Wirtschaftswissenschaftliches Prüfungsamt

Bachelor-Prüfung
Schliessende Statistik
Wintersemester 2011/12

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 + 18 + 18 + 14 + 6 + 20) 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.

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 −1 Punkt, nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).
Die Aufgabe wird insgesamt mit mindestens 0 Punkten bewertet!

wahr falsch

1. Sei $X_1, \ldots, X_n$ eine einfache Stichprobe vom Umfang n zu einer normalverteilten Zufallsvariablen Y. Dann besitzen $X_1, \ldots, X_n$ stets übereinstimmende Varianzen. ■ ☐

2. Ohne die Voraussetzung einer einfachen Stichprobe wäre bei der ML-Methode die (einfache) Darstellung der Likelihoodfunktion als Produkt von Randdichten bzw. Randwahrscheinlichkeitsfunktionen im Allgemeinen nicht möglich. ■ ☐

3. Damit eine Familie von Schätzfunktionen konsistent im quadratischen Mittel sein kann, müssen die einzelnen Schätzfunktionen auf jeden Fall erwartungstreu sein. ☐ ■

4. Ist $X_1, \ldots, X_n$ eine einfache Stichprobe vom Umfang n zu einer Zufallsvariablen Y mit $\mu := \mathrm{E}(Y)$, so ist $\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ in der Klasse der linearen erwartungstreuen Schätzer effizient für μ. ■ ☐

5. Das (symmetrische) Konfidenzintervall für den Erwartungswert μ bei bekannter Varianz σ^2 zum Konfidenzniveau $1-\alpha$ gibt einen Bereich an, in dem sich der zufällig schwankende Erwartungswert μ mit einer Wahrscheinlichkeit von mindestens $1-\alpha$ realisiert. ☐ ■

6. Mit Hilfe von Gütefunktionswerten kann man in Abhängigkeit der wahren Verteilungsparameter sowohl die Wahrscheinlichkeit für einen Fehler als auch die Wahrscheinlichkeit für eine korrekte Entscheidung leicht berechnen. Für feste Verteilungsparameter addieren sich diese beiden Wahrscheinlichkeiten stets zu 1. ■ ☐

7. Bei der einfachen Varianzanalyse basiert die Teststatistik auf dem Verhältnis von zwei Mittelwertschätzern. ☐ ■

8. Im einfachen linearen Regressionsmodell ☐ ■

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

sind Prognoseintervalle zur Vertrauenswahrscheinlichkeit $1-\alpha$ für y_0 stets kleiner als die entsprechenden Prognoseintervalle für $\mathrm{E}(y_0)$.

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 −1 Punkt, nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).
Die Aufgabe wird insgesamt mit mindestens 0 Punkten bewertet!

1. Der (approximative) t-Test zum Vergleich von Anteilswerten bzw. Erfolgswahrscheinlichkeiten kann als Spezialfall des 2-Stichproben-t-Tests für den Mittelwertvergleich bei unbekannten, aber **übereinstimmenden** Varianzen betrachtet werden, weil es sich dabei um einen Mittelwertvergleich handelt und

   (a) die Einhaltung von Anwendungsvoraussetzungen bei approximativen Tests generell nicht wichtig ist. ☐

   (b) die Varianzen von Alternativverteilungen nicht von ihrem Parameter abhängig sind. ☐

   (c) die Varianzen von Alternativverteilungen sich bei großen Stichprobenumfängen nicht mehr wesentlich unterscheiden. ☐

   (d) die Varianzen der Alternativverteilungen unter H_0 (bei Gleichheit der Parameter) übereinstimmen. ■

2. Beim zweiseitigen t-Test für den Mittelwert normalverteilter Zufallsvariablen bei unbekannter Varianz führt ein beobachteter Abstand des Stichprobenmittelwerts $\overline{x}$ zum „Sollwert“ μ_0 umso eher zur Ablehnung von H_0, je

   (a) geringer die Streuung S der Stichprobe und je geringer das Signifikanzniveau α ist. ☐

   (b) geringer die Streuung S der Stichprobe und je größer das Signifikanzniveau α ist. ■

   (c) größer die Streuung S der Stichprobe und je geringer das Signifikanzniveau α ist. ☐

   (d) größer die Streuung S der Stichprobe und je größer das Signifikanzniveau α ist. ☐

3. 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 $p = 0.06345$. 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):

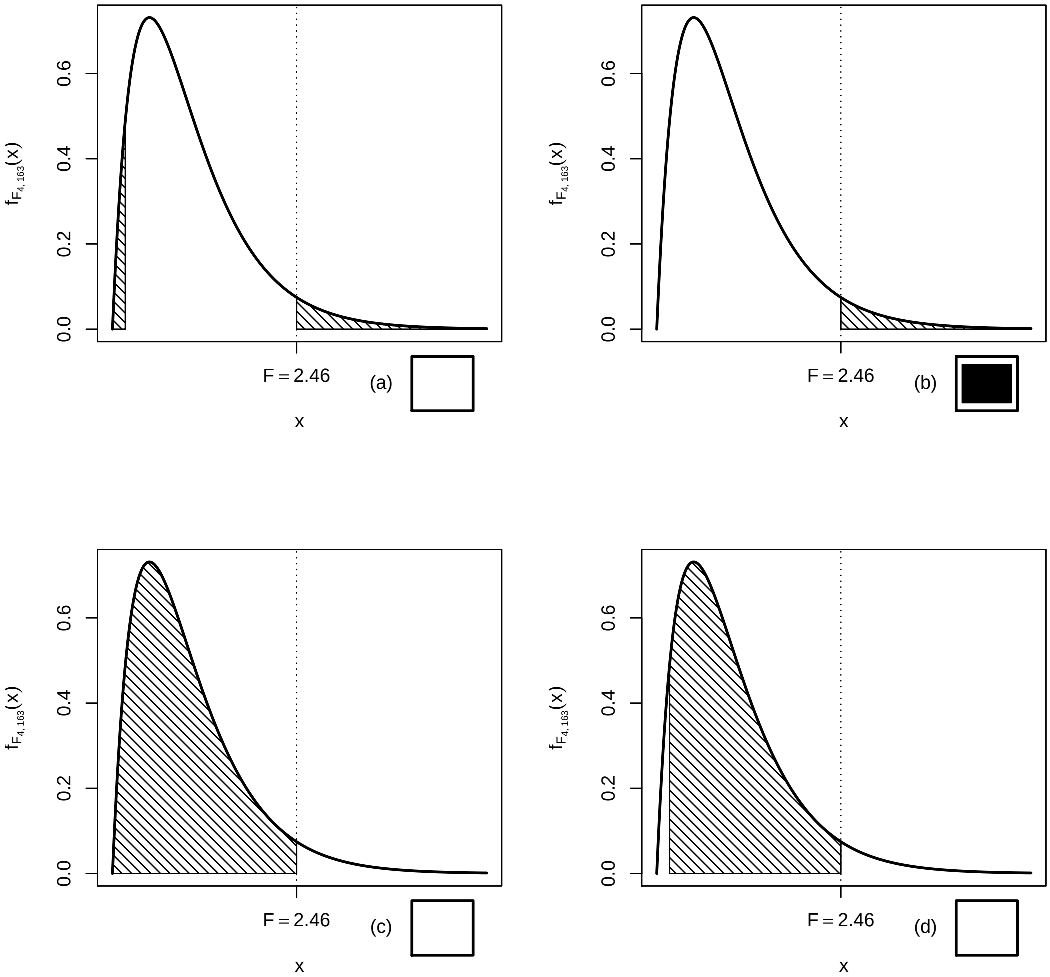
   (a) Bei beiden einseitigen Tests wird H_0 abgelehnt. ☐

   (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. ☐

4. Bei der Durchführung einer Varianzanalyse mit $k = 5$ Faktorstufen und einem Gesamtstichprobenumfang von $n = 168$ erhält man die realisierte Teststatistik $F = 2.46$. 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.



Aufgabe 3 (1 + 3 + 4 = 8 Punkte)

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

- $\widehat{\mu} = \dfrac{1}{2} \cdot \overline{X^A} + \dfrac{1}{2} \cdot \overline{X^B}$
- $\widetilde{\mu} = \dfrac{n_A}{n_A + n_B} \cdot \overline{X^A} + \dfrac{n_B}{n_A + n_B} \cdot \overline{X^B}$ und

zur Schätzung von μ betrachtet.

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

(b) Zeigen Sie, dass sowohl $\widehat{\mu}$ als auch $\widetilde{\mu}$ erwartungstreu sind für μ.

(c) Welche der beiden Schätzfunktionen $\widehat{\mu}$ und $\widetilde{\mu}$ würden Sie für $n_A = 10$ und $n_B = 20$ vorziehen? Begründen Sie Ihre Antwort.

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\overline{X^A} \sim N\left(\mu, \dfrac{\sigma^2}{n_A}\right)$, $\overline{X^B} \sim N\left(\mu, \dfrac{\sigma^2}{n_B}\right)$.

(b) Nachrechnen ergibt $\mathrm{E}(\widehat{\mu}) = \mathrm{E}(\widetilde{\mu}) = \mu$ für alle $\mu \in \mathbb{R}$, damit sind beide Schätzfunktionen erwartungstreu für μ.

(c) Man erhält $\mathrm{Var}(\widehat{\mu}) = \frac{3}{80}\sigma^2$ und $\mathrm{Var}(\widetilde{\mu}) = \frac{1}{30}\sigma^2$, also ist $\widetilde{\mu}$ wirksamer als $\widehat{\mu}$ und deshalb vorzuziehen.

Aufgabe 4 (6 + 2 = 8 Punkte)

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

$$f_Y(y) = \begin{cases} 2ay \cdot e^{-a \cdot y^2} & \text{für } y \geq 0 \\ 0 & \text{sonst} \end{cases}$$

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 $\mathrm{E}(Y) = \sqrt{\frac{\pi}{4 \cdot a}}$ 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) $\widehat{a}_{ML} = \frac{n}{\sum_{i=1}^{n} x_i^2} \left(= \frac{1}{\overline{x^2}} \right)$.

(b) $\widehat{a}_{MM} = \frac{\pi}{4\overline{x}^2}$.

Aufgabe 5 (7 + 2 + 5 + 4 = 18 Punkte)

Bei der Abfüllung von Parfum-Flacons weiß man aus langjähriger Erfahrung, dass die verwendete Maschine eine Varianz von $0.3^2 = 0.09[ml^2]$ für die abgefüllte Parfummenge hat. Nach einer Inventur hat der Hersteller den Verdacht, dass die Abfüllanlage im Mittel mehr als die auf dem Produkt ausgezeichneten $30[ml]$ in die Flacons einfüllt. Dies soll mit einem statistischen Test überprüft werden. Hierzu werden der Produktion 25 Flaschen entnommen, deren gemessene Füllmengen $x_1, \ldots, x_{25}$ als Realisation einer einfachen Stichprobe vom Umfang 25 zur annahmegemäß $N(\mu, 0.3^2[ml^2])$-verteilten Abfüllmenge betrachtet werden kann. Als Stichprobenmittelwert ergibt sich dabei

$$\overline{x} = \frac{1}{25}\sum_{i=1}^{25} x_i = 30.083[ml] \ .$$

(a) Testen Sie zum Signifikanzniveau $\alpha = 0.1$ **(!)**, 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 1. Art, falls $\mu = 29.95[ml]$ beträgt?

(d) Geben Sie auf Basis der Stichprobenrealisation ein zweiseitiges Konfidenzintervall für die mittlere Abfüllmenge zur Sicherheitswahrscheinlichkeit $1 - \alpha = 0.95$ an.

Ergebnisse (ohne Begründung/Rechenweg):

(a) $N = 1.383 \in (1.282, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

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

(b) p-Wert: 0.0838

(c) Fehlerwahrscheinlichkeit 1. Art für $\mu = 29.95$: 0.017

(d) Realisiertes symmetrisches Konfidenzintervall zum Konfidenzniveau $1-\alpha$: $[29.965, 30.201]$

Aufgabe 6 (10 + 8 = 18 Punkte)

Es werde angenommen, dass die maximale Zugkraft (bis zur Zerstörung) Y^A eines speziellen Seilfabrikats normalverteilt sei mit unbekanntem Erwartungswert μ_A und unbekannter Varianz σ_A^2. Nach Umstellung der Produktion auf ein neues Basismaterial wird weiterhin von einer normalverteilten maximalen Zugkraft Y^B (mit Erwartungswert μ_B und Varianz σ_B^2) ausgegangen. Es soll überprüft werden, ob die mittlere Belastbarkeit (maximale Zugkraft bis zur Zerstörung) durch Verwendung des neuen Materials verringert worden ist.
Aus einem Belastungstest von $n_A = 10$ Seilen mit altem und $n_B = 8$ Seilen mit neuem Basismaterial erhielt man Realisationen jeweils voneinander unabhängiger einfacher Stichproben $X_1^A, \dots, X_{10}^A$ zu Y^A sowie $X_1^B, \dots, X_8^B$ zu Y^B und hieraus die zugehörigen Mittelwerte $\overline{x^A} = 152.11$ bzw. $\overline{x^B} = 140.44$ sowie die Stichprobenvarianzen $s_{Y^A}^2 = 97.02$ bzw. $s_{Y^B}^2 = 166.41$.

(a) Testen Sie unter der Annahme $\sigma_A^2 = \sigma_B^2$ zum Signifikanzniveau $\alpha = 0.05$ die Hypothese, dass die mittlere Belastbarkeit durch Verwendung des neuen Materials verringert worden ist. 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) $t = 2.18 \in (1.746, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

Der Test kann also die Vermutung, dass die mittlere Belastbarkeit durch Verwendung des neuen Materials verringert worden ist, bestätigen.

(b) $F = 0.583 \notin [0, 0.304) \cup (3.677, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

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

Aufgabe 7 (11 + 3 = 14 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 $n = 100$ zu einer Geom(0.4)-verteilten Zufallsvariablen ausgehen kann. Die Stichprobeninformation liege in Form der folgenden Häufigkeitsverteilung vor:

i	1	2	3	4
a_i	0	1	2	≥ 3
n_i	30	17	18	35

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

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

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

(b) Zum Test, ob die angegebene Häufigkeitsverteilung als Stichprobenrealisation zu *irgendeiner* geometrischen Verteilung Geom(p) (für ein beliebiges $p \in (0,1)$) plausibel ist, wurde der Verteilungsparameter p mit Hilfe einer ML-Schätzung aus den wie oben klassierten Daten (zu $\widehat{p} = 0.2915$) geschätzt und damit die (neue) Teststatistik $\chi^2 = 1.4499$ berechnet. Zu welchem Ergebnis kommt dieser Test? Begründen Sie Ihre Antwort durch die Angabe des zugehörigen kritischen Bereichs.

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

$n \backslash p$	0.01	0.025	0.05	0.5	0.9	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
96	66.730	70.783	74.401	95.334	114.131	119.871	125.000	131.141
97	67.562	71.642	75.282	96.334	115.223	120.990	126.141	132.309
98	68.396	72.501	76.164	97.334	116.315	122.108	127.282	133.476
99	69.230	73.361	77.046	98.334	117.407	123.225	128.422	134.642
100	70.065	74.222	77.929	99.334	118.498	124.342	129.561	135.807

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\chi^2 = 13.7547 \in (7.815, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

Die Nullhypothese einer Geom(0.4)-Verteilung muss also abgelehnt werden.

(b) $\chi^2 = 1.4499 \notin (5.991, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

Die Nullhypothese einer Geom(p)-Verteilung (für ein beliebiges $p \in (0,1)$) kann also nicht abgelehnt werden.

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

Zur Erklärung der Verkaufspreise von Häusern y_i (in kanadischen Dollar) durch die Grundstücksfläche x_i (in Quadratfuß) 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 einiger Hausverkäufe wurde das lineare Modell mit der Statistik-Software R wie folgt geschätzt:

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

Residuals:
   Min     1Q Median     3Q    Max
-26039 -11595  -3756  10056  38634

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 34886.756  11816.597   2.952  0.01122 *
x               8.331      2.418   3.445  0.00435 **
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 18590 on 13 degrees of freedom
Multiple R-squared:  0.4773,       Adjusted R-squared:  0.4371
F-statistic: 11.87 on 1 and 13 DF,  p-value: 0.004348
```

(a) Wie viele Hausverkäufe 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 Verkaufspreise wird durch das lineare Modell erklärt?

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

(f) Welchen Verkaufspreis prognostiziert das Modell für ein Haus mit Grundstücksfläche 8000 (in Quadratfuß)?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $n = 15$

(b) $\widehat{\beta}_1 = 34886.756$, $\widehat{\beta}_2 = 8.331$

(c) $\widehat{\sigma}^2 = 345588100$

(d) 0.4773

(e) β_2 signifikant von Null verschieden.

(f) $\widehat{y}_0 = 101534.756$

Aufgabe 9 (6 + 3 + 3 + 5 + 3 = 20 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 $n = 20$ wurden bereits die folgenden Zwischenwerte errechnet:

$$\sum_{i=1}^{20} y_i = 195.6; \quad \sum_{i=1}^{20} y_i^2 = 1966.005; \quad \sum_{i=1}^{20} x_i = 101.925;$$

$$\sum_{i=1}^{20} x_i^2 = 559.677; \quad \sum_{i=1}^{20} x_i \cdot y_i = 1037.522; \quad \sum_{i=1}^{20} \hat{y}_i^2 = 1954.12.$$

(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 $\alpha = 0.01$, ob β_1 signifikant positiv ist. Fassen Sie das Ergebnis auch in einem Antwortsatz zusammen.

(e) Geben Sie ein symmetrisches Konfidenzintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.99$ für β_2 an.

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\widehat{\beta}_1 = 4.628$, $\widehat{\beta}_2 = 1.011$

(b) $\widehat{\sigma}^2 = 0.6584$

(c) $\widehat{\sigma^2}_{\widehat{\beta}_1} = 0.4572$, $\widehat{\sigma^2}_{\widehat{\beta}_2} = 0.01634$

(d) $t = 6.844 \in (2.552, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

β_1 ist also signifikant positiv.

(e) $[0.6432, 1.379]$