

Testatfragen zu Vorlesung I

01. Nenne den Gegenstand der Statistik in den Humanwissenschaften.
02. “Am 31.12.2025 werden 160 Einwohner von Leipzig zufällig ausgewählt. In Leipzig kommt 2025 durchschnittlich ein Neugeborenes auf 141 Einwohner. Also ist mindestens eine der ausgewählten Personen im Jahr 2025 geboren.” Ist dieser Schluß korrekt? Begründe die Antwort.
03. Kann man aus der Verteilung eines Merkmals in einer Stichprobe auf die Verteilung desselben Merkmals in der Gesamtpopulation schließen?
04. Nach welchen Kriterien beurteilt man die Aussagekraft einer Stichprobe?
05. Was ist eine binäre Variable? Nenne außerdem zwei Beispiele.
06. Sollte das Geburtsjahr in einer Studie mit $N = 40$ Teilnehmern als kategoriale, ordinale Variable oder als metrische Variable aufgefaßt werden?
07. Welche Streuungsmaße gehören zum Stichproben-Mittelwert?
08. Nenne die Definition der empirischen Standardabweichung.
09. Warum ist der Stichproben-Mittelwert empfindlich gegen Ausreißer, der Stichproben-Median aber nicht?
10. Für eine Stichprobe wird ein Boxplot erstellt. Welche Abmessungen hat darin die Box selbst?

Software

Im Kurs wird das frei verfügbare Statistik-Softwarepaket jamovi verwendet (<https://www.jamovi.org/>) (02.10.2026). Wir empfehlen, das Paket zum persönlichen Gebrauch herunterzuladen und zu installieren. Eine Dokumentation ist derzeit nur online unter <https://docs.jamovi.org/> verfügbar (02.10.2026).

Literaturhinweise zum Kurs

- Skriptenheft: *Methodensammlung zur Auswertung klinischer und epidemiologischer Daten*. IMISE-Report Nr. 2/2007 (und neuere Auflagen)
- Dubben, H.-H.; Beck-Bornholdt, H.-P.: *Der Hund, der Eier legt*. Rowohlt Taschenbuch Verlag; Reinbek bei Hamburg 2020, 11. Aufl. (Nachdruck der vollständig überarbeiteten und erweiterten Neuausgabe 2006)

Anmerkungen

- * *Vorarbeiten für diesen Kurs*: Der Kurs wurde ursprünglich von Prof. Götz Gelbrich (jetzt Universität Würzburg) konzipiert und von Dr. Meinhard Mende (Universität Leipzig bis April 2021) laufend weiterentwickelt. Ich stütze mich sowohl inhaltlich als auch methodisch auf die Vorarbeiten beider Wissenschaftler.
- * *Verfälschte Zitate* (p. 1-04): Nach [BARKE (2004)] ist das Churchill zugeschriebene Zitat sehr wahrscheinlich eine Erfindung der deutschen Propaganda im Zweiten Weltkrieg. Dem H. G. Wells zugeschriebenen Zitat liegt wenigstens eine von ihm selbst veröffentlichte, jedoch wesentlich vorsichtiger und ausgewogener formulierte Textpassage zugrunde. Diese Passage ist nachträglich bis zur Unrichtigkeit verändert worden, siehe [TANKARD JR. (1979)].
- * *Grafik* auf p. 1-07: Verändert nach Josef Shinar, Haifa.
- * *Grundkonzepte der Statistik* (pp. 1-08 ff.): Mit dem ersten Grundkonzept beschäftigen sich die Vorlesungen I und II, mit dem dritten die Vorlesungen III, IV und VII, mit dem vierten die Vorlesungen V, VI und VIII, während das zweite Grundkonzept in allen Vorlesungen durchläuft.

- * *Kuh im Dorfteich* (p. 1-08): Die Widerristhöhe einer erwachsenen Kuh liegt zwischen 1.20 m und 1.60 m.
- * *Beispiel 1.1.* (p. 1-09): Die Daten für Leipzig kommen von statistik.leipzig.de/statcity (02.10.2026), die für Görlitz von www.goerlitz.de/Statistische.Zahlen.html (02.10.2026).
- * *Beispiel 1.3.* (p. 1-12): Der Body-Mass-Index (BMI) ist definiert als Körpergewicht geteilt durchs Quadrat der Körpergröße. Für unsere Rechnung haben wir die Kategorie “Übergewicht” mit den noch stärkeren Adipositas-Kategorien zusammengefaßt. Die LIFE-ADULT-Studie wird in [LÖFFLER ET AL. (2015)] beschrieben; die Daten stammen aus dem LIFE-Datenportal ldp.life.uni-leipzig.de (30.08.2021).
- * *“Self-reported items”* (p. 1-14): Per Selbstauskunft erhobene Daten können auch dann sehr unzuverlässig sein, wenn sie objektive Tatsachen betreffen. Ein Beispiel: [MAUKONEN ET AL. (2018)] haben 57 Studien analysiert, in denen BMI-Werte aus exakten Messungen mit denen verglichen wurden, die sich aus der Selbstauskunft über Größe und Gewicht ergaben. In der Mehrzahl der Studien wurde durch die Selbstauskunft der Anteil der Personen mit Übergewicht merklich unterschätzt, und zwar bis zu 9% .
- * *Beispiel 1.5.* (p. 1-16): Die Daten wurden aus [REYNER/HORNE (2013)], p. 95, Fig. 1 (a), abgelesen. Der Welch-Test wurde mit empirischen Standardabweichungen $\hat{\sigma}' = 2.08$ (Gruppe A) und $\hat{\sigma}'' = 2.0$ (Gruppe B) ausgeführt.
- * *Beispiel 1.6.* (p. 1-18): Die mittlere Verbesserung wurde nach [STÖCKEL ET AL. (2011)], p. 192, berechnet, und zwar wurden für beide Trainingsgruppen jeweils die Zeitdifferenzen Posttest minus Pretest für einhändiges Dribbling mit rechter und linker Hand gemittelt.
- * *Kategoriale Variablen* (p. 1-19): Man bezeichnet sie auch als Faktoren oder nominal bzw. ordinal skalierte Variablen.
- * *Metrische Variablen* (p. 1-19): Sie werden auch als stetige bzw. kontinuierlich skalierte Variablen bezeichnet. Kategorische, ordinale Variablen mit hinreichend vielen Kategorien (beispielsweise Prüfungsnoten) kann man auch als metrische Variablen behandeln.
- * *HADS-Fragebogen* (p. 1-20): Der zuerst in [ZIGMOND/SNAITH (1983)] publizierte Fragebogen gehört mittlerweile zu den Standardinstrumenten zur Beurteilung der Lebensqualität bei Krankheit bzw. medizinischer Behandlung. Mit 14 Fragen soll der Zustand des Probanden bezüglich Angst und Depression erfaßt werden (“HADS” ist ein Akronym für “Hospital Anxiety and Depression Scale”).
- * *Kenngrößen für Stichproben* (oder endliche Meßreihen) (pp. 1-24 ff.): Um sie von den jeweiligen Kenngrößen für die Gesamtpopulation zu unterscheiden, bezeichnen wir die Kenngrößen für Stichproben im gesamten Kurs durch Buchstaben mit Dach.
- * *Empirische Standardabweichung* (p. 1-24): Der Nenner ($N - 1$) hat damit zu tun, daß für die Mittelbildung unter der Wurzel nur noch ($N - 1$) Freiheitsgrade zur Verfügung stehen (ein Freiheitsgrad wurde schon bei der Berechnung von $\hat{\mu}$ “verbraucht”).
- * *Quartile* (p. 1-26): auch “Viertelswerte”, von lateinisch “quartilatus”: in vier Teile geteilt. In Definition 1.20. werden $\hat{Q}_{25}$ und $\hat{Q}_{75}$ als Mediane der beiden geordneten Stichproben-Hälften erklärt.
- * *Boxplot* (p. 1-27): Die Konventionen über die Länge der Antennen und die Unterscheidung von Ausreißern und extremen Ausreißern wurden von dem amerikanischen Statistiker J. W. Tukey vorgeschlagen. Achtung: nicht alle Autoren halten sich daran.
- * *Empirische Verteilungsfunktion* (p. 1-29): Mathematisch muß eine empirische Verteilungsfunktion folgende Eigenschaften besitzen: monoton wachsend, rechtsseitig stetig, $\lim_{x \rightarrow -\infty} \hat{F}(x) = 0$ und $\lim_{x \rightarrow \infty} \hat{F}(x) = 1$. Andere Bezeichnungen: “kumulative Verteilungsfunktion” bzw. “ecdf”.

* *Definition 1.20.a* (p. 1-30): Wegen der rechtsseitigen Stetigkeit der empirischen Verteilungsfunktion werden die Minima in der Definition stets angenommen. Ist die Anzahl der Messungen $N = 2k + 1$ ungerade, so stimmt der Stichproben-Median mit dem 2. Quartil überein: $\hat{m} = \hat{Q}_{50}$.

Quellenverzeichnis

- [BARKE (2004)] Barke, W.: *“Ich glaube nur der Statistik, die ich selbst gefälscht habe ...”* Statistisches Monatsheft Baden-Württemberg 11/2004 (2004) : 50 – 53
- [LÖFFLER ET AL. (2015)] Loeffler, M.; Engel, C.; Ahnert, P.; et al.: *The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany.* BMC Public Health **15** (2015) : 691
- [MAUKONEN ET AL. (2018)] Maukonen, M.; Männistö, S.; Tolonen, H.: *A comparison of measured versus self-reported anthropometrics for assessing obesity in adults: a literature review.* Scand. J. Public Health **46** (2018) : 565 – 579
- [REYNER/HORNE (2013)] Reyner, L. A.; Horne, J. A.: *Sleep restriction and serving accuracy in performance tennis players, and effects of caffeine.* Physiology & Behavior **120** (2013) : 93 – 96
- [STÖCKEL ET AL. (2011)] Stöckel, T.; Weigelt, M.; Krug, J.: *Acquisition of a complex basketball-dribbling task in school children as a function of bilateral practice order.* Res. Q. Exerc. Sport **82** (2011) : 188 – 197
- [TANKARD JR. (1979)] Tankard jr., J. W.: *The H. G. Wells quote of statistics: a question of accuracy.* Historia Mathematica **6** (1979) : 30 – 33
- [ZIGMOND/SNAITH (1983)] Zigmond, A. S.; Snaith, R. P.: *The hospital anxiety and depression scale.* Acta Psych. Scand. **67** (1983) : 361 – 370

Bildnachweis

- * p. 1-01: Foto: Stefan Brending, veröffentlicht unter Lizenz Creative Commons CC-BY-SA-3.0 (verändert)
- * p. 1-09: Foto: Marco Verch, veröffentlicht unter Lizenz Creative Commons CC-BY 2.0
- * p. 1-10: Foto: Steffen Prößdorf, veröffentlicht unter Lizenz Creative Commons CC-BY-SA-4.0
- * p. 1-12: Logo des Leipziger Forschungszentrums für Zivilisationskrankheiten, www.life.uni-leipzig.de/de/life_forschungszentrum.html (16.07.2021)
- * p. 1-14: Grafik: Karen Arnold, veröffentlicht unter Lizenz Creative Commons CC0
- * p. 1-16: Foto: Charlie Cowins, veröffentlicht unter Lizenz Creative Commons Attribution 2.0 Generic
- * p. 1-18: pxhere.com/de/photo/870138 (16.07.2021), veröffentlicht unter Lizenz Creative Commons CC0