Course # VTPMD 6105
Biostatistics for Health Sciences
Fall Semester – 2018
3 credits

Time	Tuesday and Thursday; 8:30am to 9:45am
Location	TBD

Course Directors

Karyn Havas DVM, PhD, MS
Section Chief, Infectious Disease Epidemiology
Master of Public Health Program
S2-005E Schurman Hall
Ithaca, NY 14853
t: 607.253.3344
e: kah47@cornell.edu
Office hours by appointment

Heather Grab, PhD
Post-Doctoral Associate
Department of Entomology
4128 Comstock Hall
Ithaca, NY 14853
e: hlc66@cornell.edu
Office hours by appointment

Graduate Teaching Assistants

Alison Stout, DVM
e. aek68@cornell.edu

Kristina Ceres
e. kc649@cornell.edu

Textbook:
Recommended:
Intuitive Biostatistics. Harvey Motulsky, 4th Ed. 2018. ISBN: 09780190643560
Statistics for Life Sciences. Myra Samuels, 4th Ed. 2010. ISBN: 9780321652805

Additional:
Biostatistics: the Bare Essentials. Geoffrey Norman, 4th Ed. 2014. ISBN: 9781607951780
The R Book. Michael J Crawley, 2nd Ed. 2012. Online:
    https://onlinelibrary.wiley.com/doi/book/10.1002/9781118448908
The Little SAS Book: A Primer. Lora D. Delwiche, 5th Ed. 2010. Online:
    https://newcatalog.library.cornell.edu/catalog/8024405
SPSS Statistics for Dummies. Keith McCormick, 3rd Ed. 2015. Online:
    https://newcatalog.library.cornell.edu/catalog/9295068

Course Overview
This course teaches statistical concepts and application for health related data analysis. The course will relate health and biological sciences data back to Gaussian, non-Gaussian, Poisson, Binomial, and other distributions. Topics in descriptive statistics include summary measures, measure of association, concepts related to data distribution, and confidence intervals. Topics related to analytical analysis include categorical data, parametric and non-parametric population comparisons and correlation, and regression techniques. Students will be exposed to a variety of software packages, but will also be expected to calculate the simple statistical approaches as well. Emphasizes the understanding of statistical concepts and application and the structure of health and biological data.

Learning Outcomes
In this course, you will be developing an understanding of the theory and science of statistics and its routine application. The overarching learning outcomes are:
- Understand the concept of probability density functions, test statistics, probability values.
- Use data visualization techniques to assess the validity of summary statistics
- Apply appropriate statistical tests to compare summary statistics based on the type of data and to identify statistically relevant associations.
- Understand the relationship between p-values and confidence intervals and the strengths and weaknesses of each statistical measure
- Describe the principle of variance and how ordinary least square concepts are related to the analysis of variance
- Use linear regression to conduct a multivariable analysis
- Become familiar with commonly used software programs for statistical analysis

Core competencies and indicators below for the MPH Concentrations:

Foundational Competencies: Evidence-based Approaches to Public Health
3.3 Analyze quantitative and qualitative data using biostatistics, informatics, computer-based programming and software, as appropriate

      3.3.3 Use statistical software to analyze health-related data

      3.3.4 Use technology to visualize and present data

3.4 Interpret results of data analysis for public health research, policy or practice

      3.4.1 Analyze and interpret quantitative data

      3.4.2 Analyze and interpret qualitative data

      3.4.3 Draw appropriate inferences from epidemiologic data

Infectious Disease Epidemiology (IDE) Concentration Competencies
1.2 Understand and apply different epidemiological and statistical methodologies and assessment techniques to infectious disease events and programs.

      1.2.2 Exhibit the ability to manage data, conduct basic analyses to identify differences between samples using parametric, non-parametric, and interpret survival and regression technique results.

      1.2.3 Calculate descriptive measures of risk, measures of association, and conduct comparisons between populations to understand disease.

VTPMD 6105 Biostatistics for Health Sciences

Cornell University

1.2.4 Have exposure to data visualization and mapping techniques.

Food Systems & Health (FSH) Outcomes Concentration Competencies
1.1 Identify and investigate food safety risks using data collection and analysis to assess linkages between food system pathogens, food security, and health.

1.1.2 Collect, analyze, and interpret relevant foodborne pathogen data - FDSC

Course Information

All lecture materials, assignments, assignment keys, exams, exam keys, and other information will be posted via Blackboard. Paper handouts will not be provided. All assignments will need to be turned in as a **PDF** or **Word** document to Blackboard or Canvas as well.

Attendance Policy

Each student is allowed two unexcused absences. Additional absences may impact your class participation grade. The class will be fast paced, and the content will provide the foundation required to meet the learning outcomes of the MPH program or to prepare you for further instruction in statistics. It is recommended you make every effort to attend. Additional absences may be allowed for unforeseen circumstances and illness.

Process of Assessment

Weekly homework will make up 60% of the final grade. There will be 10 homework assignments in total. These will be comprised of short problem sets that will require both hand calculation and software based solutions to   practice concepts learned during the week. The homework assignments will cumulatively assess IDE Competency 1.2 and FSH Competency 1.1. These will be posted on Blackboard or Canvas after class on Thursday and due the following Monday or Tuesday at midnight. Pay attention to the due date posted in Blackboard or Canvas. There will be no homework the week of the midterm, the week of Thanksgiving break, or the final week of classes.

Participation is assessed by class attendance and will make up 5% of the final grade. See the attendance policy in this syllabus for further guidance.

On October 4th there will be an in-class midterm exam. It will be cumulative to that point of the semester. It will cover everything from August 23rd through September 27th. It will be worth 15% of your grade and assesses components of IDE Competency 1.2 and FSH Competency 1.1.

The remaining 20% of your grade will be assessed using a final project. This cumulative project will be on an assigned data set. You will use multiple techniques to describe, assess, and interpret the data set. The final class period will be made available for additional questions related to this class project. You will receive information on this final project on November 8th and will put together a list of hypotheses you will test as part of your project. This will be due the following Wednesday at midnight and will contribute 10% to your final project grade. The final project will be due on Friday December 7th at midnight.

Grading

Grading in this course is by a letter grade. Grades of "I" (incomplete) will be given strictly according to

University policy. If a student chooses to drop the course it is the student's responsibility to work with the registrar to drop the course. ***The final Add/Drop date is October 18, 2018.***

MPH Competence Assessment

MPH students will upload their pamphlets into CORE as part of their professional portfolio. They will then self-assess and receive an assessment from the faculty on this course. The levels of competence are Developing competence, Basic competence, Competence, and Exemplary competence. This is not a part of your grades for the assignments or the course, but rather is part of a cumulative assessment process for MPH students. We expect all students to graduate at the level of competence or exemplary competence for each Concentration competency. The rubrics are available in the CORE website.

Students should click on the "CORE link" on each MPH course's Blackboard/Canvas site at the beginning of the semester to enroll in all course CORE pages.

Cornell University Code of Academic Integrity

Each student in this course is expected to abide by the Cornell University Code of Academic Integrity. Any work submitted by a student for academic credit will be the student's own work. We expect that students will also not plagiarize or self-plagiarize. Each assignment should be each student's individual work. You are encouraged to study together and to discuss information and concepts with other students. You can give "consulting" help to or receive "consulting" help from other students. For example, you can discuss your approaches for completing problems, but the calculations should be done independently and the do files for the software written individually. However, this permissible cooperation should never involve one student having possession of a copy of all or part of a course assignment done by someone else, in the form of an email, an electronic file, or a hard copy. Using someone else's assignment as a template for completing your assignment is strictly forbidden.

Classroom Etiquette

We commit to starting and ending classes on time. Please be prompt and do not leave before the lecture is over unless it is necessary. Cell phones should be silenced, headphones removed, and use of a computer is expected to support your classroom experience. No one should be replying to email, on social media, or conducting non-academic use of the internet. If a student's use of technology becomes distracting, the instructor reserves the right to ban its use in class. Please wait until the lecture is over before you begin to gather your belongings. This large classroom is only one of the many "meetings" you will attend in your professional life. Practicing basic etiquette will make it a more positive experience for everyone.

Diversity and Inclusion

We are committed to the Cornell University legacy of diversity summarized by our founding mission of, "Any Person, Any Study". We aim to create an environment that allows our differences to foster our education and engage our learning. We are committed to honoring the unique contributions of each student and expect all students to respectfully listen and dialogue with those that have a different experience to share. We are committed to disability and religious accommodations. The student must make the instructors aware of this need so we can assist.

https://diversity.cornell.edu/

Resources

R: www.r-project.org
Stata: www.stata.com
SAS: www.sas.com
SPSS: www.ibm.com/products/spss-statistics
Khan Academy Statistics and Probability: www.khanacademy.org/math/statistics-probability
Biostatistics with R: An introduction to statistics through biological data
    https://link.springer.com/book/10.1007%2F978-1-4614-1302-8
Biostatistics in public health using STATA:
    https://ebookcentral.proquest.com/lib/cornell/detail.action?docID=4460057
Biostatistics by example using SAS studio: https://proquest.safaribooksonline.com/9781629604930
Biostatistics using JMP: A practical guide: https://proquest.safaribooksonline.com/9781635262414
Cornell library: www.library.cornell.edu
Cornell Wellness Services
    https://health.cornell.edu/
    https://health.cornell.edu/services/counseling-psychiatry/lets-talk

Schedule

Lecturer	Weekday	Date	Topic	Assignment/Exam
Grab/Havas	Th	23-Aug	Course overview, software access, excel	Pre-eval due, excel
Havas	Tu	28-Aug	Theory of statistical distributions	
Grab	Th	30-Aug	Data summary statistics and visualization	HW 1 - Due SEP 4 @ midnight
Havas	Tu	4-Sep	Data variance, standard deviation, and standard error, and n-1	
Grab	Th	6-Sep	Intro to T-distribution, Student's T	HW 2 -Due SEP 10 @ midnight
Grab	Tu	11-Sep	Paired T, equal and unequal variance	
Grab	Th	13-Sep	Software example R/STATA	HW 3 Due SEP 17 @ midnight
Havas	Tu	18-Sep	One way ANOVA, F stat	
Havas	Th	20-Sep	Two-way comparisons + application	HW 4 Due SEP 24 @ midnight
Havas	Tu	25-Sep	ANOVA example R/STATA	
TA	Th	27-Sep	Confidence intervals	HW 5 Due OCT 1 @ midnight
Havas	Tu	2-Oct	Categorical data and binomial data (proportions)	
--	Th	4-Oct	**MIDTERM TAKE HOME (extra period)**	Mid-Term exam covers through confidence intervals
--	Tu	9-Oct	*FALL BREAK*	
Havas	Th	11-Oct	Chi-squared distribution and test	HW 6 Due Oct 15 @ midnight
Grab	Tu	16-Oct	Fisher's exact test	
Grab	Th	18-Oct	Categorical data example R/STATA	HW 7 Due Oct 22 @ midnight, **ADD/DROP DEADLINE**
Grab	Tu	23-Oct	Non-parametric summary statistics and Spearman correlation	
Havas	Th	25-Oct	Mann Whitney U	HW 8 Due Oct 29 @ midnight
Grab	Tu	30-Oct	Rank Sum	
Havas	Th	1-Nov	Kruskal Wallis	HW 9 Due NOV 5 @ midnight
Havas/Grab	Tu	6-Nov	Non-parametric example R/STATA	
Havas	Th	8-Nov	Intro to regression assumptions	Final Project Posted - Hypotheses due 14 NOV @ midnight
Havas	Tu	13-Nov	Ordinary Least Squares and Pearson correlation	
Grab	Th	15-Nov	Hand calculation of F-statistic and software application	
Havas	Tu	20-Nov	Evaluating assumptions with residuals	
	Th	22-Nov	*THANKSGIVING*	
Grab	Tu	27-Nov	Goodness of fit and what to do with outliers	HW 10 Due NOV 19@ midnight
Havas	Th	29-Nov	Intro to Logistic Regression	
	Tu	4-Dec	**QUESTIONS FOR FINAL PROJECT**	Due December 7

VTPMD 6105 Biostatistics for Health Sciences

Cornell University