Aufgabenstellung und Ergebnisse zur

Bachelor-Prüfung
Schliessende Statistik
Sommersemester 2024

apl. Prof. Dr. Martin Becker

Hinweise für die Klausurteilnehmer

- 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 (= 18 + 15 + 11 + 12 + 10 + 11 + 18 + 6 + 19) 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 1–11 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.
- Alle Punkte innerhalb von Zahlen sind stets Dezimalpunkte (und keine Tausenderpunkte).

Bewertungsteil — Bitte nicht beschreiben							
Aufgabe	(a)	(b)	(c)	(d)	(e)	(f)	Σ
1							
2							
3							
4							
5							
6							
7							
8							
9							
Σ							

Aufgabe 1 (18 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!).

	wahr	falsch
1. Sei $X_1, X_2, \ldots, X_n$ eine einfache Stichprobe vom Umfang n zu einer normalverteilten Zufallsvariablen Y. Dann gilt stets $\text{Cov}(X_i, X_j) = 0$ für alle $i, j \in \{1, \ldots, n\}$ mit $i \neq j$.	■	□
2. Jede Familie $\widehat{\theta}_n, n \in \mathbb{N}$, von Schätzfunktionen für einen Parameter $\theta \in \Theta$, für die (unabhängig von θ) sowohl $\text{E}(\widehat{\theta}_n) = \theta$ für jedes $n \in \mathbb{N}$ als auch $\lim_{n\to\infty} \text{Var}(\widehat{\theta}_n) = 0$ gilt, ist konsistent im quadratischen Mittel für θ.	■	□
3. Lehnt ein zweiseitiger Chi-Quadrat-Test für die Varianz einer normalverteilten Zufallsvariablen bei unbekanntem Erwartungswert die Nullhypothese zu einem Signifikanzniveau von $\alpha = 0.01$ ab, so wird die Nullhypothese stets auch bei einem entsprechenden Test zum Signifikanzniveau von $\widetilde{\alpha} = 0.05$ verworfen.	■	□
4. Bei einem linkssseitigen Gauß-Test für den Mittelwert einer normalverteilten Zufallsvariablen bei bekannter Varianz zum Signifikanzniveau $\alpha = 0.05$ muss der wahre Mittelwert μ den hypothetischen Wert μ_0 mindestens um $\alpha = 0.05$ unterschreiten, damit eine signifikante Verringerung vorliegt.	□	■
5. Ist die Nullhypothese bei einem zweiseitigen Gauß-Test zum Signifikanzniveau α erfüllt, so wird man nach erfolgter Stichprobenziehung stets einen p-Wert von mindestens α erhalten.	□	■
6. Liegt die Teststatistik bei der Durchführung eines Gauß-Tests nicht im kritischen Bereich, so führt dies entweder zu einer korrekten Entscheidung oder zu einem Fehler 2. Art.	■	□
7. Beim Test zum Vergleich von zwei Erfolgswahrscheinlichkeiten (als Spezialfall des 2-Stichproben-t-Tests) ist unter der speziellen Annahme $p_A = p_B$ die Voraussetzung $\sigma_A^2 = \sigma_B^2$ automatisch erfüllt.	■	□
8. Beim Chi-Quadrat-Unabhängigkeitstest ist bei einem Stichprobenumfang von $n = 200$ zur Konstruktion des kritischen Bereichs stets eine $\chi^2(199)$-Verteilung zu verwenden..	□	■
9. Im einfachen linearen Regressionsmodell $$y_i = \beta_1 + \beta_2 \cdot x_i + u_i, \quad u_i \overset{\text{iid}}{\sim} N(0, \sigma^2),$$ gilt für die Punktprognose $\widehat{y}_0$ für y_0 gegeben $x_0 = \overline{x}$ stets $\widehat{y}_0 = \overline{y}$.	■	□

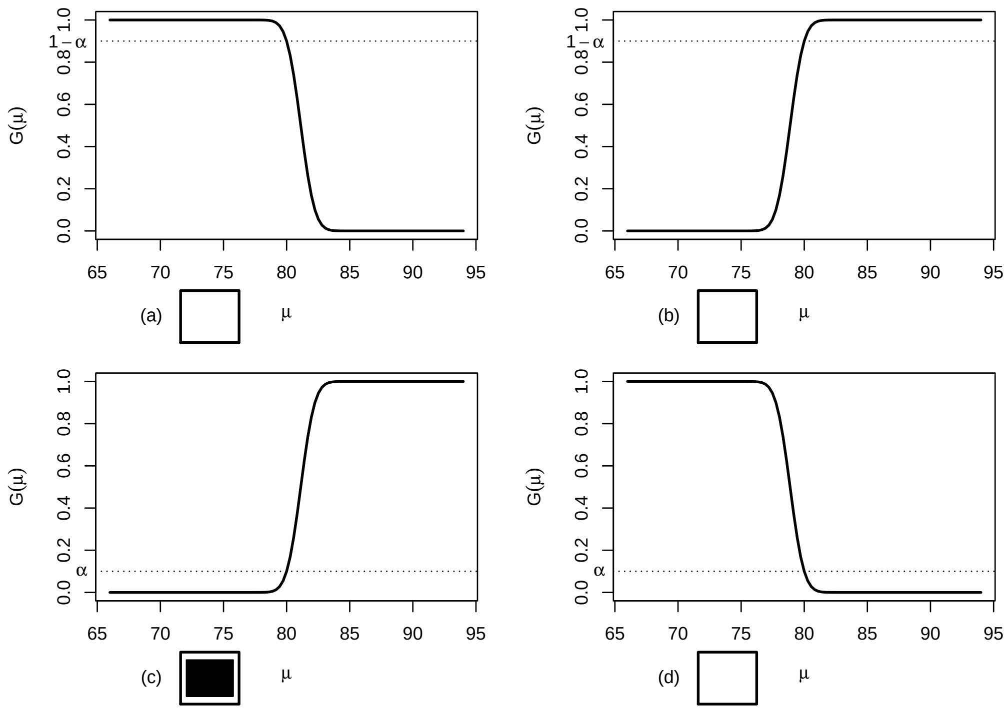
Aufgabe 2 (15 Punkte)

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. Als p-Wert zur realisierten Teststatistik eines zweiseitigen Gauß-Tests für den Mittelwert einer normalverteilten Grundgesamtheit bei bekannter Varianz (mit $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$) erhält man $p = 0.08736$. Dann gilt für das Ergebnis der einseitigen Tests (mit $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$ bzw. $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$) zum Signifikanzniveau $\alpha = 0.05$ (auf Grundlage derselben Stichprobenrealisation):
   (a) Bei beiden einseitigen Tests wird H_0 abgelehnt.
   (b) Bei keinem der beiden einseitigen Tests wird H_0 abgelehnt.
   (c) Bei genau einem der beiden einseitigen Tests wird H_0 abgelehnt. Bei welchem dies der Fall ist, hängt vom Vorzeichen der Teststatistik ab.
   (d) Auf Grundlage der vorhandenen Informationen ist noch unklar, ob bei keinem, genau einem oder beiden einseitigen Tests H_0 abgelehnt wird.

2. Auf der Grundlage einer einfachen Stichprobe $X_1, \ldots, X_{64}$ vom Umfang $n = 64$ zu einer $N(\mu, 7^2)$-verteilten Zufallsvariablen wird ein Gauß-Test zur Überprüfung der Hypothesen
$$H_0 : \mu \leq 80 \qquad \text{gegen} \qquad H_1 : \mu > 80$$
bei einem Signifikanzniveau von $\alpha = 0.1$ betrachtet.
Markieren Sie die Abbildung, welche die Gütefunktion des oben genannten Tests korrekt darstellt.



3. Konfidenzintervalle für den Erwartungswert einer normalverteilten Zufallsvariablen mit unbekannter Varianz sind umso breiter,

   (a) je größer S^2 und je größer das Konfidenzniveau $1-\alpha$ ist. ■

   (b) je größer S^2 und je kleiner das Konfidenzniveau $1-\alpha$ ist. □

   (c) je kleiner S^2 und je größer das Konfidenzniveau $1-\alpha$ ist. □

   (d) je kleiner S^2 und je kleiner das Konfidenzniveau $1-\alpha$ ist. □

4. Bei der Durchführung eines χ^2-Tests für die Varianz einer normalverteilten Zufallsvariablen bei unbekanntem Erwartungswert auf Grundlage einer einfachen Stichprobe vom Umfang n zum Signifikanzniveau 0.05 lehnen sowohl der rechtsseitige als auch der zweiseitige Test H_0 ab. Damit weiß man über die Realisation χ^2 der Teststatistik:

   (a) $\chi^2 \in [0, \chi^2_{n-1;0.025})$ □

   (b) $\chi^2 \in [\chi^2_{n-1;0.025}, \chi^2_{n-1;0.05})$ □

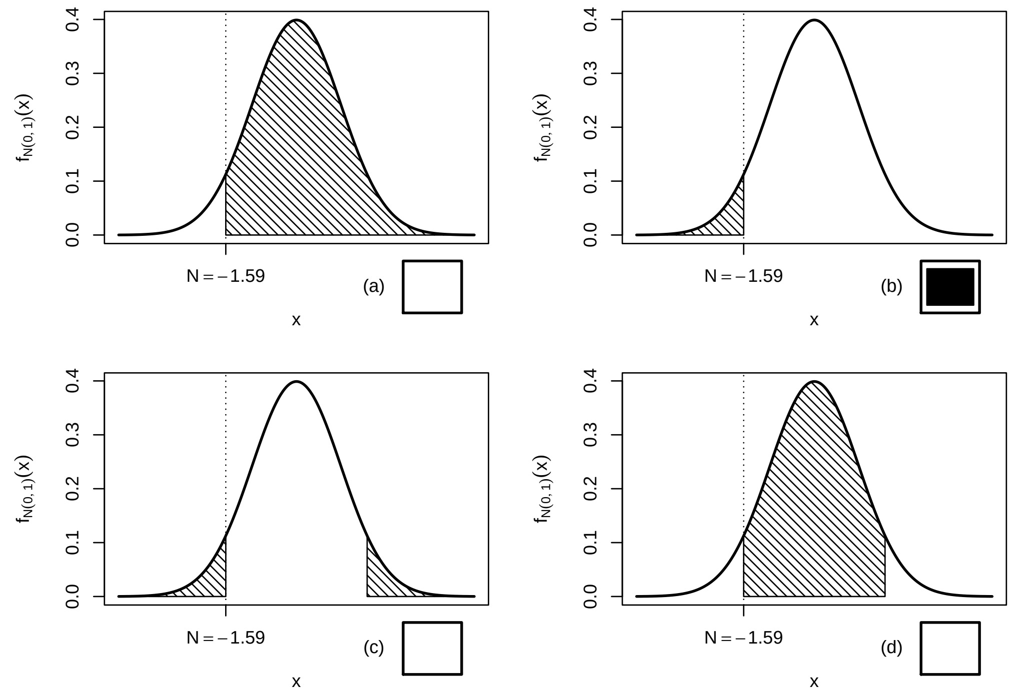
   (c) $\chi^2 \in (\chi^2_{n-1;0.95}, \chi^2_{n-1;0.975}]$ □

   (d) $\chi^2 \in (\chi^2_{n-1;0.975}, \infty)$ ■

5. Sei $X_1, \ldots, X_{25}$ eine einfache Stichprobe zu einer $N(\mu, \sigma_0^2)$-verteilten Zufallsvariablen Y mit unbekanntem Parameter μ und bekanntem $\sigma_0^2 = 4^2$. Auf der Grundlage einer Stichprobenrealisation zu dieser einfachen Stichprobe vom Umfang $n = 25$ soll

$$H_0 : \mu \geq \mu_0 = 50 \qquad \text{gegen} \qquad H_1 : \mu < \mu_0 = 50$$

   mit einem Gauß-Test getestet werden. Als realisierte Teststatistik erhält man $N = -1.59$. 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 (für $\mu = \mu_0$) darstellt.



Aufgabe 3 (6 + 3 + 2 = 11 Punkte)

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

$$f_Y(y|\lambda) = \begin{cases} \dfrac{\lambda+1}{2^{\lambda+1}} \cdot y^{\lambda} & \text{für } 0 \leq y \leq 2 \\ 0 & \text{sonst} \end{cases}$$

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

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

(b) Zeigen Sie, dass $E(Y) = \dfrac{2\lambda+2}{\lambda+2}$ gilt.

(c) Bestimmen Sie den Schätzer $\widehat{\lambda}_{MM}$ nach der Methode der Momente.

Hinweise:

- *Beachten Sie, dass Sie Teil (b) auch ohne die Bearbeitung von Teil (a) und Teil (c) ohne die Bearbeitung der Teile (a) und (b) 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) $\widehat{\lambda}_{ML} = \dfrac{n}{n\ln(2) - \sum_{i=1}^{n} \ln(x_i)} - 1 = \dfrac{1}{\ln(2) - \frac{1}{n}\sum_{i=1}^{n} \ln(x_i)} - 1$

(b) Nachweis durch Berechnung des Erwartungswerts

(c) $\widehat{\lambda}_{MM} = \dfrac{2 - 2\overline{x}}{\overline{x} - 2}$

Aufgabe 4 (7 + 2 + 3 = 12 Punkte)

Bei der Herstellung von Montagekleber weiß der Hersteller aus Erfahrung, dass die verwendete Abfüllmaschine eine Standardabweichung von $4[g]$ für die abgefüllte Menge hat. Nach einer Inventur hat der Hersteller den Verdacht, dass die Abfüllmaschine im Mittel weniger als die auf dem Produkt ausgezeichneten $440[g]$ in die Kartuschen abfüllt. Dies soll mit einem statistischen Test überprüft werden. Hierzu werden der Produktion 16 Kartuschen entnommen, deren gemessene Klebstoffmengen $x_1, \dots, x_{16}$ als Realisation einer einfachen Stichprobe vom Umfang 16 zur annahmegemäß $N(\mu, 4^2[g^2])$-verteilten abgefüllten Menge betrachtet werden können. Als Stichprobenmittelwert ergibt sich dabei

$$\overline{x} = \frac{1}{16} \sum_{i=1}^{16} x_i = 438.535[g] \; .$$

(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). Wie wäre die Entscheidung zum Test aus Teil (a) bei einem Signifikanzniveau von $\alpha = 0.10$ ausgefallen?

(c) Wie groß ist die Wahrscheinlichkeit, bei der Ziehung einer Stichprobe und der anschließenden Durchführung des Tests aus Teil (a) eine Testentscheidung zu Gunsten der Nullhypothese zu erhalten, falls die tatsächliche mittlere Abfüllmenge $438[g]$ beträgt?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $N = -1.465 \notin (-\infty, -1.645) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

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

(b) p-Wert $p = 0.0708$. Entscheidung wäre zu Gunsten der Gegenhypothese ausgefallen.

(c) $\beta(438) = 0.3594$

Aufgabe 5 (10 Punkte)

Die Leistungsfähigkeit zweier CPUs ("A" und "B") soll mit Hilfe von Benchmarks zur Leistungsmessung verglichen werden. Man nehme hierzu an, dass die erhaltenen Werte Y^A bzw. Y^B der Benchmarks zu CPU A bzw. CPU B jeweils normalverteilt seien mit den unbekannten Erwartungswerten μ_A bzw. μ_B sowie den unbekannten Varianzen σ_A^2 bzw. σ_B^2. Es soll überprüft werden, ob CPU B im Mittel geringere Benchmarkergebnisse als CPU A liefert.
Aus einer wiederholten Durchführung mit $n_A = 12$ Benchmark-Durchläufen für CPU A sowie $n_B = 16$ Durchläufen für CPU B erhielt man Realisationen jeweils voneinander unabhängiger einfacher Stichproben $X_1^A, \ldots, X_{12}^A$ zu Y^A sowie $X_1^B, \ldots, X_{16}^B$ zu Y^B und hieraus die zugehörigen Mittelwerte $\overline{x^A} = 76018$ bzw. $\overline{x^B} = 74854$ sowie die Stichproben**standardabweichungen** $s_{Y^A} = 484$ bzw. $s_{Y^B} = 625$. Testen Sie unter der Annahme $\sigma_A^2 = \sigma_B^2$ zum Signifikanzniveau $\alpha = 0.01$ **(!)** die Hypothese, dass CPU B im Mittel geringere Benchmarkergebnisse als CPU A liefert. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

Ergebnisse (ohne Begründung/Rechenweg):

$t = 5.351 \in (2.479, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!
Der Test kann also die Vermutung, dass CPU B im Mittel geringere Benchmarkergebnisse als CPU A liefert, bestätigen.

Aufgabe 6 (11 Punkte)

Mit einem Chi-Quadrat-Anpassungstest soll zum Signifikanzniveau $\alpha = 0.05$ überprüft werden, ob man bei einem beobachteten Stichprobenergebnis von der Realisation einer einfachen Stichprobe vom Umfang $n = 200$ zu einer Geom(0.4)-verteilten Zufallsvariablen ausgehen kann. Die Stichprobeninformation liege in Form der folgenden Häufigkeitsverteilung vor:

a_i	0	1	2	≥ 3
n_i	100	47	25	28

Führen Sie den beschriebenen Test durch. Fassen Sie das Ergebnis auch in einem Antwortsatz zusammen.

Hinweise:

- *Die geometrische Verteilung mit Parameter* $p = 0.4$ *hat den Träger* $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$ *und die Wahrscheinlichkeitsfunktion*

$$p_{\text{Geom}(0.4)} : \mathbb{N}_0 \to [0,1]; p_{\text{Geom}(0.4)}(i) = (1 - 0.4)^i \cdot 0.4 \ .$$

- *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):

$\chi^2 = 10.8703 \in (7.815, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!
Die Nullhypothese einer Geom(0.4)-Verteilung muss also abgelehnt werden.

Aufgabe 7 (14 + 4 = 18 Punkte)

Um zu untersuchen, ob es einen generellen Zusammenhang zwischen der Fachsemesterzahl und der Leistung in der schriftlichen Prüfung gibt, wurden die Teilnehmenden einer Statistik-Klausur im Sommersemester 2023 in 3 Gruppen eingeteilt (Gruppe 1: erstes Fachsemester, Gruppe 2: zweites Fachsemester, Gruppe 3: mindestens drittes Fachsemester). Aus den erreichten Punktzahlen der entsprechenden Klausur wurden auf Basis dieser Einteilung die folgenden Daten aggregiert:

j (Gruppe)	n_j	$\overline{x}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_{j,i}$	$\sum_{i=1}^{n_j} x_{j,i}^2$	s_j^2
1	40	77.92	273105	775.49
2	141	78.41	945988	565.01
3	59	76.14	362910	359.82

(a) Überprüfen Sie mit einer einfachen Varianzanalyse (Signifikanzniveau $\alpha = 0.05$) unter der Annahme, dass die erreichten Punktzahlen Realisierungen von unabhängig $N(\mu_j, \sigma^2)$-verteilten Zufallsvariablen $X_{j,i} (1 \leq j \leq 3, 1 \leq i \leq n_j)$ sind, ob die Fachsemesterzahl einen Einfluss auf die (mittlere) erreichte Punktzahl hat. Fassen Sie das Ergebnis der Varianzanalyse auch in einem Antwortsatz zusammen.

(b) Erläutern Sie im Hinblick auf den Unterschied zwischen s_1^2 und s_2^2 mit kurzer Begründung (1–2 Sätze), inwiefern die Gültigkeit der zur Anwendung der Varianzanalyse getroffenen Annahme der Varianzgleichheit in den zugehörigen Gruppen zu hinterfragen ist. *(Hinweis: $F_{39,140;0.025} = 0.58$, $F_{39,140;0.975} = 1.602$)*

Hinweis: Verwenden Sie für Teil (a) den folgenden Tabellenausschnitt mit 0.95*-Quantilen von $F(m, n)$-Verteilungen:*

$n \backslash m$	1	2	3	4	5	235	236	237	238	239
1	161.448	199.500	215.707	224.583	230.162	253.772	253.774	253.776	253.779	253.781
2	18.513	19.000	19.164	19.247	19.296	19.491	19.491	19.492	19.492	19.492
3	10.128	9.552	9.277	9.117	9.013	8.538	8.538	8.538	8.538	8.538
4	7.709	6.944	6.591	6.388	6.256	5.643	5.643	5.643	5.643	5.643
5	6.608	5.786	5.409	5.192	5.050	4.382	4.382	4.382	4.382	4.382
235	3.881	3.034	2.643	2.410	2.252	1.240	1.240	1.239	1.239	1.239
236	3.881	3.034	2.643	2.410	2.252	1.240	1.239	1.239	1.239	1.239
237	3.881	3.034	2.643	2.410	2.252	1.239	1.239	1.239	1.239	1.238
238	3.881	3.034	2.643	2.410	2.252	1.239	1.239	1.238	1.238	1.238
239	3.881	3.034	2.642	2.409	2.252	1.239	1.238	1.238	1.238	1.238

Ergebnisse (ohne Begründung/Rechenweg):

(a) $F = 0.196 \notin (3.034, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

Die einfache Varianzanalyse kommt also zum Ergebnis, dass die Fachsemesterzahl keinen signifikanten ($\alpha = 0.05$) Einfluss auf die (mittlere) erreichte Punktzahl hat.

(b) Es gibt keine statistische Evidenz gegen die Gültigkeit der angenommenen Varianzgleichheit.

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

Zur Erklärung des (durchschnittlichen) Tagesgeld-Zinssatzes y_i durch die Euro-Short-Term-Rate x_i unterstellt man die Gültigkeit eines Zusammenhangs im Sinne des folgenden linearen Modells:

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

Aus den Daten zu 12 aufeinanderfolgenden Monaten wurde das lineare Modell mit der Statistik-Software R wie folgt geschätzt:

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

Residuals:
       Min         1Q     Median         3Q        Max
-1.917e-04 -3.378e-05  1.074e-05  8.553e-05  1.855e-04

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.0016064  0.0003889   4.130  0.00204 **
x           0.2274817  0.0751318   3.028  0.01273 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.000118 on 10 degrees of freedom
Multiple R-squared:  0.4783,       Adjusted R-squared:  0.4261
F-statistic: 9.167 on 1 and 10 DF,  p-value: 0.01273
```

(a) Geben Sie die realisierten Kleinst-Quadrate-Schätzwerte für β_1 und β_2 an.

(b) Geben Sie den realisierten Schätzwert für σ^2 an.

(c) Welcher Anteil der Gesamtvarianz des (durchschnittlichen) Tagesgeld-Zinssatzes wird durch das lineare Modell erklärt?

(d) Entscheiden Sie mit Hilfe des zugehörigen p-Werts zum Signifikanzniveau $\alpha = 0.05$, ob β_1 signifikant von Null verschieden ist.

(e) Entscheiden Sie mit Hilfe des zugehörigen p-Werts zum Signifikanzniveau $\alpha = 0.01$, ob β_2 signifikant positiv ist.

(f) Welchen (durchschnittlichen) Tagesgeldzinssatz prognostiziert das Modell in einem Monat mit einer Euro-Short-Term-Rate von 0.03?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\widehat{\beta}_1 = 0.0016064$, $\widehat{\beta}_2 = 0.2274817$

(b) $\widehat{\sigma^2} = 1.3924 \cdot 10^{-8}$

(c) 0.4783

(d) β_1 ist signifikant von Null verschieden.

(e) β_2 ist signifikant positiv.

(f) 0.0084309

Aufgabe 9 (6 + 2 + 3 + 3 + 5 = 19 Punkte)

Zur Schätzung eines einfachen linearen Regressionsmodells

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

aus einer Stichprobe vom Umfang $n = 16$ wurden bereits die folgenden Zwischenwerte errechnet:

$$\sum_{i=1}^{16} y_i = 28.355; \quad \sum_{i=1}^{16} y_i^2 = 642.333; \quad \sum_{i=1}^{16} x_i = 68.427;$$

$$\sum_{i=1}^{16} x_i^2 = 337.911; \quad \sum_{i=1}^{16} x_i \cdot y_i = 1.053$$

(a) Schätzen Sie β_1 und β_2 mit Hilfe der Kleinst-Quadrate-Methode.

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

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

(d) Geben Sie ein symmetrisches Konfidenzintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.99$ für β_1 an.

(e) Geben Sie ein Prognoseintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für $\mathrm{E}(y_0)$ gegeben $x_0 = 3$ an.

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\widehat{\beta}_1 = 13.1298$, $\widehat{\beta}_2 = -2.6557$

(b) $\widehat{\sigma}^2 = 19.4877$

(c) $\widehat{\sigma^2}_{\widehat{\beta}_1} = 9.092$, $\widehat{\sigma^2}_{\widehat{\beta}_2} = 0.4305$

(d) $[4.153, 22.106]$

(e) $[2.19, 8.136]$