

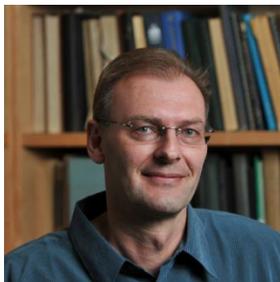
Monday Workshops

Introduction to Statistics

Statistics is the science of learning from data. This workshop will help you to develop the skills you need to analyze data and to communicate your findings. There won't be many formulas in the workshop; rather, we will develop the key ideas of statistical thinking that are essential for learning from data.

We will discuss the main tools for descriptive statistics which are essential for exploring data, with an emphasis on visualizing information. We will explain the important ideas about sampling and conducting experiments. Then we will look over some important rules of probability and discuss normal approximation and the central limit theorem. We will show you the important concepts and pitfalls of regression and how to do inference with confidence intervals and tests of hypotheses. You will learn how to analyze categorical data and discuss one-way analysis of variance. Finally, we will look at reproducibility, data snooping and the multiple testing fallacy, and how to account for multiple comparisons. These issues have become particularly important in the era of big data.

Broadly, there are three main reasons why statistical literacy is essential in data science: First, it provides the skills to assess whether the data are sufficient to answer the questions at hand. Second, it establishes a rigorous framework for quantifying uncertainty. And finally, it provides techniques for effectively communicating the findings of your analyses. This workshop equips you with the important tools in all of these areas. It is the statistical foundation on which the recent exciting advances in machine learning are built.



About the Instructor: Professor Guenther Walther studied mathematics, economics, and computer science at the University of Karlsruhe in Germany and received his Ph.D. in Statistics from UC Berkeley in 1994. His research has focused on statistical methodology for detection problems, shape-restricted inference, and mixture analysis, and on statistical problems in astrophysics and in flow cytometry. He received a Terman fellowship, a NSF CAREER award, and the Distinguished Teaching Award of the Dean of Humanities and Sciences at Stanford. He has served on the editorial boards of the Journal of Computational and Graphical Statistics, the Journal of the Royal Statistical Society, the Annals of Statistics, the Annals of Applied Statistics, and Statistical Science. He was program co-chair of the 2006 Annual Meeting of the Institute of Mathematical Statistics and served on the executive committee of IMS from 1998 to 2012.