Kausale Effektschätzung - ANCOVA

Aufgabe 1: Datensatz und Deskriptivstatistiken

```r
str(spf2)
```

```
'data.frame':	289 obs. of  3 variables:
 $ schule: Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 2 ...
 $ HISEI : num  26.6 28.5 30.5 75.5 70.1 ...
 $ hoeren: num  226 252 439 413 573 ...
```

```r
# Verteilung auf die Schulformen
table(spf2$schule)
```

```

  0   1 
 77 212 
```

```r
# Unbedingter Mittelwert der Kompetenz `hören`
mean(spf2$hoeren)
```

```
[1] 322.9552
```

```r
# Gruppenspezifische Kompetenz-Werte
tapply(spf2$hoeren, spf2$schule, mean)
```

```
       0        1 
263.8902 344.4081 
```

```r
# Gruppenspezifische Werte auf der Kovariate
tapply(spf2$HISEI, spf2$schule, mean)
```

```
       0        1 
35.68519 39.58231 
```

Aufgabe 2: Prima Facie Effekt der Hörkompetenz

```
## t-Test
m1 <- lm(hoeren ~ schule, spf2)
summary(m1)$coef
```

```
             Estimate Std. Error   t value     Pr(>|t|)
(Intercept) 263.89023   12.71918 20.747417 4.808921e-59
schule1      80.51785   14.85047  5.421905 1.252209e-07
```

$\widehat{PFE} = 80.52$

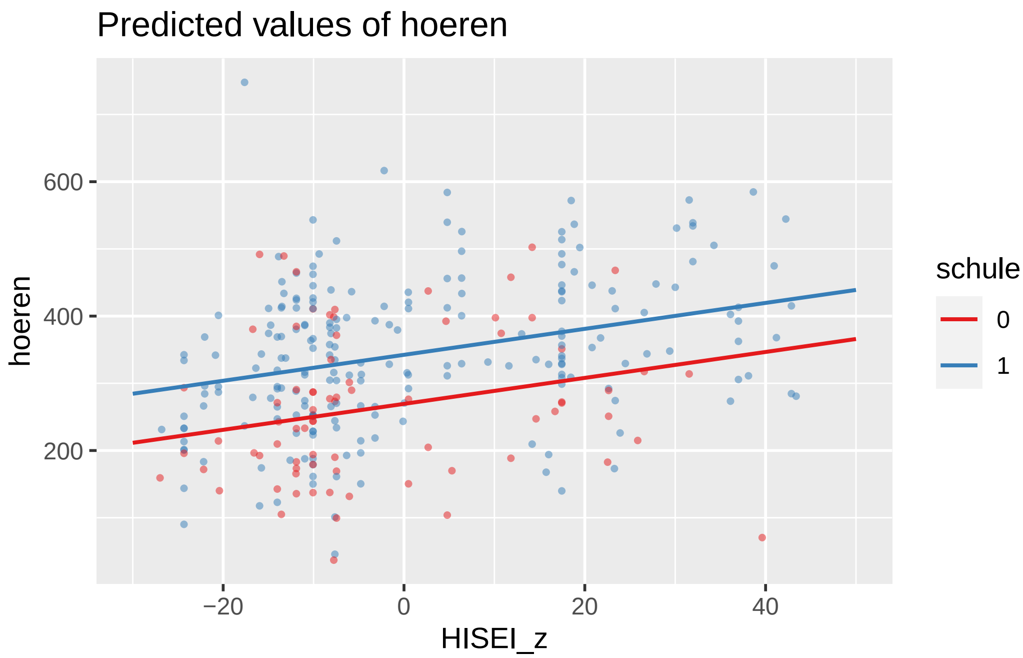
Aufgabe 3: Traditionelle ANCOVA

```
# Zentrierung der Kovariate am Gesamtmittelwert
library(jtools)
spf2$HISEI_z <- center(data = spf2$HISEI)
# trad. ANCOVA
m2 <- lm(hoeren ~ HISEI_z + schule, spf2)
summary(m2)$coef
```

```
              Estimate Std. Error   t value     Pr(>|t|)
(Intercept) 269.409676 12.2070746 22.069962 1.039432e-63
HISEI_z       1.930697  0.3656082  5.280781 2.550353e-07
schule1      72.993695 14.2715132  5.114643 5.768561e-07
```

Aufgabe 4: Streudiagramm

```
library(sjPlot)
plot_model(m2, type = "pred",
           terms = c("HISEI_z", "schule"),
           show.data = TRUE, dot.size = 1, # graph. Parameter
           ci.lvl = NA)
```



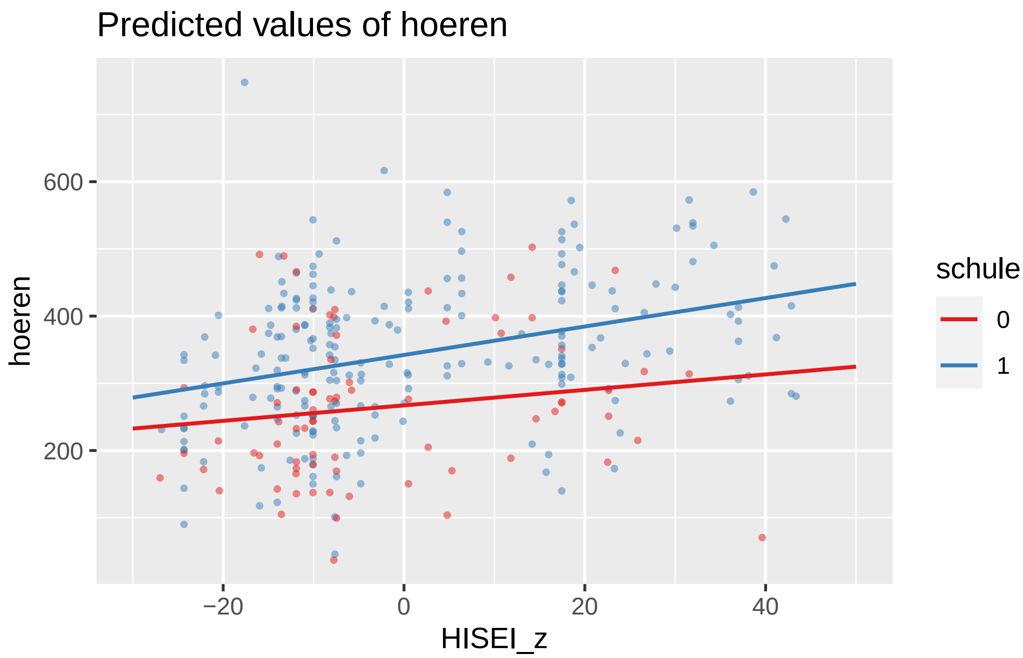
Effekt ist Abstand zwischen den Geraden ($\widehat{ATE} = \alpha_2 = 72.99$)

Aufgabe 5: Generalisierte ANCOVA

```
m3 <- lm(hoeren ~ HISEI_z * schule, spf2)
summary(m3)$coef
```

```
                   Estimate Std. Error   t value     Pr(>|t|)
(Intercept)     267.1818702 12.3917244 21.561315 8.039943e-62
HISEI_z           1.1514127  0.8334636  1.381479 1.682137e-01
schule1          75.0287695 14.4028985  5.209283 3.640344e-07
HISEI_z:schule1   0.9649005  0.9274274  1.040405 2.990334e-01
```

```
plot_model(m3, type = "pred",
           terms = c("HISEI_z", "schule"),
           show.data = TRUE, dot.size = 1, # graph. Parameter
           ci.lvl = NA)
```



ATE und ATT berechnen

Für die Berechnung des ATT brauchen wir den Mittelwert der Kovariaten in der Treatmentgruppe ($E(Z \mid X = 1)$). Für den ATE benutzen wir dass $E(Z) = 0$, da wir unsere Kovariate am Gesamtmittelwert zentriert haben.

```
# Mittelwert der zentrierten Kovariaten in Treatmentgruppe
mean(spf2$HISEI_z[spf2$schule==1])
```

```
[1] 1.038332
```

$$\begin{aligned}\widehat{ATE} &= \alpha_2 = 75.03 \\ \widehat{ATT}_{X=1} &= \alpha_2 + \alpha_3 \cdot E(Z|X = 1) \\ &= 75.0287695 + 0.9649005 \cdot 1.038332 = 76.03\end{aligned}$$

Aufgabe 6: Bedingter Effekt

Effekt bei minimalem sozioökonomischem Status:

```
# Minimaler soz. Status
min(spf2$HISEI_z)
```

```
[1] -26.98398
```

$$
\begin{aligned}
E[g_1(Z = -26.98)] &= \alpha_2 + \alpha_3 \cdot min(Z) \\
&= 75.0287695 + 0.9649005 \cdot -26.98398 = 48.99
\end{aligned}
$$

Aufgabe 7: EffectLiteR

```
library(EffectLiteR)                # Paket laden
# Modell ohne Kovariate
effectLite(y = "hoeren", x = "schule", control = "0", data = spf2)
```

```
--------------------- Message  ---------------------

 -- model converged succesfully --


--------------------- Variables  ---------------------

Outcome variable Y:  hoeren
Treatment variable X:  schule   (Reference group:  0)

Levels of Treatment Variable X
   X   schule (original)   Indicator
   0                   0       I_X=0
   1                   1       I_X=1


 --------------------- Regression Model ---------------------

 E(Y|X) = g0() + g1()*I_X=1
  g0() = g000
  g1() = g100
```

```
 Intercept Function g0()  [Reference group: 0]

   Coefficient   Estimate       SE   Est./SE   p-value
          g000     263.89   12.422    21.243         0

 Effect Function g1()   [schule: 1 vs. 0]

   Coefficient   Estimate       SE   Est./SE   p-value
          g100     80.518   14.612     5.511         0


--------------------- Cell Counts  ---------------------


Cell Counts

This table shows cell counts including missings.
See also output under lavaan results for number of observations
actually used in the analysis.

schule   0   1

        77 212


--------------------- Main Hypotheses ---------------------

H0: No average effects: E[g1()] = 0

                     Wald Chi-Square   df    p-value
No average effects              30.4    1   3.58e-08


 --------------------- Adjusted Means ---------------------

            Estimate      SE   Est./SE
Adj.Mean0        264   12.42      21.2
Adj.Mean1        344    7.69      44.8


 --------------------- Average Effects ---------------------
```

```
          Estimate     SE   Est./SE    p-value   Effect Size
E[g1()]       80.5   14.6      5.51   3.58e-08         0.739
```

```r
# trad. ANCOVA
effectLite(y = "hoeren", x = "schule", control = "0", data = spf2,
           z = "HISEI_z", interactions = "none")
```

```
--------------------- Message  ---------------------

 -- model converged succesfully --


--------------------- Variables  ---------------------

Outcome variable Y:  hoeren
Treatment variable X:  schule   (Reference group:  0)
Continuous covariates in Z=(Z1): Z1=HISEI_z

Levels of Treatment Variable X
   X   schule (original)   Indicator
   0                   0       I_X=0
   1                   1       I_X=1


 --------------------- Regression Model ---------------------

 E(Y|X,Z) = g0(Z) + g1(Z)*I_X=1
  g0(Z) = g000 + g001 * Z1
  g1(Z) = g100 + g101 * Z1

 Intercept Function g0(Z)  [Reference group: 0]

   Coefficient   Estimate       SE   Est./SE   p-value
          g000    269.433   12.387    21.751         0
          g001      1.939    0.363     5.343         0

 Effect Function g1(Z)   [schule: 1 vs. 0]

   Coefficient   Estimate       SE   Est./SE   p-value
```

```
         g100     72.962   14.378     5.074          0
         g101      0.000       NA        NA         NA


--------------------- Cell Counts  ---------------------


Cell Counts

This table shows cell counts including missings.
See also output under lavaan results for number of observations
actually used in the analysis.

schule   0   1

        77 212


--------------------- Main Hypotheses ---------------------

H0: No average effects: E[g1(Z)] = 0
H0: No covariate effects in control group: g0(Z) = constant
H0: No treatment*covariate interaction: g1(Z) = constant
H0: No treatment effects: g1(Z) = 0

                                        Wald Chi-Square   df    p-value
No average effects                                 25.7    1   3.89e-07
No covariate effects in control group              28.5    1   9.15e-08
No treatment*covariate interaction                   NA    0         NA
No treatment effects                               25.7    1   3.89e-07


 --------------------- Adjusted Means ---------------------

            Estimate      SE   Est./SE
Adj.Mean0        269   12.54      21.5
Adj.Mean1        342    7.51      45.6


 --------------------- Average Effects ---------------------

           Estimate     SE   Est./SE    p-value   Effect Size
E[g1(Z)]         73   14.4      5.07   3.89e-07         0.652
```

```
 --------------------- Effects given a Treatment Condition ---------------------

               Estimate     SE   Est./SE    p-value   Effect Size
E[g1(Z)|X=0]         73   14.4      5.07   3.89e-07         0.652
E[g1(Z)|X=1]         73   14.4      5.07   3.89e-07         0.652
```

```r
# gen. ANCOVA
effectLite(y = "hoeren", x = "schule", control = "0", data = spf2,
           z = "HISEI_z")
```

```
--------------------- Message  ---------------------

 -- model converged succesfully --


--------------------- Variables  ---------------------

Outcome variable Y:  hoeren
Treatment variable X:  schule   (Reference group:  0)
Continuous covariates in Z=(Z1): Z1=HISEI_z

Levels of Treatment Variable X
   X   schule (original)   Indicator
   0                   0       I_X=0
   1                   1       I_X=1


 --------------------- Regression Model ---------------------

 E(Y|X,Z) = g0(Z) + g1(Z)*I_X=1
  g0(Z) = g000 + g001 * Z1
  g1(Z) = g100 + g101 * Z1

 Intercept Function g0(Z)  [Reference group: 0]

   Coefficient   Estimate       SE   Est./SE   p-value
          g000    267.182   12.507    21.362     0.000
```

```
        g001      1.151    0.841     1.369     0.171

 Effect Function g1(Z)   [schule: 1 vs. 0]

   Coefficient   Estimate       SE   Est./SE   p-value
          g100     75.029   14.455     5.191     0.000
          g101      0.965    0.932     1.035     0.301


--------------------- Cell Counts  ---------------------


Cell Counts

This table shows cell counts including missings.
See also output under lavaan results for number of observations
actually used in the analysis.

schule   0   1

        77 212


--------------------- Main Hypotheses ---------------------

H0: No average effects: E[g1(Z)] = 0
H0: No covariate effects in control group: g0(Z) = constant
H0: No treatment*covariate interaction: g1(Z) = constant
H0: No treatment effects: g1(Z) = 0

                                        Wald Chi-Square   df    p-value
No average effects                                26.82    1   2.23e-07
No covariate effects in control group              1.87    1   1.71e-01
No treatment*covariate interaction                 1.07    1   3.01e-01
No treatment effects                              27.05    2   1.34e-06


 --------------------- Adjusted Means ---------------------

            Estimate       SE   Est./SE
Adj.Mean0        267   12.56      21.3
Adj.Mean1        342    7.56      45.3
```

```
 --------------------- Average Effects ---------------------

           Estimate     SE   Est./SE    p-value   Effect Size
E[g1(Z)]         75   14.5      5.18   2.23e-07         0.688


 --------------------- Effects given a Treatment Condition ---------------------

               Estimate     SE   Est./SE    p-value   Effect Size
E[g1(Z)|X=0]       72.3   14.4      5.01   5.41e-07         0.663
E[g1(Z)|X=1]       76.0   14.7      5.18   2.18e-07         0.698
```