



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



PHD COURSE  
IN STATISTICS

# Specialist Course | Cycle XXXIII

June, 2018 | Campus S. Caterina

## *An Introduction to Missing Data Handling*

**Omar Paccagnella | University of Padova**

<b>Monday</b>	<b>June 25</b>	<b>10.30 – 13.30</b>	<b>Room Cucconi</b>
<b>Tuesday</b>	<b>June 26</b>	<b>10.30 – 13.30</b> <b>14.30 – 16.00</b>	<b>Room Cucconi</b> <b>Asid 17</b>
<b>Wednesday</b>	<b>June 27</b>	<b>10.30 – 13.30</b> <b>14.30 – 16.00</b>	<b>Room Cucconi</b> <b>Asid 17</b>
<b>Friday</b>	<b>June 29</b>	<b>10.30 – 13.30</b>	<b>Room Cucconi</b>

[www.stat.unipd.it/fare-ricerca/courses-201718-xxxiii-cycle](http://www.stat.unipd.it/fare-ricerca/courses-201718-xxxiii-cycle)

[www.stat.unipd.it/ricerca/dottorato-di-ricerca](http://www.stat.unipd.it/ricerca/dottorato-di-ricerca) | [phd@stat.unipd.it](mailto:phd@stat.unipd.it)

# Omar Paccagnella | University of Padova

## *An Introduction to Missing Data Handling*

### **Aims of the module:**

Empirical researches are often characterized by missing data problems. Missing data occur when data values are not available (for different reasons) and this may be found in all stages of the research or survey collection.

This short course aims at introducing students to the basic concepts of the missing data problems (indeed the nature of the missing data mechanism plays a key role in any subsequent analysis), as well as to the main statistical solutions for handling missing data. This course is mainly focused on practical consequences of the missing data problem.

### **Contents:**

- An introduction to missing data
- Missing data mechanisms: missing completely at random (MCAR), missing at random (MAR), missing not at random (MNAR).
- Traditional methods for dealing with missing data:
  - o Listwise Deletion
  - o Pairwise Deletion
  - o Mean Imputation
  - o Regression Imputation
  - o Regression Imputation with Imputing Draws
- Matching procedures (Predictive Mean Matching, Hot-Deck Imputation)
- Likelihood-based approaches to the analysis of missing data
- Multiple Imputation
- A brief introduction to models for MNAR data

### **Main references:**

Enders C.K. (2010) *Applied Missing Data Analysis*, New York, NY, The Guilford Press.

Little R.J.A. and Rubin D. B. (2002) *Statistical analysis with missing data* (2nd ed.), Hoboken, NJ, John Wiley & Sons.

Rubin D.B. (1976) Inference and missing data, *Biometrika* 63, 581–592.

Rubin D.B. (1987) *Multiple Imputation for Nonresponse in Surveys*, New York, NY, John Wiley & Sons.

Schafer J.L. (1997) *Analysis of Incomplete Multivariate Data*, London, UK, Chapman & Hall.

van Buuren S. (2012) *Flexible Imputation of Missing Data*, Boca Raton, FL, CRC Press.