

IMAGES IN SURGERY

Modules in Statistics

Welcome to the modules in statistics session. Every Tuesdays, we shall be posting a module on Statistics. There are 18 modules in the session, and we hope to post each module every Tuesday at 7am.

The modules are repository on Dropbox, and the link to each will be posted on Telegram. Just click the link to access the materials, which shall be in PDF format.

The Modules:

1. Before starting... Conceptualizing, literature searches and software programs
2. Sample size considerations
3. Formulating your questionnaire
4. Collecting your data..... Sampling techniques
5. Entering data.... Defining variables and entering the data
6. Choosing the appropriate statistics
7. Preparing data for analysis.... Recoding variables, computing new variables, creating new time variables etc
8. Preparing data for analysis II..... Exploring data with univariate analysis
9. The Chi Square tests
10. The Student T-tests.
11. ANOVA
12. ANCOVA
13. Linear Regressions
14. Logistic Regression
15. Survival Analysis
16. Reliability analysis
17. Factor Analysis
18. Final thoughts

Mode.... Simple explanation of the basic principles, which will be illustrated by examples from SPSS.

Caveat.... These modules will not replace proper statistics texts, but they will serve to introduce you to statistic procedures and take you through the steps involved in writing your thesis.

Thank You
Professor Oluwadiya KS

Module I:

Before starting: Conceptualizing, literature searches and software programs

Conceptualizing

Conceptualization is the process of searching several sources of information to identify a good research idea. That is, formulating an answerable research question that is worth investigating. Components of this process include conducting a thorough search of the peer-reviewed literature looking for research topics, justification, methodology and study design. It involves finding a research mentor and other collaborators, and assessing feasibility of the study

How Do you Find “The Right Topic”?

There is no short cut to getting the right topic. A lot depends on both your energy and imagination. You also must take your interest -short and long term- into consideration.

Your Energy: By this I mean your investment in knowledge seeking. How often do you do literature search? Without broadening your knowledge base, your options for exciting research topics will be limited. Browse through PubMed abstracts and save the ones that interests you. Get the full articles to go through them, focusing on the objectives of the articles, their methodologies and the limitations. Often, the limitations give you the opportunity to explore the same research topic in a new dimension. On the other hand, if you consider the methodology flawed, then you may design your own project with the aims of correcting the flaw. So, reading the literature is a good springboard to finding the right topic.

Other important things to do:

- Attend weekly departmental seminars in your field.
- Go to seminars and conferences at your school
- Read other thesis in your faculty

What about imagination? Sometimes, your new topic is right in front of you, but you are missing it because you fail to grasp the uniqueness of the situation. Let me explain with an actual experience. On returning from a vacation, my fellow consultant in my unit told me about an RTC that occurred two weeks earlier in which 9 out of 18 passengers sustained 11 posterior hip fractures. I asked him if he was going to report it, he said it didn't occur to him.

Why?

Because he failed to see that an RTC where 50% of the patients sustained identical injuries is a unique situation. On further investigations, we found that the vehicle was modified to include an extra row of seat, and that it was overcrowded at the time of the crash. This means that the distance between the seat rows are shorter and passengers had to sit with the thighs closely adducted: Ideal conditions for posterior hip dislocations. The paper was

accepted in record time by one of the best journals in the field. Without my imagination, this golden opportunity would have been missed!

Give wings to your imagination

Of course, you must also take your interests into considerations. What topics excites you. Which ones fire your imagination? What is the "in thing" in your Specialty? What are the sustainable areas of research in your specialties? In these, you should make use of the experience, knowledge and skills of your supervisor and other senior members of your faculty. Interaction with your consultants and your fellow students is where a lot of your ideas may come from

Talk, Talk, Talk!

Question, question, question: Ask questions!

Literature Searches

Since the last two decades, you cannot conduct a good literature search without the internet. As surgical resident who is preparing to write his or her dissertation.

As a rule, start your search in Google Scholar. Google Scholar has come a long way since its inception and it has grown to be a reliable source for scholarly articles.

For more information, see <https://scholar.google.com/intl/en/scholar/help.html>

PUBMED

The Health Sciences Library provides access to MEDLINE via the PubMed search system from the National Library of Medicine.

- PubMed covers journal articles in medicine, nursing, dentistry, veterinary medicine, and the health care system.
- It has information about journal articles published in 5,200 journals in 30 languages dating back to 1949.
- PubMed does not include information about meeting abstracts, conference proceedings, dissertations, patents, or websites.

For more information, see <http://web.wilkes.edu/elearning/pharmacy/tutorial/topic1-01.htm>

Research toolkits: Tools to make your research easier

Reference Managers

Effective use of reference managers reduces one of the greatest chores in manuscript writing - formatting references. **EndNote** is the most popular of them all, and it is covered in Chapter 17. EndNote also has a web based version, which you can use for free if your institution is registered with HINARI.

Zotero (<http://www.zotero.org/>) is a web based alternative to EndNote. According to the description on its website, Zotero is a free, easy-to-use tool to help you collect, organize,

cite, and share your research sources. It lives right where you do your work—in the web browser itself”. Previous versions were available only as Firefox add-on, but the latest version, 3.0 is available both as a standalone program and a firefox add-on.

Mendeley

Statistic Programs

- i. **Ease of use.....** SPSS, Epiinfo, Graphpad, Medcalc and Minitab emerge as the best.
- ii. **Learning curve.....**SPSS, Epiinfo, Graphpad, and Medcalc the easiest to learn
- iii. **Depth of point and click procedures.....** SPSS is the clear leader on this criterion
- iv. **Range, quality, and ease of use of statistical procedures offered.....** R, SAS, and Stata, in that order, offer the widest ranges of statistical procedures followed by SPSS
- v. **Modifiability of analytical output specifications....**SPSS seems to be the best
- vi. **Range of graphical output offered....**R is the best, but learning curve is steep.
- vii. **Speed of handling large data sets.....**SAS is the best.
- viii. **Ease and flexibility of data importation** R has best data importation ability, followed by SAS, SPSS, and Stata
- ix. **Ease of results exportation....**SPSS is by far the best
- x. **Thoroughness and interpretability of results output**No clear leader, but Graphpad is the easiest to understand because it also offers explanations.
- xi. **Ease and flexibility of data set manipulation.....**SPSS is the best
- xii. **Pricing for individuals** Epiinfo and R are free
- xiii. **Thoroughness of documentation....**Stata is the best

For more info: <http://www.prostatservices.com/statistical-consulting/articles-of-interest/a-review-of-the-top-five-statistical-software-systems>

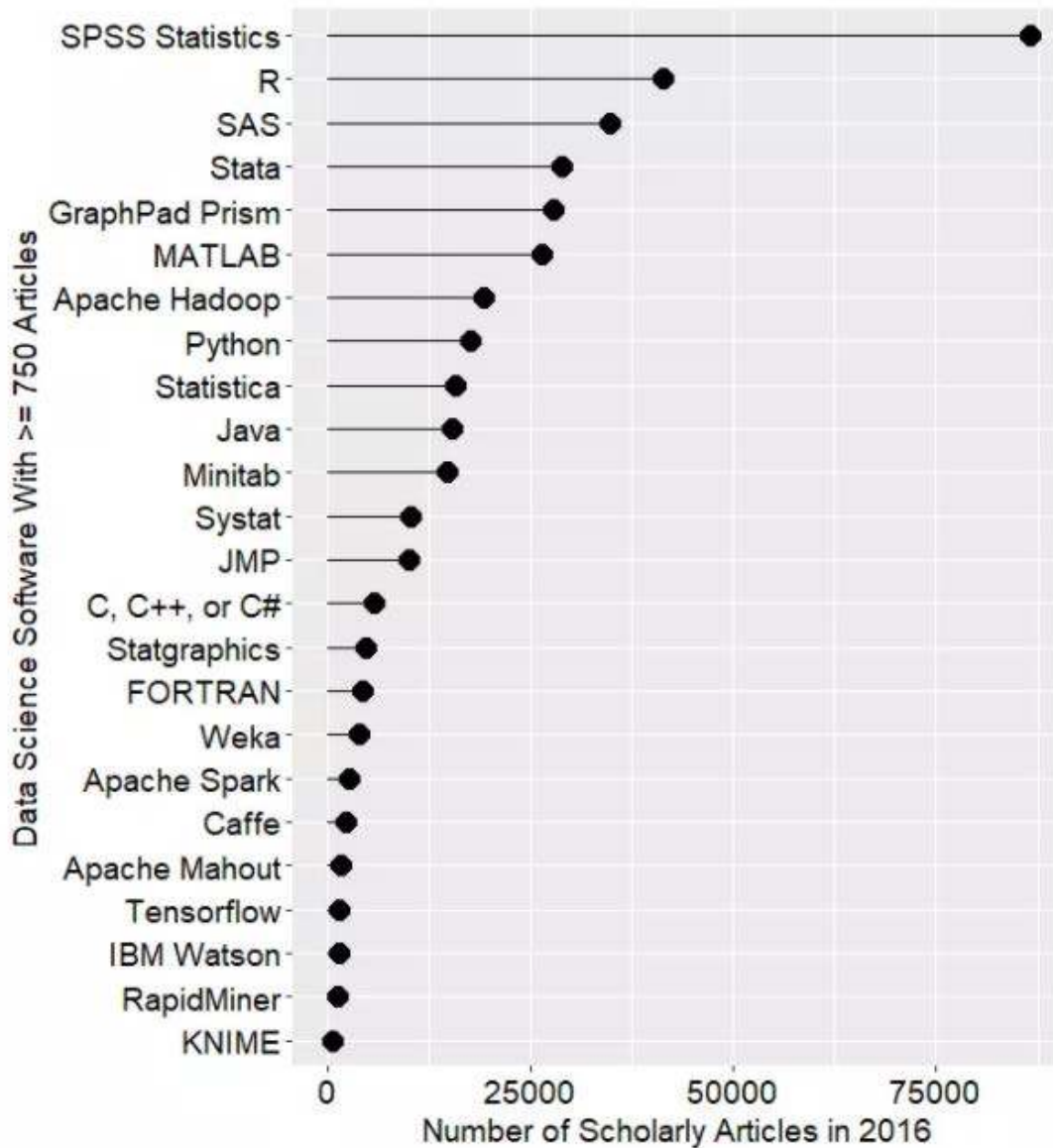


Figure 1: Popularity of Statistics software based on numbers of scholarly articles in 2016

Top free Statistics Software packages

- i. R
- ii. Epi Info: A completely free statistical software for epidemiology developed by Centers for Disease Control and Prevention (CDC). It has a GUI user interphase
- iii. PSPP: free alternative to SPSS.
- iv. GNU Octave: This tool presents an excellent alternative to Matlab.
- v. Regressit: A completely free add-in for Excel, Regressit can be used for multivariate descriptive data analysis and multiple linear regression analysis.

- vi. JASP: free and open-source graphical program for statistical analysis, designed to be easy to use, and familiar to users of SPSS. Additionally, JASP provides many Bayesian statistical methods.
- vii. WinPEPI: Can be downloaded free from <http://www.brixton-health.com>. It is different from most other statistic programs because it is not a spreadsheet. It only takes tabular data. For example, you can input the figures in a crosstab table into WinPEPI and it will give you not only the chi-square, but the Odds Ratio and the Relative Risk as well

Online (web based) Software Packages

- i. [OpenEpi Version 2.2](#): Completely free
 - ii. [SOCR -- Statistics Online Computational Resource](#): Comprehensive and free
 - iii. [SciStatCalc](#): Also free
 - iv. [MedCalc.org](#): Also free
 - v. [SISA \(Simple Interactive Statistical Analysis\)](#): Free. SISA allows you to do statistical analysis directly on the Internet
- [Statgraphics Stratus](#): Free version is limited

Useful Statistic Links

- I. <http://statpages.info/javasta3.html>: A link to very useful Statistical Books, Manuals and Journals
- II. <http://statpages.info/javasta4.html>: A link to very useful Interactive Statistical Demonstrations and Tutorials
- III. Don't know which statistical method to use? This interactive webpage may be helpful: <http://www.microsiris.com/Statistical%20Decision%20Tree/default.htm>

Learning Epidemiology (and Statistics): The Fun way

- **Studying Populations:** This is a wonderful program for those intending to learn the basic principles of epidemiology without too much hassles.. Here is the description from the program's website:

Studying Populations is computer assisted learning package for epidemiological methods. It is suitable for anyone working in health sciences including medical students, Masters students, public health practitioners and clinicians. It covers many basic epidemiological concepts in a set of over 80 interactive exercises and simulations.

- Download it from <http://medicine.dundee.ac.uk/studying-populations-epilex> and have a fun-filled adventure in learning statistics

- There is a review of the program in the International Journal of Epidemiology which can be downloaded from the journal website:
<http://ije.oxfordjournals.org/content/32/5/890.2.full>

Archiving and Note Taking Software

This group of programs allow you to capture anything you can think of from any program on your computer. The captured materials are called Notes. Notes can be formatted text, a full webpage or webpage excerpt, a photograph, or a voice memo. If you have a tablet, you can take Notes directly. What makes these software so versatile is that you can attach almost any type of files such as PDF, Word documents, Excel spreadsheets to the notes you have captured. The attachments can be fully embedded in the program, or they can be linked to the program. Finally, the notes can be sorted into folders, then tagged, annotated, edited, given comments, searched and exported as part of a notebook.

There are many of this programs on the internet, but I have only used two:

1. **Evernote** (www.evernote.com) and
2. **Microsoft OneNote.**

I am currently using the later, as it is part of the Office Suite, which I have. However, Evernote is equally good, and it is available for free download on the internet

How do I use these programs?

When a new idea strikes me, instead of scribbling on paper, I open OneNote on my convertible Lenovo tablet. I already have a folder called "*Studies*". All I need to do to start scribbling away is to create a new Note in this folder. If I get any interesting titbits from my searches on the internet or my computer, I'll just capture them into the Note. If I like, I may even use it as my initial "*Word Processor*". But if you prefer to use your native Word Processor, you can still link the two programs together. For me, the programs act like a chest drawer where I can organize my research information, and never misplace them.

Never Lose Your Data Again

Online data storage (also called cloud storage) are used primarily to backup selected files on your computer. While there are many examples, I have only used Dropbox (www.dropbox.com) and Microsoft OneDrive and Google Drive.

Dropbox: For me, Dropbox is more intuitive to use. Dropbox makes your files accessible from almost anywhere. It is designed for those who are tired of e-mailing files to themselves or who don't want to carry around flash drives.

How does it work? You must download and install Dropbox on your computer. The program creates a folder in your computer called "*Dropbox*". This is the interface for Dropbox. You simply drag and drop files into the folder, and the files are uploaded to Dropbox server immediately without any other intervention from you. In addition Dropbox also keeps track of every single change made to the contents of your storage and any changes are instantly updated to all computers linked to the account.

You can access any file you store in Dropbox from any computer, anywhere in the world as long as you have internet access. The Dropbox web interface also remembers all the changes

you make to your files and allows you to restore to any previous versions of the file. You can even un-delete files that you may have accidentally erased.

Dropbox also allows you to share whatever you want with other people. Every individual folder can be shared with other people and every member of a shared folder will be able to add, edit, and delete the contents inside but will not be able to access anything outside of that specific folder. This file sharing system is perfect for team projects. There is also a public folder that allows you to share files with non-Dropbox users.

Useful websites

www.oluwadiya.sitesled.com: My website. The section on SPSS contains some useful links on SPSS. If you are keen on learning SPSS syntax, there are useful links on this site for you.

<http://www.graphpad.com/quickcalcs/contingency1.cfm>: Online Statistics website.

<http://www.bettyjung.net/Statfxs.htm>: This **Betty C Jung's Statistical Procedure Site**. It also contains link to other useful sites.

<http://www.whichtest.info/>: This website is called **Which Test**. It is an online guide to selecting an appropriate statistical test. By following a series of instructions on the webpage, you can easily determine the appropriate statistics to use.

<https://www.statsoft.com/textbook/>: This is the grand daddy of online statistic resource. This is the description from its website: The only Internet Resource about Statistics Recommended by Encyclopedia Britannica. StatSoft has freely provided the **Electronic Statistics Textbook** as a public service for more than 17 years now.

This Textbook offers training in the understanding and application of statistics and data mining. The material was developed at the StatSoft R&D department based on many years of teaching undergraduate and graduate statistics courses and covers a wide variety of applications, including laboratory research (biomedical, agricultural, etc.), business statistics, credit scoring, forecasting, social science statistics and survey research, data mining, engineering and quality control applications, and many others.

The Electronic Textbook begins with an overview of the relevant elementary (pivotal) concepts and continues with a more in depth exploration of specific areas of statistics, organized by "modules" and accessible by buttons, representing classes of analytic techniques. A glossary of statistical terms and a list of references for further study are included.

<http://www.spsstools.net/SampleSyntax.htm>: If you want to learn SPSS syntax or you want access to syntax files that you can download and (freely) use, **Raynald's SPSS Tools** is for you. Syntax has a steep learning curve, but once you know how to use it, you can easily increase your productivity by a factor of ten! Furthermore some of SPSS's features are only available through syntax. Syntax is the real power house of SPSS.

<http://calcnnet.mth.cmich.edu/org/spss/toc.htm> : This site has some movies which can be downloaded. It is a great site for learning SPSS.

<http://www.ats.ucla.edu/stat/spss/> : Another great online learning resource for SPSS. Get your hands dirty on it!

<http://www.uccs.edu/~faculty/lbecker/SPSS/content.htm> : A well organized site, designed to help you choose the right test for your data.

<http://www.epibiostat.ucsf.edu/biostat/sampsize.html?iframe=tr#ttest> : This website contains a list of mostly free programs for sample size and power calculations.

<http://www.openepi.com> : Online Sample Size calculator.

Module Prepared by Prof Oluwadiya KS (www.oluwadiya.com)