

# HARBOUR SPACE

## UNIVERSITY

### Master's program

DS406 - Statistical Data Analysis  
Syllabus, Leah Isakov

#### 1| Dates of module

February 1, 2021 - February 19, 2021

#### 2| Abstract

Learn basic concepts of theory and application for statistical inference. Get an introduction to statistical analysis and critical thinking, including descriptive statistics, probability, sampling distributions, interval estimation, hypothesis testing and regression. Use of simulation technique for assessments of model fit and estimations. Design experiments from statistical perspectives and explore some advanced topics.

#### 3| Objectives & Learning Goals

Select an appropriate statistical technique to analyze and interpret the observed data. Students should acquire quantitative skills that they can employ and build on in flexible ways. The goal is to learn concepts and master tools for working with data and understand experiment design.

#### 4| Skills you will learn

Learn basic concepts of theory and application for statistical inference.

#### 5| Methodology

Combination of the theoretical and applied methodology.

## 6| Evaluation and grading

(40%) Four homework assignments  
(30%) Final exam  
(25%) Midterm exam  
(5%) Class participation

## 7| Outline

### Session 1

Introduction/review, data types, probability and laws of probability. Random data types. Statistics, data and statistical thinking. Data visualization. Measures of central tendency.

### Session 2

Measures of Quality of Estimators.  
MLE/ MOM.  
Central Limit Theorem and sampling distribution.

### Session 3

Large-Sample Confidence Interval for a Population Mean and Proportions.  
T-Statistics and small-sample confidence intervals for a population mean.

### Session 4

Introduction to hypothesis testing.  
Inference based on a single sample: a test of hypothesis.  
Introduction to Theory of Statistical Tests  
Likelihood Ratio Tests.

### Session 5

Comparing two population means independent and paired sampling.

### Session 6

Estimating a proportion.  
Comparing groups on categorical data.  
Chi-square tests.  
Large sample confidence interval for a population proportion.

### Session 7

Small sample test, Yates' correction, Fisher Exact test. Measures of association: Relative Risk and Odds Ratio.

### Session 8

Midterm Exam

### Session 9

Association between measurement variables.  
Correlation and regression.  
Simple Linear Regression.

### Session 10

Regression diagnostic.  
Data transformation.  
Multiple linear regression.

### Session 11

Logistic Regression

### Session 12

ANOVA

### Session 13

Non-Parametric tests

Non-parametric test about the population mean,  
comparing two populations paired test, sign test,  
Non-Parametric test for correlation.

### Session 14

More on regression diagnostic for multiple  
linear regression.  
Model building.  
Review

### Session 15

Final Exam

## 8| Bibliography

'Categorical Data Analysis" by Alan Agresti - applied classic for categorical data analysis  
"Statistical Inference" by George Casella and Roger L. Berger - more theoretical

## 9| Required Skills or Courses

Working knowledge of R/Python or SAS, basic calculus and probability theory.



**Leah Isakov**

Ph.D., Advisor at Pallantius Inc

Dr. Leah Isakov is a senior leader in the pharmaceutical industry with a unique combination of leadership and technical skills. She has worked in clinical trials for more than two decades and is known for delivering results. She has led NDA (New Drug Applications), PMA (Pre-Marketing Approvals) and BLA (Biologics License Applications). She has deep experience interacting with all the major regulatory bodies (FDA, EMEA, PMDA, Russian Ministry of Health, and Health Canada). She also has direct experience successfully managing cross-cultural international teams (USA, China, Japan and Canada). Her recent therapeutic areas include Oncology, Infectious Diseases, Cardiovascular, Asthma, Renal Failure and HIV for Phase II-IV clinical trials in drugs and biologics.

As a leader, Leah strives to be at the forefront of management practice. She incorporates data-driven decision making and quantitative risk management and focuses on building internal capabilities along with external collaborations. She believes that successful management comes from understanding the full organizational stack; that is, high-level strategy and the technical aspects that enable success. Leah has a strong grasp of the technical side from two decades of hands-on experience in analytics, protocol design, sample size calculation, SAS programming, and integrated analysis (ISS and ISE), as well as strong GCP and regulatory knowledge.

Faculty webpage: <http://harbour.space/Faculty/Data-Science/Leah-Isakov>