**Wirtschaftswissenschaftliches Prüfungsamt**

Bachelor-Prüfung
Schliessende Statistik
Wintersemester 2014/15

Aufgabenstellung und Ergebnisse

**Dr. Martin Becker**

Hinweise für die Klausurteilnehmer

- Kleben Sie bitte **sofort** Ihr Namensschild auf obige Markierung!
- Die Klausur besteht aus insgesamt 9 Aufgaben. Prüfen Sie die Vollständigkeit Ihres Exemplares nach; spätere Reklamationen können nicht berücksichtigt werden.
- Es sind insgesamt 120 Punkte (= 16 + 12 + 8 + 8 + 13 + 18 + 15 + 6 + 24) erreichbar.
- Als Hilfsmittel sind zugelassen: Taschenrechner (auch mit Grafikfähigkeit), 2 selbsterstellte DIN-A4 Blätter bzw. 4 selbsterstellte (einseitige) DIN-A4 Seiten. Benötigte Tabellen zur Normal- und $t$-Verteilung finden Sie am Ende dieses Klausurheftes.
- Bei der Korrektur werden **nur** die Lösungen auf den Seiten 1–12 berücksichtigt. Das letzte Blatt (Tabellen zur Normal- und $t$-Verteilung) darf abgetrennt werden.
- Bei mehreren Lösungsvorschlägen muss die gültige Lösung eindeutig gekennzeichnet sein.
- Mit Ausnahme der Multiple-Choice-Aufgaben muss der Lösungsweg klar ersichtlich sein.

| **Bewertungsteil — Bitte nicht beschreiben** | | | | | | | |
|---|---|---|---|---|---|---|---|
| Aufgabe | (a) | (b) | (c) | (d) | (e) | (f) | $\Sigma$ |
| 1 | | | | | | | |
| 2 | | | | | | | |
| 3 | | | | | | | |
| 4 | | | | | | | |
| 5 | | | | | | | |
| 6 | | | | | | | |
| 7 | | | | | | | |
| 8 | | | | | | | |
| 9 | | | | | | | |
| $\Sigma$ | | | | | | | |

## Aufgabe 1 (16 Punkte)

**Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.**
**Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).**

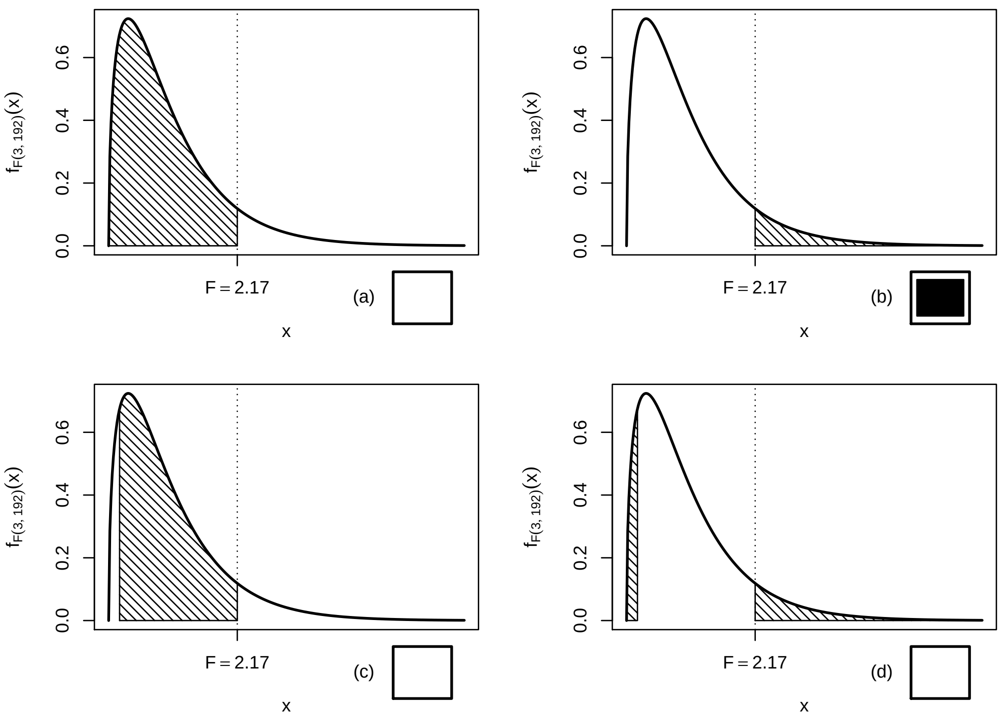
| | Aussage | wahr | falsch |
|---|---|---|---|
| 1. | Sei $X_1, \dots, X_n$ eine einfache Stichprobe vom Umfang $n$ zu einer normalverteilten Zufallsvariablen $Y$. Dann besitzen $X_1, \dots, X_n$ stets übereinstimmende Erwartungswerte und übereinstimmende Varianzen. | ■ | □ |
| 2. | Ist eine Familie von Schätzfunktionen $T_n, n \in \mathbb{N}$, konsistent im quadratischen Mittel für einen Parameter $\lambda \in \mathbb{R}$, so gilt stets $$\lim_{n\to\infty} \mathrm{Var}(T_n) = 0.$$ | ■ | □ |
| 3. | Die Breite von Konfidenzintervallen für den Erwartungswert $\mu$ einer normalverteilten Zufallsvariablen bei bekannter Varianz vergrößert sich mit wachsender Varianz. | ■ | □ |
| 4. | Das Verkleinern des Signifikanzniveaus eines Hypothesentests führt stets zu einer Vergrößerung des kritischen Bereichs. | □ | ■ |
| 5. | Gilt bei der Anwendung eines statistischen Tests für den $p$-Wert $p$ zur Teststatistik $T$ die Beziehung $p < \alpha$, so liegt $T$ im kritischen Bereich des zugehörigen Tests zum Signifikanzniveau $\alpha$. | ■ | □ |
| 6. | Nimmt ein zweiseitiger Gauß-Test die Nullhypothese zum Signifikanzniveau $\alpha = 0.05$ an, so wird $H_0$ stets auch zum Signifikanzniveau $\alpha = 0.10$ angenommen. | □ | ■ |
| 7. | Sind die Voraussetzungen zur Anwendung der einfachen Varianzanalyse erfüllt und liegen als Stichprobenrealisation jeweils 20 Beobachtungen zu 5 Faktorstufen vor, so ist die Teststatistik bei Gültigkeit der Nullhypothese $F(5, 95)$-verteilt. | □ | ■ |
| 8. | Die Anwendung der Kleinst-Quadrate-Methode bedeutet, die Summe der quadrierten horizontalen Abstände der Beobachtungspunkte zur Regressionsgeraden zu maximieren. | □ | ■ |

**Aufgabe 2** (12 Punkte)

**Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.**
**Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).**

1. Beim zweiseitigen Gauß-Test für den Erwartungswert einer normalverteilten Zufallsvariablen mit bekannter Varianz zum Signifikanzniveau $\alpha \in (0, 0.5)$
   (a) liegt die Teststatistik $N$ bei Gültigkeit von $H_0$ nie im kritischen Bereich $K$. ☐
   (b) liegt die Teststatistik $N$ bei Gültigkeit von $H_0$ mit Wahrscheinlichkeit $\alpha$ im kritischen Bereich $K$. ■
   (c) liegt die Teststatistik $N$ bei Gültigkeit von $H_0$ mit Wahrscheinlichkeit $1 - \alpha$ im kritischen Bereich $K$. ☐
   (d) liegt die Teststatistik $N$ bei Gültigkeit von $H_0$ stets im kritischen Bereich $K$. ☐

2. Bei der Durchführung einer einfachen Varianzanalyse mit $k = 4$ Faktorstufen und einem Gesamtstichprobenumfang von $n = 196$ erhält man die realisierte Teststatistik $F = 2.17$. Markieren Sie die Abbildung, welche den $p$-Wert in der beschriebenen Situation korrekt als Inhalt der schraffierten Fläche unter der Dichtefunktion der Verteilung der Teststatistik unter $H_0$ darstellt.

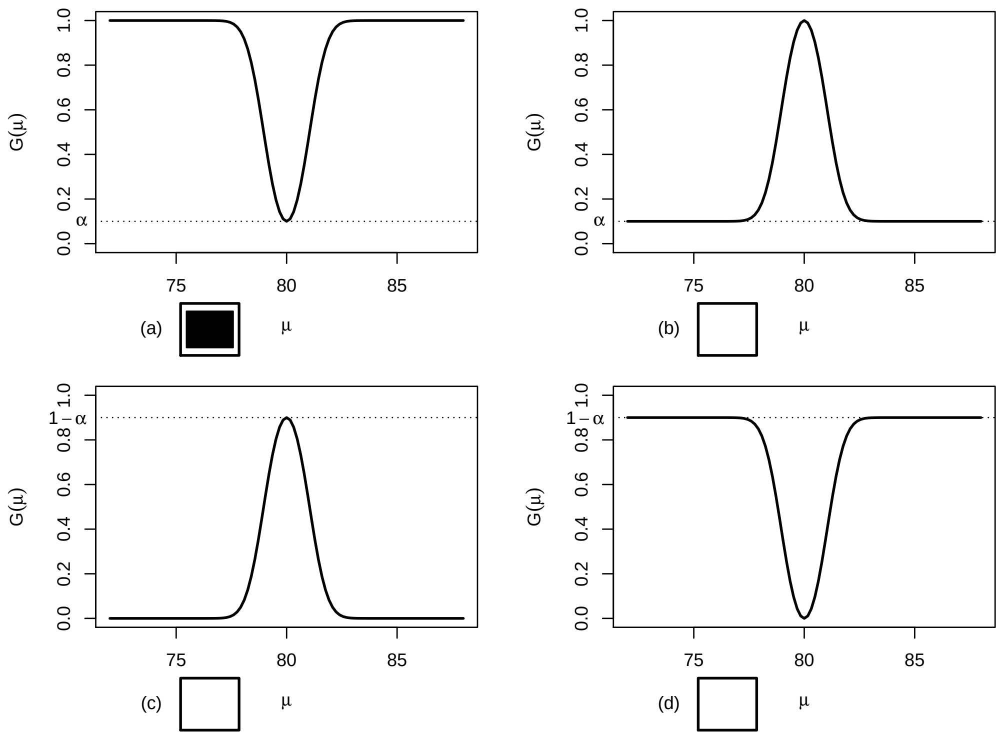


3. Beim zweiseitigen $t$-Test für den Mittelwert normalverteilter Zufallsvariablen bei unbekannter Varianz führt ein beobachteter Abstand des Stichprobenmittelwerts $\overline{x}$ zum „Sollwert“ $\mu_0$ umso eher zur Ablehnung von $H_0$, je
   (a) geringer die Streuung $s$ der Stichprobe und je geringer das Signifikanzniveau $\alpha$ ist. ☐
   (b) größer die Streuung $s$ der Stichprobe und je geringer das Signifikanzniveau $\alpha$ ist. ☐
   (c) geringer die Streuung $s$ der Stichprobe und je größer das Signifikanzniveau $\alpha$ ist. ☒
   (d) größer die Streuung $s$ der Stichprobe und je größer das Signifikanzniveau $\alpha$ ist. ☐

4. Auf der Grundlage einer einfachen Stichprobe $X_1, \dots, X_{40}$ vom Umfang $n = 40$ zu einer $N(\mu, 4^2)$-verteilten Zufallsvariablen wird ein Gauß-Test zur Überprüfung der Hypothesen

$$H_0 : \mu = 80 \qquad \text{gegen} \qquad H_1 : \mu \neq 80$$

bei einem Signifikanzniveau von $\alpha = 0.1$ betrachtet.
Markieren Sie die Abbildung, welche die Gütefunktion des oben genannten Tests korrekt darstellt.



**Aufgabe 3** (6 + 2 = 8 Punkte)

Die Verteilung einer Zufallsvariablen $Y$ sei in Abhängigkeit des unbekannten Parameters $a > 0$ durch die folgende Dichtefunktion gegeben:

$$f_Y(y|a) = \begin{cases} \frac{a^3}{2} \cdot (y+1)^2 \cdot e^{-a\cdot(y+1)} & \text{für } y > -1 \\ 0 & \text{sonst} \end{cases}$$

Der Parameter $a$ soll auf Grundlage einer einfachen Stichprobe $X_1, \ldots, X_n$ vom Umfang $n$ geschätzt werden.

(a) Bestimmen Sie den Schätzer $\widehat{a}_{ML}$ nach der Maximum-Likelihood-Methode.

(b) Man kann zeigen, dass $\mathrm{E}(Y) = \frac{3}{a} - 1$ gilt. Bestimmen Sie damit den Schätzer $\widehat{a}_{MM}$ nach der Methode der Momente.

*Hinweise:*

- *Beachten Sie, dass Sie Teil (b) auch ohne die Bearbeitung von Teil (a) lösen können.*
- *Falls sich der ML-Schätzer als lokale Extremstelle einer differenzierbaren Funktion bestimmen lässt, muss* **nicht** *überprüft werden (z.B. mit Hilfe der 2. Ableitung), ob tatsächlich eine Maximalstelle vorliegt.*

**Ergebnisse (ohne Begründung/Rechenweg):**

(a) $\widehat{a}_{ML} = \frac{3 \cdot n}{\sum_{i=1}^{n}(x_i + 1)} = \frac{3}{\overline{x} + 1}$

(b) $\widehat{a}_{MM} = \frac{3}{\overline{x} + 1}$

**Aufgabe 4** (2 + 3 + 3 = 8 Punkte)

Zu $Y \sim N(\mu, \sigma^2)$ liegen die unabhängigen einfachen Stichproben $X_1^A, \ldots, X_{10}^A$ vom Umfang 10 und $X_1^B, \ldots, X_{20}^B$ vom Umfang 20 vor. Mit $\overline{X^A} := \frac{1}{10}\sum_{i=1}^{10} X_i^A$ und $\overline{X^B} := \frac{1}{20}\sum_{i=1}^{20} X_i^B$ werden die Schätzfunktionen

- $\widehat{\mu}_1 = \dfrac{1}{2} \cdot \overline{X^A} + 2 \cdot \overline{X^B}$,
- $\widehat{\mu}_2 = \dfrac{1}{2} \cdot \overline{X^A} + \dfrac{1}{2} \cdot \overline{X^B}$ und
- $\widehat{\mu}_3 = \dfrac{1}{3} \cdot \overline{X^A} + \dfrac{2}{3} \cdot \overline{X^B}$

zur Schätzung von $\mu$ betrachtet.

(a) Wie sind $\overline{X^A}$ und $\overline{X^B}$ verteilt?

(b) Welche der Schätzfunktionen $\widehat{\mu}_1$, $\widehat{\mu}_2$ und $\widehat{\mu}_3$ sind erwartungstreu für $\mu$? Begründen Sie Ihre Antwort.

(c) Berechnen Sie zu den für $\mu$ erwartungstreuen Schätzfunktionen die zugehörige Varianz. Welche dieser Schätzfunktionen würden Sie am ehesten zur Schätzung von $\mu$ einsetzen? Begründen Sie Ihre Antwort.

**Ergebnisse (ohne Begründung/Rechenweg):**

(a) Es gilt $\overline{X^A} \sim N\big(\mu, \frac{\sigma^2}{10}\big)$ und $\overline{X^B} \sim N\big(\mu, \frac{\sigma^2}{20}\big)$.

(b) Erwartungstreu sind $\widehat{\mu}_2$ und $\widehat{\mu}_3$

(c) $\text{Var}(\widehat{\mu}_2) = \frac{3}{80} \cdot \sigma^2$, $\text{Var}(\widehat{\mu}_3) = \frac{1}{30} \cdot \sigma^2 \Rightarrow \widehat{\mu}_3$ am ehesten zur Schätzung einsetzen.

**Aufgabe 5** (7 + 2 + 4 = 13 Punkte)

Bei der Abfüllung von Sanitärsilikon weiß der Hersteller aus langjähriger Erfahrung, dass die verwendete Maschine eine Standardabweichung von $1.5[ml]$ für die abgefüllte Menge hat. Nach einer routinemäßigen Überprüfung hat der Hersteller den Verdacht, dass die Abfüllanlage im Mittel mehr als die auf dem Produkt ausgezeichneten $310[ml]$ in die Kartuschen einfüllt. Dies soll mit einem statistischen Test überprüft werden. Hierzu werden der Produktion 9 Kartuschen entnommen, deren gemessene Füllmengen $x_1, \ldots, x_9$ als Realisation einer einfachen Stichprobe vom Umfang 9 zur annahmegemäß $N(\mu, 1.5^2[ml^2])$-verteilten Abfüllmenge betrachtet werden können. Als Stichprobenmittelwert ergibt sich dabei

$$\overline{x} = \frac{1}{9}\sum_{i=1}^{9} x_i = 311.183[ml] \ .$$

(a) Testen Sie zum Signifikanzniveau $\alpha = 0.05$, ob der Verdacht der Herstellerfirma bestätigt werden kann. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

(b) Berechnen Sie den $p$-Wert zum Test aus Teil (a).

(c) Wie groß ist die Wahrscheinlichkeit des Fehlers 2. Art zu dem Test aus Teil (a), falls $\mu = 311.5[ml]$ beträgt?

**Ergebnisse (ohne Begründung/Rechenweg):**

(a) $N = 2.366 \in (1.645, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

Der Test bestätigt also den Verdacht der Herstellerfirma, dass die von der Maschine abgefüllte Menge im Mittel zu hoch ist.

(b) $p = 0.0089$

(c) $\beta(311.5) = 0.0869$

**Aufgabe 6** (10 + 8 = 18 Punkte)

Es werde angenommen, dass die (in Ladezyklen gemessene) Lebensdauer $Y^A$ eines in der laufenden Produktion eingesetzten Akku-Modells normalverteilt sei mit unbekanntem Erwartungswert $\mu_A$ und unbekannter Varianz $\sigma_A^2$. Es wird erwogen, in der Produktion zukünftig ein alternatives Modell zu verwenden, dessen Lebensdauer $Y^B$ ebenfalls als normalverteilt (mit unbekanntem Erwartungswert $\mu_B$ und unbekannter Varianz $\sigma_B^2$) angenommen werden kann. Es soll überprüft werden, ob das alternative Modell im Mittel eine höhere Lebensdauer als das aktuell verwendete Modell besitzt.
Aus einem Langzeittest mit $n_A = 11$ Akkus des aktuell verwendeten und $n_B = 9$ Exemplaren des alternativen Modells erhielt man Realisationen jeweils voneinander unabhängiger einfacher Stichproben $X_1^A, \ldots, X_{11}^A$ zu $Y^A$ sowie $X_1^B, \ldots, X_9^B$ zu $Y^B$ und hieraus die zugehörigen Mittelwerte $\overline{x^A} = 1022.27$ bzw. $\overline{x^B} = 1092.67$ sowie die Stichprobenvarianzen $s_{Y^A}^2 = 1745.42$ bzw. $s_{Y^B}^2 = 2511$.

(a) Testen Sie unter der Annahme $\sigma_A^2 = \sigma_B^2$ zum Signifikanzniveau $\alpha = 0.05$ die Hypothese, dass das alternative Modell im Mittel eine höhere Lebensdauer als das aktuell verwendete Modell besitzt. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

(b) Überprüfen Sie mit einem geeigneten Test zum Signifikanzniveau $\alpha = 0.10$, ob die in Teil (a) getroffene Annahme der Varianzgleichheit auf Grundlage der vorhandenen Stichprobeninformation verworfen werden muss. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

*Hinweis: Verwenden Sie für Teil (b) den folgenden Tabellenausschnitt mit* 0.95*-Quantilen von* $F(m,n)$*-Verteilungen sowie ggf. die Rechenregel* $F_{m,n;p} = \frac{1}{F_{n,m;1-p}}$.

| $n \backslash m$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161.448 | 199.500 | 215.707 | 224.583 | 230.162 | 233.986 | 236.768 | 238.883 | 240.543 | 241.882 |
| 2 | 18.513 | 19.000 | 19.164 | 19.247 | 19.296 | 19.330 | 19.353 | 19.371 | 19.385 | 19.396 |
| 3 | 10.128 | 9.552 | 9.277 | 9.117 | 9.013 | 8.941 | 8.887 | 8.845 | 8.812 | 8.786 |
| 4 | 7.709 | 6.944 | 6.591 | 6.388 | 6.256 | 6.163 | 6.094 | 6.041 | 5.999 | 5.964 |
| 5 | 6.608 | 5.786 | 5.409 | 5.192 | 5.050 | 4.950 | 4.876 | 4.818 | 4.772 | 4.735 |
| 6 | 5.987 | 5.143 | 4.757 | 4.534 | 4.387 | 4.284 | 4.207 | 4.147 | 4.099 | 4.060 |
| 7 | 5.591 | 4.737 | 4.347 | 4.120 | 3.972 | 3.866 | 3.787 | 3.726 | 3.677 | 3.637 |
| 8 | 5.318 | 4.459 | 4.066 | 3.838 | 3.687 | 3.581 | 3.500 | 3.438 | 3.388 | 3.347 |
| 9 | 5.117 | 4.256 | 3.863 | 3.633 | 3.482 | 3.374 | 3.293 | 3.230 | 3.179 | 3.137 |
| 10 | 4.965 | 4.103 | 3.708 | 3.478 | 3.326 | 3.217 | 3.135 | 3.072 | 3.020 | 2.978 |

**Ergebnisse (ohne Begründung/Rechenweg):**

(a) $t = -3.429 \in (-\infty, -1.734) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

Der Test kann also die Vermutung, dass das alternative Modell im Mittel eine höhere Lebensdauer als das aktuell verwendete Modell besitzt, bestätigen.

(b) $F = 0.695 \notin [0, 0.326) \cup (3.347, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

Der Test findet also keine Anzeichen für eine Verletzung der in Teil (a) angenommenen Varianzgleichheit.

**Aufgabe 7** (10 + 3 + 2 = 15 Punkte)

In einer bestimmten Multiple-Choice Aufgabe ist genau eine von vier Antwortmöglichkeiten, die mit „A“, „B“, „C“ bzw. „D“ bezeichnet sind, korrekt. Bei der Korrektur der Klausur stellt der Dozent fest, dass die von den Teilnehmern der Klausur abgegebenen Antworten wie folgt verteilt sind:

| Antwort | A | B | C | D |
|---|---|---|---|---|
| Anteil in Prozent | 32 | 18 | 22 | 28 |

Der Dozent fragt sich, ob man bei dieser Verteilung der abgegebenen Antworten davon ausgehen kann, dass sich die Teilnehmer der Klausur rein zufällig (und voneinander unabhängig) für eine der vier Antworten entschieden haben.

(a) Gehen Sie zunächst davon aus, dass 200 Studierende an der Klausur teilgenommen (und die Aufgabe bearbeitet) haben. Überprüfen Sie unter der Annahme, dass sich die abgegebenen Antworten als einfache Stichprobe auffassen lassen, mit einem geeigneten statistischen Test, ob sich die Prüflinge rein zufällig für eine der vier Antworten entschieden haben oder nicht (Signifikanzniveau $\alpha = 0.05$).

(b) Würde sich das Ergebnis des Tests aus Teil (a) ändern, wenn nur 100 Prüflinge an der Klausur teilgenommen (und die Aufgabe bearbeitet) hätten? Begründen Sie Ihre Antwort! (*Beachten Sie, dass Sie den Test nicht komplett neu durchführen müssen und insbesondere die realisierte Teststatistik relativ leicht aus dem Ergebnis des vorherigen Aufgabenteils gewinnen können!*)

(c) Ab welcher Anzahl von Klausurteilnehmern würde eine Durchführung des bereits in den Teilen (a) und (b) verwendeten Tests dazu führen, davon auszugehen, dass sich die Prüflinge *nicht* rein zufällig für eine der vier Antwortmöglichkeiten entschieden haben?

*Hinweis: Verwenden Sie den folgenden Tabellenausschnitt mit p-Quantilen von $\chi^2(n)$-Verteilungen:*

| $n \backslash p$ | 0.01 | 0.025 | 0.05 | 0.50 | 0.90 | 0.95 | 0.975 | 0.99 |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.000 | 0.001 | 0.004 | 0.455 | 2.706 | 3.841 | 5.024 | 6.635 |
| 2 | 0.020 | 0.051 | 0.103 | 1.386 | 4.605 | 5.991 | 7.378 | 9.210 |
| 3 | 0.115 | 0.216 | 0.352 | 2.366 | 6.251 | 7.815 | 9.348 | 11.345 |
| 4 | 0.297 | 0.484 | 0.711 | 3.357 | 7.779 | 9.488 | 11.143 | 13.277 |
| 5 | 0.554 | 0.831 | 1.145 | 4.351 | 9.236 | 11.070 | 12.833 | 15.086 |

**Ergebnisse (ohne Begründung/Rechenweg):**

(a) $\chi^2 = 9.28 \in (7.815, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

Der Test kommt also zum Ergebnis, dass sich die Prüflinge nicht rein zufällig für eine der vier Antwortmöglichkeiten entschieden haben.

(b) $\chi^2 = 4.64 \notin (7.815, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

Der Test kommt nun zum Ergebnis, dass sich die Prüflinge rein zufällig für eine der vier Antwortmöglichkeiten entschieden haben.

(c) Es müssen mindestens 169 Prüflinge an der Klausur teilgenommen haben.

**Aufgabe 8** (1 + 1 + 1 + 1 + 1 + 1 = 6 Punkte)

Zur Erklärung der Höhe der deutschen Warenausfuhr $y_i$ (in Milliarden Euro) durch die Höhe der deutschen Wareneinfuhr $x_i$ (in Milliarden Euro) unterstellt man die Gültigkeit eines Zusammenhangs im Sinne des folgenden linearen Modells:

$$y_i = \beta_1 + \beta_2 x_i + u_i \qquad \text{mit} \quad u_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2),\ i \in \{1, \dots, n\}$$

Aus vorliegenden Daten zu den Jahren 2007–2013 wurde das lineare Modell mit der Statistik-Software R wie folgt geschätzt:

```
Call:
lm(formula = y ~ x)

Residuals:
       1        2        3        4        5        6        7
-67.8860  22.6054  -0.4926   0.2221  -8.2959  22.5067  31.3403

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  61.5209   143.0701   0.430  0.68509
x             1.1169     0.1706   6.546  0.00125 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 36.54 on 5 degrees of freedom
Multiple R-squared:  0.8955,        Adjusted R-squared:  0.8746
F-statistic: 42.85 on 1 and 5 DF,  p-value: 0.001246
```

(a) Geben Sie die realisierten Kleinst-Quadrate-Schätzwerte für $\beta_1$ und $\beta_2$ an.

(b) Geben Sie den realisierten Schätzwert für $\sigma^2$ an.

(c) Welcher Anteil der Gesamtvarianz der Höhe der deutschen Warenausfuhr wird durch das lineare Modell erklärt?

(d) Entscheiden Sie mit Hilfe des zugehörigen $p$-Werts zum Signifikanzniveau $\alpha = 0.10$, ob $\beta_1$ signifikant positiv ist.

(e) Entscheiden Sie mit Hilfe des zugehörigen $p$-Werts zum Signifikanzniveau $\alpha = 0.05$, ob $\beta_2$ signifikant von Null verschieden ist.

(f) Welche Warenausfuhrhöhe (in Milliarden Euro) prognostiziert das Modell für ein Jahr mit einer Wareneinfuhrhöhe von 900 (in Milliarden Euro)?

**Ergebnisse (ohne Begründung/Rechenweg):**

(a) $\widehat{\beta}_1 = 61.5209$, $\widehat{\beta}_2 = 1.1169$

(b) $\widehat{\sigma^2} = 1335.1716$

(c) 0.8955

(d) $\beta_1$ ist nicht signifikant positiv.

(e) $\beta_2$ ist signifikant von Null verschieden.

(f) 1066.7309

**Aufgabe 9** (6 + 2 + 3 + 5 + 3 + 5 = 24 Punkte)

Zur Schätzung eines einfachen linearen Regressionsmodells

$$y_i = \beta_1 + \beta_2 \cdot x_i + u_i \qquad \text{mit} \quad u_i \overset{\text{iid}}{\sim} N(0, \sigma^2), \quad i \in \{1, \dots, n\}$$

aus einer Stichprobe vom Umfang $n = 30$ wurden bereits die folgenden Zwischenwerte errechnet:

$$\sum_{i=1}^{30} y_i = 382; \quad \sum_{i=1}^{30} y_i^2 = 5098.139; \quad \sum_{i=1}^{30} x_i = 148.483;$$

$$\sum_{i=1}^{30} x_i^2 = 775.668; \quad \sum_{i=1}^{30} x_i \cdot y_i = 1952.706$$

(a) Schätzen Sie $\beta_1$ und $\beta_2$ mit Hilfe der Kleinst-Quadrate-Methode.

(b) Geben Sie mit Hilfe der bekannten erwartungstreuen Schätzfunktion für $\sigma^2$ den realisierten Schätzwert für $\sigma^2$ an.

(c) Berechnen Sie $\widehat{\sigma^2}_{\widehat{\beta}_1}$ und $\widehat{\sigma^2}_{\widehat{\beta}_2}$.

(d) Überprüfen Sie zum Signifikanzniveau $\alpha = 0.01$ **(!)**, ob $\beta_2$ signifikant positiv ist. Fassen Sie das Ergebnis auch in einem Antwortsatz zusammen.

(e) Geben Sie ein symmetrisches Konfidenzintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.90$ für $\beta_1$ an.

(f) Geben Sie ein Prognoseintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für $\mathrm{E}(y_0)$ gegeben $x_0 = 5$ an.

**Ergebnisse (ohne Begründung/Rechenweg):**

(a) $\widehat{\beta}_1 = 5.201$, $\widehat{\beta}_2 = 1.522$

(b) $\widehat{\sigma}^2 = 4.985$

(c) $\widehat{\sigma^2}_{\widehat{\beta}_1} = 3.152$, $\widehat{\sigma^2}_{\widehat{\beta}_2} = 0.1219$

(d) $t = 4.36 \in (2.467, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

$\beta_2$ ist also signifikant positiv.

(e) $[2.182, 8.22]$

(f) $[11.975, 13.647]$