot. Der Anteil der Damenschuhe an den schwarzen, braunen bzw. roten Schuhen betrage 60%, 30% bzw. 90%.
Der Anteil der Damenschuhe an allen Schuhen im Lager liegt dann in (Angaben in %)

A7: (A) $[0, 60)$. (B) $[60, 63)$. (C) $[63, 66)$. (D) $[66, 100]$.

Der Anteil der Damenschuhe, die eine schwarze Farbe besitzen, an allen Damenschuhen ist dann in (Angaben in %)

A8: (A) $[0, 39)$. (B) $[39, 42)$. (C) $[42, 45)$. (D) $[45, 1]$.

Aufgabe
X sei eine diskrete Zufallsgröße mit folgender Wahrscheinlichkeitstabelle

a_j	0	1	2	3	4
$P(X = a_j)$	0.2	0	0.4	0.1	0.3.

Der Träger zu X ist dann

A9: (A) $\{0, 1, 2, 3, 4\}$. (B) $\mathbb{R}$. (C) $\{0, 2, 3, 4\}$. (D) (A)-(C) sind falsch.

$P(|X - 1| \leq 1)$ befindet sich im Intervall

A10: (A) $[0, 0.65)$. (B) $[0.65, 0.75)$. (C) $[0.75, 0.85)$. (D) $[0.85, 1]$.

$E(X^2)$ ist in

A11: (A) $(-\infty, 6.0)$. (B) $[6.0, 6.5)$. (C) $[6.5, 7.0)$. (D) $[7.0, \infty)$.

Aufgabe
Aus einer Urne mit drei Kugeln, die mit den Zahlen 1, 3 und 4 beschriftet sind, werden nacheinander zwei Kugeln ohne Zurücklegen gezogen. Sei X_1 die erste gezogene Zahl und X_2 die zweite gezogene Zahl. Weiter sei $D = \max(X_1, X_2) - \min(X_1, X_2)$, also die Differenz aus der größeren und der kleineren der beiden Zahlen X_1 und X_2.
Dann sind X_1 und X_2

A12: (A) nicht unabhängig. (B) unabhängig.

$E(D)$ liegt in

A13: (A) $(-\infty, 1.5)$. (B) $[1.5, 1.8)$. (C) $[1.8, 2.1)$. (D) $[2.1, \infty)$.

Aufgabe
X sei eine diskrete Zufallsvariable mit Verteilungsfunktion F.
Die Aussage „$F(x) < F(y)$ für alle reellen $x < y$“ ist dann

A14: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

Aufgabe
X sei stetig verteilt mit Dichtefunktion $f(x) = (\frac{1}{8} + \frac{1}{16}\, x)\, I_{[0,4]}(x)$.
$P(2.5 < X < 4.5)$ liegt in

A15: (A) $[0.45, 0.48)$. (B) $[0.48, 0.51)$. (C) $[0.51, 0.54)$. (D) $[0.54, 0.57)$.

$E(X)$ ist in

A16: (A) $[2.3, 2.4)$. (B) $[2.4, 2.5)$. (C) $[2.5, 2.6)$. (D) $[2.6, 2.7)$.

Das 0.8-Quantil zu X ist in

A17: (A) $[3.41, 3.47)$. (B) $[3.47, 3.53)$. (C) $[3.53, 3.59)$. (D) $[3.59, 3.65)$.

Aufgabe
X und Y seien unabhängige Zufallsvariablen mit E(X)=1, E(Y)=1, Var(X)=4 und Var(Y)=1. Var$(2X - Y)$ ist in

A18: (A) $(-\infty, 17.5)$. (B) $[17.5, 18.5)$. (C) $[18.5, 19.5)$. (D) $[19.5, \infty)$.

E$(Y \cdot X^2)$ ist in

A19: (A) $(-\infty, 5)$. (B) $[5, 6)$. (C) $[6, 7)$. (D) $[7, \infty)$.

Aufgabe
In einer Bevölkerung seien 20% gegen eine bestimmte Steuerreform. Eine Lokalzeitung befragt 8 zufällig ausgewählte Passanten unabhängig voneinander nach ihrer Meinung.
Die Wahrscheinlichkeit, dass höchstens ein Passant gegen die Steuerreform ist, liegt in

A20: (A) $[0, 0.42)$. (B) $[0.42, 0.45)$. (C) $[0.45, 0.48)$. (D) $[0.48, 1]$.

Aufgabe
Die Zeitdauer X zwischen dem Auftauchen zweier Kunden an einem Kundenschalter sei exponentialverteilt und betrage im Mittel 4 Minuten.
Die Wahrscheinlichkeit, dass nach dem Eintreffen eines Kunden mindestens 6 Minuten lang kein Kunde kommt, ist in

A21: (A) $[0, 0.18)$. (B) $[0.18, 0.21)$. (C) $[0.21, 0.24)$. (D) $[0.24, 1]$.

Aufgabe
Es sei $X \sim N(2, 4)$.
Das 0.975-Quantil zu X liegt in

A22: (A) $(-\infty, 5.1)$. (B) $[5.1, 5.4)$. (C) $[5.4, 5.7)$. (D) $[5.7, \infty)$.

$P(X^2 \leq 3)$ liegt in

A23: (A) $[0, 0.37)$. (B) $[0.37, 0.40)$. (C) $[0.40, 0.43)$. (D) $[0.43, 1]$.

Die Verteilung von $Y = 3 + 2 \cdot X$ ist dann

A24: (A) $N(7, 8)$. (B) $N(7, 16)$. (C) $N(5, 4)$. (D) (A)-(C) sind falsch.

Aufgabe
Die Zufallsvariable U sei auf $[0, 4]$ stetig gleichverteilt und $V = 0.5\, U^3$.
Der Wert der Verteilungsfunktion von U an der Stelle 2.5 ist dann in

A25: (A) $[0, 0.4)$. (B) $[0.4, 0.5)$. (C) $[0.5, 0.6)$. (D) $[0.6, \infty)$.

$E(V)$ liegt in

A26: (A) $(-\infty, 6.5)$. (B) $[6.5, 7.5)$. (C) $[7.5, 8.5)$. (D) $[8.5, \infty)$.

Aufgabe
Für den Zusammenhang der beiden Merkmale X = Anzahl von Personen im Haushalt und Y = Anzahl der vorhandenen Fernseher ergebe sich folgende Kontingenztabelle:

X|Y	0	1	2	3
1	0.05	0.15	0	0
2	0.05	0.20	0.15	0
3	0	0.25	0.10	0.05

Die im Mittel zu erwartende Anzahl von Fernsehern in einem Haushalt ist in

A27: (A) $[0, 1.0)$. (B) $[1.0, 1.1)$. (C) $[1.1, 1.2)$. (D) $[1.2, \infty)$.

Die im Mittel zu erwartende Anzahl von Fernsehern in einem Zweipersonenhaushalt ist in

A28: (A) $[0, 1.2)$. (B) $[1.2, 1.3)$. (C) $[1.3, 1.4)$. (D) $[1.4, \infty)$.

$\mathrm{Cov}(X, Y)$ liegt in

A29: (A) $(-\infty, 0.05)$. (B) $[0.05, 0.15)$. (C) $[0.15, 0.25)$. (D) $[0.25, \infty)$.

Aufgabe
$X_1, \ldots, X_n \sim U(0, 1)$ seien unabhängig.
Dann strebt $P(\bar{X}_n > 0.65)$ für $n \to \infty$ gegen einen Wert aus

A30: (A) $[0, 0.25)$. (B) $[0.25, 0.5)$. (C) $[0.5, 0.75)$. (D) $[0.75, 1.0]$.

Aufgabe
U und V seien unkorrelierte Zufallsvariablen mit $E(U) = E(V) = 1$, $Var(U) = 2$ und $Var(V) = 4$. Es seien $X = 2U - V$ und $Y = U + V$.
Die Aussage, dass U und V unabhängig sind, ist dann

A31: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

$\mathrm{Cov}(X, Y)$ liegt in

A32: (A) $(-\infty, 0.5)$. (B) $[0.5, 1.5)$. (C) $[1.5, 2.5)$. (D) $[2.5, \infty)$.

Aufgabe
Ein Cateringservice soll eine Hochzeitsfeier mit 100 Gästen beliefern. Es werden 75 alkoholhaltige Sektgläser für den Sektempfang angerichtet. Erfahrungsgemäß geht man davon aus, dass ein (zufällig ausgewählter) Gast mit Wahrscheinlichkeit 0.7 Sekt trinken möchte.
Die Wahrscheinlichkeit dafür, dass genügend alkoholhaltige Sektgläser zur Verfügung gestellt werden, ist dann in

A33: (A) $[0, 0.855)$. (B) $[0.855, 0.875)$. (C) $[0.875, 0.895)$. (D) $[0.895, 1]$.

Aufgabe
$X_1, \ldots, X_n \sim B(1, \pi)$ seien u.i.v. und

$$S = \frac{2X_1 + X_n}{3}, \quad T = 2(X_2 - 1), \quad U_n = \bar{X}_n(1 - \bar{X}_n), \quad V_n = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X}_n)^2.$$

Für π sind folgende Schätzer erwartungstreu:

A34: (A) S, aber nicht T. (B) T, aber nicht S. (C) weder S noch T. (D) S und T.

Für $\theta = \pi(1-\pi)$ sind folgende Schätzer konsistent:

A35: (A) U_n und V_n. (B) U_n, aber nicht V_n. (C) V_n, aber nicht U_n. (D) weder U_n noch V_n.

Aufgabe
Die Zufallsvariablen $X_1, \ldots, X_n$ seien diskret u.i.v. gemäß einer diskreten Verteilungsfamilie mit unbekanntem Parameter θ, die mit folgender Wahrscheinlichkeitstabelle beschrieben wird:

a_i	θ	$\theta+1$	$\theta+2$	$\theta+3$
$P(X=a_i)$	0.5	0.3	0.1	0.1

Der Momentenschätzer für θ ist

A36: (A) $\bar{X}_n - 0.6$. (B) $\bar{X}_n$. (C) $\bar{X}_n - 0.8$. (D) (A)-(C) sind falsch.

Aufgabe
Die Befragung von 5 zufällig ausgewählten Passanten in der Mannheimer Innenstadt nach den Ausgaben für Geburtstagsgeschenke für den Ehepartner (in Euro) ergab folgende Werte:

50, 200, 1200, 300, 120.

Zur Vereinfachung nehme man an, dass die Ausgaben unabhängig und normalverteilt sind mit Erwartungswert μ und Varianz σ^2.
Die Stichprobenvarianz liegt in (Angaben in 1000)

A37: (A) $[0, 230)$. (B) $[230, 250)$. (C) $[250, 270)$. (D) $[270, \infty)$.

Die obere Grenze des 0.95-Konfidenzintervalls für μ liegt in

A38: (A) $[0, 955)$. (B) $[955, 965)$. (C) $[965, 975)$. (D) $[975, \infty)$.

Aufgabe
Eine Firma mit 4000 Angestellten möchte mit Hilfe einer Zufallsstichprobe vom Umfang 45 die mittlere Lohnsteuer μ (in Euro) schätzen, die pro Angestellten abgeführt wird. Die Erhebung ergab als arithmetisches Mittel $\bar{x} = 640$ Euro und als Stichprobenvarianz $s^2 = 400$ Euro^2.
Die Länge des approximativen Konfidenzintervalls für μ zum Niveau 0.95 liegt in

A39: (A) $[0, 10.0)$. (B) $[10.0, 11.0)$. (C) $[11.0, 12.0)$. (D) $[12.0, \infty)$.

Aufgabe
Ein Meinungsforschungsinstitut gebe als (approximatives) 95%-Konfidenzintervall für den Stimmenanteil einer Partei [0.225, 0.275] an.
Die Anzahl der Personen, die befragt wurden, liegt in

A40: (A) $[0, 1180)$. (B) $[1180, 1240)$. (C) $[1240, 1300)$. (D) $[1300, \infty)$.

Aufgabe
Für einen Parameter θ wurde aus einer Stichprobe das realisierte 0.95-Konfidenzintervall KI=[0.472,0.756] bestimmt.
Die Wahrscheinlichkeit, dass θ in KI liegt, ist dann in

A41: (A) $[0.82, 0.87)$. (B) $[0.87, 0.92)$. (C) $[0.92, 0.97)$. (D) (A)-(C) sind falsch.

Aufgabe
Eine Maschine verpacke Kartoffeln in 3.5kg-Netze. Das abgepackte Gewicht werde als normalverteilt angenommen mit Standardabweichung $\sigma = 200$ g. In der Qualitätskontrolle beim Kunden erfolgt automatisch eine Reklamation, wenn systematisch geringere Gewichte als das Sollgewicht abgepackt werden. Routinemäßig werden bei jeder Groß-Lieferung 10 Netze entnommen und deren Gesamtgewicht gemessen. Der Toleranzbereich sei so festgelegt, dass die Wahrscheinlichkeit einer irrtümlichen Reklamation 1% beträgt.
Das Gesamtgewicht c (in kg), ab dem die Lieferung abgelehnt werden darf, liegt in

A42: (A) $(-\infty, 33.2)$. (B) $[33.2, 33.7)$. (C) $[33.7, 34.2)$. (D) $[34.2, \infty)$.

Aufgabe
Die Beobachtungen $x_1, ..., x_{20}$ können aufgefasst werden als Realisierungen von unabhängigen, normalverteilten Zufallsvariablen mit unbekanntem μ und bekanntem $\sigma^2 = 1$. Als arithmetisches Mittel der Beobachtungen wurde 2.59 ermittelt. Es ist ein Test zu $H_0 : \mu \leq 2.2$ gegen $H_1 : \mu > 2.2$ zum Signifikanzniveau $\alpha = 0.03$ durchzuführen.
Der p-Wert des Tests ist in

A43: (A) $[0, 0.027)$. (B) $[0.027, 0.032)$. (C) $[0.032, 0.037)$. (D) $[0.037, 1]$.

Die Nullhypothese ist

A44: (A) abzulehnen. (B) nicht abzulehnen.

Aufgabe
Für einen Test für das Testproblem $H_0 : \theta \leq 0$ gegen $H_1 : \theta > 0$ wurde die Gütefunktion an der Stelle $\theta = 1$ berechnet und es ergab sich ein Wert von 0.4.
Dann ist 0.4 in diesem Zusammenhang die Wahrscheinlichkeit für

A45: (A) eine richtige Entscheidung. (B) einen Fehler 2.Art. (C) einen Fehler 1.Art.

Aufgabe
Für $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ u.i.v. werde für das Testproblem $H_0 : \mu = 0$ gegen $H_1 : \mu \neq 0$ ein Gauß-Test durchgeführt. Dabei seien $n = 36$ und $\sigma^2 = 1$.
Die Verteilung der Teststatistik für $\mu = 1$ ist dann:

A46: (A) $N(0, 1)$. (B) $N(6, 1)$. (C) $N(1, 1)$. (D) (A)-(C) sind falsch.

Aufgabe
Ein Fahrzeughersteller will den Abgasausstoß eines Motorentyps überprüfen. Testfahrten ergaben folgende Emissionswerte (in mg Kohlenwasserstoffe pro 100km)

i	1	2	3	4	5	6	7	8	9	10
x_i	78	81	112	103	83	87	73	80	54	60

Es soll mit einem statistischen Test zum Signifikanzniveau $\alpha = 0.01$ gezeigt werden, dass die Abgasnorm von im Mittel weniger als 100 mg pro 100km eingehalten wird. Gehen Sie von unabhängigen Messungen und normalverteilten Emissionswerten aus.
Der Wert der Teststatistik liegt in

A47: (A) $(-\infty, -3.6)$. (B) $[-3.6, -3.3)$. (C) $[-3.3, -3.0)$. (D) $[-3.0, \infty)$.

Der kritische Wert des Tests ist in

A48: (A) $(-\infty, -2.75)$. (B) $[-2.75, -2.45)$. (C) $[-2.45, -2.15)$. (D) $[-2.15, \infty)$.

Aufgabe
Bei einer Autobahnkontrolle wurden 200 PKW hinsichtlich korrekter Bereifung kontrolliert. Dabei ergaben sich 50 Beanstandungen. Es soll geprüft werden, ob der Anteil π an Beanstandungen bei einem Niveau von 5% signifikant höher als 20% ist.
Die Nullhypothese lautet H_0 :

A49: (A) $\pi = 0.2$. (B) $\pi \geq 0.2$. (C) $\pi \neq 0.2$. (D) $\pi \leq 0.2$.

Der Wert der Teststatistik ist im Intervall

A50: (A) $(-\infty, 1.50)$. (B) $[1.50, 1.60)$. (C) $[1.60, 1.70)$. (D) $[1.70, \infty)$.

Aufgabe
Bei der Befragung von Studenten hinsichtlich der durchschnittlichen Schlafdauer ergab sich für 139 Männer (x_i) ein Mittelwert von $\bar{x} = 7.4$ Stunden bei einer geschätzten Standardabweichung von 1.3 Stunden. Für die 47 Frauen (y_j) ergab sich ein Mittelwert von $\bar{y} = 7.7$ Stunden bei einer geschätzten Standardabweichung von 0.9 Stunden. Es soll geprüft werden, ob die mittleren Schlafdauern von Männern und Frauen zum Niveau $\alpha = 0.05$ signifikant verschieden sind.
Folgender Test ist durchzuführen (in einer Zweistichprobenversion):

A51: (A) ein t-Test. (B) ein approximativer Gauß-Test. (C) ein Gauß-Test.

Der Wert der Teststatistik liegt in

A52: (A) $(-\infty, -1.9)$. (B) $[-1.9, -1.8)$. (C) $[-1.8, -1.7)$. (D) $[-1.7, \infty)$.

Die Nullhypothese wird

A53: (A) nicht abgelehnt. (B) abgelehnt.

Aufgabe
Es soll untersucht werden, ob in unterschiedlichen Landesteilen der Anteil der Arbeitnehmer, die ihr Mittagessen zu Hause einnehmen, unterschiedlich hoch ausfällt. Dazu wurden sowohl im Norden als auch im Süden des Landes jeweils 400 Personen zufällig ausgewählt und befragt. 280 Personen im Norden und 230 im Süden gaben an, das Mittag zu Hause zu essen.
Man überprüfe mit einem Test zum Niveau 0.05, ob es regionale Unterschiede bzgl. der Mittagseinnahme gibt.
Der Betrag des Wertes der Teststatistik liegt in

A54: (A) $[0, 3.80)$. (B) $[3.80, 3.90)$. (C) $[3.90, 4.00)$. (D) $[4.00, \infty)$.

Der Betrag der kritischen Werte des Tests liegt in

A55: (A) $[0, 1.72)$. (B) $[1.72, 1.82)$. (C) $[1.82, 1.92)$. (D) $[1.92, \infty)$.

Aufgabe
Im Regressionsmodell $Y_i = \alpha + \beta x_i + \varepsilon_i$ seien die $\varepsilon_i \sim N(0, \sigma^2)$ u.i.v. Für n=50 Beobachtungen (x_i, y_i) erhielt man folgende Daten

$$\bar{x} = 1.67, \quad \bar{y} = 10.46, \quad s_x^2 = 0.269, \quad s_y^2 = 25.94, \quad \frac{1}{n}\sum_{i=1}^{n} x_i y_i = 19.07.$$

Der Schätzwert für β ist in

A56: (A) $(-\infty, 6.2)$. (B) $[6.2, 6.4)$. (C) $[6.4, 6.6)$. (D) $[6.6, \infty)$.

Der Schätzwert für $\sigma_{\hat{\beta}}$ ist in

A57: (A) $[0, 0.90)$. (B) $[0.90, 1.05)$. (C) $[1.05, 1.20)$. (D) $[1.20, \infty)$.

Aufgabe
Um die Zuverlässigkeit von Maschinen vom Typ A und Typ B zu untersuchen, wurde von verschiedenen Maschinen die Ausfallzeit für Wartung (y_i, zeit, in Stunden pro Monat) in Abhängigkeit vom Alter (a_i, alter, in Jahren) und vom Typ (t_i, typ, $t_i = 1$ für Typ A und $t_i = 2$ für Typ B) mit einer multiplen linearen Regression

$$Y_i = \beta_0 + \beta_1 a_i + \beta_2 t_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \text{ u.i.v.}$$

modelliert. Die nachfolgende STATA-Ausgabe gibt das Ergebnis der Analyse an:

```
    Source |       SS       df       MS              Number of obs =      20
-----------+------------------------------           F(  2,    17) =  109.08
     Model |  2342.41577     2  1171.20789           Prob > F      =  0.0000
  Residual |  182.534226    17  10.7373074           R-squared     =  0.9277
-----------+------------------------------           Adj R-squared =  0.9192
     Total |     2524.95    19  132.892105           Root MSE      =  3.2768

------------------------------------------------------------------------------
      zeit |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------+------------------------------------------------------------------
     alter |   6.710326   .4544386    14.77   0.000     5.751545    7.669108
       typ |  -12.58514   1.712522    -7.35   0.000    -16.19824    -8.97203
     _cons |   19.54877   2.324515     8.41   0.000     14.64447    24.45306
------------------------------------------------------------------------------
```

Der Schätzwert für die im Mittel zu erwartende höhere Ausfallzeit (in Stunden pro Monat) pro zusätzlichem Einsatzjahr einer Maschine liegt in

A58: (A) $(-\infty, 6.0)$. (B) $[6.0, 6.5)$. (C) $[6.5, 7.0)$. (D) $[7.0, \infty)$.

Die Teststatistik für das Testproblem $H_0 : \beta_2 \geq -6$ gegen $H_1 : \beta_2 < -6$ liegt in

A59: (A) $(-\infty, -4.3)$. (B) $[-4.3, -4.1)$. (C) $[-4.1, -3.9)$. (D) $[-3.9, \infty)$.

Welcher Maschinentyp ist statistisch signifikant zuverlässiger ($\alpha = 0.05$), d.h. welcher Maschinentyp hat im Mittel kleinere Ausfallzeiten?

A60: (A) Typ A. (B) Typ B. (C) Beide Typen sind vergleichbar zuverlässig.

Dr. Ingo Steinke

Universität Mannheim
Lehrstuhl für Statistik

Klausur zur Veranstaltung Grundlagen der Statistik

FSS 2014, 1. Termin, 05. Juni 2014
Version: D

Die Klausur besteht aus einem losen Lösungsbogen, einem gehefteten Teil und einer Formelsammlung. Vergleichen Sie die Version von Lösungsbogen und geheftetem Teil.
Der geheftete Teil mit 10 Seiten besteht aus diesem Deckblatt und Aufgabenzetteln mit **60** Aufgaben. Überprüfen Sie bitte Ihr Exemplar auf Vollständigkeit!

Zu jeder Aufgabe sind zwei bis vier Antworten angegeben. Von den Antwortmöglichkeiten (A) bis (D) ist genau eine richtig. Ermitteln Sie dazu die Werte, auf deren Grundlage Sie Ihre Entscheidung treffen, möglichst genau. Markieren Sie Ihre Antwort bitte auf dem Lösungsbogen. Am Ende der Klausur ist nur der **unterschriebene Lösungsbogen** abzugeben.

Unbearbeitete Aussagen, d.h. wenn keine Option markiert wurde, oder undeutliche oder unmögliche Angaben zu einer Aufgabe werden mit 0 Punkten bewertet.
Korrekt bearbeitete Aufgaben werden mit 1 Punkt, nicht korrekt bearbeitete mit 0 Punkten bewertet. Die Gesamtpunktzahl ergibt sich durch Addition aller Punkte.

Für Ihre Auswahl müssen Sie die passenden Kreise auf dem Lösungsbogen ausmalen. Bei einer ungewollten Auswahl streichen Sie den nicht korrekt markierten Kreis mit einem X durch.

Illustrierende Beispiele:

Lösungsbogen

Aussage	A	B	C	D	
A1:	○	●	○	○	
A2:	○	○	●	~~●~~	
A3:	●	○	●	○	
A4:	○	~~●~~	○	~~●~~	D

Sie denken Aussage ...

- A1: B ist richtig.
- A2: C ist richtig, zuerst wurde D markiert.
- A3: A *oder* C ist richtig. Kein Punkt.
- A4: D ist richtig. Zuerst wurde D, dann aber B markiert.

In der Klausur werden 60 Punkte vergeben. Ab einer Punktzahl von 30 Punkten gilt die Klausur auf jeden Fall als bestanden. Die Bestehensquote wird auf minimal 60% begrenzt, d.h die Durchfallquote beträgt maximal 40%.

Die Bearbeitungszeit beträgt **180** Minuten.

Als Hilfsmittel ist nur ein nicht-programmierbarer Taschenrechner zugelassen.

Viel Erfolg!

Aufgabe
Beim zweimaligen Würfeln sei A das Ereignis, beim ersten Wurf eine ‚6' zu werfen, und B das Ereignis, beim zweiten Wurf eine ‚6' zu werfen.
Das Ereignis, höchstens eine ‚6' zu werfen, wird dann ausgedrückt durch

A1: (A) $\overline{A} \cup \overline{B}$. (B) $A \cup B$. (C) $\overline{A \cup B}$. (D) $A \cap B$.

Aufgabe
Für die Endrunde bei den Weltmeisterschaften im „Schnick-Schnack-Schnuck" haben sich 12 Teilnehmer qualifiziert. In der ersten Tournierphase werden aus den 12 Teilnehmern 6 Paarungen zufällig zusammengestellt, die dann gegeneinander antreten müssen.
Die Anzahl der verschiedenen Zusammenstellungen der Paarungen liegt in (Angaben in 10^6)

A2: (A) $[0, 3)$. (B) $[3, 6)$. (C) $[6, 12)$. (D) $[12, \infty)$.

Aufgabe
In einem Schuhschrank befinden sich 8 verschiedene Paare von Schuhen. Es werden zufällig zwei Schuhe (ohne Zurücklegen) dem Schuhschrank entnommen.
Die Wahrscheinlichkeit, dass beide Schuhe ein Paar bilden, liegt in

A3: (A) $[0, 0.06)$. (B) $[0.06, 0.07)$. (C) $[0.07, 0.08)$. (D) $[0.08, 1]$.

Aufgabe
Von vier Fußballspielern sei bekannt, dass Sie Elfmeter unabhängig voneinander mit Wahrscheinlichkeit 0.8, 0.7, 0.6 und 0.4 verwandeln. Jeder der vier Spieler schieße genau ein Mal.
Die Wahrscheinlichkeit, dass mindestens zwei der vier geschossenen Elfmeter nicht ins Tor gehen, liegt in

A4: (A) $[0, 0.42)$. (B) $[0.42, 0.46)$. (C) $[0.46, 0.50)$. (D) $[0.50, 1]$.

Aufgabe
A, B, C und D seien beliebige Ereignisse.
(a) Die Ungleichung $P(A \cup B) < P(A) + P(B)$ ist dann

A5: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

(b) A, B, C, D seien unabhängig. Folgende Ereignisse sind dann *nicht* stets unabhängig:

A6: (A) $A \cup C$ und $B \cap D$. (B) $A \setminus C$ und $\bar{B}$. (C) $B \cap C$ und $C \cup D$.

Aufgabe
In einem Schuhlager sind 55% der Schuhe schwarz, 35% der Schuhe braun und die restlichen Schuhe rot. Der Anteil der Damenschuhe an den schwarzen, braunen bzw. roten Schuhen betrage 60%, 40% bzw. 80%.
Der Anteil der Damenschuhe an allen Schuhen im Lager liegt dann in (Angaben in %)

A7: (A) $[0, 57)$. (B) $[57, 60)$. (C) $[60, 63)$. (D) $[63, 100]$.

Der Anteil der Damenschuhe, die eine rote Farbe besitzen, an allen Damenschuhen ist dann in (Angaben in %)

A8: (A) $[0, 16)$. (B) $[16, 19)$. (C) $[19, 22)$. (D) $[22, 100]$.

Aufgabe
X sei eine diskrete Zufallsgröße mit folgender Wahrscheinlichkeitstabelle

a_j	0	1	2	3	4
$P(X = a_j)$	0.2	0	0.1	0.4	0.3.

Der Träger zu X ist dann

A9: (A) $\mathbb{R}$. (B) $\{0, 1, 2, 3, 4\}$. (C) $\mathbb{N}_0 = \{0, 1, 2, 3, 4, 5, \dots\}$. (D) (A)-(C) sind falsch.

$P(|X - 0.5| \leq 0.5)$ befindet sich im Intervall

A10: (A) $[0, 0.25)$. (B) $[0.25, 0.35)$. (C) $[0.35, 0.45)$. (D) $[0.45, 1]$.

$E(X^2)$ ist in

A11: (A) $(-\infty, 8.0)$. (B) $[8.0, 9.0)$. (C) $[9.0, 10.0)$. (D) $[10.0, \infty)$.

Aufgabe
Aus einer Urne mit drei Kugeln, die mit den Zahlen 2, 3 und 4 beschriftet sind, werden nacheinander zwei Kugeln ohne Zurücklegen gezogen. Sei X_1 die erste gezogene Zahl und X_2 die zweite gezogene Zahl. Weiter sei $D = \max(X_1, X_2) - \min(X_1, X_2)$, also die Differenz aus der größeren und der kleineren der beiden Zahlen X_1 und X_2.
Dann sind X_1 und X_2

A12: (A) nicht unabhängig. (B) unabhängig.

$E(D)$ liegt in

A13: (A) $(-\infty, 1.1)$. (B) $[1.1, 1.3)$. (C) $[1.3, 1.5)$. (D) $[1.5, \infty)$.

Aufgabe
X sei eine stetige Zufallsvariable mit Träger $(0, 1)$ und Verteilungsfunktion F.
Die Aussage „$F(x) < F(y)$ für alle reellen $x < y$ mit $x, y \in (0, 1)$“ ist dann

A14: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

Aufgabe
X sei stetig verteilt mit Dichtefunktion $f(x) = (\frac{1}{10} + \frac{1}{25}\,x)\, I_{[0,5]}(x)$.
$P(3 < X < 5.5)$ liegt in

A15: (A) $[0.42, 0.46)$. (B) $[0.46, 0.50)$. (C) $[0.50, 0.54)$. (D) $[0.54, 0.58)$.

$E(X)$ ist in

A16: (A) $[2.6, 2.8)$. (B) $[2.8, 3.0)$. (C) $[3.0, 3.2)$. (D) $[3.2, 3.4)$.

Das 0.8-Quantil zu X ist in

A17: (A) $[4.25, 4.35)$. (B) $[4.35, 4.45)$. (C) $[4.45, 4.55)$. (D) $[4.55, 4.65)$.

Aufgabe
X und Y seien unabhängige Zufallsvariablen mit E(X)=2, E(Y)=2, Var(X)=3 und Var(Y)=1. Var$(2X - Y)$ ist in

A18: (A) $(-\infty, 12.5)$. (B) $[12.5, 13.5)$. (C) $[13.5, 14.5)$. (D) $[14.5, \infty)$.

E$(Y \cdot X^2)$ ist in

A19: (A) $(-\infty, 11)$. (B) $[11, 12)$. (C) $[12, 13)$. (D) $[13, \infty)$.

Aufgabe
In einer Bevölkerung seien 30% gegen eine bestimmte Steuerreform. Eine Lokalzeitung befragt 10 zufällig ausgewählte Passanten unabhängig voneinander nach ihrer Meinung.
Die Wahrscheinlichkeit, dass höchstens ein Passant gegen die Steuerreform ist, liegt in

A20: (A) $[0, 0.11)$. (B) $[0.11, 0.14)$. (C) $[0.14, 0.17)$. (D) $[0.17, 1]$.

Aufgabe
Die Zeitdauer X zwischen dem Auftauchen zweier Kunden an einem Kundenschalter sei exponentialverteilt und betrage im Mittel 5 Minuten.
Die Wahrscheinlichkeit, dass nach dem Eintreffen eines Kunden mindestens 7 Minuten lang kein Kunde kommt, ist in

A21: (A) $[0, 0.17)$. (B) $[0.17, 0.20)$. (C) $[0.20, 0.23)$. (D) $[0.23, 1]$.

Aufgabe
Es sei $X \sim N(2, 9)$.
Das 0.975-Quantil zu X liegt in

A22: (A) $(-\infty, 7.7)$. (B) $[7.7, 8.0)$. (C) $[8.0, 8.3)$. (D) $[8.3, \infty)$.

$P(X^2 \leq 3)$ liegt in

A23: (A) $[0, 0.29)$. (B) $[0.29, 0.32)$. (C) $[0.32, 0.35)$. (D) $[0.35, 1]$.

Die Verteilung von $Y = 3 + 2 \cdot X$ ist dann

A24: (A) $N(7, 18)$. (B) $N(3, 18)$. (C) $N(7, 36)$. (D) (A)-(C) sind falsch.

Aufgabe
Die Zufallsvariable U sei auf $[0, 5]$ stetig gleichverteilt und $V = 0.5\, U^3$.
Der Wert der Verteilungsfunktion von U an der Stelle 3 ist dann in

A25: (A) $[0, 0.65)$. (B) $[0.65, 0.75)$. (C) $[0.75, 0.85)$. (D) $[0.85, \infty)$.

$E(V)$ liegt in

A26: (A) $(-\infty, 15.0)$. (B) $[15.0, 16.0)$. (C) $[16.0, 17.0)$. (D) $[17.0, \infty)$.

Aufgabe
Für den Zusammenhang der beiden Merkmale X = Anzahl von Personen im Haushalt und Y = Anzahl der vorhandenen Fernseher ergebe sich folgende Kontingenztabelle:

X|Y	0	1	2	3
1	0.05	0.15	0	0
2	0.05	0.20	0.20	0
3	0	0.15	0.15	0.05

Die im Mittel zu erwartende Anzahl von Fernsehern in einem Haushalt ist in

A27: (A) $[0, 1.1)$. (B) $[1.1, 1.2)$. (C) $[1.2, 1.3)$. (D) $[1.3, \infty)$.

Die im Mittel zu erwartende Anzahl von Fernsehern in einem Zweipersonenhaushalt ist in

A28: (A) $[0, 1.1)$. (B) $[1.1, 1.2)$. (C) $[1.2, 1.3)$. (D) $[1.3, \infty)$.

$\mathrm{Cov}(X, Y)$ liegt in

A29: (A) $(-\infty, 0.2)$. (B) $[\,0.2, 0.3)$. (C) $[\,0.3, 0.4)$. (D) $[0.4, \infty)$.

Aufgabe
$X_1, \ldots, X_n \sim U(0, 1)$ seien unabhängig.
Dann strebt $P(\bar{X}_n \geq 0.4)$ für $n \to \infty$ gegen einen Wert aus

A30: (A) $[0, 0.25)$. (B) $[0.25, 0.5)$. (C) $[0.5, 0.75)$. (D) $[0.75, 1.0]$.

Aufgabe
U und V seien unabhängige Zufallsvariablen mit $E(U) = E(V) = 1$, $Var(U) = 3$ und $Var(V) = 2$. Es seien $X = 2U - V$ und $Y = U + V$.
Die Aussage, dass U und V unkorreliert sind, ist dann

A31: (A) stets richtig. (B) weder stets richtig noch stets falsch. (C) stets falsch.

$\mathrm{Cov}(X, Y)$ liegt in

A32: (A) $(-\infty, 1.5)$. (B) $[1.5, 2.5)$. (C) $[2.5, 3.5)$. (D) $[3.5, \infty)$.

Aufgabe
Ein Cateringservice soll eine Hochzeitsfeier mit 120 Gästen beliefern. Es werden 85 alkoholhaltige Sektgläser für den Sektempfang angerichtet. Erfahrungsgemäß geht man davon aus, dass ein (zufällig ausgewählter) Gast mit Wahrscheinlichkeit 0.7 Sekt trinken möchte.
Die Wahrscheinlichkeit dafür, dass genügend alkoholhaltige Sektgläser zur Verfügung gestellt werden, ist dann in

A33: (A) $[0, 0.605)$. (B) $[0.605, 0.625)$. (C) $[0.625, 0.645)$. (D) $[0.645, 1]$.

Aufgabe
$X_1, \ldots, X_n \sim B(1, \pi)$ seien u.i.v. und

$$S = \frac{2X_1 + X_n}{2}, \quad T = 2X_2 - X_3, \quad U_n = (1 - \bar{X}_n)\bar{X}_n, \quad V_n = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X}_n)^2.$$

Für π sind folgende Schätzer erwartungstreu:

A34: (A) S, aber nicht T. (B) T, aber nicht S. (C) weder S noch T. (D) S und T.

Für $\theta = \pi(1-\pi)$ sind folgende Schätzer konsistent:

A35: (A) U_n, aber nicht V_n. (B) weder U_n noch V_n. (C) V_n, aber nicht U_n. (D) U_n und V_n.

Aufgabe
Die Zufallsvariablen $X_1, \ldots, X_n$ seien diskret u.i.v. gemäß einer diskreten Verteilungsfamilie mit unbekanntem Parameter θ, die mit folgender Wahrscheinlichkeitstabelle beschrieben wird:

a_i	θ	$\theta+1$	$\theta+2$	$\theta+3$
$P(X=a_i)$	0.4	0.3	0.1	0.2

Der Momentenschätzer für θ ist

A36: (A) $\bar{X}_n - 1.5$. (B) $\bar{X}_n$. (C) $\bar{X}_n - 1.1$. (D) (A)-(C) sind falsch.

Aufgabe
Die Befragung von 5 zufällig ausgewählten Passanten in der Mannheimer Innenstadt nach den Ausgaben für Geburtstagsgeschenke für den Ehepartner (in Euro) ergab folgende Werte:

50, 200, 900, 300, 120.

Zur Vereinfachung nehme man an, dass die Ausgaben unabhängig und normalverteilt sind mit Erwartungswert μ und Varianz σ^2.
Die Stichprobenvarianz liegt in (Angaben in 1000)

A37: (A) $[0, 110)$. (B) $[110, 130)$. (C) $[130, 150)$. (D) $[150, \infty)$.

Die obere Grenze des 0.95-Konfidenzintervalls für μ liegt in

A38: (A) $[0, 750)$. (B) $[750, 780)$. (C) $[780, 810)$. (D) $[810, \infty)$.

Aufgabe
Eine Firma mit 4000 Angestellten möchte mit Hilfe einer Zufallsstichprobe vom Umfang 45 die mittlere Lohnsteuer μ (in Euro) schätzen, die pro Angestellten abgeführt wird. Die Erhebung ergab als arithmetisches Mittel $\bar{x} = 640$ Euro und als Stichprobenvarianz $s^2 = 256$ Euro^2.
Die Länge des approximativen Konfidenzintervalls für μ zum Niveau 0.95 liegt in

A39: (A) $[0, 9.5)$. (B) $[9.5, 9.8)$. (C) $[9.8, 10.1)$. (D) $[10.1, \infty)$.

Aufgabe
Ein Meinungsforschungsinstitut gebe als (approximatives) 95%-Konfidenzintervall für den Stimmenanteil einer Partei [0.3236, 0.3764] an.
Die Anzahl der Personen, die befragt wurden, liegt in

A40: (A) $[0, 1200)$. (B) $[1200, 1240)$. (C) $[1240, 1280)$. (D) $[1280, \infty)$.

Aufgabe
Für einen Parameter θ wurde aus einer Stichprobe das realisierte 0.95-Konfidenzintervall KI=[0.572,0.856] bestimmt.
Die Wahrscheinlichkeit, dass θ in KI liegt, ist dann in

A41: (A) $[0.82, 0.87)$. (B) $[0.87, 0.92)$. (C) $[0.92, 0.97)$. (D) (A)-(C) sind falsch.

Aufgabe
Eine Maschine verpacke Kartoffeln in 3.5kg-Netze. Das abgepackte Gewicht werde als normalverteilt angenommen mit Standardabweichung $\sigma = 200$ g. In der Qualitätskontrolle beim Kunden erfolgt automatisch eine Reklamation, wenn systematisch geringere Gewichte als das Sollgewicht abgepackt werden. Routinemäßig werden bei jeder Groß-Lieferung 20 Netze entnommen und deren Gesamtgewicht gemessen. Der Toleranzbereich sei so festgelegt, dass die Wahrscheinlichkeit einer irrtümlichen Reklamation 2% beträgt.
Das Gesamtgewicht c (in kg), ab dem die Lieferung abgelehnt werden darf, liegt in

A42: (A) $(-\infty, 68.4)$. (B) $[68.4, 68.9)$. (C) $[68.9, 69.4)$. (D) $[69.4, \infty)$.

Aufgabe
Die Beobachtungen $x_1, ..., x_{25}$ können aufgefasst werden als Realisierungen von unabhängigen, normalverteilten Zufallsvariablen mit unbekanntem μ und bekanntem $\sigma^2 = 1$. Als arithmetisches Mittel der Beobachtungen wurde 2.59 ermittelt. Es ist ein Test zu $H_0 : \mu \leq 2.2$ gegen $H_1 : \mu > 2.2$ zum Signifikanzniveau $\alpha = 0.05$ durchzuführen.
Der p-Wert des Tests ist in

A43: (A) $[0, 0.022)$. (B) $[0.022, 0.027)$. (C) $[0.027, 0.032)$. (D) $[0.032, 1]$.

Die Nullhypothese ist

A44: (A) abzulehnen. (B) nicht abzulehnen.

Aufgabe
Für einen Test für das Testproblem $H_0 : \theta \leq 0$ gegen $H_1 : \theta > 0$ wurde die Gütefunktion an der Stelle $\theta = 0.5$ berechnet und es ergab sich ein Wert von 0.3.
Dann ist 0.7 in diesem Zusammenhang die Wahrscheinlichkeit für

A45: (A) einen Fehler 1.Art. (B) einen Fehler 2.Art. (C) eine richtige Entscheidung.

Aufgabe
Für $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ u.i.v. werde für das Testproblem $H_0 : \mu = 0$ gegen $H_1 : \mu \neq 0$ ein Gauß-Test durchgeführt. Dabei seien $n = 49$ und $\sigma^2 = 1$.
Die Verteilung der Teststatistik für $\mu = 1$ ist dann:

A46: (A) $N(5, 1)$. (B) $N(3.5, 1)$. (C) $N(0, 1)$. (D) (A)-(C) sind falsch.

Aufgabe
Ein Fahrzeughersteller will den Abgasausstoß eines Motorentyps überprüfen. Testfahrten ergaben folgende Emissionswerte (in mg Kohlenwasserstoffe pro 100km)

i	1	2	3	4	5	6	7	8
x_i	78	81	112	103	83	97	83	90

Es soll mit einem statistischen Test zum Signifikanzniveau $\alpha = 0.01$ gezeigt werden, dass die Abgasnorm von im Mittel weniger als 100 mg pro 100km eingehalten wird. Gehen Sie von unabhängigen Messungen und normalverteilten Emissionswerten aus.
Der Wert der Teststatistik liegt in

A47: (A) $(-\infty, -2.6)$. (B) $[-2.6, -2.4)$. (C) $[-2.4, -2.2)$. (D) $[-2.2, \infty)$.

Der kritische Wert des Tests ist in

A48: (A) $(-\infty, -2.75)$. (B) $[-2.75, -2.45)$. (C) $[-2.45, -2.15)$. (D) $[-2.15, \infty)$.

Aufgabe
Bei einer Autobahnkontrolle wurden 300 PKW hinsichtlich korrekter Bereifung kontrolliert. Dabei ergaben sich 70 Beanstandungen. Es soll geprüft werden, ob der Anteil π an Beanstandungen bei einem Niveau von 5% signifikant höher als 20% ist.
Die Nullhypothese lautet H_0 :

A49: (A) $\pi = 0.2$. (B) $\pi \geq 0.2$. (C) $\pi \leq 0.2$. (D) $\pi \neq 0.2$.

Der Wert der Teststatistik ist im Intervall

A50: (A) $(-\infty, 1.3)$. (B) $[1.3, 1.4)$. (C) $[1.4, 1.5)$. (D) $[1.5, \infty)$.

Aufgabe
Bei der Befragung von Studenten hinsichtlich der durchschnittlichen Schlafdauer ergab sich für 139 Männer (x_i) ein Mittelwert von $\bar{x} = 7.4$ Stunden bei einer geschätzten Standardabweichung von 1.3 Stunden. Für die 87 Frauen (y_j) ergab sich ein Mittelwert von $\bar{y} = 7.7$ Stunden bei einer geschätzten Standardabweichung von 0.9 Stunden. Es soll geprüft werden, ob die mittleren Schlafdauern von Männern und Frauen zum Niveau $\alpha = 0.05$ signifikant verschieden sind.
Folgender Test ist durchzuführen (in einer Zweistichprobenversion):

A51: (A) ein Gauß-Test. (B) ein approximativer Gauß-Test. (C) ein t-Test.

Der Wert der Teststatistik liegt in

A52: (A) $(-\infty, -2.20)$. (B) $[-2.20, -2.10)$. (C) $[-2.10, -2.00)$. (D) $[-2.00, \infty)$.

Die Nullhypothese wird

A53: (A) nicht abgelehnt. (B) abgelehnt.

Aufgabe
Es soll untersucht werden, ob in unterschiedlichen Landesteilen der Anteil der Arbeitnehmer, die ihr Mittagessen zu Hause einnehmen, unterschiedlich hoch ausfällt. Dazu wurden sowohl im Norden als auch im Süden des Landes jeweils 400 Personen zufällig ausgewählt und befragt. 280 Personen im Norden und 230 im Süden gaben an, das Mittag zu Hause zu essen.
Man überprüfe mit einem Test zum Niveau 0.05, ob es regionale Unterschiede bzgl. der Mittagseinnahme gibt.
Der Betrag des Wertes der Teststatistik liegt in

A54: (A) $[0, 3.65)$. (B) $[3.65, 3.75)$. (C) $[3.75, 3.85)$. (D) $[3.85, \infty)$.

Der Betrag der kritischen Werte des Tests liegt in

A55: (A) $[0, 1.70)$. (B) $[1.70, 1.80)$. (C) $[1.80, 1.90)$. (D) $[1.90, \infty)$.

Aufgabe
Im Regressionsmodell $Y_i = \alpha + \beta x_i + \varepsilon_i$ seien die $\varepsilon_i \sim N(0, \sigma^2)$ u.i.v. Für n=60 Beobachtungen (x_i, y_i) erhielt man folgende Daten

$$\bar{x} = 1.67, \quad \bar{y} = 10.46, \quad s_x^2 = 0.269, \quad s_y^2 = 25.94, \quad \frac{1}{n}\sum_{i=1}^{n} x_i y_i = 19.07.$$

Der Schätzwert für β ist in

A56: (A) $(-\infty, 5.70)$. (B) $[5.70, 5.85)$. (C) $[5.85, 6.00)$. (D) $[6.00, \infty)$.

Der Schätzwert für $\sigma_{\hat{\beta}}$ ist in

A57: (A) $[0, 1.05)$. (B) $[1.05, 1.15)$. (C) $[1.15, 1.25)$. (D) $[1.25, \infty)$.

Aufgabe
Um die Zuverlässigkeit von Maschinen vom Typ A und Typ B zu untersuchen, wurde von verschiedenen Maschinen die Ausfallzeit für Wartung (y_i, zeit, in Stunden pro Monat) in Abhängigkeit vom Alter (a_i, alter, in Jahren) und vom Typ (t_i, typ, $t_i = 1$ für Typ A und $t_i = 2$ für Typ B) mit einer multiplen linearen Regression

$$Y_i = \beta_0 + \beta_1 a_i + \beta_2 t_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \text{ u.i.v.}$$

modelliert. Die nachfolgende STATA-Ausgabe gibt das Ergebnis der Analyse an:

```
   Source |       SS       df       MS              Number of obs =      18
----------+------------------------------           F(  2,    15) =   77.11
    Model |  1791.73108     2  895.865539           Prob > F      =  0.0000
 Residual |  174.268921    15  11.6179281           R-squared     =  0.9114
----------+------------------------------           Adj R-squared =  0.8995
    Total |        1966    17  115.647059           Root MSE      =  3.4085

----------------------------------------------------------------------------
     zeit |      Coef.   Std. Err.       t    P>|t|      [95% Conf. Interval]
----------+-----------------------------------------------------------------
    alter |   6.322355   .5092307    12.42   0.000     5.236956    7.407755
      typ |  -12.55151   1.917668    -6.55   0.000    -16.63892   -8.464094
    _cons |   20.19006   2.550613     7.92   0.000     14.75356    25.62657
----------------------------------------------------------------------------
```

Der Schätzwert für die im Mittel zu erwartende höhere Ausfallzeit (in Stunden pro Monat) pro zusätzlichem Einsatzjahr einer Maschine liegt in

A58: (A) $(-\infty, 5.0)$. (B) $[5.0, 5.5)$. (C) $[5.5, 6.0)$. (D) $[6.0, \infty)$.

Die Teststatistik für das Testproblem $H_0 : \beta_2 \geq -8$ gegen $H_1 : \beta_2 < -8$ liegt in

A59: (A) $(-\infty, -2.6)$. (B) $[-2.6, -2.4)$. (C) $[-2.4, -2.2)$. (D) $[-2.2, \infty)$.

Welcher Maschinentyp ist statistisch signifikant zuverlässiger ($\alpha = 0.05$), d.h. welcher Maschinentyp hat im Mittel kleinere Ausfallzeiten?

A60: (A) Typ A. (B) Beide Typen sind vergleichbar zuverlässig. (C) Typ B.