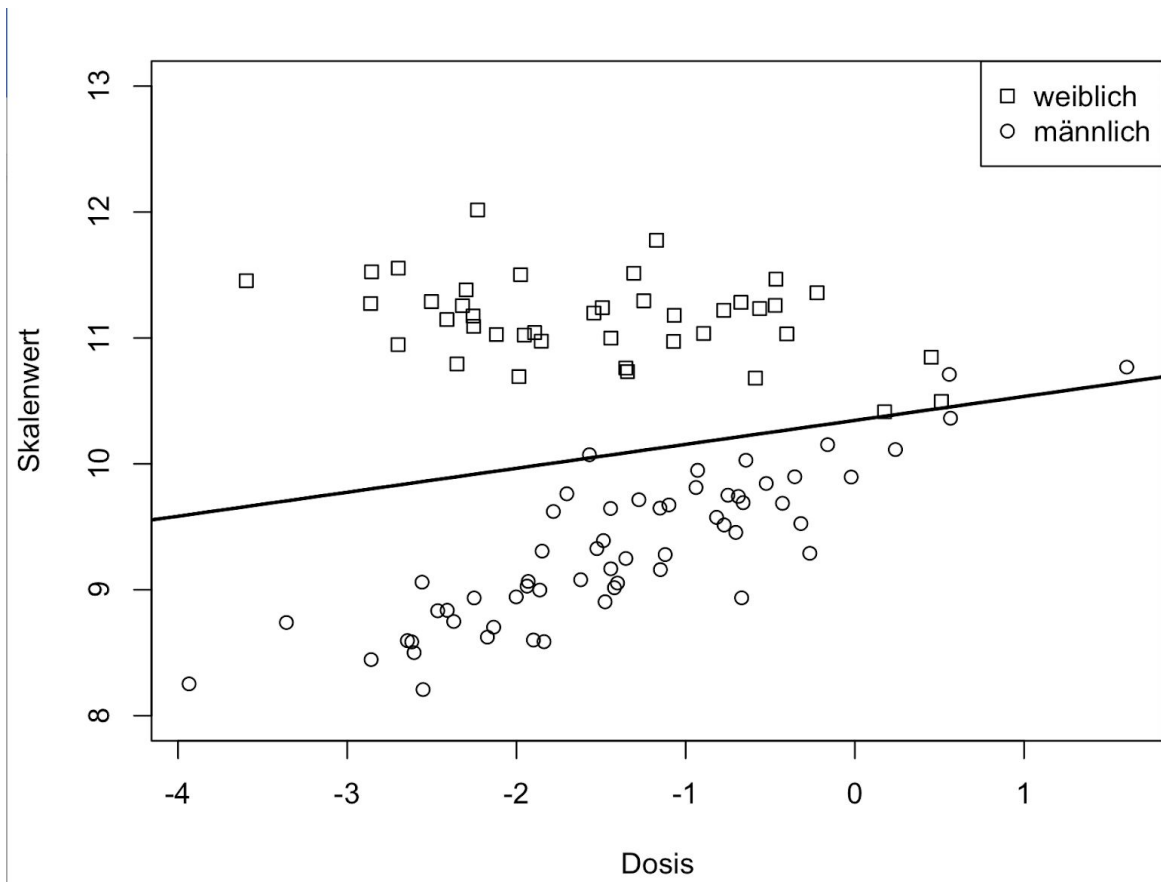


Aufgabe 1) 13 Punkte



Scatterplot separat für Frauen und Männer, für Frauen ein leicht negativer Zusammenhang, für Männer ein stark positiver. Dazwischen lag eine Regressionsgrade

y= Wert auf Depressionsskala

x= Medikamentendosis

z=Geschlecht

Angenommen wird lineare Regressionsgleichung $Y = \beta_0 + \beta_1 \cdot x$

- A) Wieso könnte die oben gezeigte Darstellung problematisch für das Modell sein?
- B) Welche Modellgleichung würden sie stattdessen schätzen? (Regressionsgleichung, ohne Annahmen, Definieren Sie die Variablen)
- C) Welche Vorzeichen haben die Beta-Koeffizienten aus b? (4 Punkte)
- D) Beschreiben Sie kurz das Prinzip (o.ä.) der Varianzhomogenität in Bezug auf Modell in A) und B)

Aufgabe 2)

- 3 günstige Eigenschaften von Schätzern nennen. (3 Punkte)
- Für 2 deren technische Bedeutung sowie ihre Bedeutung für die Interpretation von Daten nennen (4 Punkte)

Aufgabe 3)

- Geben Sie die Formeln für die 3 Gauß-Markov-Annahmen an und beschreiben Sie kurz ihre inhaltliche Bedeutung (6 Punkte)
- Geben Sie für 2 der 3 Annahmen jeweils ein Beispiel an, in dem die Annahme verletzt ist
- wofür braucht man die NV-Annahme der Fehler? Ist sie für die Konsistenz der Schätzer nötig? (3 Punkte)
- Welche Eigenschaft muss die Kreuzproduktmatrix $X'X$ haben?
- Wie kann man mit $X'X$ die Standardabweichung von beta bestimmen?

Aufgabe 4)

- Outputinterpretation: Kleine Tücke: Eine x-Variable war "centered", $\text{Alkohol_centered} = \text{Alkohol} - 0.5$. Dies muss bei weiteren Aufgabenteilen für die Rechnung berücksichtigt werden. Eine andere Variable war Geschlecht (männlich = 0, weiblich = 1), $Y = \text{Reaktionszeit}$
- Zweiseitiges KI für ein beta rechnen
- Wert für eine Person berechnen mit bestimmter Ausprägung in Prädiktor 1 + 2 (weiblich, Alkoholgehalt = 0,6)
- Scatterplot + Normal-quantilplot sind gegeben (Heteroskedastizität + keine NV der Fehler, aber $n \gg 20$). Wofür braucht man die plots generell? Was bedeuten Sie in diesem Fall, auch für die Interpretierbarkeit der Koeffizienten im output?
- Bedeutet ein kleiner p-Wert immer, dass der Effekt inhaltlich bedeutsam ist? (3 Punkte)