

# STAT 3011 (006) Fall 2022 Exam 1 Information

Yu Yang

## 1 General Information About the Exam

1. **Exam 1 will be held on Thursday, October 20th, 5:00 pm - 6:30 pm, at Anderson Hall 310 (West Bank).**
2. Exam 1 has 10 to 16 multiple choice questions, (worth 30%-40% of the exam score), (including one question asking "did you circle your multiple-choice answers on the last page of the exam?" See item 9) and 3-5 short answers questions with sub-questions (worth 60% - 70% of the exam score). Exam 1 will be based on Chapter 1, 2, 4, 5, 6, 7 that we learned in class, lab, and homework.
3. You can bring one letter-sized (8.5" by 11") cheat sheet either typed or hand-written to Exam 1. You may use both sides of your cheat sheet. You can use any content from Canvas/homework/lecture, etc. **Please write your name on both sides of your cheat sheet. Please note that your cheat sheet will be collected with your exam and will not be returned. If you want to keep your cheat sheet, please make a photocopy of it before the exam.**
4. Bring a scientific or graphic calculator. Phone/computer calculators are NOT allowed. If you have a graphing calculator with a color screen, please plan to seat in the front. I do not want to wrongly accuse you of using a cell phone during the exam and it is hard to tell from far away.
5. Bring your U-card or photo ID. Please have your ID out while taking the exam.
6. I will provide scratch paper if needed. Do not bring your own.
7. When you finish your exam :
  - Hand in your exam & cheat sheet to the teaching staff (instructor/TAs).
  - Show your ID.
  - Find your name on the class roster and write your initial.
  - Please DO NOT discuss the exam with other students when you are in line waiting to hand in your exam.
8. Sharing cheat sheets and/or sharing a calculator during the exam is NOT allowed.
9. Please circle your multiple choice answer on the last page of the exam. Failing to do so will result 0 for your multiple choice part.
10. All necessary R command results will be provided. But not all are useful.

## 2 Topics Covered

1. Chapter 1 Introduction
  - Know definitions of sample, population, statistic, parameter and can identify them from a given scenario.
2. Chapter 2 Exploring Data
  - Identify different types of data (categorical, quantitative (discrete, continuous)).
  - Know shapes/pattern/characteristics of distributions (uniform, skewed, symmetric, bell-shaped, uni-modal, etc.).
  - Know relationships between mean and median in different shapes of distribution.

- Understand summaries of categorical data (frequency table, pie chart, bar plot).
- Understand numerical & graphical summaries of quantitative data (stem-and-leaf plot, histogram, boxplot, 5-number summary).
- Know how to construct graphical summaries using R (histogram, boxplot, side-by-side boxplot, etc.) with appropriate titles.
- Calculate mean, median, variance, standard deviation, and range by hand for a small data set and understand their properties.

### 3. Chapter 4 Gathering Data

- Know the definition of explanatory variable, response variable, and identify each from a given example.
- Distinguish between types of study (randomized experiment vs observational study) and whether we can establish a cause-and-effect relationship.

### 4. Chapter 5 Probability

- Identify sample space and events in a given example.
- Understand and calculate special events such as complimentary event, intersection of two events, union of two events, and conditional probabilities based on the given tables and scenarios.
- Know how to use general addition rule.
- Know how to use addition rule for disjoint events.
- Know how to use general multiplication rule.
- Know how to use multiplication rule for two independent events.
- Understand meaning of two events being independent and verify/check the independence of two events.

### 5. Chapter 6 Probability Distribution

- Understand/identify two types of random variables (discrete/continuous)
- Construct a probability distribution of a discrete random variable and calculate the probability / mean /variance /standard deviation
- Understand properties of the continuous probability distribution (such as  $P(X = a) = 0$  for any  $a$ ).
- Understand properties of normal distribution/standard normal distribution.
- Understand the effect of changing the mean and standard deviation of normal distributions.
- Understand and use the 68-95-99.7 rule to calculate approximate probabilities of normal distributions.
- Calculate  $z$ -score of  $x$ , and interpret.
- Compare  $z$ -scores to determine which observation from two different distributions is better/worse.
- Find probabilities standardizing  $x$  into standard normal distribution, and then use `pnorm()` or `1-pnorm()` command.
- Find a value of  $z$ -score/quantile that marks  $p$ -th percentile using R command `qnorm()` and un-standardization, and then interpret.
- Identify the shape of distribution using Q-Q plot (normal,right-skewed/left-skewed/heavy tailed).
- Know conditions of binomial probability distribution.
- Calculate binomial distribution probabilities.
- Use R commands `dbinom(x, size, prob)` and `pbinom(q, size, prob)` to calculate binomial distribution probabilities.

- Calculate the mean and variance/standard deviation of binomial distribution using formulas.

## 6. Chapter 7 Sampling Distribution

- Understand the meaning of the sampling distribution of a sample statistic and distinguish it from population distribution/probability distribution.
- Understand properties of, and know how to calculate the mean and the standard deviation of the sampling distribution of a sample proportion.
- Know when the sampling distribution of the sample proportion is approximately normal (sample size is large enough) (Central Limit Theorem).
- Understand the effect of sample size ( $n$ ) on the sampling distribution of the sample mean (shape, center, spread).
- Can apply the 68-95-99.7 rule to the sampling distribution of a sample proportion.
- Understand definitions/properties of the mean and the standard deviation of the sampling distribution of the sample mean.
- Understand the effect of sample size on the sampling distribution of the sample mean (shape, center, spread).
- Know when the shape of the sampling distribution of a sample mean is normal / approximately normal;
  - (a) when population distribution is normal,
  - (b) when population is unknown/not normal (if  $n$  is large, by CLT).
- Find the probability of a sampling distribution of a sample mean by calculating the  $z$ -score and by using `pnorm()`.
- Find a particular sample mean value that marks the  $p$ -th percentile using the  $z$ -score that marks the  $p$ -th percentile in the standard normal distribution and unstandardization.