Aufgabenstellung und Ergebnisse zur

Bachelor-Prüfung
Schliessende Statistik
Sommersemester 2022

PD 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 (= 16 + 12 + 6 + 10 + 19 + 18 + 12 + 6 + 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)	Σ
1							
2							
3							
4							
5							
6							
7							
8							
9							
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.	Weisen die Stichprobenzufallsvariablen $X_1, \dots, X_n$ unterschiedliche Verteilungen auf, so kann es sich bei $X_1, \dots, X_n$ nicht um eine einfache Stichprobe handeln.	X	
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 für alle $n \in \mathbb{N}$ stets $\text{Var}(T_n) = 0$.		X
3.	Lehnt ein rechtsseitiger Chi-Quadrat-Test für die Varianz einer normalverteilten Zufallsvariablen bei unbekanntem Erwartungswert 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.		X
4.	Ist die Nullhypothese H_0 tatsächlich falsch, so wird man bei der Anwendung eines statistischen Tests zum Signifikanzniveau 0.05 nur mit einer Wahrscheinlichkeit von höchstens 5% eine Stichprobenrealisation erhalten, die nicht zu einer Ablehnung von H_0 führt.		X
5.	Beim zweiseitigen Gauß-Test für den Erwartungswert bei bekannter Varianz (mit $H_0 : \mu = \mu_0$) ist der p-Wert (bei gleichbleibendem Stichprobenumfang) umso niedriger, je größer der Abstand $\lvert\overline{X} - \mu_0\rvert$ ausfällt.	X	
6.	Das Verkleinern des Signifikanzniveaus α führt bei sämtlichen in der Veranstaltung besprochenen Hypothesentests stets zu einer Verkleinerung des kritischen Bereichs.	X	
7.	Sind die Voraussetzungen zur Anwendung der einfachen Varianzanalyse erfüllt und liegen als Stichprobenrealisation jeweils 30 Beobachtungen zu 5 Faktorstufen vor, so ist die Teststatistik bei Gültigkeit der Nullhypothese $F(5, 145)$-verteilt.		X
8.	Im einfachen linearen Regressionsmodell $y_i = \beta_1 + \beta_2 \cdot x_i + u_i, \quad u_i \overset{\text{iid}}{\sim} N(0, \sigma^2), \quad i = 1, \dots, n,$ kann die Summe der quadrierten Residuen $\sum_{i=1}^{n} \widehat{u}_i^2$ (mit $\widehat{u}_i = y_i - (\widehat{\beta}_1 + \widehat{\beta}_2 \cdot x_i)$) nicht dadurch verkleinert werden, dass die KQ-Schätzer $\widehat{\beta}_1$ und $\widehat{\beta}_2$ durch andere (reelle) Zahlen ersetzt werden.	X	

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 $X_1, \dots, X_{36}$ eine einfache Stichprobe vom Umfang 36 zu Y mit $Y \sim N(300, 18^2)$. Dann gilt für die Verteilung von $\overline{X} = \frac{1}{36} \sum_{i=1}^{36} X_i$:

   (a) $\overline{X} \sim N\left(300, 3^2\right)$ ☒

   (b) $\overline{X} \sim N\left(300, 6^2\right)$ ☐

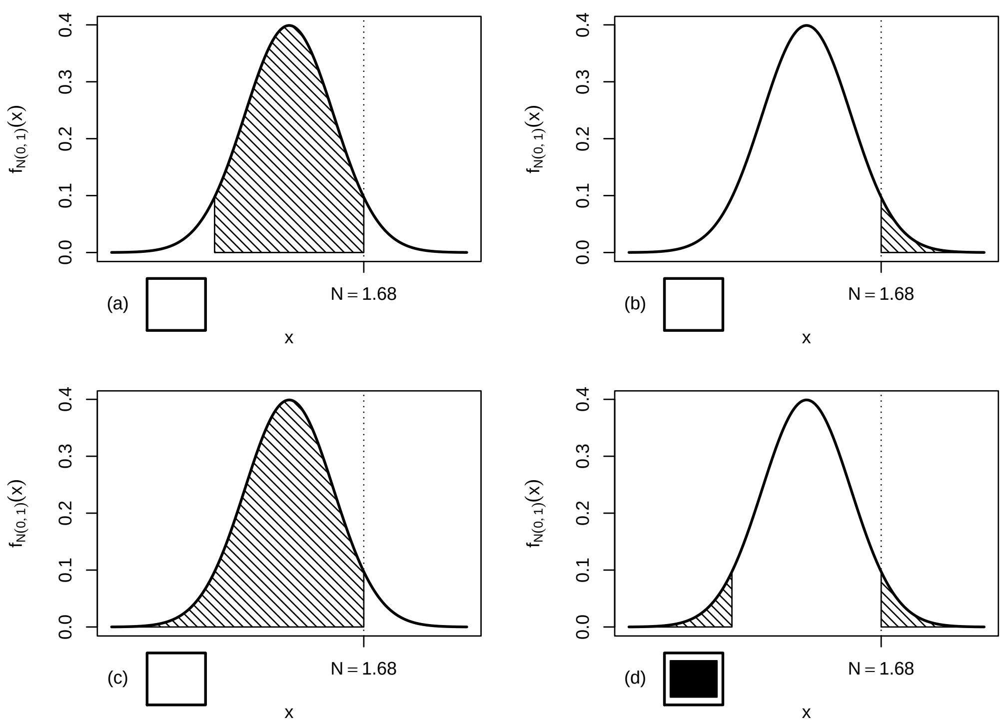
   (c) $\overline{X} \sim N\left(300, 9^2\right)$ ☐

   (d) $\overline{X} \sim N\left(300, 18^2\right)$ ☐

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

$$H_0 : \mu = \mu_0 = 30 \qquad \text{gegen} \qquad H_1 : \mu \neq \mu_0 = 30$$

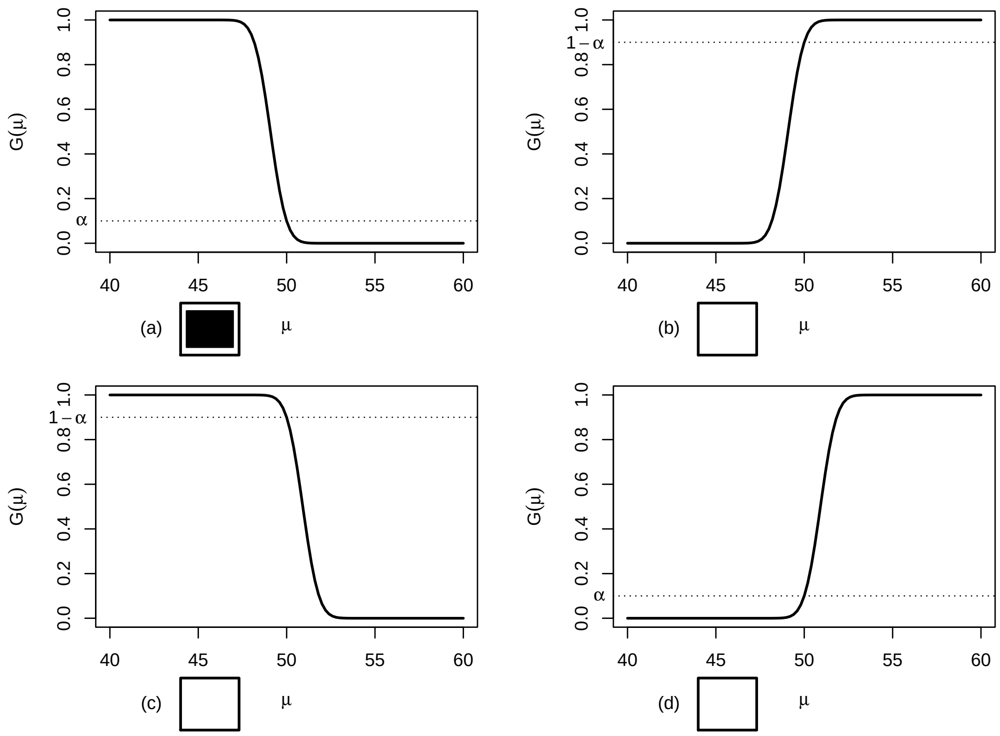
   mit einem Gauß-Test getestet werden. Als realisierte Teststatistik erhält man $N = 1.68$. Markieren Sie die Abbildung, welche den p-Wert in der beschriebenen Situation korrekt als Inhalt der schraffierten Fläche unter der Dichtefunktion der Verteilung der Teststatistik unter H_0 (für $\mu = \mu_0$) darstellt.



3. Auf der Grundlage einer einfachen Stichprobe $X_1, \ldots, X_{49}$ vom Umfang $n = 49$ zu einer $N(\mu, 5^2)$-verteilten Zufallsvariablen wird ein Gauß-Test zur Überprüfung der Hypothesen

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

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 linksssseitigen Gauß-Tests für den Mittelwert einer normalverteilten Grundgesamtheit bei bekannter 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 zweiseitigen (mit $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$) Tests (bei unverändertem Signifikanzniveau $\alpha = 0.05$):

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

   (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, der zweiseitige Test lehnt H_0 nicht ab. ☐

   (d) Der rechtsseitige Test lehnt H_0 nicht ab, der zweiseitige Test lehnt H_0 ab. ☐

Aufgabe 3 (4 + 2 = 6 Punkte)

In Abhängigkeit eines unbekannten Parameters $p > 0$ seien der Erwartungswert und die Varianz einer Zufallsvariablen Y gegeben durch $\mathrm{E}(Y) = \dfrac{1}{p}$ sowie $\mathrm{Var}(Y) = \dfrac{1}{p^2}$. Für $n \in \mathbb{N}$ sei $X_1, \ldots, X_n$ eine einfache Stichprobe vom Umfang n zu Y.

(a) Zeigen Sie, dass die Schätzfunktionen

$$T_n(X_1, \ldots, X_n) := \frac{1}{3n} \sum_{i=1}^{n} X_i^2$$

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

(b) Geben Sie für **die Varianz von** Y erwartungstreue Schätzfunktionen $\widetilde{T}_n(X_1, \ldots, X_n)$ an.

Ergebnisse (ohne Begründung/Rechenweg):

(a) Beweis durch Berechnung des Erwartungswerts von $T_n(X_1, \ldots, X_n)$.

(b) $\widetilde{T}_n(X_1, \ldots, X_n) = \dfrac{1}{2n} \displaystyle\sum_{i=1}^{n} X_i^2$.

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

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

$$f_Y(y|b) = \begin{cases} \dfrac{4 \cdot b^4}{y^5} & \text{für } y \geq b \\ 0 & \text{sonst} \end{cases}$$

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

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

(b) Zeigen Sie, dass $E(Y) = \dfrac{4}{3} \cdot b$ gilt.

(c) Bestimmen Sie den Schätzer $\widehat{b}_{MM}$ 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) $\widehat{b}_{ML} = \min\{x_1, \ldots, x_n\}$

(b) Nachweis durch Berechnung des Erwartungswerts

(c) $\widehat{b}_{MM} = \dfrac{3}{4} \cdot \overline{x}$

Aufgabe 5 (7 + 2 + 3 + 7 = 19 Punkte)

Bei der Befüllung von Propangasflaschen weiß der Hersteller aus Erfahrung, dass die verwendete Abfüllanlage eine Standardabweichung von $0.2[kg]$ für die abgefüllte Menge hat. Nach einer Inventur hat der Hersteller den Verdacht, dass die Abfüllanlage im Mittel mehr als die auf dem Produkt ausgezeichneten $11[kg]$ in die Gasflaschen einfüllt. Dies soll mit einem statistischen Test überprüft werden. Hierzu werden der Produktion 16 Gasflaschen entnommen, deren gemessene Füllmengen $x_1, \ldots, x_{16}$ als Realisation einer einfachen Stichprobe vom Umfang 16 zur annahmegemäß $N(\mu, 0.2^2[kg^2])$-verteilten Abfüllmenge betrachtet werden können. Als Stichprobenmittelwert ergibt sich dabei

$$\overline{x} = \frac{1}{16} \sum_{i=1}^{16} x_i = 11.107[kg] \; .$$

(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 $11.15[kg]$ beträgt?

(d) Überprüfen Sie unter Verwendung der Varianzschätzung $s^2 = 0.04085$ mit einem geeigneten Test zum Signifikanzniveau $\alpha = 0.05$, ob die oben getroffene Annahme $\sigma^2 = 0.2^2$ aus statistischer Sicht zu verwerfen ist. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

Hinweis: Verwenden Sie für Teil (d) 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
11	3.053	3.816	4.575	10.341	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	11.340	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	12.340	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	13.339	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	14.339	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	15.338	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	16.338	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	17.338	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	18.338	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	19.337	28.412	31.410	34.170	37.566

Ergebnisse (ohne Begründung/Rechenweg):

(a) $N = 2.14 \in (1.645, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

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

(b) p-Wert $p = 0.0162$. Entscheidung wäre zu Gunsten der Nullhypothese ausgefallen.

(c) $\beta(11.15) = 0.0869$

(d) $\chi^2 = 15.319 \notin [0, 6.262) \cup (27.488, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!
Die getroffene Annahme $\sigma^2 = 0.2^2$ muss also nicht verworfen werden.

Aufgabe 6 (10 + 8 = 18 Punkte)

Ein Anbieter von Powerstations muss sich bei der Entwicklung und Herstellung eines neuen Modells zwischen zwei verschiedenen Lithium-Ionen-Rundzellen ("Typ A" und "Typ B") entscheiden. Nominell weist der Lieferant für Rundzellen des Typs B verglichen mit Typ A eine höhere entnehmbare Energie aus. Man nehme an, dass die tatsächlich entnehmbare Energie Y^A bzw. Y^B einer Rundzelle vom jeweiligen Typ (A bzw. B) jeweils normalverteilt sei 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 Rundzellen vom Typ B verglichen mit Typ A im Mittel tatsächlich eine höhere entnehmbare Energie liefern.
Aus einer Energiemessung von $n_A = 12$ Testexemplaren des Typs A sowie $n_B = 14$ Testexemplaren des Typs B erhielt man Realisationen jeweils voneinander unabhängiger einfacher Stichproben $X_1^A, \ldots, X_{12}^A$ zu Y^A sowie $X_1^B, \ldots, X_{14}^B$ zu Y^B und hieraus die zugehörigen Mittelwerte $\overline{x^A} = 18.25$ bzw. $\overline{x^B} = 19.214$ sowie die Stichprobenvarianzen $s_{Y^A}^2 = 0.932$ bzw. $s_{Y^B}^2 = 0.797$.

(a) Testen Sie unter der Annahme $\sigma_A^2 = \sigma_B^2$ zum Signifikanzniveau $\alpha = 0.05$ die Hypothese, dass Rundzellen vom Typ B verglichen mit Typ A im Mittel tatsächlich eine höhere entnehmbare Energie liefern. 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* $F_{m,n;p} = \frac{1}{F_{n,m;1-p}}$.

$n \backslash m$	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):

(a) $t = -2.643 \in (-\infty, -1.711) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

Der Test kann also die Vermutung, dass Rundzellen vom Typ B verglichen mit Typ A im Mittel tatsächlich eine höhere entnehmbare Energie liefern, bestätigen.

(b) $F = 1.169 \notin [0, 0.362) \cup (2.635, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

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

Aufgabe 7 (12 Punkte)

Die in einem bestimmten Chip-Fertigungsprozess hergestellten Halbleiter werden nach ihrer Fertigstellung einer Kontrolle unterzogen und anhand dieser Kontrolle einer der Qualitätsstufen Q_1, Q_2, Q_3 und Q_4 zugeordnet. Aufgrund langjähriger Erfahrung nimmt man als Wahrscheinlichkeiten für die einzelnen Qualitätsstufen an:

Qualitätsstufe	Q_1	Q_2	Q_3	Q_4
Wahrscheinlichkeit	0.45	0.27	0.15	0.13

Nach einer Umstellung des Fertigungsprozesses soll untersucht werden, ob sich die Verteilung der produzierten Chips auf die unterschiedlichen Qualitätsstufen geändert hat. Hierzu wurden der laufenden Produktion zufällig 200 Chips entnommen und deren Qualitätsstufen festgestellt. Die festgestellten Qualitätsstufen, die sich annahmegemäß als Realisation einer einfachen Stichprobe vom Umfang 200 auffassen lassen sollen, besitzen die folgende Häufigkeitsverteilung:

Qualitätsstufe	Q_1	Q_2	Q_3	Q_4
(absolute) Häufigkeit	84	65	18	33

Untersuchen Sie auf dieser Grundlage mit einem geeigneten statistischen Test zum Signifikanzniveau $\alpha = 0.05$, ob sich die Wahrscheinlichkeitsverteilung auf die Qualitätsstufen nach der Umstellung des Fertigungsprozesses geändert hat.

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):

$\chi^2 = 9.326 \in (7.815, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!
Die Verteilung der Qualitätsstufen hat sich also signifikant verändert.

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

Zur Erklärung des Verbrauchs alter PKW-Modelle aus dem Modelljahr 1973/74 y_i (in [l/100 km]) durch den Hubraum x_i (in [cm^3]) 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 eines Automobil-Magazins aus dem Jahr 1974 wurde das lineare Modell mit der Statistik-Software R wie folgt geschätzt:

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

Residuals:
    Min      1Q  Median      3Q     Max
-1.7174 -0.9476  0.3819  0.4017  2.1528

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 8.0372122  1.1492536   6.993 0.000213 ***
x           0.0026947  0.0007709   3.496 0.010052 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.331 on 7 degrees of freedom
Multiple R-squared:  0.6358,        Adjusted R-squared:  0.5838
F-statistic: 12.22 on 1 and 7 DF,  p-value: 0.01005
```

(a) Wie viele PKW-Modelle gingen in die Schätzung ein?

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

(c) Welcher Anteil der Gesamtvarianz des Verbrauchs alter PKW-Modelle aus dem Modelljahr 1973/74 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 Verbrauch (in [l/100 km]) prognostiziert das Modell für einen PKW aus dem Modelljahr 1973/74 mit einem Hubraum von 1200 (in [cm^3])?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $n = 9$

(b) $\widehat{\beta}_1 = 8.0372122$, $\widehat{\beta}_2 = 0.0026947$

(c) 0.6358

(d) β_1 ist signifikant von Null verschieden.

(e) β_2 ist signifikant positiv.

(f) 11.2709

Aufgabe 9 (6 + 2 + 3 + 5 + 5 = 21 Punkte)

Zur Schätzung eines einfachen linearen Regressionsmodells

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

aus einer Stichprobe vom Umfang $n = 25$ wurden bereits die folgenden Zwischenwerte errechnet:

$$\sum_{i=1}^{25} y_i = 81.058; \quad \sum_{i=1}^{25} y_i^2 = 494.131; \quad \sum_{i=1}^{25} x_i = 113.742;$$

$$\sum_{i=1}^{25} x_i^2 = 587.459; \quad \sum_{i=1}^{25} x_i \cdot y_i = 309.304$$

(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) Überprüfen Sie zum Signifikanzniveau $\alpha = 0.05$, ob β_2 signifikant negativ ist. Fassen Sie das Ergebnis auch in einem Antwortsatz zusammen.

(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) $\widehat{\beta}_1 = 7.1104$, $\widehat{\beta}_2 = -0.85018$

(b) $\widehat{\sigma}^2 = 7.8586$

(c) $\widehat{\sigma^2}_{\widehat{\beta}_1} = 2.6394$, $\widehat{\sigma^2}_{\widehat{\beta}_2} = 0.11232$

(d) $t = -2.537 \in (-\infty, -1.714) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

β_2 ist also signifikant negativ.

(e) $[2.979, 6.141]$