

Statistical Literacy Workshop

What is Statistics?

The scientific method is central to experimental research. It provides a suite of tools that allows us to gain knowledge of the world from acquired data. It is only via the *interpretation* of what the data *means* that we are able to make any meaningful conclusions about the real world.

Statistics describes the process of drawing conclusions from data and understanding the uncertainty of those conclusions. Seen in this light, statistics is essential to every stage of the scientific method - from *collecting* and *describing* data to *inferring* something about the real world. Unfortunately, it is often misunderstood and misused.

For further details, visit our web-site at www.science-craft.com.

This workshop is instructed by either Dr. Rick Scavetta or Dr. Irina Czogiel. Rick can be contacted directly via email at rick.scavetta@science-craft.com.

What is Statistical Literacy?

The often over-heard question "*What test should I use?*" reflects a poor understanding of statistics as a static & authoritative black box that is applied *after* results are obtained. It is essentially statistical illiteracy, a failure to properly *understand* and *use* statistics as a part of scientific inquiry. The *Statistical Literacy* workshop aims to provide the foundation to become statistically literate.

This is similar to how an experimental biologist needs to have an intuitive feel for the nuances of an experimental method. Without understanding the limitations of a method, results cannot be properly interpreted; nor can problems be identified and resolved. Results are accepted at face value without considering where they come from.

The *Statistically Literacy* workshop does *not* focus on providing a flow-chart of statistical tests which students can mindlessly follow.

What will be covered?

The workshop consists of three sections: *Collecting*, *Describing* & *Inferring*. Practical tools and the theoretical background to understand how they work will be introduced. Mathematical equations will be used when helpful and are provided in the reference book, but they will not be the focus of the workshop. Our goal is to enlighten via intuitive understanding, not confuse via math.

A short list of workshop topics is provided below.

1. *Collecting*

- Defining hypotheses
- Random sampling
- Sampling bias
- Systematic error versus random error
- Understanding the scientific process and the role of statistics
- The process of conducting statistics, which acts as a guide for the rest of the workshop

2. Describing

- Understanding the different types of data scientists will encounter
- Different measurements for location and spread
- Matching location and spread measurements

3. Inferring

- Commonly used probability distributions
- Estimation and Hypothesis testing
- Test statistics and p-values
- Unifying concepts underlying different statistical tests
- Variations of the t-test (one and two-sample, plus different conditions)
- Fundamentals of linear regression
- ANOVA and χ^2 test (time permitting)
- Non-parametric tests

What Learning Instruments will be used?

This workshop will engage students on several levels. We will make use of web-based interactive applications, various data-sets, animations, a reference book, quizzes, in-class discussions and research scenarios.

Who is the Target Audience?

This workshop is targeted towards early-stage doctoral students in the life sciences. It is not possible to cover all aspects of statistics in two days, nor go into a deep discussion of specific topics. Rather, this workshop is meant as a refresher for university-level introductory statistics courses. It is not a comprehensive survey of all techniques students may encounter, nor is it a custom consultation service to advise students on their specific research questions.

Preparations (What to bring to the workshop)

Participants should bring their own laptops (or arrange for a loan) and have the following cross-platform software pre-installed:

R – v3.2 or later (www.r-project.org)

RStudio – v0.99 or later (www.rstudio.org)

We will not be programming in R. RStudio will be used to demonstrate how to conduct statistical tests in R if time permits. However, students should be equipped with laptops to use the interactive web apps.

Prerequisites (What you require for participation)

There are no prerequisites for this workshop. It is intended for young doctoral students in the life sciences and assumes no prior working knowledge of statistics or R.