Universität des Saarlandes
Lehrstab Statistik
Dr. Martin Becker

13. Übungsblatt zur Vorlesung
Schließende Statistik WS 2016/17

Aufgabe 48

Ein Landwirt möchte den Zusammenhang zwischen eingesetzter Düngemenge X [kg/ha] und Ertrag Y [t Weizen/ha] erfassen. Während der letzten 12 Jahre hat er die folgenden Beobachtungen auf seinen Äckern gemacht:

Jahr	Ertrag y_i [t Weizen/ha]	eingesetzte Düngemenge x_i [kg/ha]
1999	25.1	180
2000	28.5	220
2001	41.4	430
2002	38.8	370
2003	46.9	510
2004	49.3	800
2005	50.6	900
2006	50.4	750
2007	49.7	680
2008	48.9	620
2009	50.6	700
2010	49.0	690

Es werde die lineare Beziehung $y_i = \beta_1 + \beta_2 x_i + u_i$ mit $u_i \overset{\text{iid}}{\sim} N(0, \sigma^2)$ zugrunde gelegt.

(a) Schätzen Sie β_1 und β_2 mit Hilfe der Kleinst-Quadrate-Methode.

(b) Berechnen Sie das Bestimmtheitsmaß R^2.

(c) Geben Sie mit Hilfe der bekannten erwartungstreuen Schätzfunktion für σ^2 den realisierten Schätzwert für σ^2 an.

(d) Berechnen Sie $\widehat{\sigma^2}_{\widehat{\beta}_1}$ und $\widehat{\sigma^2}_{\widehat{\beta}_2}$.

(e) Testen Sie zum Signifikanzniveau $\alpha = 0.05$, ob die eingesetzte Düngemenge einen Einfluss auf den Ertrag hat.

(f) Geben Sie ein symmetrisches Konfidenzintervall zur Vertrauenswahrscheinlichkeit $1-\alpha = 0.95$ für β_1 an.

(g) Geben Sie ein Prognoseintervall zur Vertrauenswahrscheinlichkeit $1-\alpha = 0.90$ für die in 2011 geerntete Menge y_0 an, falls für 2011 ein Düngemitteleinsatz von $x_0 = 910$ erfolgt ist.

(h) Geben Sie ein Prognoseintervall zur Vertrauenswahrscheinlichkeit $1-\alpha = 0.90$ für den Erwartungswert der in 2011 zu erntenden Menge $\mathrm{E}(y_0)$ an, falls für 2011 ein Düngemitteleinsatz von $x_0 = 910$ erfolgt ist.

Aufgabe 49

In einer einfachen Stichprobe von $n = 42$ britischen Haushalten wurden die wöchentlichen Ausgaben für Kleidung (`y`) und das jeweilige Haushaltseinkommen (`x`) erhoben (jeweils in £). Die Schätzung des einfachen linearen Regressionsmodells

$$y_i = \beta_1 + \beta_2 x_i + u_i \qquad \text{mit} \quad u_i \overset{\text{iid}}{\sim} N(0, \sigma^2),\ i = 1, \ldots, 42$$

mit der Statistik-Software R produzierte folgende Ausgabe:

```
Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-19.796  -6.174  -3.158   5.663  28.739

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.38913    3.83720   0.101  0.91973
x            0.07383    0.02645   2.791  0.00801 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.26 on 40 degrees of freedom
Multiple R-squared:  0.163,       Adjusted R-squared:  0.1421
F-statistic:  7.79 on 1 and 40 DF,  p-value: 0.008012
```

(a) Geben Sie die realisierten Kleinst-Quadrate-Schätzwerte für β_1 und β_2 an.

(b) Geben Sie den realisierten Schätzwert für σ^2 an.

(c) Welcher Anteil der Gesamtvarianz der Ausgaben für Kleidung wird durch das lineare Modell erklärt?

(d) Entscheiden Sie mit Hilfe des zugehörigen p-Werts zum Signifikanzniveau $\alpha = 0.10$, ob β_1 signifikant von Null verschieden ist.

(e) Geben Sie ein Konfidenzintervall zur Vertrauenswahrscheinlichkeit $1 - \alpha = 0.95$ für β_1 an.

(f) Überprüfen Sie mit einem geeigneten Test zum Niveau $\alpha = 0.01$, ob das Haushaltseinkommen einen signifikant positiven Einfluss auf die Ausgaben für Kleidung hat.

(g) Welche wöchentlichen Ausgaben für Kleidung (in £) prognostiziert das Modell für einen Haushalt mit einem Haushaltseinkommen von 150 £?