ARVO EDUCATION COURSE

Saturday, May 2, 2020
8am – 12pm
Baltimore Convention Center
Baltimore, MD

The nuts and bolts of data analysis for vision research

Organizers

Alison Abraham, PhD, MS, MHS, Wilmer Eye Institute, Johns Hopkins School of Medicine (Lead) and Maureen G. Maguire, PhD, FARVO, Ophthalmology, University of Pennsylvania

Description

Have you ever wondered what to do with both eyes in your dataset? Or how to obtain an average visual acuity or refractive error? Data from vision research can stymie researchers trying to make the most out of the hard work of collecting it. This course will teach simple and accessible tools for understanding and appropriately analyzing data with a focus on data exploration, two sample comparisons and data presentation.

We will discuss ways to include both eyes in studies/experiments and in analysis and show the harm when the correlation is not accommodated appropriately. We will examine data from vision function tests to understand the statistical problems that arise and work through practical examples of how to handle the non-normally distributed or bounded data that are common in vision research. The course will introduce analytic tools and give researchers information on when more advanced methods are needed. Lastly, we will show how data visualization methods and techniques can be leveraged to make research presentations more accessible and impactful.

Following the course, attendees will understand the challenges and appropriate handling of vision research data, choose appropriate statistical tools to make the most of the information, and be able to present the data and results in a meaningful and accessible way.

Prerequisite: Attendees should be familiar with basic statistical analysis techniques. Advanced training in statistics is not required.

Learning objectives

After attending this course, participants will be able to:

- Discuss the measurements most often used in vision research and how to analyze them appropriately
- Explain the problems that arise with improper handling of ocular data
- Perform simple statistical analysis for the various type of correlated ocular data including continuous data (normal or skewed distributed), binary data or ordinal data
- Use graphical and statistical tools to understand vision data and chose the best vision metric for your question
- Use data visualization best practices to improve the impact and accessibility of data reports and presentations

Agenda

Presenters and presentations may change.

8 – 8:10am Welcome: Course aims and introductions

Alison Abraham, PhD, MS, MHS, Wilmer Eye Institute, Johns Hopkins University, Baltimore, MD

The presentation will introduce the course, speakers and content. We will discuss the aims of the course, the outline of the content that will be covered including introductions to the speakers, and the intended audience. This short welcome session will prepare attendees with an overview and familiarize them with the speakers.

8:10 – 8:40am Session 1: Review of commonly used measures in the study of eye disease and health and how to do simple arithmetic

Maureen G. Maguire, PhD, FARVO, University of Pennsylvania, Philadelphia, PA

We will review some commonly used measures of visual function and ocular morphology and the meaning of the measurement scales. Limitations of the measurement scales will be highlighted and some conventions for handling the limitations will be presented. Methods for calculation of summary statistics and for comparison of the measurements between two groups will be provided. Measures to be covered include visual acuity (LogMAR, Snellen fractions, ETDRS letter scores), refractive error (spherical equivalent, vector notation), ETDRS scale for diabetic retinopathy, and retinopathy of prematurity (location, extent, severity, plus disease).

8:40 – 9:20am Session 2: The basics of correlated data: continuous data

Bernard Rosner, PhD, Harvard Medical School and School of Public Health, Boston, MA

Virtually all presentations of continuous data in published work includes estimates of means and standard deviations (or standard errors) for one or more groups. However, if one uses the eye as the unit of analysis it is not straightforward as to how to do this. If one uses standard software (e.g., PROC MEANS OF SAS), then standard errors will be over-estimated. In addition, if some subjects present information on 2 eyes and others on a single eye then both the mean and the standard deviation will not be efficiently estimated with standard software. Similar issues arise in estimating correlation coefficients between two continuous eye-specific variables or estimating regression coefficients predicting one eye-specific variable as a function of one or more person- or eye-specific variables.

Thus, the goals of this lecture are the following:

- estimation of means and standard deviations of eye-specific variables both for (a) balanced designs (2 eyes per person) and for (b) unbalanced designs (one or two eyes per person).
- estimation of Pearson and rank correlation coefficients between two-eye-specific variables
- Using the two-sample t-test for comparing means of a continuous eye-specific variable between two groups
- Using the Wilcoxon rank sum test to compare average ranks of a continuous non-normally distributed variables between two groups.
- Using simple linear regression to assess the association between an eye-specific outcome variable and one or more person- or eye-specific covariates.
- Examples of all these procedures will be provided based on real ophthalmologic data.

9:20 – 10am Session 3: The basics of correlated data: binary/categorical data

Xiangrong Kong, PhD, Wilmer Eye Institute, Johns Hopkins University, Baltimore, MD

The goal of this presentation is to introduce fundamental statistical concepts, methods and computational tools in describing and analyzing correlated binary data commonly encountered in vision research. The talk will be organized by different research questions that can arise in data involving correlated binary/categorical measurements, including: to quantify

the level of correlation (association) between two binary variables, to test association between two paired binary variables, and to estimate the associations of exposure variables with a correlated binary outcome. Analysis of categorical data (more than 2 categories) will also be introduced.

Participants are expected to fulfil the following learning objectives:

- To have a better understanding of the concepts of different measures of association for categorical data and their differences and relationships.
- To be able to interpret and use correctly the odds ratio, risk ratio, or risk difference, i.e. the measures of association.
- To have a basic knowledge of the statistical method and software tool that can be used to compare a binary variable between eyes, accounting for the correlation between the two eyes.
- To have a basic knowledge of the statistical regression modeling frameworks and software tools that can be used to identify exposure variables associated with a binary outcome measured at eye level.

10 – 10:20am Morning break

10:20 – 11am Session 4: The basics of correlated data: diagnostic data

Gui-shuang Ying, PhD, School of Medicine, University of Pennsylvania, Philadelphia, PA

This presentation will start with general basic statistics for evaluating the performance of a diagnostic test for a disease using sensitivity, specificity, and area under receiver operating characteristic (ROC) curve etc. The presentation will then describe statistical analysis approaches for evaluating the performance of a diagnostic test/imaging for ocular disease when the test/imaging is performed in both eyes of a subject either cross-sectionally or longitudinally. The presentation will cover statistical issues on how to deal with unables for ocular test/imaging, the inter-eye correlations of test results from both eyes of a subject, the situation that requires analysis at subject level or at ocular level for a ocular test that provides either continuous measure or categorical measure. The presentation will demonstrate the analyses, with SAS statistical software, for the real studies of ocular testing/imaging for screening ocular diseases or vision problems.

11 – 11:40am Session 5: Visualizing data and grabbing attention: how to broaden your audience and send them away without eye strain

Jiangxia Wang, MS, MA, Department of Biostatistics, Johns Hopkins University, Baltimore, MD

The presentation will show how data visualization methods and techniques can be leveraged to make research presentations more accessible and impactful.

We will discuss:

- Data visualization principles and use data visualization best practices
- Pitfalls of poor data presentations and ways in which graphics can misrepresent data
- Design principles related to maximizing visibility for all audiences
- Examples of poor design to highlight the impact of various design decisions on the message conveyed from tables and figures.

- Following the session, attendees will have better awareness of basic principles of data visualization and be better equipped to present data and results to maximize information transfer and minimize confusion and obfuscation.

11:40am – 12pm Discussion