

Statistik Vorlesung

22. Mai 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	13		
2	09		
3	14		
4	12		
5	11		
6	14		
7	07		
8	10		
9	10		
Summe	100		
Note			

1. Zu Jahresbeginn 2026 war das Alter der Bevölkerung in Österreich (in Tausend Personen) etwa wie in folgender Tabelle angegeben verteilt.

Alter (in Jahren)	[0; 15[	[15; 30[	[30; 60[	[60; 100]
Personen (in Tausend)	1.304	1.643	3.680	2.592

Berechnen Sie näherungsweise

- a) (5 Punkte) das Medianalter der Bevölkerung und
- b) (3 Punkte) das Durchschnittsalter (arithmetisches Mittel),
- c) (5 Punkte) sowie den Variationskoeffizienten.

Lösung:

- a) 43,553
- b) 45,526
- c) 0,553

2. Ein Unternehmen sendet seine Vertreter zu den Kunden. Die Wegstrecken sind der folgenden Tabelle zu entnehmen.

von	nach	Strecke (in km)
A	B	90
A	C	30
C	D	40
D	A	20

Berechnen Sie jeweils die Durchschnittsgeschwindigkeit der Vertreter in den folgenden Fällen.
Hinweis: Geschwindigkeit = Wegstrecke / Zeit.

- a) (3 Punkte) Vertreter V_1 fährt von A nach B mit durchschnittlich 60 km/h und die gleiche Strecke zurück mit 80 km/h.
- b) (4 Punkte) Vertreter V_2 fährt von A nach C mit durchschnittlich 40 km/h, von C nach D mit 55 km/h und wieder zurück zu A mit 65 km/h.
- c) (2 Punkte) Vertreter V_3 fährt durchschnittlich mit 70 km/h von A nach E , mit 65 km/h von E nach F und zurück nach A mit 45 km/h. Für die erste Teilstrecke (AE) benötigte er eine Stunde, für die zweite Teilstrecke (EF) 1,5 Stunden und für die dritte (FA) eine halbe Stunde.

Lösung:

- a) 68,57
- b) 50,421
- c) 63,333

3. Ein Spargelhof verkauft weißen und grünen Spargel. Die Wahrscheinlichkeit, dass ein Bund weiß ist, beträgt p . Die Wahrscheinlichkeit, dass ein Bund holzig ist, beträgt insgesamt 18 %. Weiters gilt folgendes:

$$P(\text{holzig} \mid \text{weiß}) = 0,30,$$

$$P(\text{holzig} \mid \text{grün}) = 0,08.$$

Ein Kunde wählt zufällig einen Bund aus.

- a) (4 Punkte) Bestimmen Sie p , also die Wahrscheinlichkeit für weißen Spargel.
- b) (3 Punkte) Wie groß ist die Wahrscheinlichkeit, dass ein Bund grün ist, wenn er nicht holzig ist?
- c) (3 Punkte) Wie groß ist die Wahrscheinlichkeit, einen grünen und holzigen Bund auszuwählen?
- d) (4 Punkte) Ein Kunde kauft 5 Bund Spargel. Wie groß ist die Wahrscheinlichkeit, dass mindestens einer holzig ist?

Lösung:

- a) 0,455
- b) 0,612
- c) 0,044
- d) 0,629

4. Erdbeeren werden in 500 g-Becher abgefüllt. Das Füllgewicht sei normalverteilt mit einem Erwartungswert von $\mu = 515$ g und einer Standardabweichung von $\sigma = 20$ g.

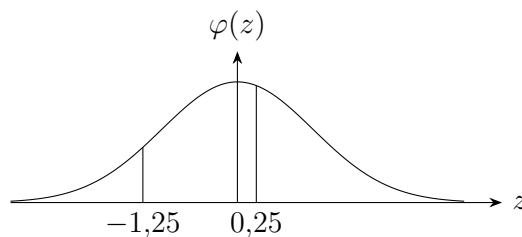
- a) (3 Punkte) Wie groß ist die Wahrscheinlichkeit, dass ein Becher zwischen 490 g und 520 g wiegt?
- b) (2 Punkte) Skizzieren Sie dazu die Dichtefunktion einer standardnormalverteilten Zufallsgröße und kennzeichnen Sie darin die standardisierten Intervallgrenzen aus Punkt a).
- c) (2 Punkte) Bestimmen Sie den Wert der Verteilungsfunktion an der Stelle 515 und interpretieren Sie diesen.
- d) (3 Punkte) Bestimmen Sie die Zahl d derart, dass

$$P(X > d) = 0,9.$$

- e) (2 Punkte) Die Becher werden unabhängig voneinander in Kisten zu je 12 Stück verpackt. Es wird das durchschnittliche Füllgewicht eines Bechers je Kiste berechnet. Wie weit weicht dieses durchschnittliche Füllgewicht vom erwarteten Füllgewicht eines Bechers ($\mu = 515$) im Mittel ab?

Lösung:

- a) 0,493
- b)



- c) 0,5; Erwartungswert
- d) 489,369
- e) 5,774

5. Im Rahmen einer Qualitätsüberprüfung wird die Abfüllmenge einer automatisierten Getränkeanlage untersucht. Aus früheren Analysen ist bekannt, dass die Standardabweichung der Abfüllmenge $\sigma = 6$ ml beträgt. Es soll ein 95%-Konfidenzintervall für die durchschnittliche Füllmenge der abgefüllten Flaschen bestimmt werden.
- a) (5 Punkte) Wie viele Flaschen müssen überprüft werden, damit die Länge des zweiseitigen 95%-Konfidenzintervalls höchstens 4 ml beträgt?
 - b) (2 Punkte) Wie verändert sich der erforderliche Stichprobenumfang, wenn statt eines 95%-ein 99%-Konfidenzintervall gleicher Länge bestimmt werden soll? Beschreiben Sie die Veränderung verbal, eine Rechnung ist nicht erforderlich.
 - c) (2 Punkte) Hat der Stichprobenmittelwert $\bar{x}$ Einfluss auf die Länge des Konfidenzintervalls? Begründen Sie.
 - d) (2 Punkte) Durch eine Verbesserung der Abfüllanlage wird die Standardabweichung auf $\sigma = 3$ ml gesenkt.
Welche Auswirkungen hat dies auf die Länge eines Konfidenzintervalls bei vorgegebenem Stichprobenumfang?
Welche Auswirkungen hat dies bei vorgegebener Intervalllänge auf den erforderlichen Stichprobenumfang?
Erläutern Sie die Zusammenhänge verbal.

Lösung:

- a) 35
- b) Es ist ein größerer Stichprobenumfang n erforderlich.
- c) Nein. Länge ist nur abhängig vom Abweichungsterm.
- d) Bei vorgegebenem n halbiert sich die Länge des Konfidenzintervalls.
Für eine vorgegebene Intervalllänge genügt ein kleinerer Stichprobenumfang n .

6. Es soll die Ausdauerleistung untersucht werden. Dazu absolvieren 8 Teilnehmer sowohl vor als auch nach einer vierwöchigen Trainingsphase einen Ausdauerstest, bei dem die Laufzeit für eine festgelegte Strecke gemessen wird. Für jede Person wurde die Differenz (in Minuten) zwischen der Laufzeit nach dem Training (Y) und der Laufzeit vor dem Training (X) berechnet ($d_i = y_i - x_i$):

Teilnehmer-ID	1	2	3	4	5	6	7	8
Differenz (in Minuten)	-2	-1	-3	1	-1	0	-2	0

Lässt sich zeigen, dass das neue Trainingsprogramm die Laufzeit signifikant verkürzt?

Die Differenzen können als Realisationen einer normalverteilten Zufallsgröße angenommen werden.

- (2 Punkte) Welcher Test ist anzuwenden? Begründen Sie Ihre Entscheidung.
- (2 Punkte) Formulieren Sie die Hypothesen.
- (6 Punkte) Führen Sie den Test zum Signifikanzniveau von $\alpha = 0,05$ durch.
- (2 Punkte) Wie lautet Ihre Testentscheidung und was bedeutet diese im gegebenen Kontext?
- (2 Punkte) Würde sich Ihre Testentscheidung ändern, wenn statt eines einseitigen ein zweiseitiger Test zum Signifikanzniveau $\alpha = 0,05$ durchgeführt werden würde? Begründen Sie Ihre Antwort.

Lösung:

- t-Test für Differenzen; Begründung: Test auf Lage, Zweistichprobentest, Normalverteilung, verbundene Stichproben.
- $H_0 : \mu_y - \mu_x \geq 0$
 $H_1 : \mu_y - \mu_x < 0$
- $t_0 = -2,16$; $K =] - \infty; -1,895[$
- H_1 bestätigt; Es konnte gezeigt werden (zum Signifikanzniveau $\alpha = 0,05$), dass das Trainingsprogramm die Ausdauerleistung verbessert.
- Ja, man würde die H_0 beibehalten, denn dann wäre $t_0 \notin K =] - \infty; -2,365 \cup] 2,365; \infty[$.

7. Die Spiel & Sport AG möchte den Umsatz (in 1.000 Euro) für eine neue Filiale schätzen und führt eine Regressionsanalyse durch. Als erklärende Merkmale werden herangezogen: das Anlagevermögen (AV) für Anlagen wie Maschinen und Gebäude (in 1.000 Euro), die Anzahl der Mitarbeiter (Mitarbeiter), der Standort (Standort) mit den Ausprägungen West, Nord und Süd, sowie die Artikelgruppe (Sparte) mit den Ausprägungen Spiel und Sport.

	Koeffizienten	Standardfehler	t-Statistik	p-Wert
Achsenabschnitt	1429,74	452,89	3,16	0,002
AV	5,68	1,99	2,85	0,01
Mitarbeiter	7	0,11	63,64	<0,001
Standort Nord	1559,38	397,11	3,93	<0,001
Standort Süd	-1310,25	561,57	-2,33	0,02
Sparte Sport	814,06	45,93	17,72	<0,001

- a) (4 Punkte) Wie hoch ist der erwartete Umsatz einer neuen Filiale mit einem Anlagevermögen von 300.000 Euro, mit 50 Mitarbeitern am Standort West in der Sparte Sport?
- b) (3 Punkte) (unabhängig von a)) Wie ändert sich der erwartete Umsatz, wenn eine Filiale geplant wird, die bei sonst gleichen Bedingungen 20 Mitarbeiter mehr beschäftigt und der Standort Nord statt Süd ist?

Lösung:

- a) 4.297.800 Euro
- b) 3.009.630 Euro

8. (10 Punkte)

Im Rahmen einer Studie wird eine Stichprobe von insgesamt $n = 100$ Filialen einer internationalen Modekette betrachtet. Untersucht werden zentrale betriebswirtschaftliche Kennzahlen wie Umsatz, Marketingbudget, Mitarbeiteranzahl, Kundenzufriedenheit (gemessen im Schulnotensystem von 1 = sehr gut bis 5 = nicht genügend) sowie die Geschäftsfläche der einzelnen Standorte. Ziel ist es, Zusammenhänge zwischen diesen Variablen zu analysieren und deren Einfluss auf den Unternehmenserfolg zu bewerten.

Aus Gründen der Übersichtlichkeit wird im Folgenden lediglich der Beginn des Datensatzes dargestellt.

Filiale	Umsatz (Tsd. €)	Marketingbudget (Tsd. €)	Mitarbeiter	Kundenzufriedenheit (1–5)	Fläche (m ²)
A	250	20	10	3	800
B	300	25	12	2	950
C	280	18	11	4	900
D	200	15	8	5	700
E	350	30	14	1	1.100
⋮	⋮	⋮	⋮	⋮	⋮

- a) Im Bericht zur Studie wird das arithmetische Mittel der Geschäftsfläche mit 500 m² angegeben. Später stellt sich jedoch heraus, dass bei Filiale E ein Datenfehler vorliegt: Die Filiale ist nur 110 m² statt der angegebenen 1.100 m² groß. Wie groß ist das korrigierte arithmetische Mittel der Geschäftsfläche?

- b) Welches Zusammenhangsmaß eignet sich zur Untersuchung des Zusammenhangs zwischen dem Umsatz und der Kundenzufriedenheit? Begründen Sie Ihre Wahl kurz.

- c) Im Rahmen der Studie kommt es aufgrund eines Softwarefehlers dazu, dass zufällig eine der 100 Beobachtungen im Datensatz verloren geht. Jede Beobachtung fällt dabei mit gleicher Wahrscheinlichkeit aus. Wie groß ist die Wahrscheinlichkeit, dass sich dadurch die Spannweite des Merkmals Umsatz nicht verändert, wenn keine zwei Filialen im gesamten Datensatz einen gleich hohen Umsatz aufweisen?

- d) Beim Jarque-Bera-Test für das Merkmal Marketingbudget ergibt sich ein p-Wert von 0,02. Welche Schlussfolgerung lässt sich aus diesem Ergebnis über die Verteilung des Marketingbudgets ziehen, wenn das Signifikanzniveau bei 0,05 liegt?

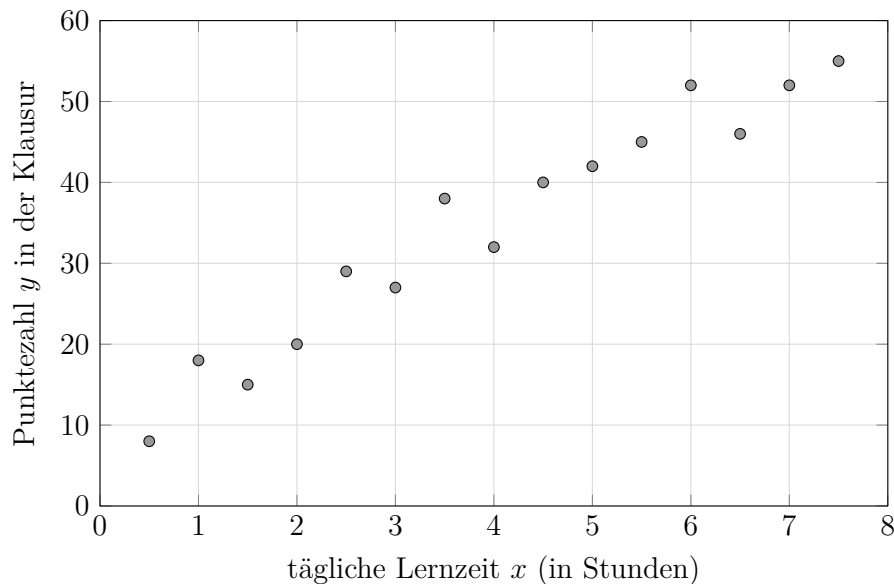
- e) Zur Analyse des Zusammenhangs zwischen dem Marketingbudget und dem Umsatz wird eine lineare Regression durchgeführt. Dabei ergibt sich die Regressionsgerade

$$y = 120 + 7x$$

wobei x das Marketingbudget (in Tsd. €) und y den Umsatz (in Tsd. €) beschreibt. Bestimmen Sie das Residuum dieser Regressionsanalyse für die Filiale A.

Lösung: 490,1; Spearman-Rangkorrelation, weil ordinales und metrisches Merkmal; 0,98; H_0 verwerfen, keine Normalverteilung; -10

9. (10 Punkte) In einer Studie wurde untersucht, ob ein Zusammenhang zwischen der täglichen Lernzeit für Wirtschaftsmathematik (in Stunden) und der erreichten Punktezahl in einer Wirtschaftsmathematik-Klausur besteht. Für 15 Studierende wurden die folgenden Paare (Lernzeit x , Punktezahl y) erhoben und in einem Streudiagramm dargestellt:



- a) Kreuzen Sie die beiden Aussagen an, die durch das Streudiagramm am besten gestützt werden.
- ☐ Es liegt ein negativer linearer Zusammenhang zwischen Lernzeit und Punktezahl vor.
 - ☐ Studierende mit höherer täglicher Lernzeit erzielen tendenziell mehr Punkte in der Klausur.
 - ☐ Zwischen Lernzeit und Punktezahl besteht praktisch kein Zusammenhang.
 - ☐ Für Lernzeiten zwischen 2 und 5 Stunden liegen die Punkte überwiegend zwischen 20 und 42.
- b) Aus dem Datensatz wurde berechnet $\bar{x} = 4$, $\bar{y} = 34,6$. Weiters wurde der Parameter $\hat{\beta}_0 = 9$ nach der Methode der kleinsten Quadrate geschätzt. Zeichnen Sie die geschätzte Regressionsgerade in das obige Diagramm ein.
- c) Für die geschätzte Regressionsgerade wurde ein Bestimmtheitsmaß von 0,946 ermittelt. Berechnen Sie den Pearson'schen Korrelationskoeffizienten.
- d) Wie viele der im Streudiagramm dargestellten Studierenden haben mindestens 40 Punkte in der Klausur erreicht?
- e) Beurteilen Sie anhand des Streudiagramms, ob es Beobachtungen gibt, die deutlich vom übrigen Punktmuster abweichen (Ausreißer). Kreuzen Sie die Antwort an, die das Streudiagramm am besten beschreibt.
- ☐ Es gibt mindestens einen offensichtlichen Ausreißer.
 - ☐ Es gibt keine Beobachtung, die deutlich als Ausreißer hervortritt.
 - ☐ Das lässt sich aus diesem Streudiagramm nicht erkennen.
 - ☐ Es gibt mehrere Ausreißer, die vor allem bei sehr geringer Lernzeit besonders hohe Punktzahlen erreichen.

Lösung: (2, 4); $y = 9 + 6,4x$; 0,973; 7; 2