ST 370 Calendar (Monday/Wednesday)

Fall 2018

**Note**: Exam 1 will cover topics 1-3, exam 2 will cover topics 4-6, and the final will be cumulative.

Green items are completed outside of class and blue items are the intended in-class items.

PCM stands for pre-class module (more detail for the modules are given after the calendar). SCL stands for short class lecture (usually intended to be about 20 minutes). LCL stands for long class lecture (may be a significant portion of the class period). ICA stands for in class activity. OA stands for online assessment.

| Dates | Lessons, Tasks, Assignments, etc. | Other Reminders |
| --- | --- | --- |
| **Week 1**  **August 22-26** | * PCM 0: Orientation to course * Intro to Course (15 minutes)   + Brief review of syllabus   + Discussion of course structure   + Prepare students for different experience * ICA: Icebreaker (20 minutes) * LCL: Overview Example – 3 major elements | Class begins Wednesday |
| **Week 2**  **August 27- September 2** | * PCM1a: Statistical process/paradigm, Review of descriptive statistics, Data collection terminology, SRS, Exploratory data analysis * SCL: Issues with not randomizing and their effect on statistics (thinking quick sampling activity with houses or boxes) * ICA1a: * PCM1b: Random variables, Discrete random variables, Histograms, Interpretation of probability, Basic probability laws, Simulation ideas * SCL: * ICA1b: |  |
| **Week 3**  **September 3-9** | * PCM1c: Limitations of point estimation, CI ideas for proportion, Interpretation of a CI * SCL: * ICA1c: | *No class Monday* |
| **Week 4**  **September 10-16** | * OA1 * ICA1b (New case study) * PCM2a |  |
| **Week 5**  **September 17-23** | * PCM2b * PCM2c |  |
| **Week 6**  **September 24-30** | * OA2 * ICA2b (New case study) |  |
| **Week 7**  **October 1-7** |  |  |
| **Week 8**  **October 8-14** |  |  |
| **Week 9**  **October 15-21** |  | Drop deadline Oct 19? |
| **Week 10**  **October 22-28** |  |  |
| **Week 11**  **October 29-November 4** |  |  |
| **Week 12**  **November 5-11** |  |  |
| **Week 13**  **November 12-18** |  |  |
| **Week 14**  **November 19-25** |  | *No class Wednesday* |
| **Week 15**  **November 26-December 2** |  |  |
| **Week 16**  **December 3-December 7** |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Notes | Dates | M/W class | Topic | T/Th class | Topic |
| Start on W | Aug 22-24 | 1 | Intro | 1 | Intro |
|  | Aug 27-31 | 2 | Unit 1, Unit 1 | 2 | Topic 1, Topic 1 |
| No class M | Sep 4-7 | 1 | Unit 1 | 2 | Topic 1, Topic 1 CS |
|  | Sep 10-14 | 2 | Unit 2, Unit 2 | 2 | Topic 2, Topic 2 |
|  | Sep 17-21 | 2 | Unit 2, DI, | 2 | Topic 2, Topic 2 CS |
|  | Sep 24-28 | 2 | Unit 3, Unit 3 | 2 | Topic 3, Topic 3 |
| No class ThF | Oct 1-3 | 2 | Unit 3 CS, Unit 4 | 1 | Topic 4 |
|  | Oct 8-12 | 2 | Exam 1, Unit 4 CS | 2 | Topic 4, Exam 1 |
|  | Oct 15-19 | 2 | Unit 5, Unit 5 | 2 | Topic 4, Topic 5 |
|  | Oct 22-26 | 2 | Unit 5 CS, Unit 6 | 2 | Topic 5, Topic 5 |
|  | Oct 29-2 | 2 | Unit 6, Unit 7 | 2 | Topic 6, Topic 6 |
|  | Nov 5-9 | 2 | Exam 2, Unit 7 | 2 | Topic 7, Exam 2 |
|  | Nov 12-16 | 2 | Unit 7, Unit 8 | 2 | Topic 7, Topic 7 |
| No class WThF | Nov 19-20 | 1 | Unit 8 | 1 | Topic 8, Topic 8 |
|  | Nov 26-30 | 2 | Unit 8 CS, Unit 9 | 2 | Topic 8, Topic 9 |
|  | Dec 3-7 | 2 | Unit 9, Unit 9 CS | 2 | Topic 9, Topic 9 |
| Total Number |  | 29 |  | 29 |  |

Drop deadline Oct 19th

Exam 1 covers topics 1-3 (and DI day)

Exam 2 covers topics 4-6

Final cumulative

Unit 1: One-Sample CI for a Proportion/Probability (3 classes)

Unit 2: Two-Sample CI for a Difference in Proportions/Probabilities (3 classes, 1 one which is a CS)

Direct instruction lecture: Wald vs Wilson CI/Poisson distribution with Small Lambda, Small n (45 min)

Unit 3: Two-Sample HT for a Difference in Proportions/Probabilities (3 classes)

Unit 4: Chi-Square Test for Statistical Independence (3 classes, 1 one which is a CS)

Unit 5: Paired Sample CI for a Mean (3 classes, 1 one which is a CS)

Unit 6: Two-Sample HT for a Difference in Means (2 classes)

Unit 7: One-Way ANOVA (3 classes)

Unit 8: Two-Way ANOVA (2 classes, 1 one which is a CS)

Unit 9: Simple Linear Regression (3 classes, 1 one which is a CS)

Pre-Class Module and Instruction Details

|  |  |
| --- | --- |
| Unit 1 | One-Sample CI for a Proportion/Probability |
| PCM1a | 1 Course overview and concept review  1.1 What is statistics?  1.1.1 Statistical process  1.1.2 Statistical paradigm  1.3 Exploratory data analysis  1.3.1 Use as a data validation tool  1.3.2 Review of descriptive statistics  1.3.2.1 Variable types  1.3.2.2 Notation  2 Principles of data collection  2.1 Basic terminology  2.2.1 Simple random sample |
| PCM1b | 1 Course overview and concept review  1.3 Exploratory data analysis  1.3.4 Applied review of graphical tools  1.3.4.1 Histograms  3 Probability  3.1 Probability rules  3.1.1 P(Omega) = 1  3.1.2 0 <= P(A) <= 1  3.1.8 Interpretation of probability (relative frequency)  3.2 Random variables  3.2.1 Discrete RVs  3.2.1.1 Functions defining the distribution (PMF, CDF)  3.3 Sampling distributions  3.3.3 Simulation demonstration |
| PCM1c | 4 Inference  4.1 Confidence interval concepts  4.1.1 Why aren’t point estimators enough?  4.1.2 Margin of Error  4.1.3 Relation to sampling distribution  4.1.5 Interpretation of observed interval  4.2 CI applications  4.2.3 CIs for a proportion  4.2.3.1 Basic Z |
| Unit 2 | Two-Sample CI for a Difference in Proportions/Probabilities |
| PCM2a | 1 Course overview and concept review  1.1 What is statistics?  1.1.1 Statistical process  2 Principles of data collection  2.4 Observational studies vs designed experiments  2.4.1 Definitions  2.4.2 OS and DE specific terminology  2.5 Principles of DE  2.5.1 Sources of variability in response variable  2.5.2 Randomization  2.5.2.1 Contrast random selection and random allocation  2.5.3 Replication  2.6 Common DE Methods  2.6.1 CRD |
| PCM2b | 3 Probability  3.1 Probability rules  3.1.1 P(Omega) = 1  3.1.2 0 <= P(A) <= 1  3.1.5 Complement rule  3.2 Random variables  3.2.1 Discrete RVs  3.2.1.1 Functions defining the distribution (PMF, CDF)  3.2.1.2 Finding probabilities using PMF or CDF  3.2.1.3 Expected value  3.2.1.4 Variance  3.2.1.5 Binomial Distribution  3.2.1.5.1 Binomial Experiment  3.2.1.5.2 PMF, Mean, Variance |
| PCM2c | 4 Inference  4.1 Confidence intervals concepts  4.1.1 Why aren’t point estimators enough?  4.1.2 Margin of Error  4.1.3 Relation to sampling distribution  4.1.4 Confidence idea  4.1.5 Interpretation of observed interval  4.1.6 Checking conditions for inference  4.1.7 Scope of inference (relation to bias, etc.)  4.2 CI applications  4.2.4 CIs for the difference of proportions  4.2.4.1 Basic Z |
| DI | Wald vs Wilson CI/Poisson Distribution |
| In class | 4 Inference  4.1 Confidence intervals concepts  4.1.4 Confidence idea  4.1.5 Interpretation of observed interval  4.1.6 Checking conditions for inference  4.2 CI applications  4.2.3 CIs for a proportion  4.2.3.1 Basic Z  4.2.3.2 Wilson  4.2.4 CIs for the difference of proportions  4.2.4.1 Basic Z  4.2.4.2 Wilson  3 Probability  3.1 Probability rules  3.1.1 P(Omega) = 1  3.1.2 0 <= P(A) <= 1  3.1.5 Complement rule  3.2 Random variables  3.2.1 Discrete RVs  3.2.1.1 Functions defining the distribution (PMF, CDF)  3.2.1.2 Finding probabilities using PMF or CDF  3.2.1.3 Expected value  3.2.1.4 Variance  3.2.1.6 Other named discrete RV  3.2.1.6.1 (Not all covered) Uniform, Hypergeometric, Negative Binomial, Poisson |

|  |  |
| --- | --- |
| Unit 3 | Two-Sample HT for a Difference in Proportions/Probabilities |
| PCM3a | 1 Course overview and concept review  1.1 What is statistics?  1.1.1 Statistical process  1.3 Exploratory data analysis  1.3.3 Applied review of numerical summaries  1.3.3.1 Measures of center  1.3.3.3 Description of shape  1.3.3.5 Summaries across groups  1.3.3.6 Idea of point estimation  2 Principles of data collection  2.2 Sampling methods  2.2.3 Convenience sample  2.2.6 Benefits and drawbacks  2.3 Sampling Bias  2.3.1 Relationship to random vs. nonrandom selection |
| PCM3b | 3 Probability  3.2 Random variables  3.2.2 Continuous RVs  3.2.2.1 Functions defining the distribution (PDF, CDF)  3.2.2.5 Normal RVs  3.2.2.5.1 PDF, mean, variance  3.2.2.5.2 Empirical rule idea  3.3 Sampling distributions  3.3.1 Definition/purpose  3.3.2 CLT |
| PCM3c | 4 Inference  4.3 Hypothesis testing concepts  4.3.1 Hypotheses  4.3.2 Conclusions and possible errors  4.3.3. Significance level  4.3.4 Relation to sampling distributions  4.3.5 P-value  4.3.7 Interpretation  4.3.8 Scope of inference (relation to bias, etc.) |

|  |  |
| --- | --- |
| Unit 4 | Chi-Square Test for Statistical Independence |
| PCM4a | 1 Course overview and concept review  1.1 What is statistics?  1.1.1 Statistical process  1.3 Exploratory data analysis  1.3.4 Applied review of graphical tools  1.3.4.2 Bar charts  1.3.4.2.1 Stacked/grouped  2 Principles of data collection  2.2 Sampling methods  2.2.4 Volunteer response sampling  2.2.5 Importance of randomization  2.2.6 Relationship between and comparison of methods (benefits and drawbacks)  2.2.7 Sampling designs: qualifications on conclusions and ethical considerations  2.3 Sampling Bias  2.3.1 Relationship to random vs. nonrandom selection |
| PCM4b | 3 Probability  3.1 Probability rules  3.1.6 Conditional probability and statistical independence  3.1.7 Bayes theorem  3.2 Random variables  3.2.1 Discrete RVs  3.2.1.1 Functions defining the distribution (PMF, CDF)  3.2.1.2 Finding probabilities using PMF or CDF  3.2.2 Continuous RVs  3.2.2.1 Functions defining the distribution (PDF, CDF)  3.2.2.2 Finding probabilities using the PDF or CDF  3.2.2.3 Expected value  3.2.2.6 Other named continuous RV  3.2.2.6.1 Uniform, Exponential, Gamma, Chi-Squared, Weibull, Lognormal, Beta  3.3 Sampling distributions  3.3.1 Definition/purpose  3.3.2 CLT  3.3.4 T, Chi-square, F |
| PCM4c | 4 Inference  4.3 Hypothesis testing concepts  4.3.1 Hypotheses  4.3.2 Conclusions and possible errors  4.3.3. Significance level  4.3.4 Relation to sampling distributions  4.3.5 P-value  4.3.6 Checking conditions for inference  4.3.7 Interpretation  4.3.8 Scope of inference (relation to bias, etc.)  4.4 HT applications  4.4.5 HT for independence (chi-square test) |
| Unit 5 | Paired Sample CI for a Mean |
| PCM5a | 1 Course overview and concept review  1.1 What is statistics?  1.1.1 Statistical process  1.3 Exploratory data analysis  1.3.2 Review of descriptive statistics  1.3.2.1 Variable types  1.3.2.2 Notation  2 Principles of data collection  2.5 Principles of DE  2.5.1 Sources of variability in response variable  2.5.1.1 Local error control  2.5.1.2 Principles of blocking  2.5.1.3 Matched-pairs  2.5.2 Randomization  2.5.3 Replication  2.5.4 Issues that arise from not having randomization, replication, and/or error control  2.6 Common DE Methods  2.6.2 RCBD  2.6.3 Benefits and drawbacks of DE methods |
| PCM5b | 3 Probability  3.1 Probability rules  3.1.6 Conditional probability and statistical independence  3.2 Random variables  3.2.2 Continuous RVs  3.2.2.1 Functions defining the distribution (PDF, CDF)  3.2.2.2 Finding probabilities using PDF or CDF  3.2.2.3 Expected value  3.2.2.4 Variance  3.2.1.5 Normal distribution  3.3 Sampling distributions  3.3.1 Definition/purpose  3.3.4 T, Chi-square, F |
| PCM5c | 1 Course overview and concept review  1.3 Exploratory data analysis  1.3.3 Applied review of numerical summaries  1.3.3.3 Description of shape  1.3.4 Applied review of graphical tools  1.3.4.1 Histograms  1.3.4.3 Box plots  2 Principles of data collection  2.4 Observational studies versus designed experiments  2.4.3 Compare and contrast study methods/conclusions  2.4.4 Causality vs related  4 Inference  4.1 Confidence intervals concepts  4.1.1 Why aren’t point estimators enough?  4.1.2 Margin of Error  4.1.3 Relation to sampling distribution  4.1.4 Confidence idea  4.1.5 Interpretation of observed interval  4.1.6 Checking conditions for inference  4.1.7 Scope of inference (relation to bias, etc.)  4.2 CI applications  4.2.1 CI for a single mean (paired t only) |
| Unit 6 | Two-Sample HT for a Difference in Means |
| PCM6a | 1 Course overview and concept review  1.1 What is statistics?  1.1.1 Statistical process  2 Principles of data collection  2.2 Sampling methods  2.2.2 Stratified random sampling  2.2.5 Importance of randomization  2.2.6 Relationship and comparison of methods  2.3 Sampling Bias  2.4 Observation studies versus designed experiments |
| PCM6b | 3 Probability  3.2 Random variables  3.2.2 Continuous RVs  3.2.2.1 Functions defining the distribution (PDF, CDF)  3.2.2.2 Finding probabilities using PDF or CDF  3.2.2.3 Expected value  3.2.2.5 Normal RVs  3.3 Sampling distributions  3.3.1 Definition/purpose  3.3.2 CLT  3.3.4 T, Chi-square, F |
| PCM6c | 1 Course overview and concept review  1.3 Exploratory data analysis  1.3.3 Applied review of numerical summaries  1.3.3.1 Measures of center  1.3.3.2 Measures of spread  1.3.3.3 Description of shape  1.3.3.5 Summaries across groups  1.3.4 Applied review of graphical tools  1.3.4.1 Histograms  1.3.4.3 Box plots  1.3.4.3.1 Side-by-side  2 Principles of data collection  2.2 Sampling methods  2.2.7 Sampling designs: qualifications on conclusions and ethical considerations  2.4 Observational studies versus designed experiments  2.4.3 Compare and contrast study methods/conclusions  2.4.4 Causality vs related  4 Inference  4.3 Hypothesis testing concepts  4.3.1 Hypotheses  4.3.2 Conclusions and possible errors  4.3.3. Significance level  4.3.4 Relation to sampling distributions  4.3.5 P-value  4.3.6 Checking conditions for inference  4.3.7 Interpretation  4.3.8 Scope of inference (relation to bias, etc.)  4.3.9 Compare/contrast with CI  4.4 HT applications  4.4.2 HTs for the difference in means |
| Unit 7 | One-Way ANOVA |
| PCM7a | 1 Course overview and concept review  1.1 What is statistics?  1.1.1 Statistical process  1.3 Exploratory data analysis  1.3.3 Applied review of numerical summaries  1.3.3.1 Measures of center  1.3.3.2 Measures of spread  1.3.3.3 Description of shape  1.3.3.5 Summaries across groups  1.3.3.6 Idea of point estimation  1.3.4 Applied review of graphical tools  1.3.4.3 Box plots  1.3.4.3.1 Side-by-side  2 Principles of data collection  2.5 Principles of DE  2.5.2 Randomization  2.5.3 Replication  2.6 Common DE Methods  2.6.1 CRD |
| PCM7b | 3 Probability  3.2 Random variables  3.2.2 Continuous RVs  3.2.2.1 Functions defining the distribution (PDF, CDF)  3.2.2.5 Normal RVs  3.3 Sampling distributions  3.3.1 Definition/purpose  3.3.4 T, Chi-square, F |
| PCM7c | 2 Principles of data collection  2.4 Observational studies versus designed experiments  2.4.3 Compare and contrast study methods/conclusions  2.4.4 Causality vs related  4 Inference  4.3 Hypothesis testing concepts  4.3.1 Hypotheses  4.3.2 Conclusions and possible errors  4.3.3. Significance level  4.3.4 Relation to sampling distributions  4.3.5 P-value  4.3.6 Checking conditions for inference  4.3.6.1 Remedies when assumptions are violated  4.3.7 Interpretation  4.3.8 Scope of inference (relation to bias, etc.)  4.4 HT applications  4.4.6 One-factor ANOVA  4.4.6.1 Global hypotheses  4.4.6.2 Sums of squares/mean squares  4.4.6.4 Relation to t-tests |
| Unit 8 | Two-Way ANOVA |
| PCM8a | 1 Course overview and concept review  1.1 What is statistics?  1.1.1 Statistical process  1.3 Exploratory data analysis  1.3.3 Applied review of numerical summaries  1.3.3.5 Summaries across groups  2 Principles of data collection  2.5 Principles of DE  2.5.1 Sources of variability in response variable  2.5.2 Randomization  2.5.3 Replication  2.6 Common DE Methods  2.6.1 CRD |
| PCM8b | 3 Probability  3.2 Random variables  3.2.2 Continuous RVs  3.2.2.1 Functions defining the distribution (PDF, CDF)  3.2.2.5 Normal RVs  3.3 Sampling distributions  3.3.1 Definition/purpose  3.3.4 T, Chi-square, F |
| PCM8c | 2 Principles of data collection  2.4 Observational studies versus designed experiments  2.4.3 Compare and contrast study methods/conclusions  2.4.4 Causality vs related  4 Inference  4.3 Hypothesis testing concepts  4.3.1 Hypotheses  4.3.2 Conclusions and possible errors  4.3.3. Significance level  4.3.4 Relation to sampling distributions  4.3.5 P-value  4.3.6 Checking conditions for inference  4.3.6.1 Remedies when assumptions are violated  4.3.7 Interpretation  4.3.8 Scope of inference (relation to bias, etc.)  4.4 HT applications  4.4.7 Multi-factor ANOVA  4.4.7.1 Global hypotheses  4.4.7.2 Tests for model effects  4.4.7.2 Sums of squares/mean squares  4.4.7.4 Main effects  4.4.7.5 Interaction effects  4.4.7.6 Interaction plots |
| Unit 9 | Simple Linear Regression |
| PCM9a | 1 Course overview and concept review  1.1 What is statistics?  1.1.1 Statistical process  1.3 Exploratory data analysis  1.3.1 Use as a data validation tool  1.3.3 Applied review of numerical summaries  1.3.3.2 Measures of spread  1.3.3.4 Diagnosing outliers  1.3.4 Applied review of graphical tools  1.3.4.4 Scatterplots  1.3.4.4.1 Grouping/coloring  2 Principles of data collection  2.2 Sampling methods  2.2.1 SRS  2.2.3 Convenience sample  2.2.4 Volunteer response sampling  2.2.5 Importance of randomization  2.2.6 Benefits and drawbacks |
| PCM9b | 3 Probability  3.2 Random variables  3.2.2 Continuous RVs  3.2.2.1 Functions defining the distribution (PDF, CDF)  3.2.2.2 Finding probabilities using PDF or CDF  3.2.2.3 Expected value  3.2.2.5 Normal RVs  3.2.2.5.1 PDF, mean, variance  3.3 Sampling distributions  3.3.1 Definition/purpose  3.3.2 CLT  3.3.4 T, Chi-square, F |
| PCM9c | 4 Inference  4.1 Confidence intervals concepts  4.1.1 Why aren’t point estimators enough?  4.1.2 Margin of Error  4.1.3 Relation to sampling distribution  4.1.4 Confidence idea  4.1.5 Interpretation of observed interval  4.1.6 Checking conditions for inference  4.1.7 Scope of inference (relation to bias, etc.)  4.2 CI applications  4.2.5 CIs for slope terms  4.3 Hypothesis testing concepts  4.3.1 Hypotheses  4.3.2 Conclusions and possible errors  4.3.3. Significance level  4.3.4 Relation to sampling distributions  4.3.5 P-value  4.3.7 Interpretation  4.4. HT applications  4.4.8 Simple linear regression  4.4.8.1 HTs for “slope” terms (SLR)  4.4.8.2 Fitted values  4.4.8.3 Interpretations of model terms |

Notes about modules:

Course level LOs:

1. construct and interpret common numerical and graphical summaries of data
   * Use of JMP to read in data in unit 2, create bar graph, grouped bar graph
   * Hit same thing in unit 4
   * Unit 5 numeric summaries, qqplots (just interpret, not understand all that goes into creating it), histogram, boxplots
   * Unit 6/7/8 side-by-side box plots, qqplots, means/variances for groups
   * Unit 9 scatterplots, correlation
2. recognize and account for sources of variability and bias in experiments and observational studies
   * throughout
3. use statistical distributions to draw inference and quantify uncertainty
   * throughout
4. determine and conduct valid statistical inference
   * throughout
5. draw appropriate real world conclusions given the sources of data (Scope of Inference)
   * Unit 2, 3, 6, 7, 9

For sampling distributions we aren’t going to worry as much about theory but understanding that knowing a predictable pattern for statistics is what allows us to make inference.

Pattern will tend to be normal for many situations – but not all!

State distributions as fact and give some simulation examples.

Bead box for “C” day in Unit 2.