SC2000 Freshmen Exemption Test Information

Syllabus

This weekly schedule outlines the main topics that will be covered throughout the course.

- **Week 1:** Descriptive Statistics
- **Week 2:** Descriptive Statistics & Probability Theory
- **Week 3:** Probability Theory
- **Week 4:** Discrete Probability Distribution
- **Week 5:** Discrete & Continuous Probability Distribution
- **Week 6:** Continuous Probability Distribution
- **Week 7:** Continuous Probability Distribution
- **Week 8:** Sampling Distributions
- **Week 9:** Large-Sample Estimation
- **Week 10:** Large-Sample Tests of Hypotheses
- **Week 11:** Large-Sample Tests of Hypotheses
- **Week 12:** Inference from Small Samples
- **Week 13:** Linear Regression

Reference Books

The primary recommended textbooks for this course are:

Online Statistics Education: An Interactive Multimedia Course of Study (David M. Lane et a): Free Statistics Book . Relevant chapters for SC2000 are:

- **Ch. 1–4**: Introduction, Descriptive Statistics, Graphing Data, Summarizing Distributions
- **Ch. 5–7**: Probability, Discrete & Continuous Distributions, Normal Distribution
- **Ch. 9–11**: Sampling Distributions, Estimation, Hypothesis Testing
- **Ch. 12–13**: Testing Means, Statistical Power
- **Ch 14:** Regression

b) Introduction To Probability And Statistics Metric Edition 15th Edition William Mendenhall| Robert J. Beaver| Barbara M. Beaver SBN-13: 9780357114469 | ISBN-10: 0357114469 ©2020

Additionally, we refer to the following book:

Title: *Miller & Freund's Probability and Statistics for Engineers, 9th Edition*

Authors: R.A. Johnson, I. Miller, and J. Freund

The course topics generally correspond to the following sections in the book:

- **Descriptive Statistics:** Chapters **1 and 2** covering introduction to statistics, data presentation, and measures of central tendency and variation.
- **Probability Theory:** Chapters **3** covering sample space, events, conditional probability, and Bayes' theorem etc.
- **Probability Distributions:** Chapters **4 and 5** covering discrete and continuous random variables, including Bernoulli, Binomial, Poisson, Uniform, Exponential, and Normal distributions.
- **Sampling Distributions:** Chapter **6** explaining the central limit theorem and the sampling distributions of the mean and proportion.
- **Estimation & Inference:** Chapters **7 and 10** on point estimates, confidence intervals, and hypothesis testing for large and small samples (including t-distributions).
- **Simple Linear Regression and Correlation:** Chapter **11** focused on bivariate data, Pearson correlation, and fitting a linear regression model.

Sample Questions

Descriptive Statistics

1. A dataset has the following observations: 2.5, 2.7, 2.9, 3.0, 3.1, 3.4, 3.6, 3.8, 4.2, 4.5, 4.8, 5.0.

   Find the 25th, 50th, and 75th percentiles.

2. Construct a box plot for the following droplet sizes: {2.1, 2.8, 2.9, 3.5, 3.6, 5.1, 5.3, 7.9, 8.9}.

Probability & Bayes' Rule

3. **Medical Test Accuracy (Bayes' Rule)**
   A rare disease affects 1 in 1,000 people. A diagnostic test has:

- Sensitivity (true positive rate) = 99%

- Specificity (true negative rate) = 95%
  If a randomly selected person tests positive, what is the probability that they actually have the disease?

4. A box contains four coins with values of 1, 2, 5, and 10 units. Three coins are selected at random without replacement. What is the probability that the total value of the selected coins is **at least 12 units**?

5. A biased die is rolled 50 times and the number of twos observed is 10. Estimate the probability of rolling exactly three twos in the next 10 rolls.

Discrete Distributions

6. A random variable has probabilities P(X=0)=0.2, P(X=1)=0.4, P(X=2)=0.3, P(X=3)=0.1. Find:
   (a) P(X ≥ 2)
   (b) E[X] and Var(X).

Continuous Distributions

7. The waiting time in a queue follows an exponential distribution with mean 4 minutes.
   (a) Find the variance.
   (b) Find the probability of waiting at least 10 minutes.

Normal Distribution

8. Test scores follow a normal distribution with mean 75 and standard deviation 10. What score marks the cutoff for the top 2.5% of students?

9. If X ~ N(16.2, 1.5625), compute P(X > 16.8).

Sampling Distributions

10. Suppose the weight of apples follows a normal distribution with μ=200g and σ=30g. If a sample of 36 apples is taken, what is the probability that the sample mean weight is greater than 210g?

11. The lifetime of a type of bulb follows a normal distribution with μ=1200 hours and σ=200 hours. Find the probability that the average lifetime of a sample of 25 bulbs is less than 1150 hours.

12. Assume scores of a test are normally distributed with the mean 920 and standard deviation 150. If 9 scores are randomly selected, find the probability that they have a mean between 914 and 929.

Estimation

12. A sample of 64 students has an average exam score of 70 with a standard deviation of 12. Construct a 95% confidence interval for the population mean.

13. In a survey of 400 voters, 220 supported a candidate. Construct a 99% confidence interval for the proportion of support.

Large Hypothesis Testing

14. A factory claims that the average lifetime of its batteries is 500 hours. A sample of 36 batteries has a mean lifetime of 485 hours with σ=60 hours (assume normality). At the 5% significance level, test the manufacturer's claim.

15. Two teaching methods are compared. Group A (n=40) has mean = 75, σ=10. Group B (n=36) has mean = 78, σ=12. Test at the 5% level whether there is a significant difference in mean scores.

Small Hypothesis Testing

16. Suppose you do a study of acupuncture to determine how effective it is in relieving pain. You measure sensory rates for 15 subjects with the results given. Use the sample data to construct a 95% confidence interval for the mean sensory rate for the population (assumed normal) from which you took the data. The data are given below.

    7.7, 8.6, 8.0, 7.7, 7.1, 8.2, 7.2, 9.5, 9.8, 6.9, 9.1, 7.6, 7.8, 9.6, 5.4

17. A hypothesis test is to be performed to determine whether the mean waiting time during peak hours for customers in a supermarket has increased from the previous mean waiting

time of 8.2 minutes. Previous experience indicates that the waiting time follows a normal distribution with standard deviation equal to 3.8 minutes. To test the hypothesis, a random sample of 25 customers will be selected yields mean = 9.75.