How to Select Your Biostatistics Requirement

Department of Biostatistics

Options for the Biostatistics Requirement

- **Biostatistics options (choose one)**
  - **Statistical Reasoning in Public Health** (611-612)
  - **Statistical Methods in Public Health** (621-623 required, 624 is optional)
  - **Methods in Biostatistics** (651-654)

Why are there 3 options?

- **Heterogeneous student population**
  - Diverse backgrounds
  - Differing quantitative backgrounds
  - Varying needs
- **What are the different desired skills?**
  - Critical statistical reasoning and thinking
  - Understanding of statistical methods and techniques
  - Skills in performing data analysis
  - Understanding of statistical theory

Why would you want to develop data analysis skills?

- If you want to develop skills to have <u>hands-on</u> experience in using a data set to do research or program evaluation. For example:
  - Assess relationships between risk factors and disease status
  - Evaluate a health program or treatment regimen
  - Analyze outcome or performance between two groups or programs over time
  - etc.

Types of Students

- “**Consumer**” - wants to develop skills for critical reading of the literature and reviewing of research proposals
- “**User**” - wants to develop additional computational skills and hands-on experience in analyzing data sets (data analysis skills)
- “**Advanced**” - has more advanced mathematical skills and wants to understand statistical techniques in more depth (theoretical underpinnings)

Description of the Options

- For the consumer**:** 3 credits x 2 terms = 6 credits

  Statistical Reasoning in Public Health (611-612)
  - 2 terms; 2 lectures per week; no labs;minimal computing
- For the user**:** 4 credits x 4 terms = 16 credits

  Statistical Methods in Public Health (621-623)
  - 4 terms; 2 lectures per week; 1 lab; other sessions
  - statistical computing using Stata statistical analysis package
- For the advanced**:** 4 credits x 4 terms = 16 credits

  Methods in Biostatistics (651-654)
  - 4 terms; 2 lectures per week; 1 lab; statistical computing using R statistical analysis package

What topics are covered in all 3 options?

- The topics of **all 3 options**:
  - Causal reasoning
  - Summarizing data: exploratory data analysis, tables and graphs
  - Probability concepts and distributions
  - Hypothesis testing and confidence intervals
  - p-values and statistical significance
  - Sample size and power
  - Linear and logistic multivariable regression analysis
  - Survival analysis and Cox regression analysis

How do the options differ?

- **Biostatistics 611-612** involves minimal calculation/computing.
- **Biostatistics 621-624** and **Biostatistics 651-654** teach the tools and techniques of data analysis. Both sequences uses computers and statistical analysis packages.
- **Biostatistics 651-654** explains statistical techniques in more depth and requires students to have more advanced mathematical skills.

How do students typically distribute across the sequences?

- **Statistical Reasoning in Public Health (611- 612)**
  - 130+ students on campus
  - 220+ students on line
- **Statistical Methods in Public Health (621-624)**
  - 450+ students

How do students typically distribute across the sequences?

- **Methods in Biostatistics (651-654)**
  - 50+ students

How does a student choose the introductory sequence?

- The Departments have requirements for non-MPH degree candidates.
- The **MPH student** will need to assess his/her own:
  - Mathematical skills and aptitude based on
    - Familiarity with mathematics, algebra
    - Performance in previous quantitative courses
  - Professional needs or ambitions

But HOW will I choose the best option for me?

- Would you like an overview of biostatistical concepts and methods in two terms with minimal focus on computing and calculations and limited hands-on data analysis? If YES → **Statistical Reasoning (611-612)**

But HOW will I choose the best option for me?

- Are you seeking the ability to conduct, or actively participate in, the design and data analysis for a public health practice or research program? If YES → **Statistical Methods in Public Health (621- 624)**

But HOW will I choose the best option for me? (continued)

- If you seek design and data analysis skills, do you have a working knowledge of linear algebra and multivariate calculus from your previous training? If YES → **Methods in Biostatistics (651-654)**

Example: Characteristics Associated with Risk of Event

- Multivariable Cox Proportional Hazards Model

	Hazard Ratio for Event (95% CI)	p-value
Low SES (vs. High SES)	2.7 (1.6 – 6.3)	0.04
Male (vs. Female)	4.5 (0.4 – 13.2)	0.21

- 611-612 – Interpret results
- 621-624 – Interpret results and calculate from a data set using Stata
- 651-654 – Interpret results, calculate from a data set, plus theoretical development

FAQ1:
Who takes Biostat 651-654?

- Students whose interests or main professional goals are to analyze data (Biostatistics graduate students, other PhD student earning joint MHS degree)
- Students with strong mathematical abilities who recently have had a year of calculus and a course in linear algebra

FAQ2: I had calculus 15 years ago. Could I still take Biostat 651-654?

- Possibly. Some review and self-study may be necessary.
- You should be able to:
  - Perform algebraic manipulations.
  - Graph an exponential function.
  - Find values that minimize a function by setting the first derivatives equal to zero.
  - Perform an integration.
  - Find the product of AB where A is a 2x3 matrix and B is a 3x2 matrix.

FAQ3: I am seriously considering applying to a doctoral program….

- What option should I take if I plan to apply to a JHBSPH doctoral program next year?
- Check the course requirements for doctoral students in the Department of interest. Many programs require Biostat 621-624.

FAQ4: How comfortable must I feel with math or computers?

- A recent randomized study by Boyd indicated:
  - Variables associated with **good performance** in Biostat 621 were:
    - Comfort with mathematical concepts
    - Comfort with computers
    - Not employed > 10 hours per week
  - Variable associated with **decreased performance** in Biostat 621:
    - Belief: “ I think that I will need a tutor”
  - English as a native language is ***not*** a predictor of performance

FAQ5: Could I switch sequences during the school year?

- No, this is not possible. Although the sequences cover roughly similar topics, the topics are not taught in necessarily the same order or time frame.
- If you decide to drop Biostat 621-623 at the end of the first term, you must take Biostat 611-612 during the June Summer Institute in order to complete your requirements within the academic year (or the next year).

FAQ6: I'm really not sure of my career plans…..

- Should I take Biostat 621-624 just in case I find a job that requires data analysis skills?
- Learning data analysis skills is not like learning how to ride a bicycle. If you don't use the skills, you *lose* them.
- It would be preferable to take a data analysis course nearer the time that you accept the job.

FAQ7: Suppose I want more at the end of Biostat 611-612?

- Suppose I would like to gain additional data analysis skills?
- There are three 1- week intensive data analysis workshops offered during the Winter and Summer Graduate Institutes in Epidemiology and Biostatistics.

FAQ8: Why is Stata used in Biostat 621-624 rather than SAS?

- Biggest reason: Stata can be purchased inexpensively ($195) by students for use on their own computers. See www.stata.com/gpdirect
- One can perform the same procedures in Stata as in SAS.
- The graphics abilities of Stata are better than those of SAS.
- Stata has good manuals and useful Help features.

FAQ9: Why isn’t Biostat 624 required as part of the option?

- Biostat 621-623 covers methods through multivariable regression procedures.
- Biostat 624 provides:
  - Concentrated review of statistical methods
  - Some advanced topics (e.g., data analysis for correlated observations)
  - Data analysis project of your choosing to pull it all together!