



PhUSE EU Connect 2018,
Frankfurt

The Propensity Score Matching



RW03

05 November 2018 / Stephanie Fechtner





Agenda

- // Background
- // **Propensity Score Matching**
- // Example – The Framingham Heart Study
 - // Implementation in SAS
 - // Implementation in R
- // Summary and discussion



Background



Background

- // Randomized controlled trial = gold standard for evaluating the effect of an intervention
 - // Balanced sample characteristics (on average)

- // But: what if the application of a randomized controlled trial is not possible or inconceivable?
 - // E.g.: smoking status, inpatient vs. outpatient rehabilitation,...

- // Problem: Observational study → biased estimates due to unbalanced covariates possible

- ➡ One Solution: The Propensity Score Matching



Propensity Score Matching



The Propensity Score E_i

Rosenbaum and Rubin (1983)

$$E_i = P(Z_i | X_i=x_i) = \frac{\exp(x_i\beta)}{1+\exp(x_i\beta)} = \frac{1}{1+\exp(-x_i\beta)} \iff \log[E_i/(1-E_i)] = x_i\beta$$

// Propensity Score = Balancing Score

// Z_i : binary treatment condition

// $Z_i=1 \rightarrow$ patient i in treatment group

// $Z_i=0 \rightarrow$ patient i in control group

// Covariates X_i : Variables which are associated as potential confounders regarding the main analysis

// β : Vector of regression parameters $\rightarrow$ estimated by maximum likelihood estimator

// E_i estimated by a logistic regression model for each patient $\rightarrow$ used for matching

$\rightarrow$ Matching is reduced to finding partners with similar propensity scores

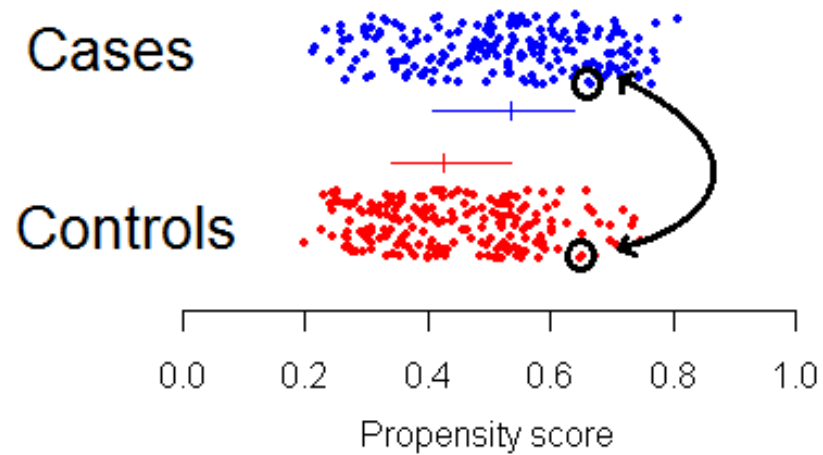


Propensity Score Matching

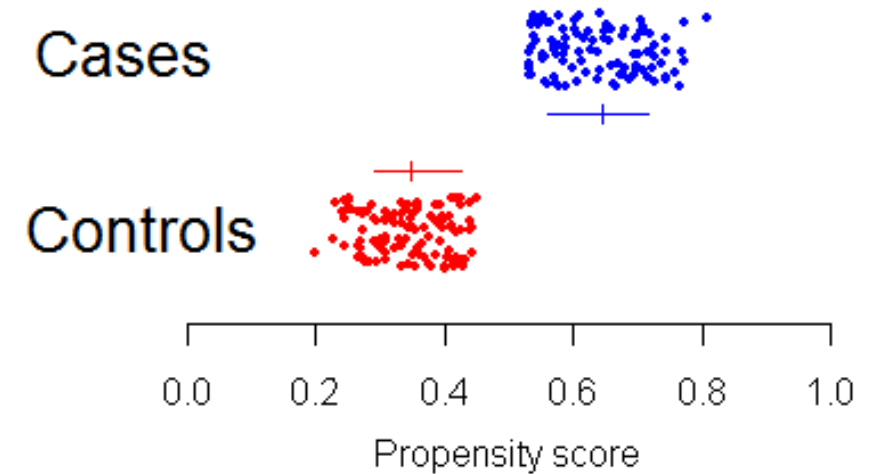
Creation of two groups with similar propensity scores

- // Different matching methods (Nearest neighbor, Exact matching, optimal matching,...)
- // Proportion of cases and controls (1:1, 2:1,...)
- // With or without replacement
- // Caliper (maximal distance of the propensity scores) as an additional restriction
 - // Useable sample size decreases
 - // Similarity of cases and controls increases
 - // Optimal choice = 0.2 of the standard deviation of the propensity scores (based on Monte Carlo simulations performed by Austin (2011))
- // Sufficient overlapping of the propensity scores necessary

Propensity Score Matching



→ Matching possible & useful



→ Matching not useful (or not possible)



Propensity Score matched samples

Characteristics of propensity score matched samples

- // Balanced covariates of cases and controls → Balancing needs to be checked:
 - // (Weighted) descriptive statistics: Mean, median, standard deviation,...
 - // (Weighted) t-tests or (weighted) chi-square tests
 - // Test for balance: Overall balance test

- // Same statistical approaches applicable as those employed in randomized controlled trials

Example

The Framingham Heart Study



The dataset “Heart”

- // Systematic investigation of the population of Framingham (Massachusetts) began in 1948
- // Aim: detection of causes and risks of the Coronary heart disease (CHD)
- // Still ongoing
- // Data of the first 5209 adult patients included in Sashelp dataset „Heart“
- // 17 variables related to CHD

Question: Has smoking a relevant impact on the development of CHD?

Sashelp.heart --- Framingham Heart Study

The CONTENTS Procedure

Variables in Creation Order			
#	Variable	Type	Len Label
1	Status	Char	5
2	DeathCause	Char	26 Cause of Death
3	AgeCHDdiag	Num	8 Age CHD Diagnosed
4	Sex	Char	6
5	AgeAtStart	Num	8 Age at Start
6	Height	Num	8
7	Weight	Num	8
8	Diastolic	Num	8
9	Systolic	Num	8
10	MRW	Num	8 Metropolitan Relative Weight
11	Smoking	Num	8
12	AgeAtDeath	Num	8 Age at Death
13	Cholesterol	Num	8
14	Chol_Status	Char	10 Cholesterol Status
15	BP_Status	Char	7 Blood Pressure Status
16	Weight_Status	Char	11 Weight Status
17	Smoking_Status	Char	17 Smoking Status

The dataset “Heart”

	Smoker (n = 2093)	Non-Smoker (n = 3116)	P value
Sex: female, %	35.93	68.07	< 0.01
AgeAtStart, mean (SD)	42.16 (8.04)	45.35 (8.68)	< 0.01
Chol_Status			0.15
% Borderline	38.20	35.85	
% Desirable	27.80	27.77	
% High	34.00	36.38	
Diastolic, mean (SD)	83.92 (12.33)	86.33 (13.30)	< 0.01
Systolic, mean (SD)	133.50 (20.90)	139.20 (25.21)	< 0.01
BP_Status			< 0.01
% High	38.84	46.66	
% Normal	43.72	39.41	
% Optimal	17.44	13.93	
Weight_Status			< 0.01
% Normal	35.79	23.26	
% Overweight	59.76	73.92	
% Underweight	4.45	2.83	

Baseline characteristics of smokers and non-smokers:

- // Smoker = person who smokes >5 cigarettes per day
- // Non-smoker = person who smokes ≤5 cigarettes per day
- // SD = Standard deviation
- // P values resulting from a t-test or a chi-squared test

→ Investigation of the influence of the smoking status on CHD with `heart` not useful



Application of the Propensity Score Matching

- // Treatment group = Smoker (person who smokes >5 cigarettes per day)
- // Control group = Non-smoker (person who smokes ≤ 5 cigarettes per day)
- // Estimation of propensity scores
 - // Logistic regression model
 - // Nearest neighbor method
 - // With replacement $\rightarrow$ weights will be calculated for each person
 - // Ratio of 2:1 $\rightarrow$ 2 non-smokers for 1 smoker
 - // Caliper = 0.2



The propensity score matching in SAS

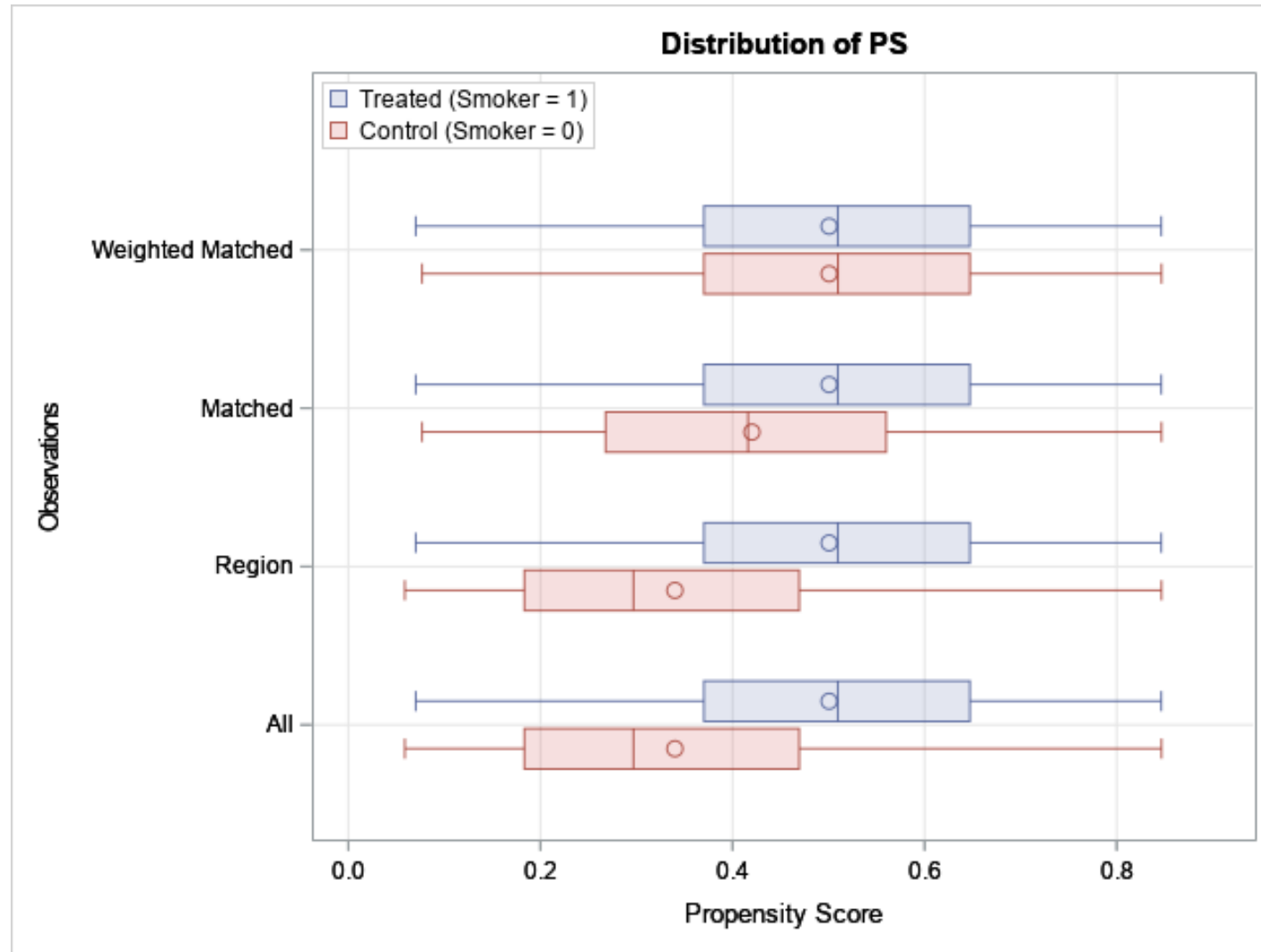
The procedure PROC PSMATCH

- // Define the class variables
- // Psmode1 specifies logistic regression model for computing the propensity scores
- // Match defines the matching method
- // Assess differences in propensity scores (=PS) before and after the matching and create boxplots

```
ods graphics on;
proc psmatch data=heart region=allobs;
  class Smoker Sex Chol_Status BP_Status Weight_Status;
  psmode1 Smoker(Treated='1')= Sex Chol_Status BP_Status Weight_Status AgeAtStart Diastolic Systolic;
  match method=replace(k=2) distance=ps caliper=0.2;
  assess ps / plots=(boxplot);
  output out(obs=match)=m_heart matchid=_MatchID;
run;
ods graphics off;
```

- // m_heart contains the data of the matched samples

The propensity score matching in SAS





The propensity score matching in SAS

	Smoker (n = 2047)	Non-Smoker (n = 1656)	P value
Sex: female, %	35.56	36.47	0.55
AgeAtStart, mean (SD)	42.11 (8.02)	41.87 (9.37)	0.37
Chol_Status			0.054
% Borderline	38.25	35.83	
% Desirable	27.80	31.17	
% High	33.95	33.00	
Diastolic, mean (SD)	83.98 (12.35)	83.74 (13.27)	0.56
Systolic, mean (SD)	133.59 (20.99)	132.48 (22.41)	0.10
BP_Status			0.67
% High	38.89	37.81	
% Normal	43.77	43.92	
% Optimal	17.34	18.27	
Weight_Status			0.66
% Normal	35.86	36.27	
% Overweight	59.65	58.70	
% Underweight	4.49	5.03	

Weighted characteristics of matched smokers and non-smokers:

// Mean = Weighted mean

// SD = Weighted standard deviation

// P values resulting from weighted t-tests or weighted chi-squared tests

→ Investigation of the influence of the smoking status on CHD with `m_heart` useful



The propensity score matching in R

The R-package MatchIt()

- // formula includes the binary treatment indicator and the covariates
- // method defines the matching method
- // distance defines the method used to estimate the distance measure ("logit"=logistic regression)

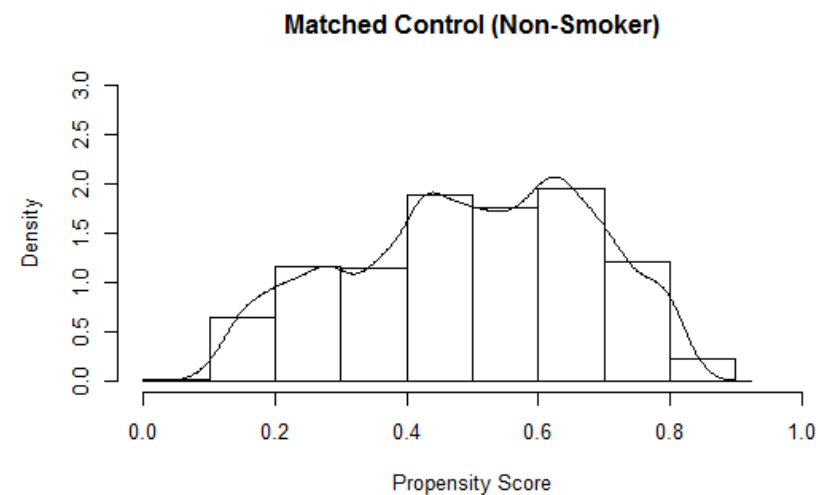
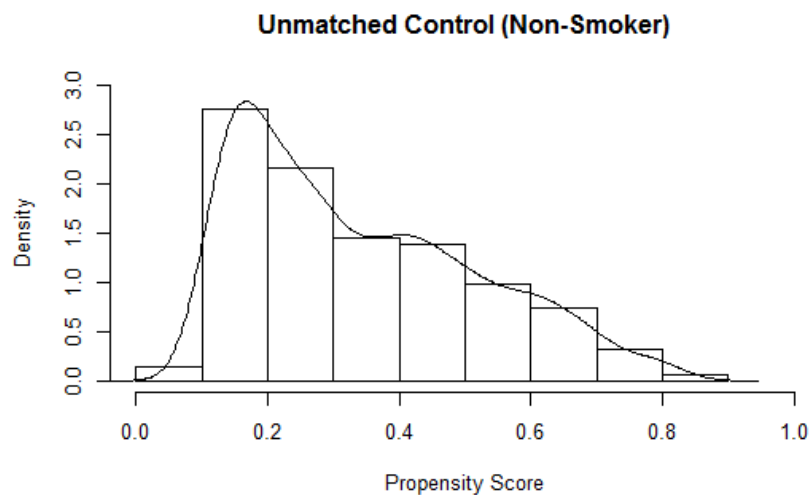
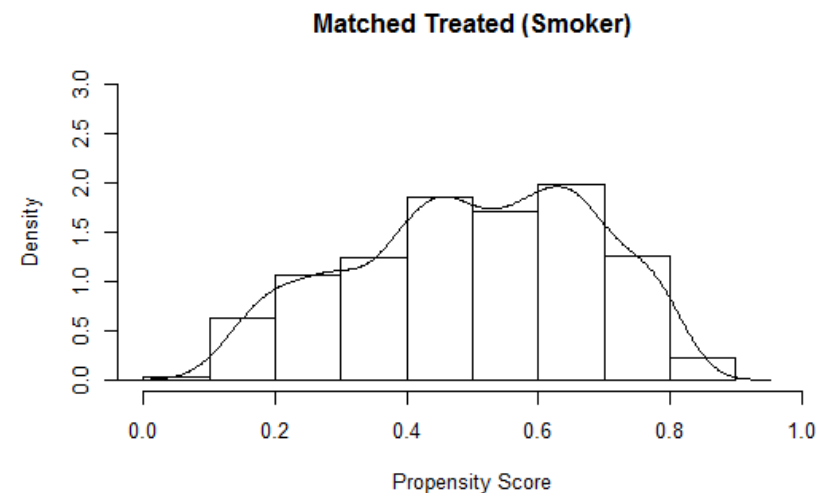
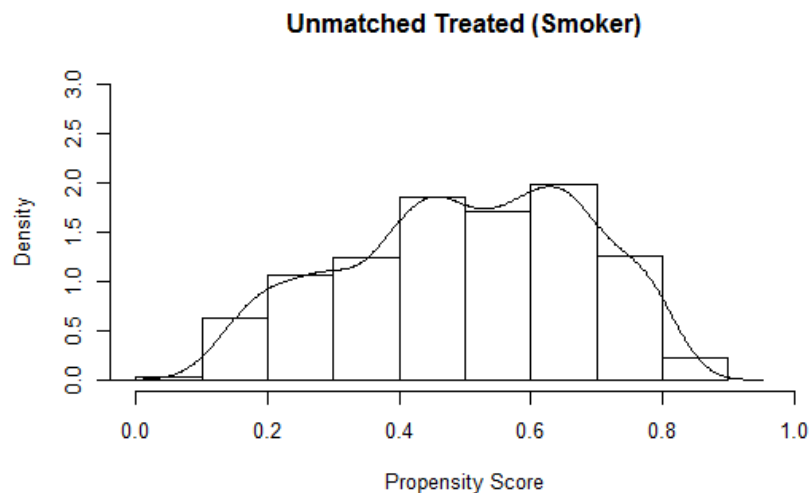
```
library(MatchIt)
formel <- "Smoker ~ as.factor(Sex) + AgeAtStart + Diastolic + Systolic + as.factor(Chol_Status) +
as.factor(BP_Status) + as.factor(Weight_Status)"

match_heart <- matchit(as.formula(formel), method="nearest", distance="logit", replace=TRUE,
ratio=2, caliper=0.2, discard="none", data=heart)

m.all <- cbind(heart, match_heart$distance, match_heart$weights)
m.heart <- subset(m.all, weights!=0)
```

- // Propensity scores (=distance) and weights of the persons saved in match_heart
- // Weights of nonmatched persons = 0 → matched persons saved in dataset m.heart

The propensity score matching in R



The propensity score matching in R

	Smoker (n = 2047)	Non-Smoker (n = 1804)	P value
Sex: female, %	35.56	35.74	0.91
AgeAtStart, mean (SD)	42.11 (8.02)	42.18 (8.40)	0.78
Chol_Status			0.52
% Borderline	38.25	39.28	
% Desirable	27.80	28.51	
% High	33.95	32.22	
Diastolic, mean (SD)	83.98 (12.35)	84.09 (11.91)	0.77
Systolic, mean (SD)	133.59 (20.99)	133.54 (20.00)	0.94
BP_Status			0.76
% High	38.89	39.30	
% Normal	43.77	44.26	
% Optimal	17.34	16.44	
Weight_Status			0.54
% Normal	35.86	34.83	
% Overweight	59.65	59.99	
% Underweight	4.49	5.18	

Weighted characteristics of matched smokers and non-smokers:

// Mean = Weighted mean

// SD = Weighted standard deviation

// P values resulting from weighted t-tests or weighted chi-squared tests

➔ Investigation of the influence of the smoking status on CHD with `m.heart` useful



Summary and discussion

Summary and discussion

Creation of one treatment and one control group that have similar covariate values → subsequent comparisons not confounded by differences in covariate distributions

Advantages	Disadvantages
- A large number of confounders can be considered	- Selective loss of study participants
- No specification of the functional form between covariates and outcome necessary	- Groups are balanced in all observed covariates (not in unobserved)
- Separation of confounder adjustment and outcome analyses	- No estimation of the effects of the confounders or interactions

References

- // Austin PC (2011). *Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies*. Pharm Stat 10: 150-161.
- // Cepeda MS, Boston R, Farrar JT, Strom BL (2003): *Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders*. Am J Epidemiol. 158(3): 280-287.
- // Ho DE, Imai K, King G, Stuart EA (2011). *MatchIt: nonparametric preprocessing for parametric causal inference*. J Stat Softw 42:1–28.
- // NHLBI – National Heart, Lung and Blood Institute (2015). *Coronary Heart Disease*. Bethesda, Maryland.
<https://www.nhlbi.nih.gov/health-topics/coronary-heart-disease> (from 26.09.2018)
- // R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- // Rosenbaum PR, Rubin DB (1983). *The central role of the propensity score in observational studies for causal effects*. Biometrika 70:41–55.
- // SAS Institute Inc. (2017a). *Sashelp Data Sets*. Cary, NC, USA.
- // SAS Institute Inc. (2017b). *SAS/STAT 14.3 User's Guide, The PSMATCH Procedure*. Cary, NC, USA.
- // Zanutto EL (2006). *A comparison of Propensity Score and Linear Regression Analysis of Complex Survey Data*. Journal of Data Science 4: 67-91.



Thank you!



Questions?

