

Kausale Effektschätzung - Propensity Score Methods II

Vorbereitung

```
set.seed(12345)
```

Aufgabe 1: PS Matching

Erstellen Sie mit `matchit()` ein PS-Matching-Modell mit additiven Effekten (keine Interaktionen oder Polynome) mit den Kovariaten Geschlecht (`sex`), Vorliebe für Mathe (`matVL`) und dem Englisch-Vortest (`eng_pre`). Verwenden Sie dabei den Nearest Neighbor Approach und ansonsten die Standardeinstellungen.

```
library(MatchIt)
ps <- matchit(group ~ sex + matVL + eng_pre, data = dat,
              distance = "glm", link = "logit",
              method = "nearest")
```

Aufgabe 2

Extrahieren Sie den Datensatz der gematchten Personen, erstellen Sie eine PS-Logit Variable und zentrieren Sie diese am Gesamtmittelwert.

```
psdat <- match.data(ps)
str(psdat)
```

Classes 'matchdata' and 'data.frame': 96 obs. of 8 variables:

```
$ sex      : num  1 1 1 1 1 1 1 1 1 1 ...
$ matVL    : num  0 1 1 1 1 1 1 1 1 1 ...
$ eng_pre  : num  0.74 0.59 0.73 0.65 0.63 0.77 0.61 0.64 0.63 0.65 ...
$ eng_post : num  0.84 0.84 0.9 0.25 0.69 0.9 0.35 0.89 0.67 0.74 ...
$ group    : num  0 0 0 0 0 0 0 0 0 0 ...
$ distance : num  0.282 0.702 0.54 0.636 0.659 ...
$ weights  : num  1 1 1 1 1 1 1 1 1 1 ...
$ subclass: Factor w/ 48 levels "1","2","3","4",...: 10 39 9 20 21 40 7 4 3 19 ...
- attr(*, "distance")= chr "distance"
- attr(*, "weights")= chr "weights"
- attr(*, "subclass")= chr "subclass"
```

```
library(psych)
psdat$logit <- logit(psdat$distance)
library(jtools)
psdat <- center(psdat, var = "logit")
```

Aufgabe 3

Berechnen Sie den ATE und ATT auf Basis des gematchten Datensatzes ohne zusätzliche Kontrolle von PS oder Kovariaten (*t*-Test).

```
m1 <- lm(eng_post ~ group, data = psdat)
coef(m1)
```

```
(Intercept)      group
0.71104167 -0.08833333
```

```
confint(m1)[2,]
```

```
      2.5 %      97.5 %
-0.167053649 -0.009613017
```

$$\widehat{ATE} = \widehat{ATT} = -0.09$$

Aufgabe 4

Berechnen Sie den ATE und den ATT mithilfe einer generalisierten ANCOVA und dem PS-logit als Kovariate.

```
m2 <- lm(eng_post ~ logit*group, data = psdat)
coef(m2)
```

```
(Intercept)      logit      group logit:group
0.69609427 -0.04723870 -0.03998038 -0.05833383
```

```
# Gruppenmittelwerte der Kovariate logit für Berechnung des ATTs ausrechnen
aggregate(logit ~ group, data = psdat, mean)
```

	group	logit
1	0	-0.3164228
2	1	0.3164228

$$\widehat{ATE} = \alpha_2 = -0.040$$

$$\begin{aligned}\widehat{ATT} &= \alpha_2 + \alpha_3 \cdot \text{logit}_{\text{english}=1} \\ &= -0.040 - 0.058 \cdot 0.316 = -0.061\end{aligned}$$

Aufgabe 5: PS-Gewichtung

Berechnen Sie für den vollständigen Datensatz die Gewichte zur Schätzung des ATE durch Propensity-Score Gewichtung.

```
ps2 <- matchit(group ~ sex + matVL + eng_pre, data = dat,
               distance = "logit",
               method = "subclass")
ps2dat <- match.data(ps2)
ps2dat$ATEweight <- ifelse(test = ps2dat$group == 1,
                           yes = 1/ps2dat$distance,
                           no = 1/(1-ps2dat$distance))
summary(ps2dat$ATEweight)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.067	1.380	1.614	2.047	2.190	10.308

Aufgabe 6: ATE

Ermitteln Sie den geschätzten Gesamt-ATE unter Berücksichtigung der unter 5) erstellten Gewichte.

```
wATE <- lm(eng_post ~ group, data = ps2dat, weights = ATEweight)
coef(wATE)
```

(Intercept)	group
0.700729533	-0.003987133

$$\widehat{ATE} = -0.004$$