Proposals for BSU Third International Week

Tomasz Przechlewski (Kwidzyn/Poland)
tprzechlewski@gmail.com

#1: Quality statistical charts: introduction to data visualisation

The last stage of any serious statistical analysis is publication of the results. This–often neglected or marginally covered at most courses on statistics–stage is of a vital importance, as poor charts are prone to be misunderstood (infamous 'mutant numbers'), or worse not attract public attention at all (although they are interesting and/or important). Without doubt quality publication is a key to successful promotion.

As nowadays pictures become ubiquitous means of disseminating the results the skills to create clear and high quality statistical charts seems to be of the uttermost importance. Yet in reality one can easily find infinite examples of bad graphics, often produced by renown publisher (Economist journal for example, cf https://www.economist.com/graphic-detail/).

It this lecture I will show how to produce quality statistical graphics. I will cover dot plots, bar charts, lineplots, histograms, kernel density estimates, stripcharts, multipanel displays, scatterplots and some others.

Time: 1.5hr or more

#2: Introduction to reproducible research or how to make statistics more meaningful

I will introduce a paradigm of reproducibility of research (RR)–modern and popular approach to statistical analysis based on self-documenting statistical computing (SDS) concept. I will try to demonstrate that abandoning traditional trio of spreadsheet—editor—presentation program (Excel–Word–PP or Calc–Writer–Impress) and switching to some modern tools for statistical analysis is a feasible way to go, as (some) modern tools are not much more difficult that office software while offers greater productivity.

Using popular OpenSource Statistical Software: R and Rstudio I will demonstrate how to put RR theory into everyday practice.

Time: 1.0hr or more

#3: Statistics from the (true) beginning

In the lecture it is argued that one cannot properly understand a statistics unless one knows the details of the process by which the numbers came into being. It is furthe argued that statistics are imperfect (due to fanciful concept definitions, sloppy measures, poor samples, and erroneous computations) and one have to recognize their limitations. Yet very often courses on statistics (for economics and other social sciences in particular) envision it as a branch of mathematics, ie. they concentrate on theory, mathematical formulas in particular, software commands to compute measure, and (to a lesser extent) provide some guidelines how to interpret the results. Lecturers usually focus on clarifying the theory and computational complexities while students consider it as boring and difficult.

As the theory is complicated considerations of how real-life statistics came into being (defining concepts, measurement, missing data, etc) are usually downplayed, ie. students are presented with clean and small ('artificial' or toy) datasets.

I argue that to properly understand a statistics one should start from true beginning (defining concepts, measurement, missing data, classifications, etc), and I will illustrate my point with several real-world examples. Data from EuroStat repository (https://ec.europa.eu/eurostat/data/database) will be explained and used.

Time: 1.0hr