Aufgabenstellung und Ergebnisse zur

Bachelor-Prüfung
Schliessende Statistik
Wintersemester 2023/24

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 + 8 + 10 + 22 + 7 + 14 + 5 + 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)	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!).

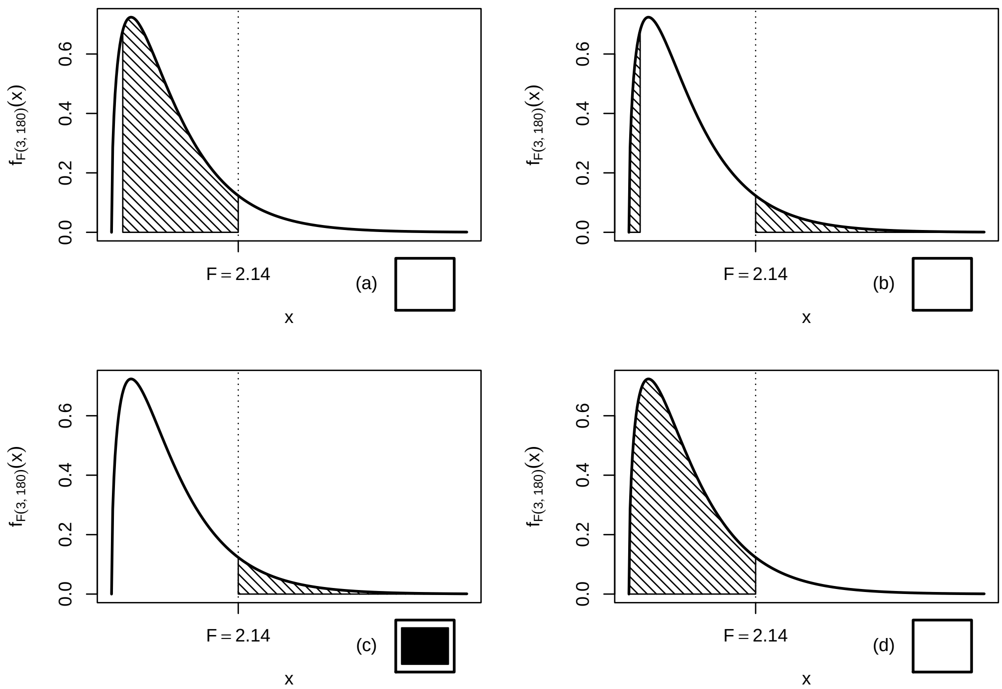
		wahr	falsch
1.	Sind die Zufallsvariablen $X_1, X_2, \dots, X_n$ nicht stochastisch unabhängig, so kann es sich bei diesen Zufallsvariablen nie um eine einfache Stichprobe vom Umfang n handeln.	■	□
2.	Ist eine Folge von Schätzfunktionen T_n, $n \in \mathbb{N}$, asymptotisch erwartungstreu für den Parameter λ, so gilt $\lim_{n\to\infty} \mathrm{Var}(T_n) = 0$.	□	■
3.	Gilt $0 < \widetilde{\alpha} < \alpha < 1$, so ist das Konfidenzintervall für den Erwartungswert zum Konfidenzniveau $1-\alpha$ stets im (zur selben Stichprobenrealisation bestimmten) Konfidenzintervall für den Erwartungswert zum Konfidenzniveau $1-\widetilde{\alpha}$ enthalten.	■	□
4.	Wird die Verletzung der Nullhypothese $\mu = 100$ bei einem tatsächlichen Erwartungswert von 102 im Mittel bei 7 von 10 Durchführungen des zugehörigen Gauß-Tests auch erkannt, so gilt für die zugehörige Gütefunktion $G(102) = 0.7$.	■	□
5.	Ist die Nullhypothese bei einem zweiseitigen Gauß-Test zum Signifikanzniveau α tatsächlich verletzt, so wird man nach erfolgter Stichprobenziehung stets einen p-Wert kleiner als α erhalten.	□	■
6.	Sind die Voraussetzungen für die exakte Anwendung eines zweiseitigen t-Differenzentests für den Vergleich zweier Erwartungswerte erfüllt, so stimmen die Verteilungen von Y^A und Y^B bei Gültigkeit von H_0 stets vollständig überein.	□	■
7.	Kann ein Chi-Quadrat-Unabhängigkeitstest die Nullhypothese zum Signifikanzniveau $\alpha = 0.01$ nicht ablehnen, so wird die Nullhypothese auch bei einem Signifikanzniveau von $\alpha = 0.05$ nicht verworfen.	□	■
8.	Mit der einfachen Varianzanalyse können mehrere mit übereinstimmenden Erwartungswerten normalverteilte Grundgesamtheiten daraufhin untersucht werden, ob ihre Varianzen (ebenfalls) übereinstimmen.	□	■
9.	Gilt bei der Schätzung eines einfachen linearen Regressionsmodells $\overline{x} = 8.3$, so ist das Prognoseintervall für $\mathrm{E}(y_0)$ gegeben $x_0 = 10$ zur Vertrauenswahrscheinlichkeit 0.95 breiter als das Prognoseintervall für $\mathrm{E}(y_0)$ gegeben $x_0 = 5$ (zur selben Vertrauenswahrscheinlichkeit).	□	■

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. Bei der Durchführung eines linksseitigen Gauß-Tests für den Erwartungswert einer normalverteilten Zufallsvariablen mit bekannter Varianz zu einem Signifikanzniveau $\alpha \in (0, 0.5)$
   (a) liegt die Teststatistik N bei Gültigkeit von H_0 stets im kritischen Bereich K. ☐
   (b) liegt die Teststatistik N bei Gültigkeit von H_0 mit einer Wahrscheinlichkeit von mindestens $1 - \alpha$ im kritischen Bereich K. ☐
   (c) liegt die Teststatistik N bei Gültigkeit von H_0 mit einer Wahrscheinlichkeit von höchstens α im kritischen Bereich K. ■
   (d) liegt die Teststatistik N bei Gültigkeit von H_0 nie im kritischen Bereich K. ☐

2. Bei der Durchführung einer einfachen Varianzanalyse mit $k = 4$ Faktorstufen und einem Gesamtstichprobenumfang von $n = 184$ erhält man die realisierte Teststatistik $F = 2.14$. 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.



3. Zur Schätzung des Parameters $\theta \geq 0$ seien für (Stichprobenumfänge) $n \geq 2$ jeweils zwei (für den Parameter θ) erwartungstreue Schätzfunktionenen T_n sowie $\widetilde{T}_n$ gegeben mit $\text{Var}(T_n) = \frac{1}{n-1}\theta$ sowie $\text{Var}(\widetilde{T}_n) = \frac{1}{n}\theta$. Damit gilt (für alle $n \geq 2$):

   (a) T_n ist wirksamer als $\widetilde{T}_n$. ☐

   (b) $\widetilde{T}_n$ ist wirksamer als T_n. ■

   (c) Keine der beiden Schätzfunktionen ist wirksamer als die jeweils andere. ☐

   (d) Auf Grundlage der gegebenen Informationen ist keine Aussage zur (gegenseitigen) Wirksamkeit möglich. ☐

4. Bei der Durchführung eines t-Tests für den Erwartungswert einer normalverteilten Zufallsvariablen bei unbekannter Varianz auf Grundlage einer einfachen Stichprobe vom Umfang n zum Signifikanzniveau α lehnen sowohl der linksseitige als auch der zweiseitige Test H_0 ab. Damit weiß man über die Realisation t der Teststatistik:

   (a) $t \in (-\infty, -t_{n-1,1-\frac{\alpha}{2}})$ ■

   (b) $t \in [-t_{n-1,1-\frac{\alpha}{2}}, -t_{n-1,1-\alpha})$ ☐

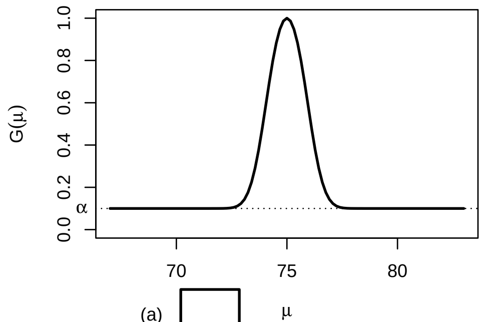
   (c) $t \in (t_{n-1,1-\alpha}, t_{n-1,1-\frac{\alpha}{2}}]$ ☐

   (d) $t \in (t_{n-1,1-\frac{\alpha}{2}}, \infty)$ ☐

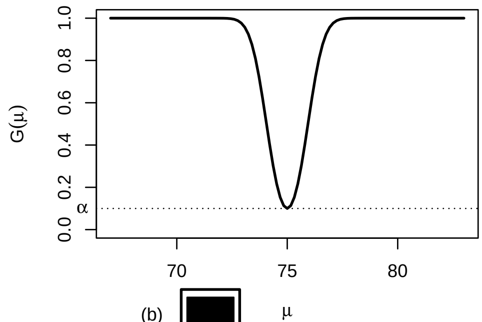
5. Auf der Grundlage einer einfachen Stichprobe $X_1, \ldots, X_{49}$ vom Umfang $n = 49$ zu einer $N(\mu, 4^2)$-verteilten Zufallsvariablen wird ein Gauß-Test zur Überprüfung der Hypothesen

$$H_0 : \mu = 75 \qquad \text{gegen} \qquad H_1 : \mu \neq 75$$

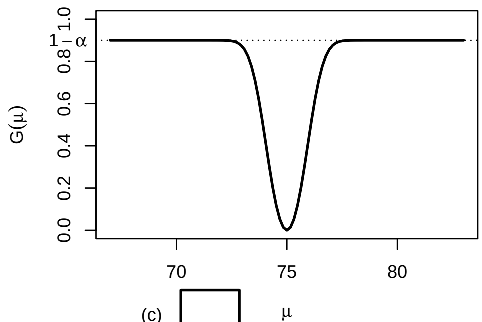
   bei einem Signifikanzniveau von $\alpha = 0.1$ betrachtet.
   Markieren Sie die Abbildung, welche die Gütefunktion des oben genannten Tests korrekt darstellt.



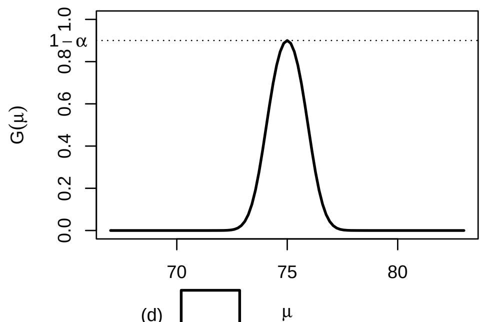
(a) ☐



(b) ■



(c) ☐



(d) ☐

Aufgabe 3 (3 + 3 + 2 = 8 Punkte)

In Abhängigkeit eines unbekannten Parameters a mit $0 \leq a \leq 4$ seien der Erwartungswert und die Varianz von Zufallsvariablen Y mit der zugehörigen Verteilung aus einer parametrischen Verteilungsfamilie gegeben durch

$$\mathrm{E}(Y) = \frac{a+4}{3} \qquad \text{sowie} \qquad \mathrm{Var}(Y) = \frac{a^2 - 4a + 16}{18}\ .$$

Für $n \in \mathbb{N}$ seien $X_1, \ldots, X_n$ eine einfache Stichprobe vom Umfang n zu Y und T_n die wie folgt definierte Schätzfunktion für a:

$$T_n(X_1, \ldots, X_n) := \left(\frac{3}{n}\sum_{i=1}^{n} X_i\right) - 4$$

(a) Berechnen Sie den Bias der Schätzfunktionen T_n für a.

(b) Berechnen Sie die Varianz der Schätzfunktionen T_n.

(c) Ist die Folge von Schätzfunktionen T_n, $n \in \mathbb{N}$, konsistent im quadratischen Mittel für a? *(Begründung erforderlich!)*

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\mathsf{Bias}(T_n(X_1, \ldots, X_n)) = 0$.

(b) $\mathrm{Var}(T_n(X_1, \ldots, X_n)) = \frac{a^2-4a+16}{2n}$.

(c) Ja.

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

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

$$f_Y(y|\lambda) = \begin{cases} \frac{1}{9} \cdot \lambda^{-3} \cdot y^2 & \text{für } 0 \leq y \leq 3 \cdot \lambda \\ 0 & \text{sonst} \end{cases}$$

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

(a) Zeigen Sie, dass $\mathrm{E}(Y) = \frac{9}{4} \cdot \lambda$ gilt.

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

(c) Bestimmen Sie den Schätzer $\widehat{\lambda}_{ML}$ 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) $\widehat{\lambda}_{MM} = \frac{4}{9} \cdot \overline{x}$

(c) $\widehat{\lambda}_{ML} = \frac{1}{3} \cdot \max\{x_1, \ldots, x_n\}$

Aufgabe 5 (3 + 7 + 2 + 3 + 7 = 22 Punkte)

Bei der Herstellung von Filamentrollen für 3D-Drucker weiß der Hersteller aus Erfahrung, dass die verwendete Maschine eine Standardabweichung von $3[g]$ für die aufgewickelte Menge hat. Nach einer Inventur hat der Hersteller den Verdacht, dass die Maschine im Mittel weniger als die auf dem Produkt ausgezeichneten $1000[g]$ auf die Rollen aufwickelt. Dies soll mit einem statistischen Test überprüft werden. Hierzu werden der Produktion 16 Rollen entnommen, deren gemessene Filamentmengen $x_1, \ldots, x_{16}$ als Realisation einer einfachen Stichprobe vom Umfang 16 zur annahmegemäß $N(\mu, 3^2[g^2])$-verteilten aufgewickelten Menge betrachtet werden können. Als Stichprobenmittelwert ergibt sich dabei

$$\overline{x} = \frac{1}{16} \sum_{i=1}^{16} x_i = 998.585[g] \ .$$

(a) Geben Sie auf Basis der Stichprobenrealisation ein zweiseitiges Konfidenzintervall für die mittlere aufgewickelte Menge zur Sicherheitswahrscheinlichkeit $1 - \alpha = 0.99$(!) an.

(b) 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.

(c) Berechnen Sie den p-Wert zum Test aus Teil (b). Wie wäre die Entscheidung zum Test aus Teil (b) bei einem Signifikanzniveau von $\alpha = 0.01$ ausgefallen?

(d) Wie groß ist die Wahrscheinlichkeit, bei der Ziehung einer Stichprobe und der anschließenden Durchführung des Tests aus Teil (b) eine Testentscheidung zu Gunsten der Nullhypothese zu erhalten, falls die tatsächliche mittlere Abfüllmenge $998[g]$ beträgt?

(e) Überprüfen Sie unter Verwendung der Varianzschätzung $s^2 = 5.946$ mit einem geeigneten Test zum Signifikanzniveau $\alpha = 0.05$, ob die oben getroffene Annahme $\sigma^2 = 3^2$ aus statistischer Sicht zu verwerfen ist. Fassen Sie das Ergebnis des Tests in einem Antwortsatz zusammen.

Hinweis: Verwenden Sie für Teil (e) 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) Realisiertes symm. Konfidenzintervall zum Konf.-niveau $1-\alpha = 0.99$: $[996.653, 1000.517]$

(b) $N = -1.887 \in (-\infty, -1.645) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

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

(c) p-Wert $p = 0.0294$. Entscheidung wäre zu Gunsten der Nullhypothese ausgefallen.

(d) $\beta(998) = 0.1539$

(e) $\chi^2 = 9.91 \notin [0, 6.262) \cup (27.488, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

Die getroffene Annahme $\sigma^2 = 3^2$ muss also laut Test nicht verworfen werden.

Aufgabe 6 (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 $X_1^A, \ldots, X_{26}^A$ vom Umfang 26 zu Y^A sowie $X_1^B, \ldots, X_{29}^B$ vom Umfang 29 zu Y^B auffassen, aus den zugehörigen Realisationen wurden bereits die Mittelwerte $\overline{x^A} = 1000.719$ bzw. $\overline{x^B} = 999.963$ sowie die Stichprobenvarianzen $s_{Y^A}^2 = 30.361$ bzw. $s_{Y^B}^2 = 19.273$ berechnet. Überprüfen Sie mit einem geeigneten Test zum Signifikanzniveau $\alpha = 0.05$, ob Messgerät A eine geringere Präzision (im Sinne einer höheren 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* $F_{m,n;p} = \frac{1}{F_{n,m;1-p}}$.

$n\backslash m$	21	22	23	24	25	26	27	28	29	30
21	2.084	2.073	2.063	2.054	2.045	2.037	2.030	2.023	2.016	2.010
22	2.059	2.048	2.038	2.028	2.020	2.012	2.004	1.997	1.990	1.984
23	2.036	2.025	2.014	2.005	1.996	1.988	1.981	1.973	1.967	1.961
24	2.015	2.003	1.993	1.984	1.975	1.967	1.959	1.952	1.945	1.939
25	1.995	1.984	1.974	1.964	1.955	1.947	1.939	1.932	1.926	1.919
26	1.978	1.966	1.956	1.946	1.938	1.929	1.921	1.914	1.907	1.901
27	1.961	1.950	1.940	1.930	1.921	1.913	1.905	1.898	1.891	1.884
28	1.946	1.935	1.924	1.915	1.906	1.897	1.889	1.882	1.875	1.869
29	1.932	1.921	1.910	1.901	1.891	1.883	1.875	1.868	1.861	1.854
30	1.919	1.908	1.897	1.887	1.878	1.870	1.862	1.854	1.847	1.841

Ergebnisse (ohne Begründung/Rechenweg):

$F = 1.575 \notin (1.906, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!
Der Test kann also die Vermutung, dass Messgerät A eine geringere Präzision (im Sinne einer höheren Streuung) als Messgerät B hat, nicht bestätigen.

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 Exp(0.4)-verteilten Zufallsvariablen ausgehen kann. Die Stichprobeninformation liege in Form der folgenden Häufigkeitsverteilung vor:

a_i	$(-\infty, 2]$	$(2, 4]$	$(4, 6]$	$(6, \infty)$
n_i	42	21	19	18

(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* Exponentialverteilung Exp(λ) (für ein beliebiges $\lambda > 0$) plausibel ist, wurde der Verteilungsparameter λ mit Hilfe einer ML-Schätzung aus den wie oben klassierten Daten (zu $\widehat{\lambda} = 0.27$) geschätzt und damit die (neue) Teststatistik $\chi^2 = 2.2607$ berechnet. Zu welchem Ergebnis kommt dieser Test? Begründen Sie Ihre Antwort durch die Angabe des zugehörigen kritischen Bereichs.

Hinweise:

- *Die Verteilungsfunktion einer* Exp(λ)*-verteilten Zufallsvariablen X ist gegeben durch:*

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 - e^{-\lambda x} & \text{für } x \geq 0 \end{cases}$$

- *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):

(a) $\chi^2 = 18.0436 \in (7.815, +\infty) = K \quad \Rightarrow \quad H_0$ wird abgelehnt!

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

(b) Unter Berücksichtigung der um 1 reduzierten Freiheitsgrade erhält man:

$\chi^2 = 2.2608 \notin (5.991, +\infty) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

Die Nullhypothese einer Exp(λ)-Verteilung (für ein beliebiges $\lambda > 0$) kann also nicht abgelehnt werden.

Aufgabe 8 (1 + 1 + 1 + 1 + 1 = 5 Punkte)

Zur Erklärung der durchschnittlichen jährlichen privaten Pro-Kopf-Konsumausgaben in Deutschland y_i (in Euro) durch die durchschnittlichen jährlichen verfügbaren Pro-Kopf-Nettoeinkommen in Deutschland x_i (in Euro) unterstellt man die Gültigkeit eines Zusammenhangs im Sinne des folgenden linearen Modells:

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

Aus vorliegenden Daten für die Jahre 2013 – 2020 wurde das lineare Modell mit der Statistik-Software R wie folgt geschätzt:

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

Residuals:
    Min      1Q  Median      3Q     Max
-780.98 -100.12   27.88  202.30  523.76

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 6463.6822  2513.0141   2.572  0.04222 *
x              0.6245     0.1135   5.501  0.00151 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 408.6 on 6 degrees of freedom
Multiple R-squared:  0.8345,       Adjusted R-squared:  0.8069
F-statistic: 30.26 on 1 and 6 DF,  p-value: 0.001514
```

(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 der durchschnittlichen jährlichen privaten Pro-Kopf-Konsumausgaben in Deutschland wird durch das lineare Modell erklärt?

(d) Entscheiden Sie mit Hilfe des zugehörigen p-Werts zum Signifikanzniveau $\alpha = 0.001$, ob β_2 signifikant positiv ist.

(e) Welche durchschnittlichen jährlichen privaten Pro-Kopf-Konsumausgaben (in Euro) prognostiziert das Modell bei einem durchschnittlichen jährlichen verfügbaren Pro-Kopf-Nettoeinkommen von 22000 (in Euro)?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\widehat{\beta}_1 = 6463.6822$, $\widehat{\beta}_2 = 0.6245$

(b) $\widehat{\sigma^2} = 166953.96$

(c) 0.8345

(d) β_2 ist signifikant positiv.

(e) 20202.6822

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:

$$\sum_{i=1}^{25} y_i = 73.65; \quad \sum_{i=1}^{25} y_i^2 = 586.79; \quad \sum_{i=1}^{25} x_i = 154.1;$$

$$\sum_{i=1}^{25} x_i^2 = 1053.69; \quad \sum_{i=1}^{25} x_i \cdot y_i = 383.11$$

(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_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 $x_0 = 4$ an.

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\widehat{\beta}_1 = 7.1537$, $\widehat{\beta}_2 = -0.68262$

(b) $\widehat{\sigma}^2 = 13.9757$

(c) $\widehat{\sigma^2}_{\widehat{\beta}_1} = 5.6738$, $\widehat{\sigma^2}_{\widehat{\beta}_2} = 0.13462$

(d) $t = -1.86 \notin (-\infty, -2.5) = K \quad \Rightarrow \quad H_0$ wird nicht abgelehnt!

β_2 ist also nicht signifikant negativ.

(e) $[2.166, 6.68]$