

Hinweis: Für die Lösung der Übung benötigen Sie möglicherweise folgende Quantile:

$$t_{0,975}^{25} = 2,060; t_{0,975}^{26} = 2,056; t_{0,975}^{30} = 2,042; t_{0,95}^{25} = 1,708; t_{0,95}^{29} = 1,699$$

und Werte der Verteilungsfunktion der t-Verteilung

$$F_{t^{26}}(0,5285) = 0,699183; F_{t^{26}}(4,8935) = 0,999978; F_{t^{26}}(0,1080) = 0,542588;$$

$$F_{t^{30}}(4,8935) = 0,999984; F_{t^{25}}(4,8935) = 0,999975;$$

Aufgabe 1:

Im Folgenden wird ein Datensatz über Pizzalieferungen in München betrachtet. Angenommen, es liegen für die Variablen „Zustellzeit“ (in Minuten), „Rechnungsbetrag“ (in €) und „Filiale“ (‘1’ = Altstadt/Glockenbach; ‘2’ = Schwabing/Maxvorstadt, ‘3’ = Neuhausen/Nymphenburg) 30 Beobachtungen vor. Im Folgenden sind die ersten 10 der 30 Beobachtungen zu sehen:

i	Zustellzeit	Rechnungsbetrag	Filiale
1	34,5	13,78	1
2	28	12,95	2
3	29	20,75	3
4	40,5	11,10	2
5	31	55,20	2
6	33,5	30,00	2
7	35	22,98	1
8	29,5	13,78	1
9	30	44,00	3
10	38	25,30	3

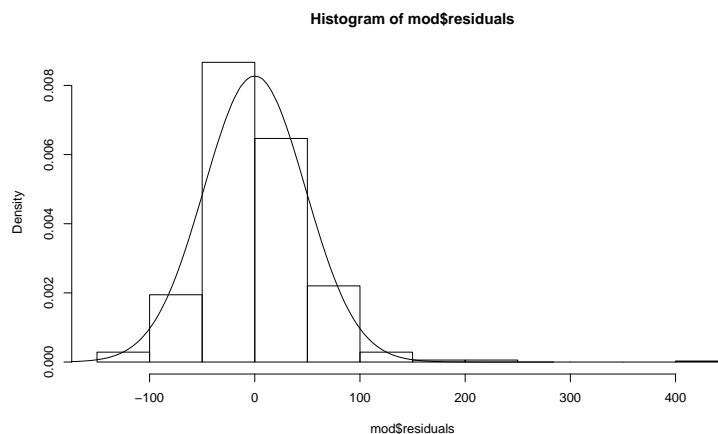
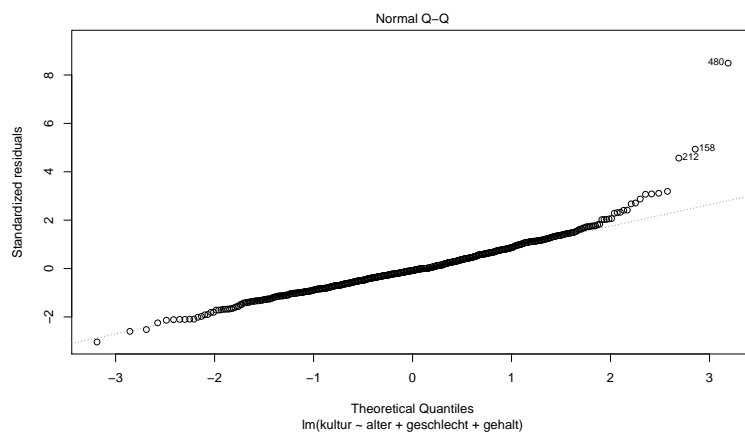
Es soll nun der Einfluss des „Rechnungsbetrags“ und der „Filiale“ auf die „Zustellzeit“ untersucht werden.

- Wie lässt sich die kategoriale Variable „Filiale“ mit Hilfe der Dummykodierung mit Filiale 1 als Referenz darstellen?
- Sie erhalten den folgenden R-Output (ausgewertet mit dem Gesamtdatensatz):
⇒ *Siehe Zusatzangabe (R-Outputs)*
 - Geben Sie die Modellgleichung an!
 - Geben Sie an, ob der geschätzte Effekt für die Variable „Rechnungsbetrag“ zum 5%-Niveau signifikant ist.
 - Bestimmen Sie ein 95%-Konfidenzintervall für den geschätzten Effekt für „Rechnungsbetrag“. Stellen Sie anschließend einen Zusammenhang zum vorher durchgeführten t-Test her.

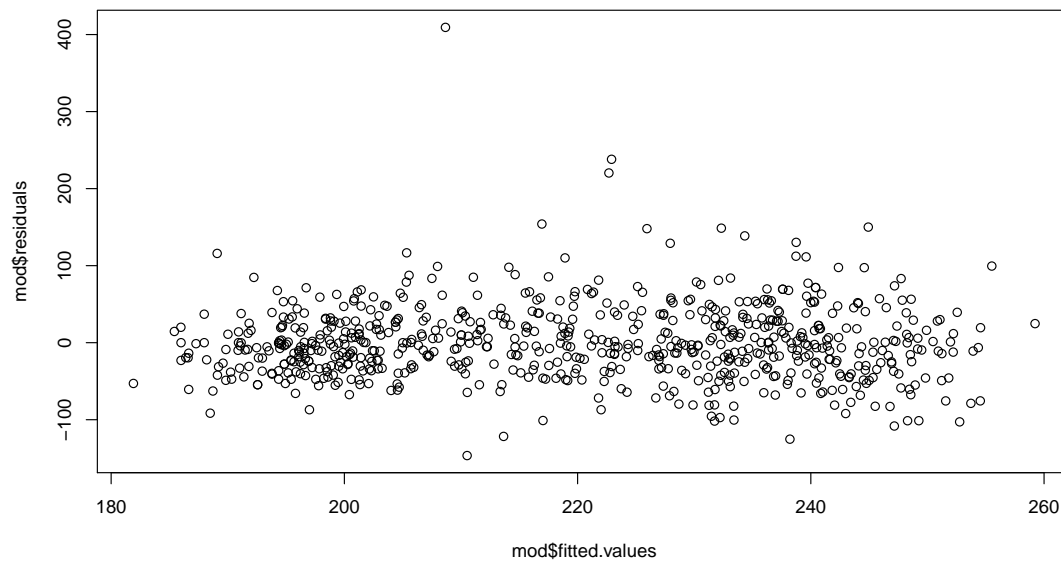
- (iv) Wie hoch ist der Erklärungsgehalt des Modells?
- c) Prognostizieren Sie basierend auf dem Modell die erwartete Zustellzeit für eine Lieferung aus der Filiale Neuhausen/Nymphenburg mit einem Rechnungsbetrag von 40€.

Aufgabe 2: Wir betrachten wieder den Datensatz ‘Theater’ von Aufgabenblatt 10. Sie erhalten folgenden R-Output: (⇒ *Siehe Zusatzangabe (R-Outputs)*)

- a) Geben Sie die Modellgleichung an!
- b) Betrachten Sie den folgenden Output der Regressionsanalyse und interpretieren Sie ihn hinsichtlich der Modellgüte und der Koeffizientenschätzer. Gehen Sie dabei stets von einem Signifikanzniveau von $\alpha = 0,05$ aus.
- c) Prognostizieren Sie basierend auf dem Modell von 2a) die Kulturausgaben für einen 40-jährigen Mann mit einem Einkommen von 60000 SFR.
- d) Welche Annahmen des Modells werden in den folgenden beiden Grafiken geprüft?



e) Welche Annahme wird mit dem folgenden Output überprüft?



f) Warum sind die unter c) und d) überprüften Annahmen von Bedeutung?