EVKOLA · STATISTIK IM STUDIUM

20 STATISTIK-ÜBUNGSAUFGABEN MIT LÖSUNGEN

Klausurnahe Aufgaben für Statistik I und II im Bachelor, etwa in BWL, Psychologie, VWL oder Sozialwissenschaften. Jede Lösung mit vollständigem Rechenweg.

- Beschreibende Statistik: Lage, Streuung, Häufigkeiten
- Zusammenhänge: Korrelation, Regression, χ^2
- Wahrscheinlichkeit: Bayes, Binomial-, Normalverteilung
- Schließende Statistik: Konfidenzintervall, t-Tests, p-Wert

Kostenlos von evkola.org · Mehr Aufgaben, Lernvideos und KI-Tutor: www.evkola.org/studium/statistik
Hinweis: Formeln und Rundungsregeln unterscheiden sich je Hochschule. Maßgeblich ist immer dein Skript.

Aufgaben

Rechne zuerst alle Aufgaben selbst, ohne in die Lösungen zu schauen. Runde Zwischenergebnisse möglichst wenig. Die Lösungen beginnen auf einer neuen Seite.

TEIL A · BESCHREIBENDE STATISTIK

1 Skalenniveaus bestimmen

Bestimme für jedes Merkmal das Skalenniveau (nominal, ordinal, intervall, verhältnis):

a. Studiengang
b. Zufriedenheit auf einer Skala von 1 (sehr unzufrieden) bis 5 (sehr zufrieden)
c. Temperatur in Grad Celsius
d. Monatliches Nettoeinkommen in Euro
e. Postleitzahl

2 Mittelwert, Median, Modus

Zehn Studierende erreichen in einer Probeklausur folgende Punkte:

12, 15, 9, 15, 18, 11, 15, 20, 7, 14

Berechne arithmetisches Mittel, Median und Modus.

3 Streuung: Spannweite, Varianz, Standardabweichung

Verwende die Daten aus Aufgabe 2. Berechne die Spannweite, die empirische Varianz (Division durch n) und die Stichprobenvarianz (Division durch n − 1) sowie die jeweilige Standardabweichung.

4 Variationskoeffizient

Produkt A verkauft im Schnitt 200 Stück pro Woche mit einer Standardabweichung von 30 Stück. Produkt B verkauft im Schnitt 50 Stück mit einer Standardabweichung von 10 Stück. Welches Produkt schwankt relativ stärker?

5 Gewichteter Mittelwert (Notendurchschnitt)

Berechne den nach ECTS-Punkten gewichteten Notendurchschnitt.

Modul	ECTS	Note
Mikroökonomie	5	1,7
Statistik I	10	2,3
Wissenschaftliches Arbeiten	5	1,0
Rechnungswesen	10	3,0

6 Geometrisches Mittel (Wachstumsraten)

Der Umsatz eines Unternehmens wächst im ersten Jahr um 10 %, sinkt im zweiten Jahr um 5 % und wächst im dritten Jahr um 20 %. Wie hoch ist die durchschnittliche jährliche Wachstumsrate?

7 Häufigkeitstabelle

50 Studierende bewerten eine Vorlesung von 1 (schlecht) bis 5 (sehr gut). Ergänze relative und kumulierte relative Häufigkeiten. Wie viel Prozent vergeben höchstens die Note 3?

Bewertung	Anzahl
1	4
2	6
3	10
4	16
5	14

TEIL B · ZUSAMMENHÄNGE

8 Korrelation nach Pearson

Fünf Studierende notieren ihre Lernstunden pro Woche (x) und ihre Klausurpunkte (y). Berechne den Korrelationskoeffizienten nach Pearson und interpretiere ihn.

Person	x (Stunden)	y (Punkte)
1	2	45
2	4	50
3	6	65
4	8	60
5	10	80

9 Lineare Regression

Bestimme mit den Daten aus Aufgabe 8 die Regressionsgerade $\hat{y} = a + b \cdot x$ nach der Methode der kleinsten Quadrate. Welche Punktzahl sagt das Modell bei 7 Lernstunden voraus? Wie groß ist das Bestimmtheitsmaß R^2?

10 Rangkorrelation nach Spearman

Zwei Dozierende bringen fünf Hausarbeiten in eine Rangfolge. Berechne den Rangkorrelationskoeffizienten nach Spearman.

Arbeit	Rang Dozent 1	Rang Dozent 2
A	1	2
B	2	1
C	3	4
D	4	3
E	5	5

11 Kontingenztabelle, Chi-Quadrat und Cramérs V

120 Studierende aus zwei Studiengängen haben eine Klausur geschrieben. Hängen Studiengang und Bestehen zusammen? Berechne die erwarteten Häufigkeiten, χ^2 und Cramérs V. Prüfe auf dem 5-%-Niveau (kritischer Wert bei 1 Freiheitsgrad: 3,841).

	bestanden	nicht bestanden	Summe
BWL	40	20	60
Psychologie	30	30	60
Summe	70	50	120

TEIL C · WAHRSCHEINLICHKEITSRECHNUNG

12 Laplace-Wahrscheinlichkeit

Du wirfst zwei faire Würfel. Wie groß ist die Wahrscheinlichkeit, dass die Augensumme genau 7 beträgt?

13 Bedingte Wahrscheinlichkeit (Satz von Bayes)

Ein Screening-Test erkennt eine Erkrankung in 90 % der Fälle (Sensitivität) und ist bei Gesunden in 95 % der Fälle negativ (Spezifität). 2 % der getesteten Personen sind erkrankt. Wie wahrscheinlich ist eine Person bei positivem Test tatsächlich erkrankt?

14 Binomialverteilung

Eine Multiple-Choice-Klausur hat 10 Fragen mit je 4 Antwortmöglichkeiten, genau eine ist richtig. Jemand rät bei jeder Frage. Wie groß ist die Wahrscheinlichkeit für (a) genau 3 richtige Antworten und (b) mindestens eine richtige Antwort?

15 Normalverteilung und z-Werte

Intelligenzwerte sind näherungsweise normalverteilt mit $\mu = 100$ und $\sigma = 15$.

a. Welcher Anteil hat einen IQ über 115?
b. Welchen z-Wert hat ein IQ von 130, und welcher Anteil liegt darüber?

TEIL D · SCHLIESSENDE STATISTIK

16 Konfidenzintervall für den Mittelwert

In einer Stichprobe von $n = 36$ Studierenden beträgt die mittlere Lernzeit pro Woche 24 Stunden bei einer Standardabweichung von $s = 6$ Stunden. Berechne das 95-%-Konfidenzintervall (t-Wert für 35 Freiheitsgrade: 2,030).

17 Hypothesen und Fehlerarten

Eine Dozentin vermutet, dass ein neues Tutorium die durchschnittliche Klausurnote verbessert.

a. Formuliere Null- und Alternativhypothese.
b. Was wäre hier ein Fehler 1. Art, was ein Fehler 2. Art?

18 t-Test für eine Stichprobe

16 Studierende eines Förderprogramms erreichen in einem standardisierten Test im Mittel 105 Punkte ($s = 12$). Der bekannte Durchschnitt aller Studierenden liegt bei 100. Unterscheidet sich die Gruppe auf dem 5-%-Niveau (zweiseitig)? Kritischer Wert $t(15; 0{,}975) = 2{,}131$.

19 t-Test für zwei unabhängige Stichproben

Gruppe A (n = 10) lernt mit Lernvideos und erreicht im Mittel 24 Punkte (s = 4). Gruppe B (n = 10) lernt nur mit dem Skript und erreicht 20 Punkte (s = 5). Gehe von gleichen Varianzen aus. Berechne t und entscheide auf dem 5-%-Niveau, einmal zweiseitig (kritischer Wert 2,101) und einmal einseitig mit H_1: $\mu_A > \mu_B$ (kritischer Wert 1,734), jeweils bei 18 Freiheitsgraden.

20 p-Wert richtig interpretieren

Ein Test ergibt p = 0,03 bei α = 0,05. Welche Aussagen sind richtig?

a. H_0 wird verworfen.
b. Die Wahrscheinlichkeit, dass H_0 wahr ist, beträgt 3 %.
c. Wenn H_0 wahr wäre, träte ein mindestens so extremes Ergebnis mit 3 % Wahrscheinlichkeit auf.
d. Der Effekt ist groß und praktisch bedeutsam.

Lösungen

1 Skalenniveaus bestimmen

a. **nominal** – nur Kategorien ohne Reihenfolge
b. **ordinal** – Rangfolge, aber Abstände nicht sicher gleich
c. **intervall** – gleiche Abstände, aber kein natürlicher Nullpunkt (0 °C ist nicht „keine Temperatur“)
d. **verhältnis** – natürlicher Nullpunkt, Verhältnisse sind sinnvoll („doppelt so viel“)
e. **nominal** – Zahlen dienen nur als Name; ein Mittelwert von Postleitzahlen ist sinnlos

2 Mittelwert, Median, Modus

Mittelwert: 136 / 10 = **13,6 Punkte**.

Sortiert: 7, 9, 11, 12, 14, 15, 15, 15, 18, 20. Bei gerader Anzahl liegt der Median zwischen dem 5. und 6. Wert: (14 + 15) / 2 = **14,5 Punkte**.

Modus (häufigster Wert): **15 Punkte** (dreimal).

3 Streuung: Spannweite, Varianz, Standardabweichung

Spannweite: 20 − 7 = **13 Punkte**.

Summe der quadrierten Abweichungen vom Mittelwert 13,6: **140,4**.

Empirische Varianz: 140,4 / 10 = **14,04**, Standardabweichung √14,04 ≈ **3,75 Punkte**.

Stichprobenvarianz: 140,4 / 9 = **15,60**, Standardabweichung ≈ **3,95 Punkte**.

> Klausurtipp: Achte darauf, welche Formel dein Skript verlangt. Beschreibst du nur die vorliegenden Daten, nimmst du meist n; schätzt du die Streuung einer Grundgesamtheit, nimmst du n − 1.

4 Variationskoeffizient

Variationskoeffizient = Standardabweichung / Mittelwert.

A: 30 / 200 = **0,15** (15 %). B: 10 / 50 = **0,20** (20 %).

Produkt B schwankt relativ stärker, obwohl seine absolute Standardabweichung kleiner ist.

5 Gewichteter Mittelwert (Notendurchschnitt)

Summe Note × ECTS: 5·1,7 + 10·2,3 + 5·1,0 + 10·3,0 = 66,5. Summe ECTS: 30.

Durchschnitt: 66,5 / 30 ≈ **2,22**. Der ungewichtete Mittelwert wäre 2,00 und damit falsch.

6 Geometrisches Mittel (Wachstumsraten)

Wachstumsfaktoren multiplizieren: 1,10 · 0,95 · 1,20 = 1,2540.

Geometrisches Mittel: $1{,}2540^{1/3} - 1 \approx$ **7,84 %** pro Jahr.

Das arithmetische Mittel der Raten (8,33 %) überschätzt das Wachstum, weil Raten sich multiplizieren und nicht addieren.

7 Häufigkeitstabelle

Bewertung	h	f	F (kumuliert)
1	4	0,08	0,08
2	6	0,12	0,20
3	10	0,20	0,40
4	16	0,32	0,72
5	14	0,28	1,00

Höchstens 3 vergeben F(3) = **0,40 = 40 %**.

8 Korrelation nach Pearson

Mittelwerte: $\bar{x} = 6$, $\bar{y} = 60$.

$\Sigma(x - \bar{x})(y - \bar{y}) = 160$; $\Sigma(x - \bar{x})^2 = 40$; $\Sigma(y - \bar{y})^2 = 750$.

$r = 160 / \sqrt{(40 \cdot 750)} \approx$ **0,924**.

Sehr starker positiver linearer Zusammenhang. Achtung: Korrelation beweist keine Kausalität, und fünf Personen sind eine sehr kleine Stichprobe.

9 Lineare Regression

Steigung $b = \Sigma(x - \bar{x})(y - \bar{y}) / \Sigma(x - \bar{x})^2 = 160 / 40 =$ **4**.

Achsenabschnitt $a = \bar{y} - b \cdot \bar{x} = 60 - 4 \cdot 6 =$ **36**. Gerade: $\hat{y} = 36 + 4 \cdot x$.

Bei 7 Stunden: $\hat{y} = 36 + 4 \cdot 7 =$ **64 Punkte**.

$R^2 = r^2 \approx$ **0,853**: Rund 85 % der Streuung der Punkte wird durch die Lernzeit erklärt.

10 Rangkorrelation nach Spearman

Differenzen d: −1, 1, −1, 1, 0 → $\Sigma d^2 =$ **4**.

$\rho = 1 - 6 \cdot \Sigma d^2 / (n \cdot (n^2 - 1)) = 1 - 6 \cdot 4 / (5 \cdot 24) =$ **0,8**. Die beiden Urteile stimmen stark überein.

11 Kontingenztabelle, Chi-Quadrat und Cramérs V

Erwartet = Zeilensumme · Spaltensumme / Gesamt: BWL 35 / 25, Psychologie 35 / 25.

$\chi^2 = \Sigma(\text{beobachtet} - \text{erwartet})^2 / \text{erwartet} = 25/35 + 25/25 + 25/35 + 25/25 \approx$ **3,43**.

Da 3,43 < 3,841, ist der Zusammenhang auf dem 5-%-Niveau **nicht signifikant**.

Cramérs $V = \sqrt{(\chi^2 / (n \cdot (\min(\text{Zeilen}, \text{Spalten}) - 1)))} = \sqrt{(3{,}43 / 120)} \approx$ **0,17**: ein schwacher Zusammenhang.

12 Laplace-Wahrscheinlichkeit

36 gleich wahrscheinliche Ergebnisse. Günstig: (1,6), (2,5), (3,4), (4,3), (5,2), (6,1) = 6 Ergebnisse.

$P = 6 / 36 =$ **1/6 ≈ 16,7 %**.

13 Bedingte Wahrscheinlichkeit (Satz von Bayes)

$P(\text{positiv}) = 0{,}90 \cdot 0{,}02 + 0{,}05 \cdot 0{,}98 = 0{,}067$.

$P(\text{krank} \mid \text{positiv}) = 0{,}90 \cdot 0{,}02 / 0{,}067 \approx$ **26,9 %**.

Trotz guter Testwerte ist nur gut jede vierte positiv getestete Person krank, weil die Erkrankung selten ist. Dieser Effekt ist ein Klausurklassiker.

14 Binomialverteilung

$X \sim B(n = 10; p = 0{,}25)$.

(a) $P(X = 3) = C(10,3) \cdot 0{,}25^3 \cdot 0{,}75^7 = 120 \cdot 0{,}015625 \cdot 0{,}1335 \approx$ **25,0 %**.

(b) $P(X \geq 1) = 1 - P(X = 0) = 1 - 0{,}75^{10} \approx$ **94,4 %**.

15 Normalverteilung und z-Werte

(a) $z = (115 - 100) / 15 = 1$. $P(Z > 1) = 1 - \Phi(1) \approx 1 - 0{,}8413 =$ **15,9 %**.

(b) $z = (130 - 100) / 15 =$ **2**. $P(Z > 2) = 1 - \Phi(2) \approx 1 - 0{,}9772 =$ **2,3 %**.

16 Konfidenzintervall für den Mittelwert

Standardfehler: $s / \sqrt{n} = 6 / 6 = 1$.

KI: $24 \pm 2{,}030 \cdot 1 =$ **[21,97; 26,03] Stunden**.

Mit der Normalverteilungsnäherung (1,96) ergibt sich [22,04; 25,96]. Prüfe, welche Variante dein Skript verlangt.

17 Hypothesen und Fehlerarten

(a) Bei Schulnoten bedeutet „besser" einen kleineren Wert. H_0: Das Tutorium verbessert die Note nicht, $\mu_{Tutorium} \geq \mu_{ohne}$. H_1: Mit Tutorium ist die durchschnittliche Note besser, $\mu_{Tutorium} < \mu_{ohne}$.

(b) **Fehler 1. Art**: Man schließt auf eine Verbesserung, obwohl das Tutorium in Wahrheit nichts bringt (H_0 fälschlich verworfen; Wahrscheinlichkeit höchstens α). **Fehler 2. Art**: Man findet keinen Effekt, obwohl das Tutorium tatsächlich hilft (H_0 fälschlich beibehalten; Wahrscheinlichkeit β).

18 t-Test für eine Stichprobe

$t = (\bar{x} - \mu_0) / (s / \sqrt{n}) = (105 - 100) / (12 / 4) = 5 / 3 \approx$ **1,667**.

Da $|1{,}667| < 2{,}131$, wird H_0 **nicht verworfen**. Der Unterschied ist bei dieser Stichprobengröße nicht signifikant. Das heißt nicht, dass es keinen Unterschied gibt, sondern nur, dass die Daten nicht ausreichen, um ihn nachzuweisen.

19 t-Test für zwei unabhängige Stichproben

Gepoolte Varianz: $s^2_p = (9 \cdot 16 + 9 \cdot 25) / 18 = 20{,}5$.

Standardfehler: $\sqrt{(20{,}5 \cdot (1/10 + 1/10))} \approx 2{,}025$. $t = (24 - 20) / 2{,}025 \approx$ **1,975**.

Zweiseitig: $1{,}975 < 2{,}101 \rightarrow$ **nicht signifikant**. Einseitig: $1{,}975 > 1{,}734 \rightarrow$ **signifikant**.

Deshalb muss die Richtung der Hypothese *vor* der Datenerhebung feststehen und darf nicht nachträglich passend gewählt werden.

20 p-Wert richtig interpretieren

Richtig: a und c.

b ist falsch: Der p-Wert ist keine Wahrscheinlichkeit für die Hypothese, sondern für die Daten unter der Annahme von H_0. d ist falsch: Signifikanz sagt nichts über die Größe des Effekts. Dafür gibt es Effektstärken wie Cohens d.

Wie lief es? Notiere dir die Aufgaben, bei denen du unsicher warst. Genau dort lohnt sich das Nacharbeiten.

Im Statistik-Kurs auf evkola gibt es zu jedem Thema Lernvideos, weitere Übungsaufgaben und einen KI-Tutor, der Rückfragen beantwortet. Für Studierende als **Semesterpass: 6 Monate für einmalig 99 €**, ohne Abo.

www.evkola.org/studium/statistik