Lösungen zur Übung 6

1. a)

$$\mathbf{X}^t\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 7 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 5 \\ 1 & 7 \end{pmatrix} = \begin{pmatrix} 4 & 16 \\ 16 & 84 \end{pmatrix}$$

$$\mathbf{X}^t\mathbf{Y} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 7 \end{pmatrix} \begin{pmatrix} 34 \\ 47 \\ 55 \\ 64 \end{pmatrix} = \begin{pmatrix} 200 \\ 898 \end{pmatrix}$$

$$\mathbf{Y}^t\mathbf{Y} = \begin{pmatrix} 34 & 47 & 55 & 64 \end{pmatrix} \begin{pmatrix} 34 \\ 47 \\ 55 \\ 64 \end{pmatrix} = 10486$$

$$(\mathbf{X}^t\mathbf{X})^{-1} = \frac{1}{80} \begin{pmatrix} 84 & -16 \\ -16 & 4 \end{pmatrix} = \begin{pmatrix} \frac{21}{20} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{1}{20} \end{pmatrix}$$

b)

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \frac{21}{20} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{1}{20} \end{pmatrix} \begin{pmatrix} 200 \\ 898 \end{pmatrix} = \begin{pmatrix} 30.4 \\ 4.9 \end{pmatrix}$$

c)

```
Call: lm(formula = y ~ x)
Residuals:
    1   2   3    4
 -1.3 1.9 0.1 -0.7

Coefficients:
              Value Std. Error t value Pr(>|t|)
(Intercept) 30.4000  1.7450    17.4213  0.0033
          x  4.9000  0.3808    12.8680  0.0060


Residual standard error: 1.703 on 2 degrees of freedom
Multiple R-Squared: 0.9881
F-statistic: 165.6 on 1 and 2 degrees of freedom, the p-value is 0.005985
```

2. a) einfache lineare Regression durch den Nullpunkt

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \qquad \boldsymbol{\beta} = \beta_1$$

b) quadratische Regression

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix} \qquad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

c) 1-Weg-Varianzanalyse, 2-Stichprobenproblem

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix} \qquad \boldsymbol{\beta} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

3. a) ja, pro zusätzliche Kuh steigt das Einkommen um 20 \$.

```
 Call: lm(formula = I ~ C, data = farm)
Residuals:
    Min     1Q Median    3Q   Max
 -204.7 -80.02  15.48 54.57 284.4

Coefficients:
               Value Std. Error  t value Pr(>|t|)
(Intercept) 694.0189  50.0393    13.8695   0.0000
          C  20.1109   4.7251     4.2562   0.0005
```

b)

```
> xx_data.frame(C=c(0,20,mean(farm$C)))
> pred_predict(mod1,xx)
>
sepred_summary(mod1)$sigma*sqrt(1/20+(xx-mean(farm$C))^2/(19*var(farm$C)))

> vo_pred+sepred*qt(0.975,18)
> vu_pred+sepred*qt(0.025,18)
> vi_cbind(vu,vo)
> vi
% Vertrauensintervall fuer E(I)
C=    0:  588.9  799.1
C=   20:  971.4 1221.1
C= 8.85:  814.3  929.7
> seprog_summary(mod1)$sigma*sqrt(1+1/20+(xx-mean(farm$C))^2/(19*var(farm$C)))
```

```
> uu_pred+seprog*qt(0.025,18)
> pu_pred+seprog*qt(0.025,18)
> po_pred+seprog*qt(0.975,18)
> p_cbind(pu,po)
> p
% Prognoseintervall fuer I:
C=   0: 415.3  972.8
C=  20: 809.4 1383.0
C=8.85: 607.4 1136.6
```

c)

```
Call: lm(formula = I ~ A, data = farm)
Coefficients:
               Value Std. Error  t value Pr(>|t|)
(Intercept) 868.7363 105.9796     8.1972   0.0000
          A   0.0234   0.7066     0.0331   0.9740   nicht sign.

Call: lm(formula = A ~ C, data = farm)
Coefficients:
                Value Std. Error   t value  Pr(>|t|)
(Intercept)  191.0590   18.6410    10.2494    0.0000
          C   -5.8259    1.7602    -3.3097    0.0039  sign.


Call: lm(formula = I ~ A + C, data = farm)
Coefficients:
               Value Std. Error  t value Pr(>|t|)
(Intercept) 285.4572  81.3793     3.5077   0.0027
          A   2.1384   0.3936     5.4336   0.0000 sign.
          C  32.5690   3.7276     8.7372   0.0000 sign.
```

Die Einnahmequelle Land wird nur erkennbar, wenn die Anzahl Kühe kontrolliert ist, d. h. Gleiches mit Gleichem verglichen wird.