

Vertiefende statistische Verfahren

3. Übungsblatt SS 2024

Allgemeine Information

Alle Aufgaben sind mit R zu lösen, wenn nicht explizit anders angegeben. Die Berechnungen sollen nachvollziehbar und dokumentiert sein. Um die vollständige Punktezahl zu erreichen, müssen alle Ergebnisse und Fragen entsprechend interpretiert bzw. beantwortet werden. Code alleine ist nicht ausreichend! Die Abgabe erfolgt über Moodle entsprechend der Abgaberichtlinien als pdf und Rmd File. Bitte inkludieren Sie namentlich alle beteiligten Gruppenmitglieder sowohl im Bericht als auch im Source Code. Die jeweiligen Datensätze die für diese Übung relevant sind finden Sie ebenfalls in Moodle.

1 Einfaktorielle ANOVA (händisch) [2P]

Ein Sportwissenschaftler möchte sehen, ob es einen Unterschied in der Gewichtszunahme von Sportlern gibt, die einer von drei speziellen Diäten folgen. Die Athleten werden nach dem Zufallsprinzip drei Gruppen zugewiesen und unterziehen sich für 6 Wochen der jeweiligen Diät. Die Gewichtszunahmen (in Pfund) sind angegeben. Nehmen Sie an, dass die Gewichtszunahmen normalverteilt sind und die Varianzen gleich sind. Kann der Sportwissenschaftler bei einem Signifikanzniveau von 0,05 schlussfolgern, dass es einen Unterschied zwischen den Diäten gibt? Führen Sie die ANOVA händisch durch und überprüfen Sie das Ergebnis mit `summary(aov(...))`.

Diät	Messwerte
A	3, 6, 7, 4
B	10, 12, 11, 14, 8, 6
C	8, 3, 2, 5

2 Einfaktorielle ANOVA [2P]

Analysieren Sie die Daten zu niedrigem Geburtsgewicht von Neugeborenen in `birthwt_aov.xlsx`. Untersuchen Sie ob die ethnische Zugehörigkeit der Mutter (Variable `ethnic`, wobei 1 = "white", 2 = "black", 3 = "other") einen Einfluss auf das Geburtsgewicht von Neugeborenen hat (Variable `bwt`). Beachten Sie dabei folgende Punkte:

- Überprüfen Sie die Voraussetzungen.
- Verwenden Sie eine geeignete grafische Darstellung der Daten / Ergebnisse.
- Gibt es signifikante Unterschiede zwischen den Gruppen, wenn ja zwischen welchen Gruppen?
- Achten Sie auf eine "statistisch korrekte" Formulierung des Ergebnisses.

3 Einfaktorielle ANOVA [2P]

Verwenden Sie den bereits bekannten `Framingham.sav` Datensatz. Analysieren Sie ob es Unterschiede im BMI in Abhängigkeit von der Schulbildung gibt. Achten Sie auf Ausreißer und fehlende Daten (`NaN`, `NA's`).

- i) Überprüfen Sie die Voraussetzungen.
- ii) Verwenden Sie eine geeignete grafische Darstellung der Daten / Ergebnisse.
- iii) Gibt es signifikante Unterschiede zwischen den Bildungsstufen, wenn ja zwischen welchen Stufen?
- iv) Achten Sie auf eine “statistisch korrekte” Formulierung des Ergebnisses.

4 Mehrfaktorielle ANOVA [2P]

Sie führen eine Studie bezüglich des Einflusses unterschiedlicher Diäten und Aktivitätslevels auf den Erfolg bei der Gewichtsabnahme durch. Jedem Probanden wird eine Diät und ein Aktivitätslevel zugewiesen und die Differenz zum Ausgangsgewicht nach 2 Monaten gemessen (in kg). Die Daten der Studie sind in `weightloss.sav` zusammengefasst. Analysieren Sie ob die beiden Faktoren einen Einfluss auf das Gewicht haben.

- i) Überprüfen Sie die Voraussetzungen.
- ii) Verwenden Sie eine geeignete grafische Darstellung der Daten / Ergebnisse.
- iii) Gibt es signifikante Haupteffekte sowie Interaktionseffekte (inkl. Interaktions-Plot)?
- iv) Achten Sie auf eine “statistisch korrekte” Formulierung des Ergebnisses.

5 Mehrfaktorielle ANOVA [2P]

Verwenden Sie den erneut den `Framingham.sav` Datensatz. Analysieren Sie den systolischen Blutdruck `sysbp` abhängig von Geschlecht und Bildungsstufe. Achten Sie auf Ausreißer und fehlende Daten (`NaN`, `NA's`).

- i) Überprüfen Sie die Voraussetzungen.
- ii) Verwenden Sie eine geeignete grafische Darstellung der Daten / Ergebnisse.
- iii) Gibt es signifikante Haupteffekte sowie Interaktionseffekte (inkl. Interaktions-Plot)?
- iv) Achten Sie auf eine “statistisch korrekte” Formulierung des Ergebnisses.