

IGP 304 – Statistics for Biomedical Research

Instructors: Chun Li (chun.li@vanderbilt.edu; 515D Light Hall; 322-2884)
 Leena Choi (leena.choi@vanderbilt.edu; D-2213 MCN; 343-3497)
 Assistant: TBA
 Lectures: MWF 2:00-2:55 pm (206 PRB)
 Office hours: MWF after class
 Course page: <http://biostat.mc.vanderbilt.edu/IntroBiostatCourse2007>

The **goals** of the course are: (1) understand basic concepts, ideas, and techniques often used in statistics, especially biostatistics; (2) develop appreciation of (i) variation, (ii) importance of design to the overall quality of a study, (iii) impact of assumptions on data analysis and interpretation, and (iv) artifacts and caveats in data analysis and interpretation; (3) carry out simple exploratory/graphical/formal/diagnostic analysis; and (4) know when and where to seek statisticians' help.

Statistics is best learned by analyzing real data. **Please bring your problems to the class.** The problems can be issues in study design or data analysis or result interpretation, and they can be from your own research, papers you read, news reports, etc. If the data come from your ongoing projects and you don't want to disclose all the information, you can bring your data to me and I will remove the identifying information before presenting to the class.

Your grade will be based on five homework assignments (30%), a mid-term exam (25%), a final exam (25%), and quality participation in class discussions (20%). Before each lecture, read the relevant chapters in EMS and other assigned readings, and **come prepared to discuss.**

Textbook and reading materials:

- [EMS] Kirkwood BR, Sterne JAC (2003) *Essential Medical Statistics*, 2nd ed. Blackwell Publishers. \$58.95. ISBN: 0865428719. We will cover most chapters, with additional materials from other sources. Corrections and four datasets are available at book web site www.blackwellpublishing.com/essentialmedstats.
- [SN] Altman DG, Bland JM et al. (1994-2006) *Statistics Notes*. British Medical Journal. Fifty-two valuable short articles on misuses and correct uses of statistics in medical research. Links at www-users.york.ac.uk/~mb55/pubs/pbstnote.htm.
- [LHSP] Dallal G. *Little Handbook of Statistical Practice*: www.StatisticalPractice.com
- [SP] *Statistical Problems to Document and to Avoid*. A collection of misuses of statistics. <http://biostat.mc.vanderbilt.edu/ManuscriptChecklist>.

Other useful books (not required; asterisks — highly recommended to read):

- Altman DG (1990) *Practical Statistics for Medical Research*. Chapman & Hall/CRC. [A very good book on medical statistics. The second edition should come out soon.]
- *Altman DG, Machin D, Bryant TN, Gardner MJ (2000) *Statistics with Confidence*, 2nd ed. Blackwell Publishers. [Short essays on the advantages of using confidence intervals. The book comes with software Confidence Interval Analysis (CIA).]

- Armitage P, Berry G, Matthews JNS (2001) *Statistical Methods in Medical Research*, 4th ed. Blackwell Publishers. [This book is quite comprehensive, covering more materials than a semester's course. It may serve as a reference book, but definitely not a cookbook.]
- Bland M (2000) *An Introduction to Medical Statistics*, 3rd ed. Oxford University Press. [Another popular introductory book on medical statistics.]
- *Motulsky H (1995) *Intuitive Biostatistics*. Oxford University Press. [This book covers basic materials in biostatistics and explains the basic concepts very well.]
- Rosner B (2005) *Fundamentals of Biostatistics*, 6th ed. Duxbury Press. [Old style of teaching statistics. Lots of examples from many medical fields.]
- *Freedman D, Pisani R, and Purves R (1997) *Statistics*, 3rd ed. W. W. Norton & Company. [This is a very good introduction to statistics, without being technical.]
- Moore DS, Notz WI (2005) *Statistics: Concepts and Controversies*, 6th ed. W. H. Freeman. [Another very good non-technical introduction to statistics.]

Free online resources:

- Glossary of Statistical Terms: biostat.mc.vanderbilt.edu/twiki/pub/Main/ClinStat/glossary.pdf
- HyperStat Online (by David Lane): davidmlane.com/hyperstat/ [An online book on basic statistics, with some useful links.]
- SticiGui (by Philip Stark): www.stat.berkeley.edu/~stark/SticiGui/ [An online non-technical textbook on introductory statistics, with applets to allow interactive learning.]
- Electronic Statistics Textbook (by StatSoft): www.statsoft.com/textbook/stathome.html
- Electronic Encyclopedia of Statistical Examples and Exercises (EESSEE): www.whfreeman.com/eesee/ [A good site for learning statistics with lots of examples. Developed by folks at the Ohio State University.]

Software:

- Intercooled Stata 9: Powerful, with a good graphical user interface. UCLA has a good support site: www.ats.ucla.edu/stat/stata/. [\$89 for a year and \$145 for life, through GradPlan at www.stata.com/order/new/edu/gradplans/gp-campus.html. Buy Small Stata for \$45 if you have to pay by yourself. Stata also is available at the [College of Arts & Science Microcomputer Labs](#).]
- R 2.4 (optional, free from www.r-project.org): Powerful, versatile, and actively maintained and updated. It may require a steeper learning curve than Stata and SPSS, but the effort will pay off later on. To get a feel, go to "Help" → "Manuals (in PDF)" → "An Introduction to R", and try the commands in "A sample session". The Department of Biostatistics has a free R Clinic every Thursday (biostat.mc.vanderbilt.edu/RClinic).

We will use Windows versions of these packages. If you install them on Unix/Linux or Mac OS, let us know; the graphical interface may be different. Bringing a laptop to class is optional.

Comments on other software packages:

- SAS: The oldest survivor, with strong legacy. Hard to learn and extend, with outdated structure.
- SPSS: Have "standard" methods and good graphical user interface. However, it is difficult to extend beyond the "standard" methods.
- Honorable mention: Epi Info (free from CDC), SUDAAN, S-plus.