6.1. Betrachten Sie das inhomogene Modell $y_t = \alpha + \beta_1 x_{t1} + \cdots + \beta_K x_{tK} + u_t$. Zeigen Sie, dass die F-Test Statistik für $H_0: \beta_1 = \cdots = \beta_K = 0$ gegeben ist durch

$$F = \frac{T-K-1}{K}\frac{(y-\imath\bar{y})'(y-\imath\bar{y}) - \hat{u}'\hat{u}}{\hat{u}'\hat{u}} = \frac{T-K-1}{K}\frac{y'\tilde{X}^p y}{y'(I_T - (\imath, X)^p)y} = \frac{T-K-1}{K}\frac{R_\star^2}{1-R_\star^2}$$

Hier bezeichnet $\tilde{X} = (I - \imath^p)X = (X - \imath\bar{X})$ die Matrix der Mittelwert-bereinigten unabhängigen Variablen und $R_\star^2$ ist das zentrierte Bestimmtheitsmaß. Beachte: $(\imath, X)^p = (\imath, \tilde{X})^p = \imath^p + \tilde{X}^p$.

6.2. Gegeben ist ein Regressionsmodell $y = X_1\beta_1 + X_2\beta_2 + u$ mit normalverteilten Fehlern. Zeigen Sie, dass die F-Teststatistik für $H_0: \beta_2 = 0 \in \mathbb{R}^s$ die Darstellung

$$F = \frac{T-K}{s}\frac{(R_1^2 - R_0^2)}{1-R_1^2}$$

hat, wobei R_0^2 das Bestimmtheitsmaß der Regression von y auf X_1bezeichnet und R_1^2 das Bestimmtheitsmaß der Regression von y auf $X = (X_1, X_2)$. Wenn X_1 ein Absolutglied enthält, dann kann man in der obigen Formel auch das jeweilige zentrierte Bestimmtheitsmaß einsetzen.

6.3. Gegeben ist eine Stichprobe $y_t \sim N(\mu, \sigma^2)$, $t = 1, \ldots, T$.

1. Konstruieren Sie ein Konfidenzintervall für das Mittel μ zum Sicherheitsniveau $(1-\alpha)$.
2. Konstruieren Sie ein Konfidenzintervall für die Varianz σ^2 zum Sicherheitsniveau $(1-\alpha)$.

6.4. Gegeben sei die QR-Zerlegung der Matrix (X, y)

$$(X, y) = (Q_1, Q_2)\begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$$

wobei die Matrizen $Q = (Q_1, Q_2)$ so partitioniert sind, dass $X = Q_1 R_{11}$ und $y = Q_1 R_{12} + Q_2 R_{22}$. Die Matrix Q hat orthonormale Spalten ($Q'Q = I_{K+1}$) und die Matrix $R_{11} \in \mathbb{R}^{K\times K}$ ist eine obere Dreiecksmatrix. Zeigen Sie, dass die OLS Schätzer in folgender Weise aus dieser QR Zerlegung berechnet werden können:

DRt.

$$\hat{y} = Q_1 R_{12},\ \hat{u} = Q_2 R_{22},\ \hat{\beta} = \boxed{R_{11}^{-1}} R_{12},\ \hat{u}'\hat{u} = R_{22}^2,\ \hat{y}'\hat{y} = R_{12}' R_{12}$$

numerisch stabil!

6.5. Schätzen einer Investitionsgleichung. Der Datensatz `DATS_03.txt` (Quelle: Hackl, 2004) enthält Beobachtungen für die Variablen GNP (nominelles GNP), INVEST (nominelle Investitionen), PC (Verbraucherpreisindex) und R (Zinssatz) für die Jahre 1968-1982. Die Daten finden Sie unter: `http://eos.tuwien.ac.at/lehre/grundlagen_der_oekonometrie/`

1. Einlesen der Daten in R:

```
investment = read.table("DatS_03.txt",header=TRUE,comment.char="#")
print(summary(investment))
attach(investment)
```

2. Berechnen Sie die realen Investitionen bzw. reales GNP und die Inflationsrate:

```
IR = 100*INVEST/CPI
GNPR = 100*GNP/CPI
PI = c(NaN,100*diff(CPI)/CPI[1:14])
t = YEAR-1967
```

   Beachten Sie, dass eine Beobachtung (für das Jahr 1968) bei der Berechnung der Inflationsrate "verloren" geht .

3. Schätzen Sie folgendes Investitionsmodell
   IR $=\beta_1+\beta_2$t$+\beta_3$GNPR$+\beta_4$R$+\beta_5$PI+U

4. Testen Sie $\beta_2 = \cdots = \beta_5 = 0$.

5. Testen Sie $\beta_3 = 1$
6. Testen Sie $\beta_4 - \beta_5 = 0$ auf zwei Arten: mit dem Wald Test und mittels eines t-Tests in einem geeignet transformierten Modell.
7. Machen Sie folgende plots (hier bezeichnet `m` das in R geschätzte Modell):
   (a) IR gegen gefittete Werte (`m$fitted`)
   (b) geschätzte Residuden (`m$resid`) gegen gefittete Werte (`m$fitted`)
   (c) geschätzte Residuen (`m$resid`) gegen die Zeit (`t`)
   (d) geschätzte Residuen zum Zeitpunkt t gegen die Residuen zum Zeitpunkt $t-1$.

6.1. unrestr. Modell:

$$y_t = \alpha + \beta_1 x_{t1} + \ldots + \beta_K x_{tK} + u_t$$

$$\Leftrightarrow y = \cancel{\alpha}\,\iota\alpha + X\beta + u = (\iota, X)\binom{\alpha}{\beta} + u$$

$$\Rightarrow \hat{u}'\hat{u} = y'(I_T - (\iota, X)^{r})\,y$$

$$rg\,(\iota, X^{*})^{r} = \cancel{\#}\,K+1 \quad \Rightarrow rg\,(I_T - (\iota, X)^{r}) = T-K-1$$

restr. Modell: $y = \iota\alpha + u \Rightarrow \tilde{\alpha} = \bar{y}$

$$\tilde{y} = \iota^{r} y = \iota\tilde{\alpha} = \iota\bar{y}$$

$$\Rightarrow \tilde{u}'\tilde{u} = (y-\tilde{y})'(y-\tilde{y}) = (y-\iota\bar{y})'(y-\iota\bar{y})' = y'(I-\iota^{r})y$$

$$\Rightarrow F = \frac{T-K-1}{K}\;\frac{\tilde{u}'\tilde{u} - \hat{u}'\hat{u}}{\hat{u}'\hat{u}}$$

$$= \frac{T-K-1}{K}\;\frac{(y-\iota\bar{y})'(y-\iota\bar{y})' - \hat{u}'\hat{u}}{\hat{u}'\hat{u}}$$

$$\tilde{u}'\tilde{u} - \hat{u}'\hat{u} = y'(I-\iota^{r})y - y'(I - \underbrace{(\iota, X^{r})}_{\iota^{r} + \tilde{X}^{r}})\,y = y'\tilde{X}^{r} y$$

$$\Rightarrow F = \frac{T-K-1}{K}\;\frac{y'\tilde{X}^{r} y}{y'(I-(\iota, X)^{r})\,y}$$

$$R_*^2 = 1 - \frac{\hat{u}'\hat{u}}{y'(I-\iota^{r})y} = \cancel{\ldots} = 1 - \frac{\hat{u}'\hat{u}}{\tilde{u}'\tilde{u}}$$

$$\Rightarrow F = \frac{T-K-1}{K}\;\cancel{\ldots}\;\frac{\tilde{u}'\tilde{u} - \hat{u}'\hat{u}}{\hat{u}'\hat{u}} = \frac{T-K-1}{K}\;\frac{1 - \frac{\hat{u}'\hat{u}}{\tilde{u}'\tilde{u}}}{\frac{\hat{u}'\hat{u}}{\tilde{u}'\tilde{u}}} = \frac{T-K-1}{K}\;\frac{R_*^2}{1-R_*^2}.$$

6.2 unrestr.: $y = (X_1, X_2)\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} + u \quad \Rightarrow R_0^2 = 1 - \frac{\hat{u}'\hat{u}}{y'y}$

restr.: $y = X_1\beta_1 + u \quad \Rightarrow R_1^2 = 1 - \frac{\tilde{u}'\tilde{u}}{y'y}$

$$\Rightarrow F = \frac{T-K}{S}\,\frac{\tilde{u}'\tilde{u} - \hat{u}'\hat{u}}{\hat{u}'\hat{u}} = \frac{T-K}{S}\,\frac{\frac{\tilde{u}'\tilde{u}}{y'y} - \frac{\hat{u}'\hat{u}}{y'y}}{\frac{\hat{u}'\hat{u}}{y'y}}$$

$$= \frac{T-K}{S}\,\frac{1-R_0^2 - 1 + R_1^2}{1-R_0^2} = \frac{T-K}{S}\,\frac{R_1^2 - R_0^2}{1-R_0^2}$$

Für zentrierten Bestimmtheitsmaße analog.

6.3 $y \sim N(\mu, \sigma^2 I_T)$

-) $\Rightarrow \frac{\cancel{\;\;}}{\sqrt{T}} \frac{\overline{X}-\mu}{S} \sim t_{T-1}$

$$\Rightarrow P\left[-t_{T-1;\,1-\frac{\alpha}{2}} \le \frac{\cancel{\;\;}}{\sqrt{T}} \frac{\overline{X}-\mu}{S} \le t_{T-1;\,1-\frac{\alpha}{2}}\right] = 1-\alpha$$

$$P\left[\overline{X} - \overset{\wedge}{\sqrt{T}} \cdot S \cdot t_{T-1;\,1-\frac{\alpha}{2}} \le \mu \le \overline{X} + \overset{\wedge}{\sqrt{T}} \cdot S \cdot t_{T-1;\,1-\frac{\alpha}{2}}\right]$$

~~$\Rightarrow KI = [\overline{X} - \sqrt{T} \cdot S \cdot t$~~

-) $(T-1)\frac{S^2}{\sigma^2} \sim \chi^2_{T-1}$

$$\Rightarrow P\left[(T-1)\frac{S^2}{\sigma^2} \le \chi^2_{T-1;\,1-\alpha}\right] = 1-\alpha$$

$$P\left[(T-1)\frac{S^2}{\chi^2_{T-1;\,1-\alpha}} \le \sigma^2\right] = 1-\alpha$$

$$\Rightarrow KI = \left(-\infty,\ (T-1)\frac{S^2}{\chi^2_{T-1;\,1-\alpha}}\right).$$

6.4 $(X, y) = (Q_1, Q_2)\begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22}\end{pmatrix}$

$$\hat\beta = X^{+} y = (X'X)^{-1}X'y = (R_{11}' \underbrace{Q_1'Q_1}_{I} R_{11})^{-1} R_{11}' \cancel{R_{12}} \underbrace{\underbrace{Q_1'(Q_1}_{I} R_{12} + Q_2 R_{21})}_{0}$$

$$= R_{11}^{-1} \underbrace{(R_{11}')^{-1} R_{11}'}_{I} \cancel{Q_2} R_{12} = R_{11}^{-1} R_{12}$$

$$\Rightarrow \hat y = X\hat\beta = Q_1 R_{11} R_{11}^{-1} R_{12} = Q_1 R_{12}$$

$$\hat u = y - \hat y = Q_1 R_{12} + Q_2 R_{22} - Q_1 R_{12} = Q_2 R_{22}$$

$$\hat u'\hat u = \underset{\in \mathbb{R}}{R_{22}'} \underbrace{Q_2' Q_2}_{I} R_{22} = R_{22}^2$$

$$\hat y'\hat y = R_{12}' \underbrace{Q_1' Q_1}_{I} R_{12} = R_{12}' R_{12}.$$

1. Einlesen der Daten

```
> investment <- read.table("DatS_03.txt", header=TRUE, comment.char="#")
> summary(investment)
      YEAR           GNP            INVEST          CPI               R
 Min.   :1968   Min.   : 873.4   Min.   :133.3   Min.   : 82.54   Min.   : 4.500
 1st Qu.:1972   1st Qu.:1131.8   1st Qu.:180.7   1st Qu.: 98.00   1st Qu.: 5.480
 Median :1975   Median :1549.2   Median :229.8   Median :125.79   Median : 6.250
 Mean   :1975   Mean   :1748.6   Mean   :276.0   Mean   :131.40   Mean   : 7.453
 3rd Qu.:1978   3rd Qu.:2290.8   3rd Qu.:394.4   3rd Qu.:156.92   3rd Qu.: 9.055
 Max.   :1982   Max.   :3057.5   Max.   :471.5   Max.   :207.23   Max.   :13.420
```

2. Berechnung der realen Größen

```
> IR <- 100*INVEST/CPI
> GNPR <- 100*GNP/CPI
> PI <- c(NaN,100*diff(CPI)/CPI[1:14])
> t <- YEAR-1967
```

3. Modellschätzung: $\mathrm{IR} = \beta_1 + \beta_2\mathrm{t} + \beta_3\mathrm{GNPR} + \beta_4\mathrm{R} + \beta_5\mathrm{PI} + \mathrm{U}$

```
> m1<-lm(IR ~ t + GNPR + R + PI)
```

```
> summary(m1)

Call:
lm(formula = IR ~ t + GNPR + R + PI)

Residuals:
     Min       1Q   Median       3Q      Max
-9.89101 -3.08000 -0.05605  2.26175  7.12772

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -518.47989   46.62878 -11.119 1.47e-06 ***
t            -17.60271    1.72784 -10.188 3.06e-06 ***
GNPR           0.68355    0.04672  14.630 1.40e-07 ***
R             -1.79110    1.07039  -1.673    0.129
PI            -0.43430    1.15909  -0.375    0.717
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.656 on 9 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.98,	Adjusted R-squared: 0.9711
F-statistic: 110.4 on 4 and 9 DF,  p-value: 1.218e-07
```

4. Test auf $\beta_2 = \cdots = \beta_5 = 0$

```
F-statistic: 110.4 on 4 and 9 DF,  p-value: 1.218e-07
```

Also wird die Nullhypothese abgelehnt.

5. Test auf $\beta_3 = 1$

```
> koeff <- coef(summary(m1))
> tau <- koeff[3,"t value"] - 1/koeff[3,"Std. Error"]
> 2*pt(abs(tau),9,lower.tail=FALSE)
[1] 8.149428e-05
```

Die Nullhypothese wird also abgelehnt.

6. Test auf $\beta_4 - \beta_5 = 0$

a) Wald-Test

```
> R<-matrix(c(0,0,0,1,-1),nrow=1)
> r<-0
> b<-R%*%coef(m1)-r
> vardach<-vcov(m1)

> W<-b^2/(R%*%vardach%*%t(R))

> pf(W,1,10,lower.tail=FALSE)
          [,1]
[1,] 0.4966356
```

Die Nullhypothese wird also angenommen.

b) transformiertes Modell: $\mathrm{IR} = \beta_1 + \beta_2\mathrm{t} + \beta_3\mathrm{GNPR} + (\beta_4\text{-}\beta_5)\mathrm{R} + \beta_5(\mathrm{R+PI}) + \mathrm{U}$

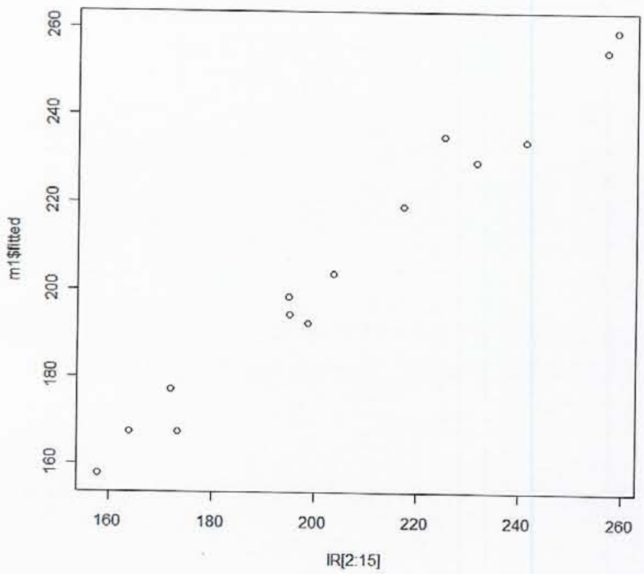
```
> H<-R+PI

> m2<-lm(IR ~ t + R + H)
> coef(summary(m2))
               Estimate Std. Error     t value     Pr(>|t|)
(Intercept) 158.4012934  27.424532  5.77589768 0.0001787321
t             5.9253040   2.983971  1.98571128 0.0751551695
R            -1.2331462   9.083347 -0.13575902 0.8947057582
H             0.4889543   5.465993  0.08945388 0.9304874561
```

Der Koeffizient von H ist nicht signifikant, die Nullhypothese wird angenommen.

7. Plots

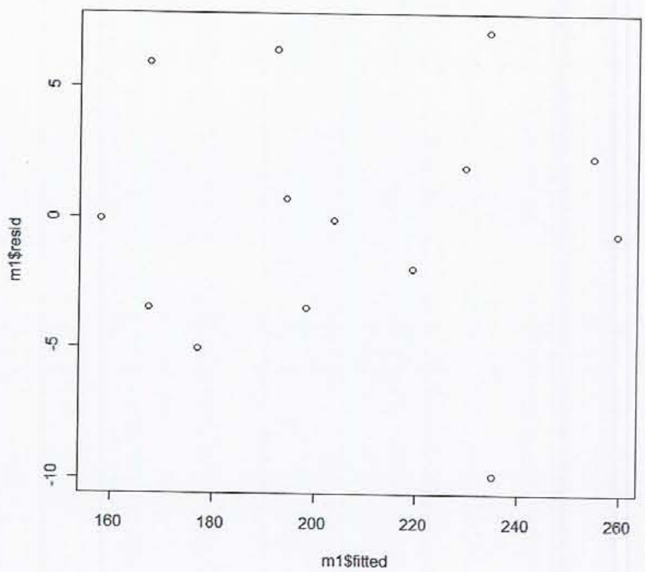
a) IR gegen gefittete Werte

```
> plot(IR[2:15],m1$fitted)
```



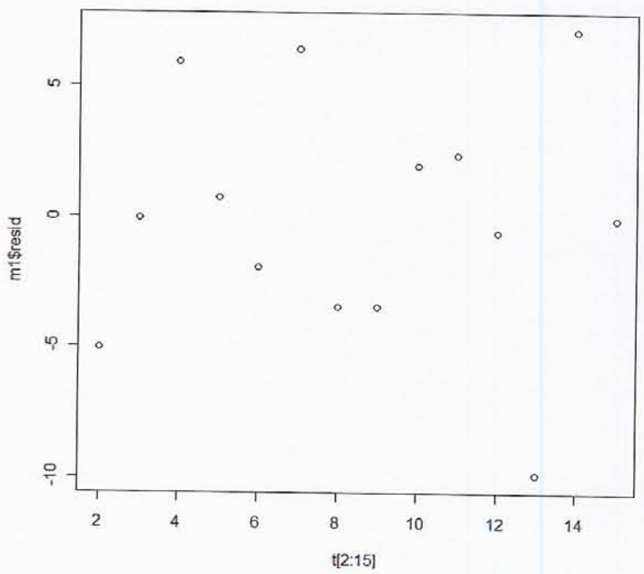
b) geschätzte Residuen gegen gefittete Werte

```
> plot(m1$fitted,m1$resid)
```



c) geschätzte Residuen gegen t

```
> plot(t[2:15],m1$resid)
```



d) geschätzte Residuen zu t gegen gesch. Res. zu t-1

```
> plot(m1$resid[2:14],m1$resid[1:13])
```