

# Research Methods for Political Science PO3110 (TCD)

HT: Tutorial 3 - Week 4

Letícia Meniconi Barbabela

University College Dublin,  
<https://github.com/letmeni/research-methods>

11-12 February 2020

# Today's topics

- Homework correction and doubts:
  - Difference between t-score and z-score (Ex. 2 d);
  - Post-treatment bias;
- Residuals.

# Homework correction

- Open your homework assignment;
- Have SPSS ready and "Data\_HT02.sav" and "Data HT03 Adams.sav" loaded.

# Post-treatment bias

- Gary King explains that it occurs in two situations:
  - ① when controlling away for the consequences of treatment and;
  - ② when causal ordering among predictors is ambiguous or wrong.
- E.g.: when doing a regression you should control for confounders, that is, variables that affect both the dependent and the independent variable:
  - When a variable is influenced by the independent variable, including it in the model as a control is a case of "post-treatment bias".
  - Results from such a model are unreliable because the estimation of the relationship between the dependent and independent variables is actually affected by the relationship between the control you included and the independent variable;
  - It is a problem because you cannot really separate these effects, that is, the coefficient for the independent variable will be biased.

# Post-treatment bias

- Gary King explains that it occurs in two situations:
  - ① when controlling away for the consequences of treatment and;
  - ② when causal ordering among predictors is ambiguous or wrong.
- E.g.: when doing a regression you should control for confounders, that is, variables that affect both the dependent and the independent variable:
  - When a variable is influenced by the independent variable, including it in the model as a control is a case of "post-treatment bias".
  - Results from such a model are unreliable because the estimation of the relationship between the dependent and independent variables is actually affected by the relationship between the control you included and the independent variable;
  - It is a problem because you cannot really separate these effects, that is, the coefficient for the independent variable will be biased.

# Post-treatment bias

- Gary King explains that it occurs in two situations:
  - ① when controlling away for the consequences of treatment and;
  - ② when causal ordering among predictors is ambiguous or wrong.
- E.g.: when doing a regression you should control for confounders, that is, variables that affect both the dependent and the independent variable:
  - When a variable is influenced by the independent variable, including it in the model as a control is a case of "post-treatment bias".
  - Results from such a model are unreliable because the estimation of the relationship between the dependent and independent variables is actually affected by the relationship between the control you included and the independent variable;
  - It is a problem because you cannot really separate these effects, that is, the coefficient for the independent variable will be biased.

# Post-treatment bias

- Gary King explains that it occurs in two situations:
  - ① when controlling away for the consequences of treatment and;
  - ② when causal ordering among predictors is ambiguous or wrong.
- E.g.: when doing a regression you should control for confounders, that is, variables that affect both the dependent and the independent variable:
  - When a variable is influenced by the independent variable, including it in the model as a control is a case of "post-treatment bias".
  - Results from such a model are unreliable because the estimation of the relationship between the dependent and independent variables is actually affected by the relationship between the control you included and the independent variable;
  - It is a problem because you cannot really separate these effects, that is, the coefficient for the independent variable will be biased.

# Post-treatment bias

- Gary King explains that it occurs in two situations:
  - ① when controlling away for the consequences of treatment and;
  - ② when causal ordering among predictors is ambiguous or wrong.
- E.g.: when doing a regression you should control for confounders, that is, variables that affect both the dependent and the independent variable:
  - When a variable is influenced by the independent variable, including it in the model as a control is a case of "post-treatment bias".
  - Results from such a model are unreliable because the estimation of the relationship between the dependent and independent variables is actually affected by the relationship between the control you included and the independent variable;
  - It is a problem because you cannot really separate these effects, that is, the coefficient for the independent variable will be biased.



# Post-treatment bias

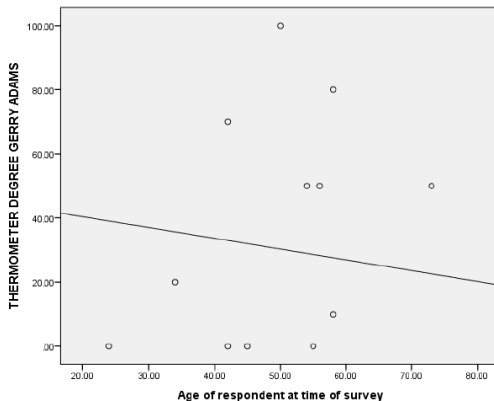
- Gary King explains that it occurs in two situations:
  - ① when controlling away for the consequences of treatment and;
  - ② when causal ordering among predictors is ambiguous or wrong.
- E.g.: when doing a regression you should control for confounders, that is, variables that affect both the dependent and the independent variable:
  - When a variable is influenced by the independent variable, including it in the model as a control is a case of "post-treatment bias".
  - Results from such a model are unreliable because the estimation of the relationship between the dependent and independent variables is actually affected by the relationship between the control you included and the independent variable;
  - It is a problem because you cannot really separate these effects, that is, the coefficient for the independent variable will be biased.

# Difference between t-score and z-score

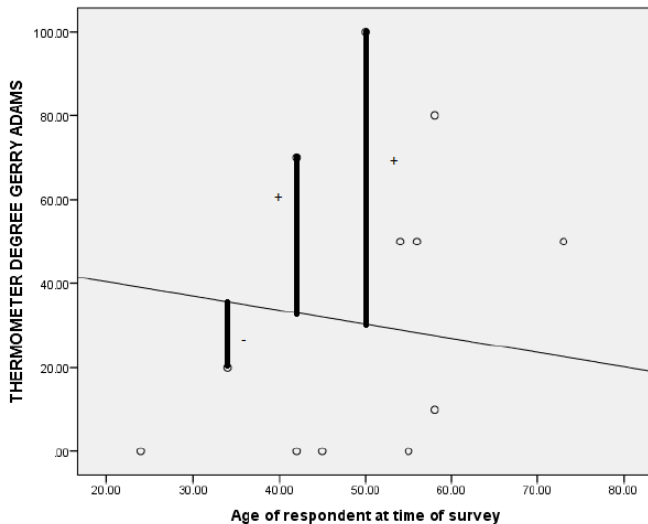
- Both have a similar application to hypothesis testing and constructing a CI;
- T-score is used when sample is below 30 and standard deviation of sample is unknown;
- In samples larger than 30 the t-distribution resembles the z-distribution;
- For 95% confidence level we know the z-score is 1.96;
- For the t-score, we need to know the degrees of freedom ( $n-1$ ) and check a t-table to figure out what the value is.

## What are residuals?

The residuals are defined as the deviance between the observed and the predicted values. The graph below displays a selection of cases from the dataset and the regression line:



# Residuals for individual observations



# Calculating the Residual Sum of Squares

ANOVA<sup>a</sup>

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	29045.763	1	29045.763	31.650	.000 <sup>b</sup>
	Residual	942504.215	1027	917.726		
	Total	971549.979	1028			

a. Dependent Variable: THERMOMETER DEGREE GERRY ADAMS

b. Predictors: (Constant), Age of respondent at time of survey

Formula:

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad TSS = \sum_{i=1}^n (y_i - \bar{y})^2$$

$$R^2 = 1 - \frac{RSS}{TSS}$$

# References

- Field, A (2013) *Discovering Statistics Using SPSS*. 4th edition. London:Sage
- HT 2019 Slides at <http://andrsalvi.github.io/research-methods>