LiMo Sheet 3 Exercise 3

Christoph Stiehm

17 November 2016

a)

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}} = \frac{Cov(x,y)}{Var(x)} \tag{1}$$

Kovarianz und Varianz können aus $\hat{\Sigma}$ entnommen werden.

$$\Rightarrow \hat{\beta}_1 = \frac{16.52447}{23.25701} = 0.71052 \tag{2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 * \bar{x} \Rightarrow \hat{\beta}_0 = 70.67662 - 0.71052 * 178.9221 = -56.45111 \tag{3}$$

b)

SS_{Total}

$$Var(y) = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \Leftrightarrow \sum_{i=1}^{n}(y_i - \bar{y})^2 = Var(y)*(n-1) \Rightarrow \sum_{i=1}^{n}(y_i - \bar{y})^2 = 50.56418*76 = 3842.878 \tag{4}$$

SS_x

$$Var(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \Leftrightarrow \sum_{i=1}^{n}(x_i - \bar{x})^2 = Var(x)*(n-1) \Rightarrow \sum_{i=1}^{n}(x_i - \bar{x})^2 = 23.25701*76 = 1767.533 \tag{5}$$

$SS_{Regression}$

$$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \hat{\beta}_1^2 * \sum_{i=1}^{n}(x_i - \bar{x})^2 \Rightarrow 0.5048325 * 1790.79 = 892.308 \tag{6}$$

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \Rightarrow \frac{904.0489}{3893.442} = 0.2321979 \tag{7}$$

Kontrolle/Alternative

$$\rho = \ = \frac{s_{xy}^2}{s_x^2 s_y^2} = \frac{cov(x_i, y_i)}{\sqrt{\sigma_x^2 \sigma_y^2}} \Rightarrow \frac{16.52447}{34.29244} = 0.4818691 \quad \rho^2 = 0.2321979 = R^2 \tag{8}$$

23 % der Streuung des Gewichts y können durch die Größe/Höhe x erklärt werden -> Moderate Anpassungsqualität.

c)

$$\hat{Var}(\hat{\beta}_1) = \frac{\hat{\sigma^2}}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{9}$$

σ^2 über SS_{Error} berechnen, $SS_{Error} = R^2 * SS_{Total}$

$$\begin{aligned}\hat{\sigma^2} &= \frac{1}{n-2}\left(\sum_{i=1}^{n}(y_i - \bar{y})^2 - \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}\right) \\ &= \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2 - \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{n-2} \\ &= \frac{2950.57}{75} = 39.34093\end{aligned} \tag{10}$$

Einsetzen

$$\Rightarrow \hat{Var}(\hat{\beta}_1) = \frac{39.34093}{1767.533} = 0.02225754 \tag{11}$$

Konfidenzintervall (Überdeckungswahrscheinlichkeit) berechnen

$$\hat{\beta}_1 \pm t_{\alpha/2,n-2} \frac{s}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \Rightarrow 0.7105157 \pm -1.992102 * 0.1491896 \tag{12}$$

Lower 0.4133147 Upper 1.007717

R Code

```
# Sheet3_Exercise3 ---------------------------------------------------------------
Sigma <- matrix(c(23.25701, 16.52447, 16.52447, 50.56418), nrow = 2, ncol = 2)

n <- 77
barX <-  178.9221
barY <- 70.67662

#-- a)
(beta1 <- Sigma[1,2] / Sigma[1,1])
```

```
## [1] 0.7105157
```

```
(beta0 <- barY - beta1 * barX)
```

```
## [1] -56.45034
```

```
#-- b)
# SS_Total
SSt <- Sigma[2,2] * (n - 1)

# SS_Regression
SSr <- Sigma[1,1] * (n-1) * beta1^2

(R2 <- SSr / SSt)
```

```
## [1] 0.2321979
```

```
# Kontrolle, Alternative
rho <- Sigma[1,2] / sqrt(Sigma[1,1] * Sigma[2,2])
rho^2 == R2
```

```
## [1] TRUE
```

```
# SS_Error
SSe <- SSt - SSr

# \sigma^2
sigma2 <- SSe / (n-2)
(varBeta1 <- sigma2 / (Sigma[1,1] * (n - 1)))
```

```
## [1] 0.02225754
```

```
# CI
(beta1 + qt(0.025, n-2) * sqrt(varBeta1))
```

```
## [1] 0.4133147
```

```
(beta1 - qt(0.025, n-2) * sqrt(varBeta1))
```

```
## [1] 1.007717
```