Chapter 2

- Variable:
  - Categorical: nominal or ordinal
  - Numeric: continuous or discrete
- Frequency table and histogram
- Shapes of distributions:
  - Symmetric and bell-shaped
  - Skewed to the right
  - Skewed to the left
  - How to tell if a distribution is left skewed or right skewed?
    - Mean v.s. median
    - Normal quantile plot
  - Transformations to correct skewness
  - Skewness of Binomial distribution

- Descriptive statistics:
  - mean and median
  - standard deviation, IQR, and range
  - percentile and quantile
  - robustness
- Boxplots
- Effect of transformation of variables
  - How does mean, median, SD, IQR change under linear transformation $Y' = aY + b$?

Chapter 3

- Probability rules (1) to (8)
  - Disjoint v.s. independent
  - Conditional probability
- Random variables:
  - Discrete:
    - pmf, summation equals 1
    - Mean and variance
    - Binomial distribution
  - Continuous:
    - pdf, integral equals 1
    - Areas under the curve

        - Normal distribution
    - Discrete v.s. continuous: the probability of single point

- Rules for means of random variables (1) to (4)
- Binomial distribution $B(n, p)$:
    - Binary outcomes, independent trials, n is fixed, same value of p
    - pmf
    - Properties of binomial coefficients
    - The use of complement to calculate probability
    - Mean and variance

Chapter 4

- Normal distribution $N(\mu, \sigma^2)$:
    - Note that the second argument denotes the variance, rather than the SD
    - pdf
    - Areas under a normal curve: standardization and Z table
    - Inverse reading of Z table
    - Empirical rule for normal distribution
    - Assessing normality: normal quantile plot (only R implementation required)

Chapter 5

- The sampling distribution of $\bar{Y}$
    - $\mu_{\bar{Y}} = \mu$, $\sigma_{\bar{Y}} = \sigma/\sqrt{n}$
    - If $Y \sim N(\mu, \sigma^2)$, then $\bar{Y} \sim N(\mu, \sigma^2/n)$
    - Central Limit Theorem
- Normal Approximation to the Binomial Distribution
    - $Y \overset{\text{approx}}{\sim} N(np, np(1-p))$ and $\hat{p} \overset{\text{approx}}{\sim} N(p, p(1-p)/n)$
    - Continuity connection
    - $np \geq 5$ and $n(1-p) \geq 5$

Chapter 6

- Student's t distribution
    - t Table
- Confidence intervals for μ
    - Two-sided

    - One-sided: upper and lower
    - Conditions
    - Summary

      $1-\alpha$ confidence interval for μ:

      * Two-sided: $\bar{Y} \pm t_{n-1}(\alpha/2) \times \text{SE}_{\bar{Y}}$
      * Upper one-sided: $(-\infty, \bar{Y} + t_{n-1}(\alpha) \times \text{SE}_{\bar{Y}})$
      * Lower one-sided: $(\bar{Y} - t_{n-1}(\alpha) \times \text{SE}_{\bar{Y}}, \infty)$

      where

      $$\text{SE}_{\bar{Y}} = \frac{s}{\sqrt{n}}$$

- Interpretation of a confidence interval
- Planning a study to estimate μ
- Confidence intervals for $\mu_1 - \mu_2$
    - Two-sided
    - One-sided: upper and lower
    - Conditions
    - Summary

      $1-\alpha$ confidence interval for $\mu_1 - \mu_2$:

      * Two-sided: $(\bar{Y}_1 - \bar{Y}_2) \pm t_\nu(\alpha/2) \times \text{SE}_{\bar{Y}_1 - \bar{Y}_2}$
      * Upper one-sided: $(-\infty, (\bar{Y}_1 - \bar{Y}_2) + t_\nu(\alpha) \times \text{SE}_{\bar{Y}_1 - \bar{Y}_2})$
      * Lower one-sided: $((\bar{Y}_1 - \bar{Y}_2) - t_\nu(\alpha) \times \text{SE}_{\bar{Y}_1 - \bar{Y}_2}, \infty)$

      where the degrees of freedom ν will be given and

      $$\text{SE}_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Chapter 7

- Two-sample t test
    - Two-sided
    - One-sided: left and right
    - p-value and critical value
    - Conditions
    - Summary

H_0	H_A	Test statistic	Rejection region	p-value
$\mu_1 - \mu_2 = c$	$\mu_1 - \mu_2 \neq c$	$T = \frac{(\bar{Y}_1 - \bar{Y}_2) - c}{\text{SE}_{\bar{Y}_1 - \bar{Y}_2}} \overset{H_0}{\sim} t_\nu$ with	$\lvert T \rvert > t_\nu(\alpha/2)$	$2 \times P(t_\nu > \lvert T \rvert)$
$\mu_1 - \mu_2 \geq c$	$\mu_1 - \mu_2 < c$	$\nu = \frac{(\text{SE}_1^2 + \text{SE}_2^2)^2}{\text{SE}_1^4/(n_1-1) + \text{SE}_2^4/(n_2-1)}$	$T < -t_\nu(\alpha)$	$P(t_\nu < T)$
$\mu_1 - \mu_2 \leq c$	$\mu_1 - \mu_2 > c$		$T > t_\nu(\alpha)$	$P(t_\nu > T)$

- Type I and Type II errors, power
- How are H_0 and H_A chosen

Chapter 8

- One-sample t test
  - Two-sided
  - One-sided: left and right
  - p-value and critical value
  - Conditions
  - Summary

H_0	H_A	Test statistic	Rejection region	p-value
$\mu = c$	$\mu \neq c$	$T = \frac{\bar{Y} - c}{\text{SE}_{\bar{Y}}} \overset{H_0}{\sim} t_{n-1}$	$\lvert T \rvert > t_{n-1}(\alpha/2)$	$2 \times P(t_{n-1} > \lvert T \rvert)$
$\mu \geq c$	$\mu < c$		$T < -t_{n-1}(\alpha)$	$P(t_{n-1} < T)$
$\mu \leq c$	$\mu > c$		$T > t_{n-1}(\alpha)$	$P(t_{n-1} > T)$

- Paired-sample
  - One-sample for differences
  - Conditions

Chapter 9

- Confidence intervals for p
  - Two-sided
  - One-sided: upper and lower
  - Conditions
- Chi-square goodness-of-fit test
  - Null and alternative hypotheses
  - Test statistic and its null distribution
  - Conditions

Chapter 10

- Confidence intervals for $p_1 - p_2$
  - Two-sided
- Chi-square test of independence
  - Null and alternative hypotheses
  - Test statistic and its null distribution
  - Conditions

Chapter 11

- One-way ANOVA
  - ANOVA table
  - F test
  - Conditions
- Multiple comparisons
  - Bonferroni's method

Chapter 12

- Correlation coefficient
  - Interpretation
  - Test for population correlation
- Fitted regression line
  - b_0 and b_1
  - Interpretation
  - SSE and residual standard deviation
  - Coefficient of determination
  - Confidence intervals and tests for the slope
  - Conditions