

Are you in a major other than Biostatistics, but would like to learn some introductory statistics? We have three two-semester sequences for you to consider.

Each sequence covers, in large part, the same body of statistical ideas: basic distributions, descriptive statistics and graphing, hypothesis testing and confidence intervals for means and proportions, linear regression, analysis of variance, power and sample size, logistic regression, and Kaplan-Meier/time-to-event models.

With that in common, the sequences differ in credit hours, the depth to which those techniques are covered, the level of underlying statistical fundamentals that students are expected to master, and the type and use of statistical software. Here are specifics on how they differ:

PubH 6414/6415 (3/3 credits), offered online or in class. We use Excel in 6414 and SAS in 6415 as the statistical analysis software. The classroom offering also includes a separate hour-long weekly computer lab for learning the software. Students who complete 6414/6415 are expected to be able to understand and interpret all of the basic statistical techniques used in scientific journals in their field of study, as well as use basic statistics in their own research.

PubH 6450/6451 (4/4 credits), offered in class only. We use SAS in both semesters; in 6450 there is a separate hour-long weekly lab specifically for learning SAS. The statistical techniques are covered in a bit more depth compared to 6414/6415. Students are expected to master the underlying statistical fundamentals to a slightly greater degree compared to 6414/6415. 6451 also covers proportional hazards regression, which is not covered in 6414/6415 at all. Students who complete 6450/6451 should be able to apply all the techniques covered in their own research, as well as understand basic statistical methods described in scientific journals in their field of study. This sequence (both semesters) is the **minimum** pre-requisite for most additional Biostatistics classes, such as 7420 (Clinical Trials), 7430 (Methods for Correlated Data), and 7435 (Latent Variable/Path Analysis).

PubH 7401/7402 (4/4 credits), offered in class only. We use R, Stata, and SAS; there is no separate computer lab. The statistical techniques are covered in greater mathematical detail than in either 6414/6415 or 6450/6451. Students are expected to have a good working knowledge of scalar calculus (as would be obtained from 1-2 semesters of undergraduate calculus). Additional topics are covered, such as multinomial models, and advanced methods for repeated/correlated measures and for time-to-event data.

Whichever sequence you decide is correct for you, we strongly encourage you to take both semesters, not just the first one! You will be much better prepared to understand, interpret, use, and learn new statistics in your field of research over your career. Also **be sure to check with your DGS / Major Chair / faculty committee** so that you know what is required vs. recommended, and what substitutions may be accepted.

Lastly, we also offer a 1-credit class specifically on programming in SAS, **PubH 6420**. The course covers SAS programs for reading and processing data, and for descriptive and basic statistical analysis. Very little class time is spent on explaining the statistical techniques; students are expected to know those fundamentals ahead of time, so that the course can cover the programming of those techniques in SAS.