

Statistik Vorlesung

17. September 2026

Arbeitszeit: 100 Minuten

VORNAME:		MATR.NR.:	
NACHNAME:			

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

VERBOTEN: **alle sonstigen Unterlagen, Handys**

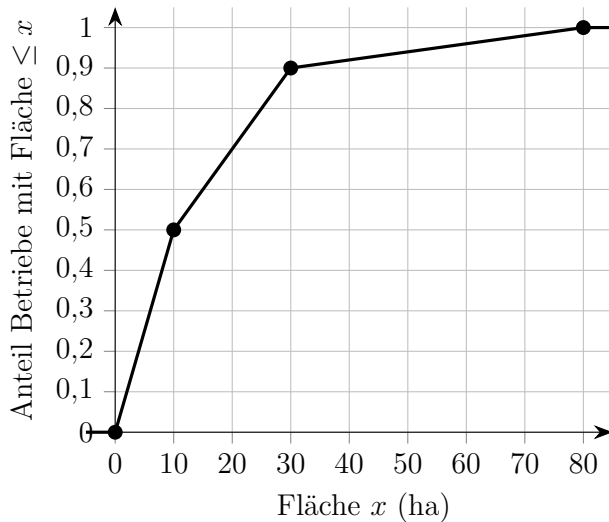
Aufgaben 8 und 9: Für jede vollständig richtig gelöste Frage gibt es 2 Punkte. Es gibt keine Teilpunkte und keine Minuspunkte.

Lösungswege müssen nachvollziehbar aufgeschrieben sein.

Aufgabe	max. Punkte	erreichte Punkte	Teilsumme
1	14		
2	08		
3	12		
4	11		
5	11		
6	14		
7	10		
8	10		
9	10		
Summe	100		
Note			

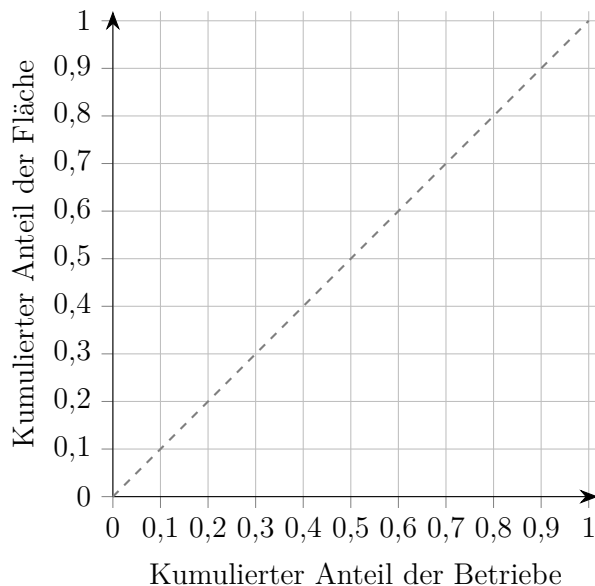
1. In einer Region gibt es insgesamt 1.000 landwirtschaftliche Betriebe. Diese werden anhand der Größe ihrer bewirtschafteten Fläche in drei Klassen eingeteilt: Kleinbetriebe (0 bis unter 10 ha), Mittelbetriebe (10 bis unter 30 ha) und Großbetriebe (30–80 ha).

- a) (3 Punkte) Ermitteln Sie basierend auf der nachfolgenden Grafik die Anzahl der Betriebe je Größenklasse sowie die relativen und kumulierten relativen Anteile. Tragen Sie alle Werte in die nachstehende Tabelle ein.



Betriebsgröße	Betriebe		
	Anzahl	rel. Anteil	kum. rel. Anteil
klein			
mittel			
groß			
Summe			—

- b) (3 Punkte) Bestimmen Sie näherungsweise die mittlere Fläche, die ein landwirtschaftlicher Betrieb bewirtschaftet.
- c) (8 Punkte) Von insgesamt 10.000 ha bewirtschafteter Fläche entfallen 20 % auf Kleinbetriebe sowie jeweils 40 % auf Mittel- und Großbetriebe. Zeichnen Sie auf Grundlage dieser Angaben und Ihrer Ergebnisse aus a) eine Lorenzkurve in das nachstehende Diagramm und berechnen Sie den Gini-Koeffizienten.



Lösung:

- a) klein: 500; 0,5; 0,5
mittel: 400; 0,4; 0,9
groß: 100; 0,1; 1
- b) $\bar{x} = 16$
- c) Punkte der Lorenzkurve: (0|0), (0,5|0,2), (0,9|0,6), (1|1); Gini: 0,42

2. Ein Online-Shop möchte untersuchen, ob die Lieferdauer mit der Kundenzufriedenheit zusammenhängt. Dazu werden bei 8 zufällig ausgewählten Bestellungen, die Lieferdauer und die Kundenzufriedenheit erfasst. Die Kundenzufriedenheit wird auf einer Skala von 1 bis 10 angegeben, wobei 1 für „sehr unzufrieden“ und 10 für „sehr zufrieden“ steht.

Bestellung	A	B	C	D	E	F	G	H
Lieferdauer (in Tagen)	2	4	1	5	3	7	6	8
Kundenzufriedenheit	10	6	7	5	9	3	8	4

- (2 Punkte) Welche Kennzahl ist geeignet, um den Zusammenhang zwischen Lieferdauer und Kundenzufriedenheit zu untersuchen? Begründen Sie Ihre Wahl.
- (2 Punkte) Die Berechnung der gesuchten Kennzahl ergibt $-0,714$. Interpretieren Sie dieses Ergebnis im Kontext der Aufgabenstellung.
- (2 Punkte) Der Online-Shop stellt seine Zufriedenheitsskala um: Statt einer Skala von 1 bis 10 wird nun eine Skala von 0 bis 20 verwendet, wobei die Reihenfolge der Kundenzufriedenheiten unverändert bleibt. Ändert sich dadurch der Wert des in Teil a) verwendete Zusammenhangsmaßes zwischen Lieferdauer und Kundenzufriedenheit? Begründen Sie Ihre Antwort.
- (2 Punkte) Es liegt eine neue Auswertung basierend auf 70 Bestellungen vor. Die Versandart wird dabei in „Express“ und „Standard“ unterschieden, die Zufriedenheit in „zufrieden“ und „unzufrieden“. Wie müssten die Einträge in der Tabelle lauten, wenn kein Zusammenhang zwischen Versandart und Zufriedenheit besteht (d.h. wenn beide Merkmale stochastisch unabhängig sind)? Vervollständigen Sie die Tabelle entsprechend.

	Zufrieden	Unzufrieden	Gesamt
Express			30
Standard			40
Gesamt	42	28	70

Lösung:

- Rangkorrelationskoeffizient. Lieferdauer ist ein kardinalskaliertes, Kundenzufriedenheit ein ordinalskaliertes Merkmal, und man orientiert sich am niedrigeren Skalenniveau.
- Es besteht ein mittlerer/starker negativer Zusammenhang zwischen der Lieferdauer und der Kundenzufriedenheit. Tendenziell geht also eine längere Lieferdauer mit einer geringeren Kundenzufriedenheit einher.
- Nein. Der Spearman-Rangkorrelationskoeffizient basiert ausschließlich auf den Rängen der Beobachtungen. Ändert sich die Rangfolge der Beobachtungen nicht, bleiben die Ränge und somit auch der Rangkorrelationskoeffizient unverändert.
- 18, 12
24, 16

3. a) Ein angehender Kunstdieb plant, ein Kunstwerk zu stehlen, und berechnet die Erfolgsaussichten seines Vorhabens. Er berücksichtigt drei Ereignisse und schätzt deren Wahrscheinlichkeiten folgenderweise ein:

- Die Wahrscheinlichkeit, mit Hammer und Pistole das Museum betreten zu können, weil seine Tasche nicht kontrolliert wird, schätzt er auf 80 Prozent.
- Die Wahrscheinlichkeit, dass die Vitrine aus Panzerglas besteht und mit seinem Hammer nicht in wenigen Minuten zerstört werden kann, schätzt er auf 10 Prozent.
- Die Wahrscheinlichkeit, dass ihn die Security laufen lässt, wenn er seine Pistole zeigt, hält er für 70 %.

Gehen Sie davon aus, dass die drei Ereignisse voneinander unabhängig sind.

- i. (3 Punkte) Wie groß ist die Wahrscheinlichkeit, den Diebstahl erfolgreich durchführen zu können?
 - ii. (3 Punkte) Er plant vier Kunstraube in jeweils unterschiedlichen Museen, bevor er sich zur Ruhe setzt. Er geht von unveränderten Wahrscheinlichkeiten aus. Wie groß ist dann die Wahrscheinlichkeit, dass er bei mindestens einem der vier Diebstähle scheitert?
- b) (Unabhängig von a)) Ein Nachahmungstäter hat andere Einschätzungen vom Verlauf eines Diebstahlversuchs und möchte den erwarteten Gewinn berechnen. Er glaubt, dass er zu 60 % nicht mit Hammer und Pistole ins Museum gelangen kann. In diesem Fall geht er erfolglos nach Hause. Dass er in das Gebäude gelangt und beim Zerschlagen der Vitrine gefasst würde, kalkuliert er mit 10 %. Die in diesem Fall zu erwartende Haftstrafe stellt für ihn einen Verlust dar, den er mit 1.000.000 bewertet. Bei Erfolg des Diebstahls beträgt der Gewinn aus dem Verkauf der Beute 40.000.000.
- i. (3 Punkte) Berechnen Sie den Erwartungswert seines Gewinns.
 - ii. (3 Punkte) Berechnen Sie den Erwartungswert seines Gewinns, wenn er es bereits mit Hammer und Pistole ins Museum geschafft hat.

Lösung:

- a) i. 0,504
- ii. 0,935
- b) i. 11.900.000
- ii. 29.750.000

4. Die Kosten für den Produktionsprozess eines Produkts sind normalverteilt mit Erwartungswert 14. Das 2,5%-Quantil beträgt 8,12.
- a) (3 Punkte) Wie groß ist die Standardabweichung?
 - b) (3 Punkte) Mit welcher Wahrscheinlichkeit entstehen Kosten über 17?
 - c) (5 Punkte) Berechnen Sie jenes symmetrisch um den Erwartungswert liegende Intervall, in dem 50 Prozent der Beobachtungen liegen werden.

Lösung:

- a) 3
- b) 0,159
- c) [11,977; 16,023]

5. a) (5 Punkte) In einer Befragung von 320 Personen in Österreich gaben 96 Personen an, nicht zu wissen, dass die Zeitumstellung auf Winterzeit Ende Oktober stattfindet. Berechnen Sie ein zweiseitiges 95-%-Konfidenzintervall für den Anteil der Personen, die das nicht wissen, und interpretieren Sie dieses!
- b) (4 Punkte) Für eine weitere Untersuchung zur Zeitumstellung soll ein 95-%-Konfidenzintervall mit einer gewünschten Länge von 0,05 berechnet werden. Es liegen keine Informationen über den Stichprobenanteil vor. Welchen Stichprobenumfang würden Sie empfehlen?
- c) (2 Punkte) Angenommen, aus früheren Umfragen ist bekannt, dass der wahre Anteil p zwischen 0,30 und 0,40 liegt. Welchen Wert für p sollte man im ungünstigsten Fall bei der Planung des Stichprobenumfangs verwenden? Begründen Sie.

Lösung:

- a) $[0,25; 0,35]$ Der Anteil der Personen, die nicht wissen, dass die Umstellung Ende Oktober stattfindet, liegt mit einer Sicherheit von 95 % im Intervall $[0,25; 0,35]$.
- b) 1.537
- c) 0,4; Das Produkt $\hat{p}(1 - \hat{p})$ ist größer für $\hat{p} = 0,4$ als für $0,3 \leq \hat{p} < 0,4$. Daher ist 0,4 als der ungünstigste Fall in der Planung de Stichprobengröße zu verwenden.

6. Immer wieder ist davon zu lesen, dass besonders Personen in Städten mehr Geld für Herbstdekoration (Kürbisse, Lichter, Deko) ausgeben als Personen am Land. Um zu überprüfen, ob Stadtbewohner tatsächlich mehr ausgeben als Landbewohner, wurden jeweils 7 Personen dazu befragt. Die Ausgaben in Euro sind:

X: Stadt (in Euro)	58	92	47	81	69	105	76
Y: Land (in Euro)	35	54	49	41	63	57	38

Es ist bekannt, dass die Ausgaben für Herbstdekoration nicht als Realisation einer normalverteilten Zufallsgröße angesehen werden können.

Lässt sich zeigen, dass Stadtbewohner mehr für Herbstdekoration ausgeben als Landbewohner?

- (3 Punkte) Welcher Test ist anzuwenden? Begründen Sie Ihre Entscheidung.
- (6 Punkte) Formulieren Sie die Hypothesen dieses Tests und berechnen Sie den Testwert.
- (3 Punkte) Bestimmen Sie den p -Wert und interpretieren Sie diesen!
- (2 Punkte) Wie ist zu einem Testniveau von 5 % zu entscheiden? Ändert sich Ihre Antwort, wenn das Testniveau 1 % beträgt?

Lösung:

- Wilcoxon-Rangsummentest; Lage, zwei Stichproben, unverbunden, nicht parametrisch
- $H_0 : c \geq 0$; $H_1 : c < 0$; $t_0 = 2,492$
- 0,006; Wahrscheinlichkeit, einen solchen Testwert oder einen extremeren zu erhalten, wenn die Nullhypothese stimmt.
- Ja, es lässt sich zeigen; Nein

7. Ein pharmazeutisches Unternehmen führt für die Zulassung eines neuen Wirkstoffes eine Studie mit 10 Personen zur zeitlichen Änderung der Wirkstoffkonzentration im Blut durch. Bei jeder Person wird genau eine Blutprobe entnommen, und zwar zu einem individuell festgelegten Zeitpunkt nach der Einnahme. Die ermittelten Daten enthalten zu jeder Person die seit Einnahme vergangene Zeit X (in Stunden), die Konzentration des Wirkstoffs im Blut Y (in $\mu\text{g/mL}$), und das Geschlecht der Person (Ausprägungen M = männlich, W = weiblich). Das Ergebnis einer durchgeführten Regressionsanalyse lautet

$$\hat{Y} = 2,68 - 0,06X - 0,3W$$

- a) (2 Punkte) Wie hoch ist die geschätzte mittlere Konzentration $\hat{Y}$ nach 2 Stunden bei einer männlichen Person?
- b) (3 Punkte) Nach wie viel Stunden haben Männer durchschnittlich die gleiche Konzentration wie Frauen nach 2 Stunden?
- c) (5 Punkte) Die Standardabweichung des Parameters für die Zeit ist 0.03. Ist dieser Parameter auf dem 5%-Niveau signifikant? Führen Sie einen entsprechenden Test durch.

Lösung:

- a) 2,56
- b) 7
- c) $H_0 : \beta_1 = 0, H_1 : \beta_1 \neq 0; t_0 = -2; M t_{7;0,975} = 2,365; K =] - \infty; -2,365[\cup] 2,365; \infty[$
 $t_0 \notin K; H_0$ nicht bestätigt. Parameter nicht signifikant.

8. (10 Punkte)

- a) Welche der folgenden Kennzahlen bleiben gleich, wenn jeder Wert eines Datensatzes um 5 erhöht wird? Kreuzen Sie die beiden zutreffenden Antworten an.

- ☐ Spannweite
☐ arithmetisches Mittel
☐ Median
☐ Modus
☐ Standardabweichung

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

Ein metrisch skaliertes Merkmal kann durch geeignete Transformation auf ein nominales, nicht aber auf ein ordinales Skalenniveau überführt werden.

☐ Richtig
 ☐ Falsch

Begründung:

- ☐ Eine metrische Skalierung kann in keine andere Skalierung überführt werden.
☐ Die Ordnung der Skalen von niedrig nach hoch ist nominal $\rightarrow$ metrisch $\rightarrow$ ordinal.
☐ Ein metrisch skaliertes Merkmal kann durch geeignete Transformation auf jedes niedrigere Skalenniveau überführt werden.

- c) Ein Forscher senkt das Signifikanzniveau α eines Hypothesentests von 5 % auf 1 %, ohne den Stichprobenumfang oder das Testverfahren zu verändern. Wie ändert sich dadurch das Risiko, einen Fehler 2. Art (Beta-Fehler) zu begehen?

Das Risiko .

- 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.

Der ____①____ misst die Stärke und Richtung des linearen Zusammenhangs zwischen zwei metrischen Variablen und nimmt Werte im Intervall ____②____ an.

①	②
☐ Kontingenzkoeffizient	☐ $[0; 1]$
☐ Pearson Korrelationskoeffizient	☐ $] - \infty; \infty[$
☐ Variationskoeffizient	☐ $[-1; 1]$

- e) Ein 95%-Konfidenzintervall für den Erwartungswert der Füllmenge einer Abfüllmaschine lautet $[498; 505]$. Es wird ein zweiseitiger Test der Hypothese

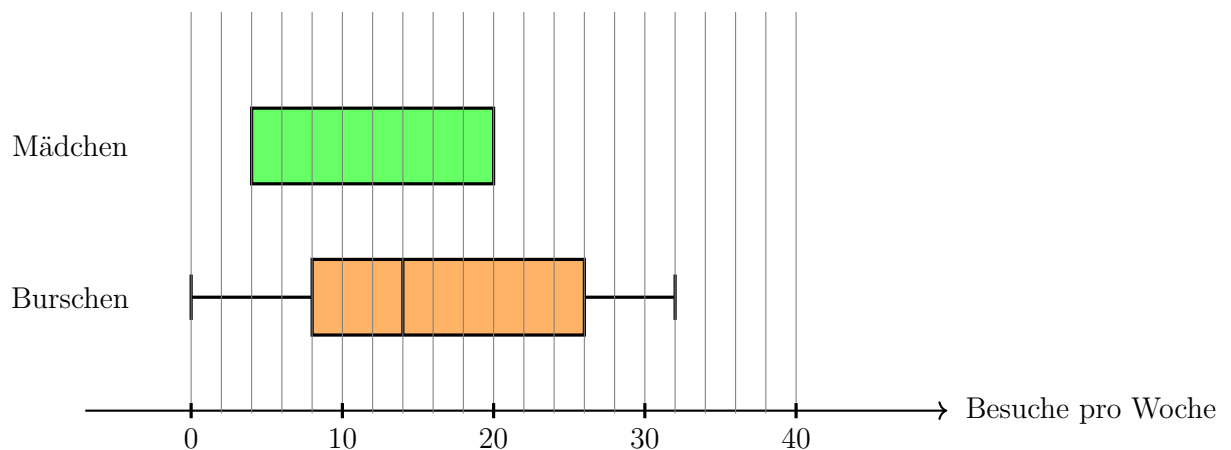
$H_0 : \mu = 500$ gegen $H_1 : \mu \neq 500$ auf dem 5%-Niveau durchgeführt. Dann lautet die Testentscheidung: Bei diesem Test wird die Nullhypothese .

Lösung: (1, 5); (F, 3); steigt; (2, 3); beibehalten

9. (10 Punkte) Die Nutzung einer bestimmten Internetplattform durch Jugendliche wird getrennt nach Mädchen und Burschen untersucht. Dabei wird erhoben, wie häufig die befragten Jugendlichen diese Plattform pro Woche besuchen. Die folgenden Boxplots stellen die Ergebnisse der Untersuchung für beide Gruppen dar. Ein Boxplot ist unvollständig und soll ergänzt werden. Für

die Gruppe der Mädchen sind folgende Daten bekannt:

Maximum	32
Spannweite	32
Median	12



- Vervollständigen Sie den Boxplot der Mädchen.
- Welche der folgenden Maßzahlen kann aus dem Boxplot der Burschen abgelesen werden. Kreuzen Sie die richtige Antwort an:
 - ☐ 5%-Quantil
 - ☐ Mittelwert
 - ☐ Standardabweichung
 - ☐ Drittes Quartil
- Aus dem Boxplot für die Burschen kann Folgendes abgelesen werden: Mindestens ein Viertel der Burschen hat eine Anzahl an Besuchen pro Woche kleiner oder gleich einem Wert a und zugleich haben mindestens drei Viertel der Burschen eine Anzahl an Besuchen größer als oder gleich diesem Wert a . Geben Sie diesen Wert a an: .
- Kreuzen Sie die beiden zutreffenden Aussagen an.
 - ☐ Der Median der Anzahl von Besuchen pro Woche ist bei den Burschen höher als bei den Mädchen.
 - ☐ Die Spannweite der wöchentlichen Nutzung der Plattform ist bei den Burschen größer als bei den Mädchen.
 - ☐ Aus der Grafik kann man ablesen, dass genauso viele Mädchen wie Burschen die Plattform wöchentlich besuchen.
 - ☐ Der Anteil der Burschen, die mehr als 20-mal pro Woche die Plattform nützen, ist zumindest gleich groß oder größer als jener der Mädchen.
 - ☐ Etwa 75 % der Mädchen und etwa 80 % der Burschen nützen die Plattform genau 20 mal pro Woche.
- Um wie viel Prozent ist der Interquartilsabstand der wöchentlichen Nutzung der Plattform bei den Burschen größer als bei den Mädchen? Um Prozent.

Lösung: Whisker von 0 bis 4 und 20 bis 32, Median 12; 4; 8; (1, 4); 12,5