

**WIRTSCHAFTSWISSENSCHAFTLICHES
PRÜFUNGSSEKRETARIAT**

FAKULTÄT FÜR EMPIRISCHE HUMANWISSENSCHAFTEN UND WIRTSCHAFTSWISSENSCHAFT
DER UNIVERSITÄT DES SAARLANDES

Von der/dem Studierenden auszufüllen (Bitte leserlich und in Blockschrift):
Schließende Statistik

Name der Prüfung: _____

Semester, dem die Prüfung zugeordnet ist: SS 2019 (z. B. WS 2015/2016, SS 2016)
(Prüfungen im Februar/April = WS; Prüfungen im August/Oktobre = SS)

Nachname, Vorname der/des Studierenden: _____

Matrikelnummer der/des Studierenden: _____

Erklärung

Hiermit erkläre ich, dass mir die für diese Prüfung relevanten Zulassungsvoraussetzungen aus der für mich geltenden Prüfungsordnung bekannt sind.

Mir ist damit bewusst, dass diese Prüfungsleistung als nicht abgelegt zählt, wenn die hierfür vorgesehenen Zulassungsvoraussetzungen nicht erfüllt sind.

Mir ist bekannt, dass die Teilnahme an der Prüfung zudem die ordnungsgemäße Anmeldung zur Prüfung voraussetzt. Die Teilnahme bei versäumter Anmeldung hat die Ungültigkeit der Prüfung zur Folge.

Zudem ist mir bekannt, dass eine nicht bestandene Prüfung zweimal wiederholt werden kann. Die Wiederholung einer bestandenen Prüfung ist nicht zulässig.

Datum: _____ **Unterschrift der/des Studierenden:** _____

Von der Prüferin/Vom Prüfer auszufüllen:

Aufgabe	Punkte	Max. Punkte	Bemerkungen
1+2		28	
3		4	
4		10	
5		13	
6		16	
7		7	
8		14	
9		7	
10		21	
Summe		120	

bestanden

Note: _____

nicht bestanden

Unterschrift der Prüferin/des Prüfers: _____

KLAUSURHEFT 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 1–26 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 . | <input type="checkbox"/> | <input type="checkbox"/> |
| 2. Für Schätzwerte $\hat{\theta}_{\text{ML}}$ nach der Maximum-Likelihood-Methode und Schätzwerte $\hat{\theta}_{\text{MM}}$ nach der Momenten-Methode (basierend auf derselben Stichprobenrealisation) gilt stets $\hat{\theta}_{\text{ML}} \geq \hat{\theta}_{\text{MM}}$. | <input type="checkbox"/> | <input type="checkbox"/> |
| 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. | <input type="checkbox"/> | <input type="checkbox"/> |
| 4. Sind $\hat{\theta}$ und $\tilde{\theta}$ zwei Schätzfunktionen zur Schätzung eines Parameters $\theta \in \Theta$ und ist $\hat{\theta}$ wirksamer als $\tilde{\theta}$, so gilt $\text{Var}(\hat{\theta}) < \text{Var}(\tilde{\theta})$ für alle $\theta \in \Theta$. | <input type="checkbox"/> | <input type="checkbox"/> |
| 5. Bei einem zweiseitigen Gauß-Test für den Mittelwert bei bekannter Varianz ist der p -Wert umso größer, je näher \bar{X} am hypothetischen Erwartungswert μ_0 liegt. | <input type="checkbox"/> | <input type="checkbox"/> |
| 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. | <input type="checkbox"/> | <input type="checkbox"/> |
| 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 $\tilde{\alpha} = 0.05$ verworfen. | <input type="checkbox"/> | <input type="checkbox"/> |
| 8. Im einfachen linearen Regressionsmodell | <input type="checkbox"/> | <input type="checkbox"/> |

$$y_i = \beta_1 + \beta_2 \cdot x_i + u_i, \quad u_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2),$$

liegt der Punkt (\bar{x}, \bar{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!).

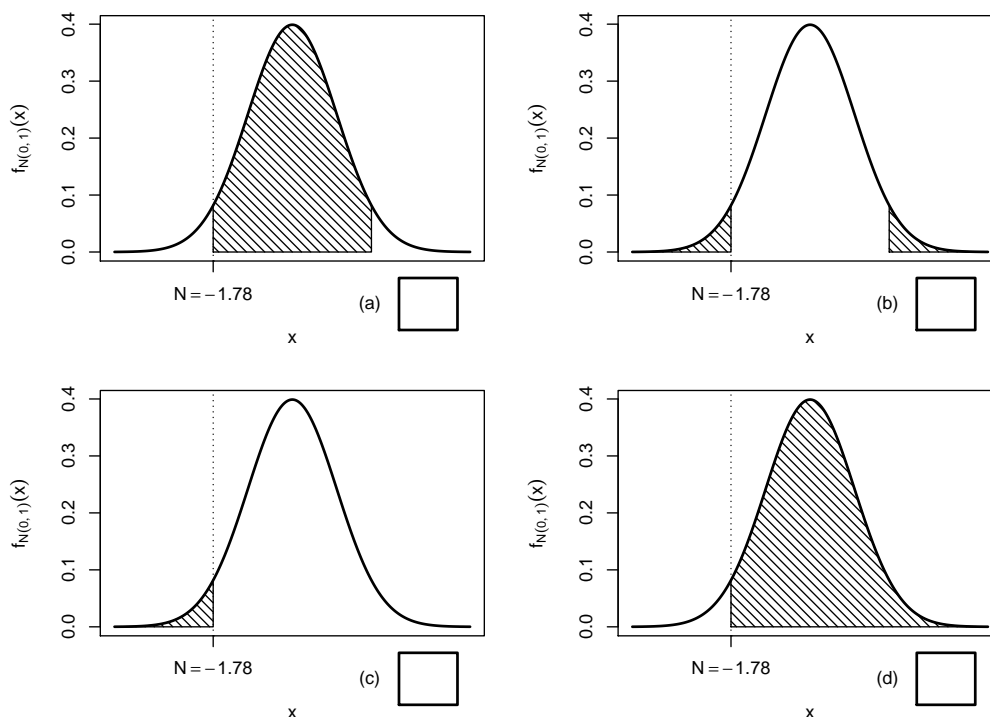
1. Es sei X_1, \dots, X_{25} eine einfache Stichprobe vom Umfang 25 zu Y mit $Y \sim N(100, 20^2)$. Dann gilt für die Verteilung von $\bar{X} = \frac{1}{25} \sum_{i=1}^{25} X_i$:

- (a) $\bar{X} \sim N(100, 2^2)$
- (b) $\bar{X} \sim N(100, 4^2)$
- (c) $\bar{X} \sim N(100, 10^2)$
- (d) $\bar{X} \sim N(100, 20^2)$

2. Sei X_1, \dots, 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

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

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.

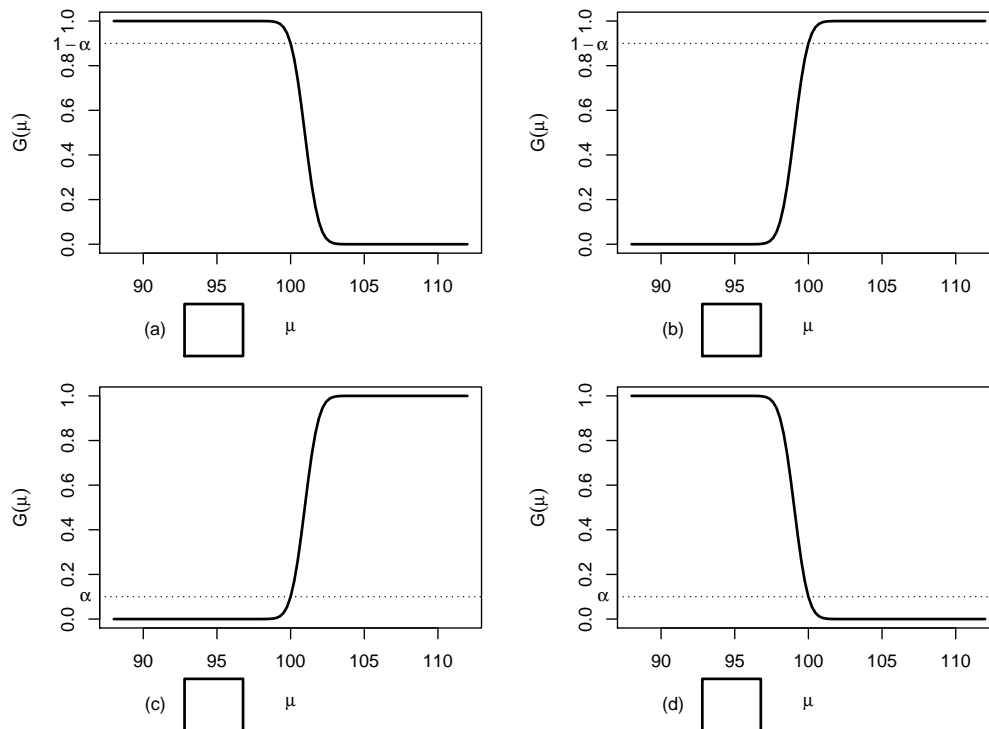


3. Auf der Grundlage einer einfachen Stichprobe X_1, \dots, X_{64} vom Umfang $n = 64$ zu einer $N(\mu, 6^2)$ -verteilten Zufallsvariablen wird ein Gauß-Test zur Überprüfung der Hypothesen

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

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

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

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



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

Die Verteilung einer Zufallsvariablen Y sei in Abhängigkeit des unbekanntes 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, \dots, X_n vom Umfang n geschätzt werden.

- (a) Zeigen Sie, dass $E(Y) = \frac{8}{5} \cdot a$ gilt.
- (b) Bestimmen Sie den Schätzer \hat{a}_{MM} nach der Methode der Momente.
- (c) Bestimmen Sie den Schätzer \hat{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.



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, \dots, 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

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

- (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?







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 $s^2 = 0.004551$ [cm²] 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 $\sigma_0^2 = 0.01$ [cm²] 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







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 unbekanntem Erwartungswerten μ_A bzw. μ_B sowie unbekanntem Varianzen σ_A^2 bzw. σ_B^2 . Die Ergebnisse der wiederholten Messungen lassen sich als (voneinander unabhängige) einfache Stichproben X_1^A, \dots, X_{17}^A vom Umfang 17 zu Y^A sowie X_1^B, \dots, X_{19}^B vom Umfang 19 zu Y^B auffassen, aus den zugehörigen Realisationen wurden bereits die Mittelwerte $\bar{x}^A = 100.167$ bzw. $\bar{x}^B = 99.456$ 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





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	$\bar{x}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{j,i}$	$\sum_{i=1}^{n_j} x_{j,i}^2$	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





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:

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

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

- Wie viele Wochenrenditen gingen in die Schätzung ein?
- Geben Sie die realisierten Kleinst-Quadrate-Schätzwerte für β_1 und β_2 an.
- Geben Sie den realisierten Schätzwert für σ^2 an.
- Welcher Anteil der Gesamtvarianz der stetigen Wochenrenditen der BASF-Aktie wird durch das lineare Modell erklärt?
- Entscheiden Sie mit Hilfe des zugehörigen p -Werts zum Signifikanzniveau $\alpha = 0.10$, ob β_1 signifikant von Null verschieden ist.
- Entscheiden Sie mit Hilfe des zugehörigen p -Werts zum Signifikanzniveau $\alpha = 0.01$, ob β_2 signifikant positiv ist.
- Welche stetige Wochenrendite der BASF-Aktie prognostiziert das Modell in einer Woche mit stetiger DAX-Rendite von 0.4 (in Prozent)?



Aufgabe 10 (6 + 2 + 2 + 3 + 3 + 5 = 21 Punkte)

Zur Schätzung eines einfachen linearen Regressionsmodells

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

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

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

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

- Schätzen Sie β_1 und β_2 mit Hilfe der Kleinst-Quadrate-Methode.
- Berechnen Sie das Bestimmtheitsmaß R^2 .
- Geben Sie mit Hilfe der bekannten erwartungstreuen Schätzfunktion für σ^2 den realisierten Schätzwert für σ^2 an.
- Berechnen Sie $\hat{\sigma}_{\hat{\beta}_1}^2$ und $\hat{\sigma}_{\hat{\beta}_2}^2$.
- Geben Sie ein symmetrisches Konfidenzintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für β_2 an.
- Geben Sie ein Prognoseintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für y_0 gegeben $x_0 = 6$ an.









Tabelle zur Standardnormalverteilung

$$F_{N(0,1)}(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt$$

$$\Phi(x) = 1 - \Phi(-x)$$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

p-Quantile der Standardnormalverteilung

$$F_{N(0,1)}(N_p) = p$$

<i>p</i>	0.85	0.90	0.95	0.975	0.99	0.995	0.9995
<i>N_p</i>	1.036	1.282	1.645	1.960	2.326	2.576	3.291

p -Quantile der $t(n)$ -Verteilungen $t_{n;p}$

$$T \sim t(n) \Rightarrow F_T(t_{n;p}) = p$$

$n \setminus p$	0.85	0.90	0.95	0.975	0.99	0.995	0.9995
1	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	1.060	1.319	1.714	2.069	2.500	2.807	3.768
24	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	1.055	1.310	1.697	2.042	2.457	2.750	3.646
31	1.054	1.309	1.696	2.040	2.453	2.744	3.633
32	1.054	1.309	1.694	2.037	2.449	2.738	3.622
33	1.053	1.308	1.692	2.035	2.445	2.733	3.611
34	1.052	1.307	1.691	2.032	2.441	2.728	3.601
35	1.052	1.306	1.690	2.030	2.438	2.724	3.591
40	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	1.045	1.296	1.671	2.000	2.390	2.660	3.460
80	1.043	1.292	1.664	1.990	2.374	2.639	3.416
100	1.042	1.290	1.660	1.984	2.364	2.626	3.390
120	1.041	1.289	1.658	1.980	2.358	2.617	3.373
140	1.040	1.288	1.656	1.977	2.353	2.611	3.361
160	1.040	1.287	1.654	1.975	2.350	2.607	3.352
180	1.039	1.286	1.653	1.973	2.347	2.603	3.345
200	1.039	1.286	1.653	1.972	2.345	2.601	3.340
250	1.039	1.285	1.651	1.969	2.341	2.596	3.330