

Topics for Final Exam

1. Exploratory data analysis

- 1.1 Variables: quantitative and qualitative
- 1.2 Distribution of a variable
- 1.3 Describing distributions: location, spread, shape, numerical summaries
- 1.4 Pictures of distributions: histograms and box plots

2. Significance Testing

- 2.1 Hypothesis testing: null hypotheses or model, test statistic, p -value.
- 2.2 Definition of statistical significance

3. Probability

- 3.1 Definition and interpretations: relative frequency interpretation and probability models.
- 3.2 Estimating probabilities through simulations
- 3.3 Independent events and conditional probabilities
- 3.4 Computing probabilities: equally likely events, independent events, mutually exclusive events, probabilities of joint events

4. Random Variables and their Distributions

- 4.1 Definition of a random variable: random experiments and sample spaces.
- 4.2 Discrete versus continuous random variable
- 4.3 Probability distribution of a discrete random variable
- 4.4 Cumulative distribution
- 4.5 Expectation of a random variable
- 4.6 The Law of Large Numbers
- 4.7 Variance and standard deviation of a random variable.

5. Sampling

- 5.1 Random sampling
- 5.2 Sampling with and without replacement

5.3 Parameters and statistics

5.4 Sampling distribution

6. Experiments

6.1 Randomized comparative experiment

6.2 Control group

6.3 Randomization test

6.4 Hypothesis testing

7. Estimation

7.1 Sampling distributions

7.2 Independent Bernoulli random variables and the Binomial distribution

7.3 The Central Limit Theorem

7.4 Estimating proportions: Confidence interval

7.5 Confidence interval for the mean

8. Goodness of Fit Test

8.1 The Chi-Squared distance

8.2 The Chi-Squared distribution

8.3 The Chi-Squared goodness of fit test

9. Association Between Categorical Variables

9.1 Two-way tables

9.2 Joint distributions; marginal distributions; independence.

9.3 The Chi-Squared test for independence

Relevant chapters and sections in the text:

- Sections 1.1 and 1.2 in Chapter 1
- Sections 4.1–4.5 in Chapter 4
- Chapter 5 on *Sampling Distributions*

- Sections 6.1 and 6.2 in Chapter 6
- Chapter 9 on *Analysis of two-way tables* and *Goodness of Fit*

Relevant chapters in the Class Notes:

Chapters 2, 3, 4, 5, 6 and Appendix A

Important Concepts

Quantitative and categorical variables; distributions; sampling; random sample; parameter; statistic; sampling distribution; probability; probability distribution; statistical significance; p -value; random variable; expected value; variance; standard deviation; Bernoulli trials; Binomial distribution; the Central Limit Theorem; the normal distribution; confidence interval estimates; Chi-Squared distance; the Chi-Squared distribution; goodness of fit; two-way tables.

Important Skills

1. Know how to estimate or compute probabilities
2. Know how to compute expected values and variances of random variables
3. Know how to set up significance tests
4. Know how to obtain confidence intervals for means and proportions
5. Know how to perform a goodness of fit test
6. Know how to perform a Chi-Squared test of independence