Florian Wickelmaier
Forschungsmethoden und Mathematische Psychologie
Fachbereich Psychologie
Universität Tübingen

Vorlesung Statistik II
SS 2026

Übungsaufgaben

Aufgabe 1. Zur Untersuchung der Einkommensverhältnisse in bäuerlichen Betrieben in Baden-Württemberg wird eine Zufallsauswahl von 200 ländlichen Gemeinden getroffen. In diesen Gemeinden werden 10% der bäuerlichen Betriebe zufällig erfasst.

a) Um welches Verfahren der Stichprobenauswahl handelt es sich hier?

b) Worin unterscheidet sich die Erhebung einer Klumpenstichprobe von der beschriebenen Vorgehensweise?

Aufgabe 2. Stichprobenziehung bei PISA 2000:

> "Die Stichprobenziehung in den Teilnehmerstaaten erfolgte nach detaillierten Vorgaben der internationalen Projektleitung. In einem ersten Schritt wurden die schulischen Systeme der Teilnehmerstaaten nach zentralen Merkmalen unterteilt, wie zum Beispiel nach Regionen (Länder, Provinzen, Kantone u.Ä.) und nach Arten von Schulen. Innerhalb dieser Unterteilungen (in Deutschland: Schulformen innerhalb der Länder) wurden dann die Schulen nach einem Zufallsverfahren ausgewählt. Auch die im zweiten Schritt getroffene Auswahl von Schülerinnen und Schülern innerhalb der Schulen erfolgte per Zufall." (aus PISA 2000: Die Studie im Überblick – Grundlagen, Methoden und Ergebnisse, S. 4)

Wie kann dieses Verfahren der Stichprobenziehung beschrieben werden?
Literaturhinweis: Die Problematik der durch eine nicht repräsentative Stichprobenziehung verzerrten Ergebnisse bei PISA 2000 in Österreich wird beschrieben in Neuwirth, E. (2006). PISA 2000: Sample weight problems in Austria. OECD Education Working Papers, No. 5, OECD Publishing. `https://doi.org/10.1787/220456725273`

Aufgabe 3. Aus den drei Grundschulen G1 (500 Schüler), G2 (200 Schüler) und G3 (300 Schüler) einer Stadt soll für eine Studie eine Stichprobe von 100 Schülern zufällig ausgewählt werden.

a) Wie groß ist die Wahrscheinlichkeit für einen Schüler aus Schule G1, G2 bzw. G3 in die Stichprobe aufgenommen zu werden, wenn in einem mehrstufigen Vorgehen zunächst zufällig eine Schule ausgewählt wird, mit gleicher Wahrscheinlichkeit für jede Schule, und anschließend in dieser Schule eine Zufallsauswahl der Schüler stattfindet?

b) Wie groß ist die Wahrscheinlichkeit für einen Schüler aus Schule G1, G2 bzw. G3 in die Stichprobe aufgenommen zu werden, wenn unmittelbar eine Zufallsauswahl aus der Gesamtmenge der Schüler erfolgt?

c) Wie müssen in einem mehrstufigen Vorgehen, wie unter a) beschrieben, die Wahrscheinlichkeiten für die Ziehung der Schule gewählt werden, sodass sich für die Wahrscheinlichkeiten, dass ein Schüler aus Schule G1, G2 bzw. G3 in die Stichprobe aufgenommen wird, jeweils gerade der in b) berechnete Wert ergibt?

d) Implementieren Sie die drei Auswahlverfahren a), b) und c) in R. Hinweis: Verwenden Sie die Funktionen `sample()` für die Zufallsauswahl und `replicate()` zur Wiederholung des Zufallsexperiments.

Aufgabe 4. In einer schriftlichen Befragung aller 1000 Gymnasiallehrer in einem Schulbezirk erhebt ein Forschungsinstitut Informationen zur Häufigkeit, mit der über den Stoff hinausgehend aktuelle politische Themen im Unterricht diskutiert werden. In den 330 beantworteten und zurückgesendeten Fragebögen wird in 22% der Fälle die Frage bejaht, ob in den letzten vier Wochen im Unterricht eine Diskussion zu aktuellen politischen Themen geführt wurde. Was lässt sich über den entsprechenden Prozentsatz in der Grundgesamtheit aussagen unter der Annahme, dass derartige Diskussionen bei denen, die den Fragebogen nicht zurück gesendet haben, in höchstens 5% der Fälle stattgefunden haben.

Aufgabe 5. Ein Arzt misst den Blutdruck bei seinen Patienten gewöhnlich zweimal hintereinander. Langjährige Erfahrung sagt ihm, dass bei einem verhältnismäßig hohen Wert in der ersten Messung die zweite Messung gewöhnlich niedriger ausfällt. Er schließt daraus, dass Patienten bei der zweiten Messung etwas entspannter sind. Sein Kollege kann ihm da überhaupt nicht zustimmen. Er meint, dass Patienten, die einen verhältnismäßig niedrigen Wert bei der ersten Messung haben, gewöhnlich einen höheren Wert bei der zweiten Messung aufweisen. Der Kollege denkt daher, Patienten seien bei der zweiten Messung nervöser. Welcher Arzt hat nun Recht?

Aufgabe 6. Wir betrachten die n-fach wiederholte Darbietung einer Reizsituation (x, s), bestehend aus einem Vergleichsreiz x und einem Standardreiz s im Rahmen der Konstanzmethode als Folge unabhängiger Bernoulli-Experimente. Zur Schätzung der über alle Darbietungen konstanten Wahrscheinlichkeit p einer "Ja"-Antwort auf die Frage, ob der Vergleichsreiz intensiver als der Standardreiz ist, wird deren relative Häufigkeit bestimmt. Berechnen Sie den Standardfehler der relativen Häufigkeit. (Hinweis: Beschreiben Sie die relative Häufigkeit als arithmetisches Mittel über Bernoulli-verteilte Indikatorvariablen.)