Aufgabenstellung und Ergebnisse zur

Bachelor-Prüfung
Schliessende Statistik
Sommersemester 2019

PD Dr. Martin Becker

Hinweise für die Klausurteilnehmer

- Die Klausur besteht aus insgesamt 10 Aufgaben. Prüfen Sie die Vollständigkeit Ihres Exemplares nach; spätere Reklamationen können nicht berücksichtigt werden.
- Es sind insgesamt 120 Punkte (= 16 + 12 + 4 + 10 + 13 + 16 + 7 + 14 + 7 + 21) 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.
- Alle Punkte innerhalb von Zahlen sind stets Dezimalpunkte (und keine Tausenderpunkte).

Bewertungsteil — Bitte nicht beschreiben								
Aufgabe	(a)	(b)	(c)	(d)	(e)	(f)	(g)	Σ
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
Σ								

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!).

		wahr	falsch
1.	Stimmen die Verteilungen von $X_1, \dots, X_n$ mit der Verteilung von Y überein, dann handelt es sich bei $X_1, \dots, X_n$ stets um eine einfache Stichprobe vom Umfang n zu Y.	☐	■
2.	Für Schätzwerte $\widehat{\theta}_{\mathrm{ML}}$ nach der Maximum-Likelihood-Methode und Schätzwerte $\widehat{\theta}_{\mathrm{MM}}$ nach der Momenten-Methode (basierend auf derselben Stichprobenrealisation) gilt stets $\widehat{\theta}_{\mathrm{ML}} \geq \widehat{\theta}_{\mathrm{MM}}$.	☐	■
3.	Bei der Anwendung der Maximum-Likelihood-Methode für stetige Verteilungsfamilien kann die Likelihoodfunktion wegen der (vorausgesetzten) Unabhängigkeit der Stichprobenzufallsvariablen als Produkt von Randdichten dargestellt werden.	■	☐
4.	Sind $\widehat{\theta}$ und $\widetilde{\theta}$ zwei Schätzfunktionen zur Schätzung eines Parameters $\theta \in \Theta$ und ist $\widehat{\theta}$ wirksamer als $\widetilde{\theta}$, so gilt $\mathrm{Var}(\widehat{\theta}) < \mathrm{Var}(\widetilde{\theta})$ für alle $\theta \in \Theta$.	☐	■
5.	Bei einem zweiseitigen Gauß-Test für den Mittelwert bei bekannter Varianz ist der p-Wert umso größer, je näher $\overline{X}$ am hypothetischen Erwartungswert μ_0 liegt.	■	☐
6.	Die Gütefunktion eines zweiseitigen Gauß-Tests gibt zu tatsächlichen Erwartungswerten $\mu \neq \mu_0$ an, mit welcher Wahrscheinlichkeit die Durchführung des Tests zu einer falschen Entscheidung führt.	☐	■
7.	Lehnt ein Chi-Quadrat-Anpassungstest die Nullhypothese zu einem Signifikanzniveau von $\alpha = 0.10$ ab, so wird die Nullhypothese stets auch bei einem entsprechenden Test zum Signifikanzniveau von $\widetilde{\alpha} = 0.05$ verworfen.	☐	■
8.	Im einfachen linearen Regressionsmodell $y_i = \beta_1 + \beta_2 \cdot x_i + u_i, \quad u_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2),$ liegt der Punkt $(\overline{x}, \overline{y})$ stets auf der nach der KQ-Methode bestimmten Regressionsgeraden.	■	☐

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. Es sei $X_1, \ldots, X_{25}$ eine einfache Stichprobe vom Umfang 25 zu Y mit $Y \sim N(100, 20^2)$. Dann gilt für die Verteilung von $\overline{X} = \frac{1}{25}\sum_{i=1}^{25} X_i$:

   (a) $\overline{X} \sim N\left(100, 2^2\right)$ ☐

   (b) $\overline{X} \sim N\left(100, 4^2\right)$ ■

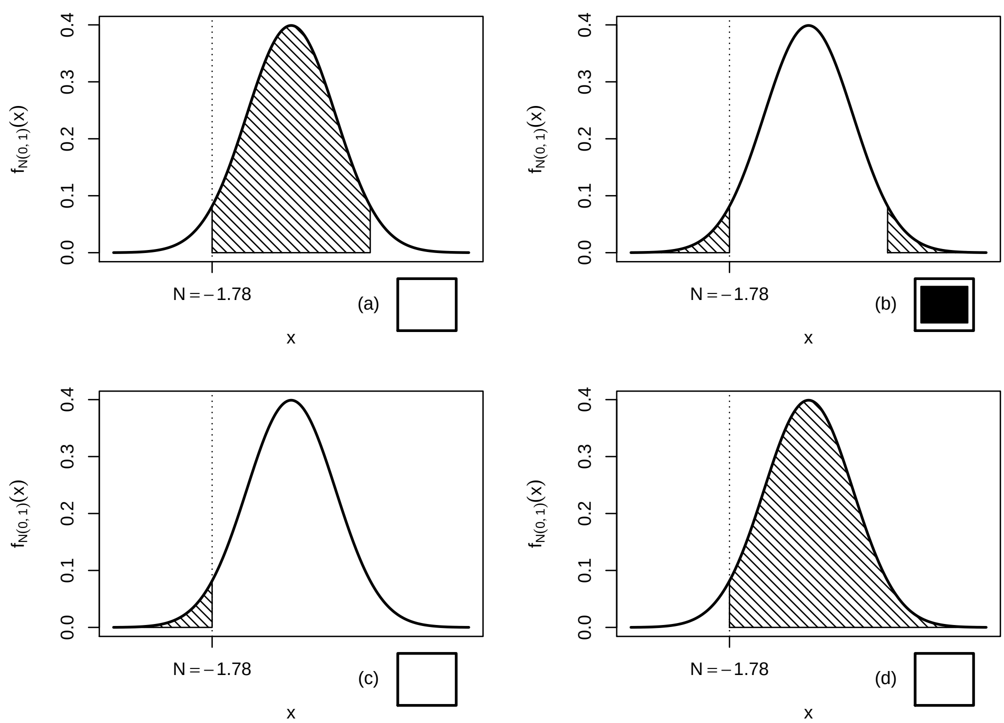
   (c) $\overline{X} \sim N\left(100, 10^2\right)$ ☐

   (d) $\overline{X} \sim N\left(100, 20^2\right)$ ☐

2. Sei $X_1, \ldots, X_{36}$ 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 $n = 36$ soll

$$H_0 : \mu = \mu_0 = 40 \qquad \text{gegen} \qquad H_1 : \mu \neq \mu_0 = 40$$

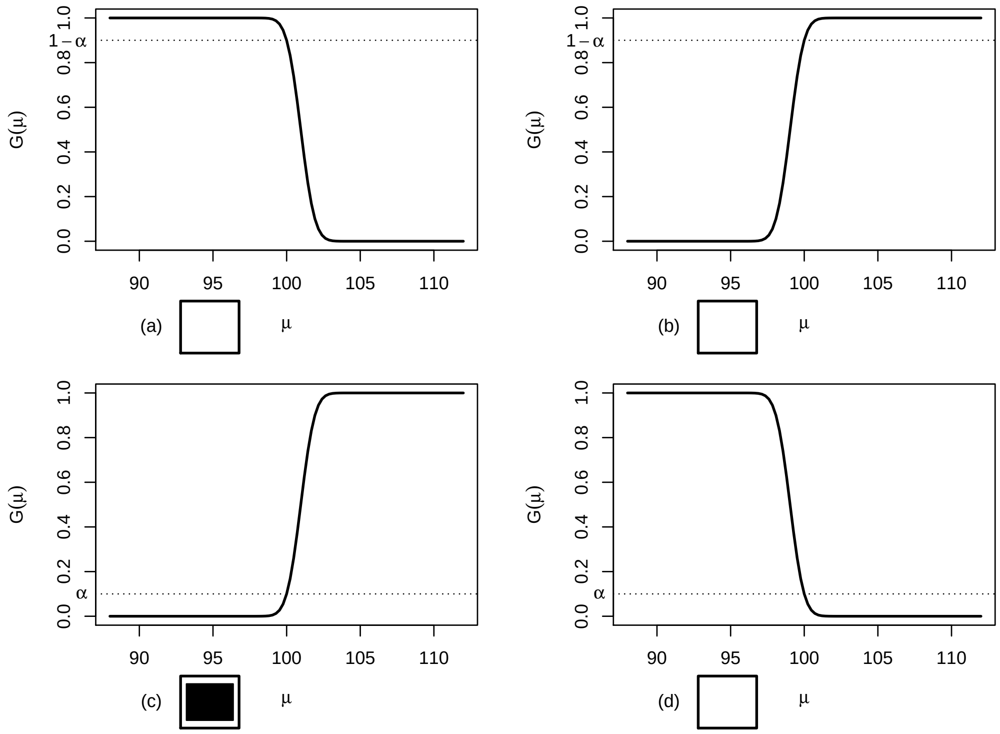
   mit einem Gauß-Test getestet werden. Als realisierte Teststatistik erhält man $N = -1.78$. Markieren Sie die Abbildung, welche den p-Wert in der beschriebenen Situation korrekt als Inhalt der schraffierten Fläche unter der Dichtefunktion der Verteilung der Teststatistik unter H_0 (für $\mu = \mu_0$) darstellt.



3. Auf der Grundlage einer einfachen Stichprobe $X_1, \ldots, X_{64}$ vom Umfang $n = 64$ zu einer $N(\mu, 6^2)$-verteilten Zufallsvariablen wird ein Gauß-Test zur Überprüfung der Hypothesen

$$H_0 : \mu \leq 100 \qquad \text{gegen} \qquad H_1 : \mu > 100$$

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



4. Bei der Durchführung eines linksseitigen t-Tests für den Mittelwert einer normalverteilten Grundgesamtheit bei unbekannter Varianz (mit $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$) wird die Nullhypothese zum Signifikanzniveau $\alpha = 0.05$ abgelehnt. Dann gilt für die Testentscheidungen des rechtsseitigen (mit $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$) und des zweiseitigen (mit $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$) Tests zum **veränderten Signifikanzniveau** $\alpha = 0.10$:

   (a) Der rechtsseitige Test lehnt H_0 nicht ab, der zweiseitige Test lehnt H_0 ab. ■

   (b) Der rechtsseitige Test lehnt H_0 nicht ab, die Entscheidung des zweiseitigen Tests kann für oder gegen H_0 ausfallen. □

   (c) Der rechtsseitige Test lehnt H_0 ab, die Entscheidung des zweiseitigen Tests kann für oder gegen H_0 ausfallen. □

   (d) Sowohl beim rechtsseitigen als auch beim zweiseitigen Test kann die Entscheidung für oder gegen H_0 ausfallen. □

Aufgabe 3 (4 Punkte)

Für $\lambda > 0$ sei $Y \sim \mathrm{Exp}(\lambda)$, es gilt also insbesondere $\mathrm{E}(Y) = \frac{1}{\lambda}$ sowie $\mathrm{Var}(Y) = \frac{1}{\lambda^2}$. $X_1, \ldots, X_n$ sei für $n \in \mathbb{N}$ eine einfache Stichprobe vom Umfang n zu Y. Untersuchen Sie, ob die Schätzfunktionen

$$T_n(X_1, \ldots, X_n) := \frac{1}{2n} \sum_{i=1}^{n} X_i^2$$

erwartungstreu für **die Varianz von** Y sind.

Ergebnisse (ohne Begründung/Rechenweg):

Die Schätzfunktionen $T_n(X_1, \ldots, X_n)$ sind erwartungstreu für die Varianz von Y.

Aufgabe 4 (3 + 1 + 6 = 10 Punkte)

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

$$f_Y(y|a) = \begin{cases} \frac{1}{4} \cdot a^{-4} \cdot y^3 & \text{für } 0 \leq y \leq 2 \cdot a \\ 0 & \text{sonst} \end{cases}$$

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

(a) Zeigen Sie, dass $\mathrm{E}(Y) = \frac{8}{5} \cdot a$ gilt.

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

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

Hinweise:

- *Beachten Sie, dass Sie die Teile (b) und (c) 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) Nachweis durch Berechnung des Erwartungswerts

(b) $\widehat{a}_{MM} = \frac{5}{8} \cdot \overline{x}$

(c) $\widehat{a}_{ML} = \frac{1}{2} \cdot \max\{x_1, \ldots, x_n\}$

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

Bei der Abfüllung von Babynahrung weiß der Hersteller aus Erfahrung, dass die verwendete Maschine eine Standardabweichung von $2[g]$ für die abgefüllte Menge hat. Nach einer Inventur hat der Hersteller den Verdacht, dass die Abfüllanlage im Mittel weniger als die auf dem Produkt ausgezeichneten $190[g]$ in die Gläser einfüllt. Dies soll mit einem statistischen Test überprüft werden. Hierzu werden der Produktion 16 Gläser entnommen, deren gemessene Füllmengen $x_1, \ldots, x_{16}$ als Realisation einer einfachen Stichprobe vom Umfang 16 zur annahmegemäß $N(\mu, 2^2[g^2])$-verteilten Abfüllmenge betrachtet werden können. Als Stichprobenmittelwert ergibt sich dabei

$$\overline{x} = \frac{1}{16} \sum_{i=1}^{16} x_i = 188.959[g] \ .$$

(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 $\alpha = 0.01$ 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 $188[g]$ beträgt?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $N = -2.082 \in (-\infty, -1.645) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

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

(b) p-Wert $p = 0.0188$. Entscheidung wäre zu Gunsten der Nullhypothese ausgefallen.

(c) $\beta(188) = 0.0091$

Aufgabe 6 (9 + 7 = 16 Punkte)

Ein Hersteller von Metallwaren produziert unter anderem Stahlstifte mit einer Soll-Länge von 8 [cm]. Es soll angenommen werden, dass die Länge der hergestellten Stahlstifte gemäß einer Normalverteilung mit unbekanntem Erwartungswert und unbekannter Varianz schwankt. Zur Überprüfung der Qualität der produzierten Stahlstifte werden 10 Stahlstifte aus dem laufenden Produktionsprozess entnommen. Die gemessenen Längen

$$7.94, 8.10, 8.03, 8.01, 8.16, 8.04, 8.14, 7.99, 8.03, 8.04$$

seien als Realisation einer einfachen Stichprobe zur normalverteilten Grundgesamtheit aufzufassen. Aus dieser Realisation wurde bereits $s^2 = 0.004551$ [cm^2] berechnet.

(a) Überprüfen Sie mit einem geeigneten Test zum Signifikanzniveau $\alpha = 0.05$, ob die tatsächliche mittlere Länge der produzierten Stahlstifte im Vergleich zur angegebenen Soll-Länge von 8 [cm] zu groß ist. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

(b) Überprüfen Sie mit einem geeigneten Test zum Signifikanzniveau $\alpha = 0.05$, ob die Varianz der Länge der produzierten Stahlstifte von der vom Hersteller angegebenen Toleranz $\sigma_0^2 = 0.01$ [cm^2] abweicht. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

Hinweis: Verwenden Sie für Teil (b) 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
6	0.872	1.237	1.635	5.348	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	6.346	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	7.344	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	8.343	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	9.342	15.987	18.307	20.483	23.209

Ergebnisse (ohne Begründung/Rechenweg):

(a) $t = 2.25 \in (1.833, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

Der Test kann also die Vermutung, dass die mittlere Länge der produzierten Stahlstifte im Vergleich zur angegebenen Soll-Länge von 8 [cm] zu groß ist, bestätigen.

(b) $\chi^2 = 4.096 \notin [0, 2.7) \cup (19.023, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

Der Test kann also die Vermutung, dass die Varianz der Länge der produzierten Stahlstifte von der vom Hersteller angegebenen Toleranz $\sigma_0^2 = 0.01$ [cm^2] abweicht, nicht bestätigen.

Aufgabe 7 (7 Punkte)

Zur Beurteilung der Präzision zweier Messgeräte A und B wird eine Referenzgröße jeweils unabhängig voneinander mit beiden Messgeräten wiederum jeweils unabhängig voneinander mehrfach gemessen. Es werde angenommen, dass die gemessenen Werte Y^A bzw. Y^B der beiden Messgeräte jeweils normalverteilt seien mit unbekannten Erwartungswerten μ_A bzw. μ_B sowie unbekannten Varianzen σ_A^2 bzw. σ_B^2. Die Ergebnisse der wiederholten Messungen lassen sich als (voneinander unabhängige) einfache Stichproben $X_1^A, \ldots, X_{17}^A$ vom Umfang 17 zu Y^A sowie $X_1^B, \ldots, X_{19}^B$ vom Umfang 19 zu Y^B auffassen, aus den zugehörigen Realisationen wurden bereits die Mittelwerte $\overline{x^A} = 100.167$ bzw. $\overline{x^B} = 99.456$ sowie die Stichprobenvarianzen $s_{Y^A}^2 = 0.762$ bzw. $s_{Y^B}^2 = 1.847$ berechnet. Überprüfen Sie mit einem geeigneten Test zum Signifikanzniveau $\alpha = 0.05$, ob Messgerät A eine höhere Präzision (im Sinne einer geringeren Streuung) als Messgerät B hat. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

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

$n \backslash m$	11	12	13	14	15	16	17	18	19	20
11	2.818	2.788	2.761	2.739	2.719	2.701	2.685	2.671	2.658	2.646
12	2.717	2.687	2.660	2.637	2.617	2.599	2.583	2.568	2.555	2.544
13	2.635	2.604	2.577	2.554	2.533	2.515	2.499	2.484	2.471	2.459
14	2.565	2.534	2.507	2.484	2.463	2.445	2.428	2.413	2.400	2.388
15	2.507	2.475	2.448	2.424	2.403	2.385	2.368	2.353	2.340	2.328
16	2.456	2.425	2.397	2.373	2.352	2.333	2.317	2.302	2.288	2.276
17	2.413	2.381	2.353	2.329	2.308	2.289	2.272	2.257	2.243	2.230
18	2.374	2.342	2.314	2.290	2.269	2.250	2.233	2.217	2.203	2.191
19	2.340	2.308	2.280	2.256	2.234	2.215	2.198	2.182	2.168	2.155
20	2.310	2.278	2.250	2.225	2.203	2.184	2.167	2.151	2.137	2.124

Ergebnisse (ohne Begründung/Rechenweg):

$F = 0.413 \in [0, 0.434) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!
Der Test kann also die Vermutung, dass Messgerät A eine höhere Präzision (im Sinne einer geringeren Streuung) als Messgerät B hat, bestätigen.

Aufgabe 8 (14 Punkte)

Um zu überprüfen, ob die Anzahl der bearbeiteten Zusatzübungsblätter einen Einfluss auf die Leistung in der schriftlichen Prüfung hat, wurden die Prüflinge einer Statistik-Klausur im Wintersemester 2017/18 in 3 Gruppen eingeteilt (Gruppe 1: kein Zusatzblatt bearbeitet, Gruppe 2: ein Zusatzblatt bearbeitet, Gruppe 3: beide Zusatzblätter bearbeitet). 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	87	81.09	615649	506.66
2	24	95.38	221282	128.07
3	51	99.29	509830	140.93

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 Anzahl der bearbeiteten Zusatzblätter einen Einfluss auf die (mittlere) erreichte Punktzahl hat. Fassen Sie das Ergebnis der Varianzanalyse auch in einem Antwortsatz zusammen.

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

$n \backslash m$	1	2	3	4	5	155	156	157	158	159
1	161.448	199.500	215.707	224.583	230.162	253.492	253.497	253.503	253.508	253.513
2	18.513	19.000	19.164	19.247	19.296	19.489	19.489	19.489	19.489	19.489
3	10.128	9.552	9.277	9.117	9.013	8.544	8.544	8.544	8.544	8.544
4	7.709	6.944	6.591	6.388	6.256	5.651	5.651	5.651	5.651	5.651
5	6.608	5.786	5.409	5.192	5.050	4.391	4.391	4.391	4.390	4.390
155	3.902	3.054	2.663	2.430	2.273	1.303	1.303	1.303	1.302	1.302
156	3.902	3.054	2.663	2.430	2.272	1.303	1.302	1.302	1.301	1.301
157	3.901	3.054	2.662	2.429	2.272	1.302	1.302	1.301	1.301	1.300
158	3.901	3.053	2.662	2.429	2.271	1.302	1.301	1.301	1.300	1.300
159	3.901	3.053	2.661	2.429	2.271	1.301	1.300	1.300	1.300	1.299

Ergebnisse (ohne Begründung/Rechenweg):

$F = 17.543 \in (3.053, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!
Die einfache Varianzanalyse kommt also zum Ergebnis, dass die Anzahl der bearbeiteten Zusatzblätter einen signifikanten ($\alpha = 0.05$) Einfluss auf die (mittlere) erreichte Punktzahl hat.

Aufgabe 9 (1 + 1 + 1 + 1 + 1 + 1 + 1 = 7 Punkte)

Zur Erklärung der stetigen Wochenrenditen der BASF-Aktie y_i (in Prozent) durch die stetigen Wochenrenditen des DAX x_i (in Prozent) 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 Daten der XETRA-Börse wurde das lineare Modell mit der Statistik-Software R wie folgt geschätzt:

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

Residuals:
    Min      1Q  Median      3Q     Max
-2.0828 -0.6635 -0.2124  0.7805  2.1818

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.5098     0.3602   1.415   0.1789
x             0.7064     0.2473   2.856   0.0127 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.313 on 14 degrees of freedom
Multiple R-squared:  0.3682,       Adjusted R-squared:  0.323
F-statistic: 8.158 on 1 and 14 DF,  p-value: 0.01269
```

(a) Wie viele Wochenrenditen gingen in die Schätzung ein?

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

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

(d) Welcher Anteil der Gesamtvarianz der stetigen Wochenrenditen der BASF-Aktie wird durch das lineare Modell erklärt?

(e) Entscheiden Sie mit Hilfe des zugehörigen p-Werts zum Signifikanzniveau $\alpha = 0.10$, ob β_1 signifikant von Null verschieden ist.

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

(g) Welche stetige Wochenrendite der BASF-Aktie prognostiziert das Modell in einer Woche mit stetiger DAX-Rendite von 0.4 (in Prozent)?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $n = 16$

(b) $\widehat{\beta}_1 = 0.5098$, $\widehat{\beta}_2 = 0.7064$

(c) $\widehat{\sigma^2} = 1.724$

(d) 0.3682

(e) β_1 ist nicht signifikant von Null verschieden.

(f) β_2 ist signifikant positiv.

(g) 0.79236

Aufgabe 10 (6 + 2 + 2 + 3 + 3 + 5 = 21 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 $n = 16$ wurden bereits die folgenden Zwischenwerte errechnet:

$$\sum_{i=1}^{16} y_i = 464.89; \quad \sum_{i=1}^{16} y_i^2 = 13947.79; \quad \sum_{i=1}^{16} x_i = 91.27;$$

$$\sum_{i=1}^{16} x_i^2 = 570.42; \quad \sum_{i=1}^{16} x_i \cdot y_i = 2792.65$$

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

(b) Berechnen Sie das Bestimmtheitsmaß R^2.

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

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

(e) Geben Sie ein symmetrisches Konfidenzintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für β_2 an.

(f) Geben Sie ein Prognoseintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für y_0 gegeben $x_0 = 6$ an.

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\widehat{\beta}_1 = 12.9275$, $\widehat{\beta}_2 = 2.8273$

(b) $R^2 = 0.90402$

(c) $\widehat{\sigma}^2 = 3.0179$

(d) $\widehat{\sigma^2}_{\widehat{\beta}_1} = 2.1615$, $\widehat{\sigma^2}_{\widehat{\beta}_2} = 0.06063$

(e) $[2.3, 3.35]$

(f) $[26.06, 33.72]$