

Statistik Vorlesung

02. Februar 2026

Arbeitszeit: 100 Minuten

VORNAME:		MATR.NR.:	
NACHNAME:			

ERLAUBT: **Skriptum des Instituts, Taschenrechner gemäß Liste**

VERBOTEN: **alle sonstigen Unterlagen, Handys**

Aufgaben 8 und 9: Für jede vollständig richtig gelöste Frage gibt es 2 Punkte. Es gibt keine Teilpunkte und keine Minuspunkte.

Lösungswege müssen nachvollziehbar aufgeschrieben sein.

Aufgabe	max. Punkte	erreichte Punkte	Teilsumme
1	10		
2	12		
3	10		
4	10		
5	16		
6	12		
7	10		
8	10		
9	10		
Summe	100		
Note			

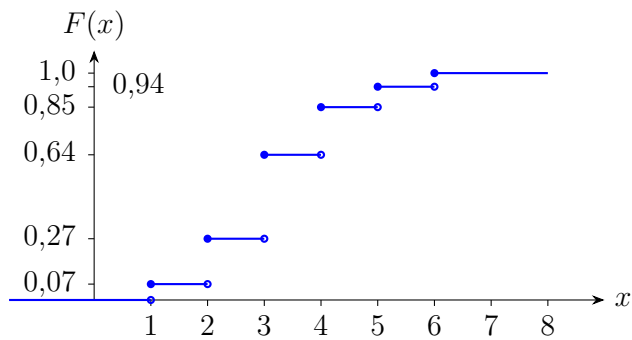
1. Wie viele Abendkleider haben Damen im Kleiderschrank? Dazu liegt folgende Statistik vor:

Anzahl der Abendkleider	1	2	3	4	5	6
Absolute Häufigkeit	21	60	111	63	27	18

- (6 Punkte) Zeichnen Sie die Verteilungsfunktion des Merkmals X (Anzahl der Abendkleider). Interpretieren Sie den Wert $F(3,5)$.
- (2 Punkte) In welchem Intervall liegen die mittleren 50 % der Daten?
- (2 Punkte) Bestimmen Sie das arithmetische Mittel.

Lösung:

a)



64 % der Damen haben höchstens 3,5 Abendkleider im Kleiderschrank.

- $[2; 4]$
- 3,23

2. a) (6 Punkte) Von 80 befragten Kindern werden sich 42 Kinder verkleiden und von 120 befragten Erwachsenen werden sich 63 Erwachsene verkleiden.

Welche beiden Merkmale werden hier untersucht? Welche Kenngröße ist geeignet, um einen Zusammenhang zwischen diesen beiden Merkmalen festzustellen? Berechnen Sie diese und interpretieren Sie Ihr Ergebnis.

- b) (6 Punkte) Von 7 Personen wurde das Alter sowie ihre Bereitschaft, sich zu verkleiden, erhoben. Die Bewertung für die Bereitschaft erfolgte nach der bekannten Notenskala (1 bedeutet sehr bereit, 5 bedeutet wenig bereit). Die folgende Tabelle gibt Auskunft über die erhobenen Daten:

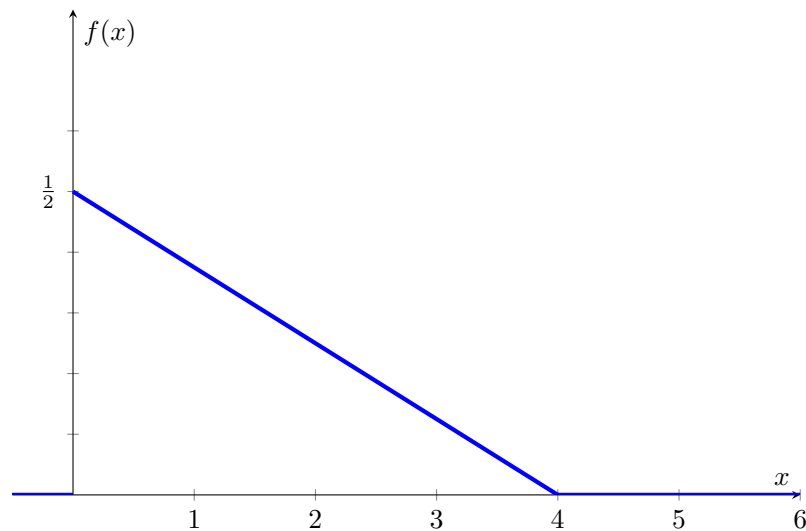
Person	A	B	C	D	E	F	G
Alter	52	19	36	24	89	12	42
Note für Bereitschaft	4	1	3	1	5	2	2

Welcher Koeffizient ist am besten geeignet, um einen Zusammenhang zwischen diesen beiden Merkmalen berechnen zu können? Berechnen Sie diesen und interpretieren Sie Ihr Ergebnis.

Lösung:

- a) korrigierter Kontingenzkoeffizient; 0; Ob sich jemand verkleidet oder nicht, ist unabhängig davon, ob man Kind oder erwachsen ist.
- b) Rangkorrelationskoeffizient, 0,786, Starker Zusammenhang, Je jünger sie sind, desto eher sind sie bereit, sich zu verkleiden.

3. Die Lebensdauer X eines elektronischen Bauteils (gemessen in Jahren) wird durch die folgende Dichtefunktion beschrieben:



- a) (2 Punkte) Zeigen Sie, dass $f(x)$ tatsächlich eine Dichtefunktion ist.
- b) (2 Punkte) Berechnen Sie die Wahrscheinlichkeit, dass ein zufällig ausgewähltes Bauteil länger als 3 Jahre hält.
- c) (3 Punkte) Der Hersteller fordert, dass höchstens 20 % der Bauteile vor Ablauf des ersten Jahres ausfallen. Überprüfen Sie, ob diese Anforderung erfüllt wird.
- d) (3 Punkte) Der Hersteller verbessert den Fertigungsprozess, sodass die Lebensdauer X des Bauteils nun im Intervall $[0; 5]$ liegt. Die Dichtefunktion behält ihre linear fallende Form bei. Zeichnen Sie die entsprechende neue Dichtekurve in obiger Grafik ein.

Lösung:

- a) Fläche unter der Dichtekurve ist 1.
- b) 0,0625
- c) 0,4375; Anforderung nicht erfüllt.
- d) Gerade verläuft durch $(0|0,4)$ und $(5|0)$

4. Die Verspätungszeit X von Flügen einer europäischen Airline kann näherungsweise als normalverteilt mit Erwartungswert $\mu = 15$ Minuten und Standardabweichung $\sigma = 10$ Minuten modelliert werden. Ein Flug gilt als „stark verspätet“, wenn er mehr als 30 Minuten zu spät ankommt. Es wird angenommen, dass die Verspätungen der Flüge unabhängig voneinander auftreten.
- a) (3 Punkte) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Flug stark verspätet ist?
 - b) (2 Punkte) Die Airline führt täglich 200 Flüge durch. Wie viele Flüge pro Tag sind im Mittel stark verspätet?
 - c) Die Airline hat außerdem statistisch ermittelt, dass bei stark verspäteten Flügen ein Gepäckstück mit Wahrscheinlichkeit von 2 % verloren geht. Bei Flügen mit Verspätung von höchstens 30 Minuten geht ein Gepäckstück nur mit Wahrscheinlichkeit von 0,3 % verloren.
 - i. (3 Punkte) Bestimmen Sie die Wahrscheinlichkeit, dass bei einem zufällig ausgewählten Flug ein Gepäckstück verloren geht.
 - ii. (2 Punkte) Ein Passagier stellt fest, dass sein Gepäck verloren gegangen ist. Wie groß ist die Wahrscheinlichkeit, dass sein Flug mehr als 30 Minuten verspätet war?

Lösung:

- a) 0,067
- b) 13,4
- c) i. 0,0041
ii. 0,3268

5. Im Rahmen einer Umfrage wurden 250 Personen gebeten, eine Selbsteinschätzung dahingehend abzugeben, ob sie einen Wiener Walzer tanzen können. 60 Personen gaben an, einen Wiener Walzer tanzen zu können.

- a) (6 Punkte) Bestimmen Sie ein zweiseitiges 95-Prozent-Konfidenzintervall für den Anteilswert p der Personen, die meinen, den Wiener Walzer tanzen zu können. Interpretieren Sie Ihr Ergebnis.
- b) (10 Punkte) Ein Tanzlehrer behauptet, dass weniger als 30 % der Österreicher den Wiener Walzer tanzen können. In einer Stichprobe von 8 Personen konnte nur eine Person den Wiener Walzer tanzen. Lässt sich damit die Behauptung des Tanzlehrers signifikant bestätigen? Testniveau $\alpha = 10 \%$.

Formulieren Sie die Hypothesen und führen Sie den entsprechenden Test durch. Wie ist demnach zu entscheiden?

Lösung:

- a) $[0,187; 0,293]$ Mit einer statistischen Sicherheit von 95 % kann darauf vertraut werden, dass der Anteil der Österreicher, die Wiener Walzer tanzen können, in diesem Intervall liegt.
- b) $H_0 : p \geq 0,3$ und $H_1 : p < 0,3$, $t_0 = 1$, $K = \{0\}$, $t_0 \notin K$, H_1 nicht bestätigt. Die Behauptung des Besitzers lässt sich nicht signifikant bestätigen.

6. Bei 6 Personen wurde das Körpergewicht (in kg) unmittelbar vor Weihnachten (X) und unmittelbar nach Weihnachten (Y) erhoben. Die Ergebnisse sind in der folgenden Tabelle dargestellt:

Person	1	2	3	4	5	6
X	78	75	90	52	93	79
Y	81	80	94	54	99	77

Es soll untersucht werden, ob das durchschnittliche Körpergewicht nach Weihnachten signifikant höher war. Die Differenzen der Gewichte können als normalverteilt angenommen werden.

- (2 Punkte) Welcher statistische Test ist geeignet? Begründen Sie Ihre Wahl.
- (10 Punkte) Formulieren Sie die Hypothesen und führen Sie den Test zu einem Niveau von $\alpha = 0,05$ durch. Interpretieren Sie das Ergebnis.

Lösung:

- t -Test für Differenzen, Test auf Lage, verbunden, normalverteilt
- $H_0 : \mu_y - \mu_x \leq 0$, $H_1 : \mu_y - \mu_x > 0$, $t_0 = 2,598$, $K =]2,015; \infty[$, H_1 bestätigt. Ja, die Weihnachtsfeiertage führten zu einer signifikanten Gewichtszunahme.

7. Eine Lehrveranstaltungsleiterin schätzt ein lineares Regressionsmodell zur Vorhersage von Klausurpunkten (Y) aus den erklärenden Variablen Lerndauer (in Stunden), Mitarbeit (Dummy-Variable: 0 = nein, 1 = ja) und Nachhilfekosten (in Euro). Die Ergebnisse sind in der folgenden Tabelle zusammengefasst:

Regressions-Statistik	
Multipler Korrelationskoeffizient	0,989
Bestimmtheitsmaß	0,978
Adjustiertes Bestimmtheitsmaß	0,945
Standardfehler	3,909
Beobachtungen	6

	Koeffizienten	Standardfehler	t-Statistik	p-Wert
Schnittpunkt	16,658	7,729	2,155	0,164
Lerndauer	2,463		7,248	0,019
Mitarbeit	21,670	4,436	4,885	0,039
Nachhilfekosten	-0,037	0,036	-1,019	0,415

- (3 Punkte) Geben Sie die geschätzte Regressionsgleichung an und interpretieren Sie den Koeffizienten der Variable Mitarbeit.
- (2 Punkte) Welche der erklärenden Variablen sind auf dem Niveau von $\alpha = 0,05$ statistisch signifikant? Interpretieren Sie dieses Ergebnis.
- (3 Punkte) Berechnen Sie den Standardfehler des geschätzten Koeffizienten der Variablen Lerndauer aus den angegebenen Werten.
- (2 Punkte) Welche Aussage lässt sich auf Basis dieser Regression über den Zusammenhang zwischen den Variablen Nachhilfekosten und Klausurpunkte treffen?

Lösung:

- $Y = 16,658 + 2,463L + 21,67M - 0,037N$; Ein Student erwartet in diesem Modell bei Mitarbeit 21,67 Klausurpunkte mehr als ohne Mitarbeit bei sonst gleichen Bedingungen.
- Lerndauer, Mitarbeit; Diese beiden Variablen sind in dem Modell wesentlich.
- 0,3398
- Die Nachhilfekosten sind, gegeben die anderen Variablen, in diesem Modell nicht signifikant und könnten aus dem Modell entfernt werden.

8. (10 Punkte)

a) Kreuzen Sie die beiden Aussagen an, die auf jeden Fall zutreffen.

Ein Streudiagramm dient zur Darstellung

- ☐ einer metrischen Variablen.
- ☐ von Verteilungen ordinaler Merkmale.
- ☐ des Zusammenhangs zwischen zwei quantitativen Merkmalen.
- ☐ des Zusammenhangs zwischen zwei nominalen Merkmalen.
- ☐ des Zusammenhangs zweier metrischer Variablen.

b) Welche der folgenden Merkmale sind dichotom?

Kreuzen Sie die beiden Aussagen an, die auf jeden Fall zutreffen.

- ☐ Anzahl der schlechten Teile einer Lieferung
- ☐ Prüfergebnis einer Gut/Schlecht-Prüfung
- ☐ Mitarbeiterzahl eines Betriebs
- ☐ Wurfresultat einer Laplace-Münze
- ☐ Haarfarbe

c) Welchen statistischen Fehler begeht man, wenn man fälschlicherweise die H_0 beibehält?

d) Ergänzen Sie die Textlücken im nachstehenden Satz durch Ankreuzen des jeweils zutreffenden Ausdrucks aus den nachstehenden Tabellen so, dass jedenfalls eine wahre Aussage entsteht.

Für eine diskrete Zufallsvariable X mit Verteilungsfunktion F gilt allgemein:

$\text{---} \textcircled{1} \text{---}$	$=$	$\text{---} \textcircled{2} \text{---}$
$\textcircled{1}$		$\textcircled{2}$
☐ $P(X < a)$		☐ $F(a)$
☐ $P(X = a)$		☐ $1 - F(a)$
☐ $P(X > a)$		☐ $F(a) - 1$

e) Markieren Sie, ob die folgende Aussage zutrifft oder nicht, und kennzeichnen Sie die jeweils passende Begründung.

Die Quadratsumme der Residuen ist bei der linearen Kleinsten-Quadrate-Regression typischerweise 0.

☐ Richtig

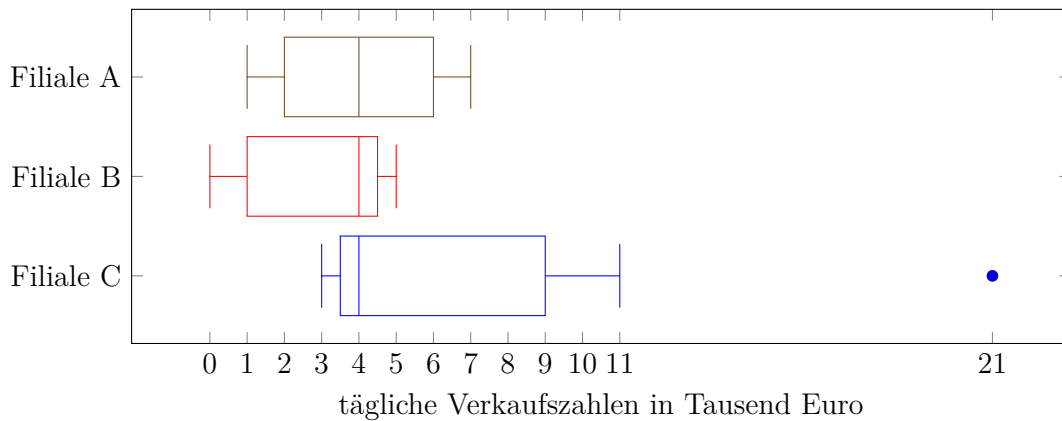
☐ Falsch

Begründung:

- ☐ Da die Residuen positive und negative Vorzeichen besitzen, heben sie sich immer auf.
- ☐ Im Allgemeinen ist die Quadratsumme der Residuen größer als 0, weil es fast immer Abweichungen zwischen beobachteten und vorhergesagten Werten gibt.
- ☐ Die Quadratsumme ist immer gleich der Anzahl der Beobachtungen minus 1.

Lösung: (3, 5); (2, 4); Fehler 2. Art; (3, 2); (F, 2)

9. (10 Punkte) In der folgenden Grafik sind drei Boxplots dargestellt. Jeder Boxplot zeigt die täglichen Verkaufszahlen einer Filiale, angegeben in Tausend Euro.



Beantworten Sie die folgenden Fragen:

- a) Für welche Filiale sind der Median und das arithmetische Mittel ungefähr gleich groß?

☐ Filiale A ☐ Filiale B ☐ Filiale C ☐ keine Filiale

- b) Welche Filiale hat den größten Mittelwert?

☐ Filiale A ☐ Filiale B ☐ Filiale C

Welche Filiale hat den kleinsten Mittelwert?

☐ Filiale A ☐ Filiale B ☐ Filiale C

- c) Wie lautet der zweitgrößte Tagesumsatz für Filiale C?

- d) Begründen Sie, warum die Daten von Filiale B nicht normalverteilt sind.

- e) Fügen Sie in die beiden Textlücken des folgenden Satzes jeweils eine Zahl ein, sodass eine wahre Aussage entsteht.

Damit ein Datenpunkt bei Filiale A ein potentieller Ausreißer nach oben ist, müsste dieser zwischen und liegen.

Lösung: A; (C, B); 11; Daten nicht symmetrisch; (12, 18)