Final Exam (40% grade)

- **Study materials:** Textbook **chapters 15 to 25, selected sections in chapters 26-27**; Lecture note week **1 to 15**; Quiz#1-5/practice
- **Policy:**
    - closed book test; no computer
    - you can bring 2 page (8.5 x 11 in) sheet of note only. A calculator is allowed. We provide you a blank scratch paper.

Understanding of the following important concepts (***Note:*** Below includes minimum knowledge. You should study thoroughly text book material and lecture notes to be well prepared for the exam):
-We will not test R code but understanding some standard regression/model output is expected.

Fundamental Statistical Concepts (prerequisite): Appendix A; Lecture note 1A

Materials Covered before Midterm: (30-40%)

Chapter 15: Introduction

1. What is the difference of experimental study and observational ?
2. Observational studies (confounding factors; prospective vs. retrospective studies)
3. Identify the response/outcome variable, explanatory/treatment factors, experiment units

Chapter 16-18: Single factor studies

1. Sample size planning:
    a. based on power and type I error or based on estimation precision
    b. balanced design, treatment levels, target treatment effect size and variability.
    c. How to design parameter changes affect the sample size (α, β, Δ, σ, CI-width) ?
2. Diagnostics of ANOVA model assumptions and remedial measures
    a. What statistical tests can be used to examine these assumptions
    b. Residual analysis
    c. Transform the data + ANOVA vs. nonparametric test
3. Understand the cell means model and the factor effects model: parameters and estimates
4. Understand the source of variation, the sum of square partition and ANOVA table decomposition: the relationship of SS, MS, df.
5. The hypothesis testing regarding factor level means: how to compute the *F*-statistics and its distribution based on corresponding df, and interpret the results.
6. Compute estimate and CI for factor level means, their paired differences or contrasts given the summary statistics (mean, sample size, ANOVA table) and critical values.
7. Understand the multiple comparison methods: Tukey, Bonferroni, and Scheffé methods
    a. How to choose the most efficient procedure for various scenarios
    b. How to interpret the output/figure for pairwise comparison

Chapter 19: Two-factor studies

1. For classic balanced two-factor study with n>1 per treatment:
    a. Full model: cell mean model and factor effects model with interaction term
    b. Model assumption, ANOVA table, hypothesis testing, estimation and interpretation
    c. How to identify interaction from treatment mean plot, and interpret interaction effect
    d. How to estimate the factor level means, treatment means and contrasts.

Chapter 20-21: two factor studies with no replication and Randomized complete block design

1. How we can test for additivity (interaction)?
2. Factor effect model (no interaction): ANOVA table and interpretation
3. How to estimate the treatment means (given a summary of factor level means)
4. Randomized complete block design:
    a. Why use blocking
    b. Factor effects model: model assumption, ANOVA table and interpretation
    c. Estimation of the model parameters and treatment effects (from data summary statistics)

Materials Covered After Midterm (60-70%) :

Chapter 22 : Analysis of Covariance

- Objective of ANCOVA analysis
- Single-factor ANCOVA factor effects model

$$Y_{ij} = \mu. + \tau_i + \gamma(X_{ij} - \overline{X}..) + \varepsilon_{ij}$$

- Equivalent regression model: how to set the indicator functions in regression
- Two additional ANCOVA assumptions

a) Linearity: For each treatment, the expected response is given by a regression line with covariate

b) Constancy (Homogeneity) of the slope: all treatment regression lines have the same slope

- How to test and estimate model parameters,
- How to estimate treatment effects (p 930) and pairwise difference

Comparison	Estimator	Variance
$\tau_1 - \tau_2$	$\hat{\tau}_1 - \hat{\tau}_2$	$\sigma^2\{\hat{\tau}_1\} + \sigma^2\{\hat{\tau}_2\} - 2\sigma\{\hat{\tau}_1, \hat{\tau}_2\}$
$\tau_1 - \tau_3 = 2\tau_1 + \tau_2$	$2\hat{\tau}_1 + \hat{\tau}_2$	$4\sigma^2\{\hat{\tau}_1\} + \sigma^2\{\hat{\tau}_2\} + 4\sigma\{\hat{\tau}_1, \hat{\tau}_2\}$
$\tau_2 - \tau_3 = \tau_1 + 2\tau_2$	$\hat{\tau}_1 + 2\hat{\tau}_2$	$\sigma^2\{\hat{\tau}_1\} + 4\sigma^2\{\hat{\tau}_2\} + 4\sigma\{\hat{\tau}_1, \hat{\tau}_2\}$

(22.16)

- How to estimate mean response with the ith treatment with covariate at mean value (adjusted mean): p931

Mean Response at $X = \overline{X}..$	Estimator	Variance
$\mu. + \tau_1$	$\hat{\mu}. + \hat{\tau}_1$	$\sigma^2\{\hat{\mu}.\} + \sigma^2\{\hat{\tau}_1\} + 2\sigma\{\hat{\mu}., \hat{\tau}_1\}$
$\mu. + \tau_2$	$\hat{\mu}. + \hat{\tau}_2$	$\sigma^2\{\hat{\mu}.\} + \sigma^2\{\hat{\tau}_2\} + 2\sigma\{\hat{\mu}., \hat{\tau}_2\}$
$\mu. + \tau_3$	$\hat{\mu}. - \hat{\tau}_1 - \hat{\tau}_2$	$\sigma^2\{\hat{\mu}.\} + \sigma^2\{\hat{\tau}_1\} + \sigma^2\{\hat{\tau}_2\} - 2\sigma\{\hat{\mu}., \hat{\tau}_1\} - 2\sigma\{\hat{\mu}., \hat{\tau}_2\} + 2\sigma\{\hat{\tau}_1, \hat{\tau}_2\}$

(22.22)

- Two-factor ANCOVA model:

$$Y_{ijk} = \mu.. + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma(X_{ijk} - \overline{X}...) + \varepsilon_{ijk} \qquad (22.26)$$
$$i = 1, \dots, a; j = 1, \dots, b; k = 1, \dots, n$$

and its equivalent regression model.

- How to test interaction and main effects.
- How to compare factor level means using regression output (p. 937)

- ANCOVA analysis with random complete block design

$$Y_{ij} = \mu.. + \rho_i + \tau_j + \gamma(X_{ij} - \overline{X}..) + \varepsilon_{ij} \qquad i = 1, \dots, n_b; j = 1, \dots, r \qquad (22.29)$$

and its equivalent regression model.

- How to test the treatment effects (22.31) and compare two treatment effects τ_1-τ_2 (22.33).

Chapter 23 : Two-Factor Studies with Unequal Sample Sizes

- How to set the regression models to estimate the model parameters
- How to test the interaction and main effects using regression approach
- Inference on factor/treatment means including confidence interval and multiple comparison adjustment
  - Point estimator and SE, pairwise difference, contrasts/linear combination of factor level means
  - Point estimator and SE, pairwise difference, contrasts/linear combination of treatment means
  - Understand how to apply formulas in **Table 23.5** (p.961-962) using treatment sample means and treatment sample size (n_{ij}) : e.g. Example in p.963 or Quiz #5

- Inferences when Treatment Means Are of Unequal Importance
  - i.e. estimate the weighted treatment Mean and factor effects by using the formulas for the linear combinations in **Table 23.5:** e.g. Example in p.971

Chapter 24 : Multi-factor Studies

- ANOVA model for three-factor studies
  - Cell mean model
  - Factor effects model
- Balanced studies:
  - SS partition of the source of variation, ANOVA table, F-test for main effects , two-way interactions and three-way interactions (Table 24.5, Table 24.6)
  - how to fill in missing component of ANOVA table, compute the *F*-statistics, find their distributions based on corresponding df, and interpret the results

  - Inference on factor level means including confidence interval and multiple comparison adjustment
- Unbalanced studies with unequal treatment sample size (n_{jik})
  - How to set up the equivalent regression models and test the model parameters such as interactions.
  - How to make inferences on factor level means
    - Use similar formulas from Table 23.5
    - Use regression output: e.g. homework problem 24.15), each of factors (A, B and C) has 2 levels and we denote with one indicator function each and fit the regression. $\mu_{1..} - \mu_{2..} = 2\alpha_1$ then as shown in Lecture 13A.

      $\hat{\mu}_{1..} - \hat{\mu}_{2..} = 2\hat{\alpha}_1$ $\qquad$ $\hat{\sigma}(\hat{\mu}_{1..} - \hat{\mu}_{2..}) = 2\hat{\sigma}(\hat{\alpha}_1)$

Chapter 25 : Random and Mixed Effects Models

- Describe the differences when using fixed effects vs. random effects & related design consideration.
- One factor study with a random factor

$$Y_{ij} = \mu_{.} + \tau_i + \epsilon_{ij} \qquad (25.38)$$

τ_i are independent $N(0, \sigma_\mu^2)$
ϵ_{ij} are are independent $N(0, \sigma^2)$

- Random factor effects model (25.38) and model assumption for the factor effects
- Correlation among observations from the same factor level: intra-class correlation and how to interpret it
  - Hypothesis testing, $H_0 : \sigma_\mu^2 = 0$ $\qquad$ $H_a : \sigma_\mu^2 > 0$

- Testing using ANOVA table and how to interpret the test results.

- Two-factor studies with random factor(s)

  a. ANOVA model **II: random** effects model (when both factors are random)

b. ANOVA model **III**: **mixed** effects model (when A is fixed/B is random ; or when A is random/B is fixed)

For either model II or III,
- ANOVA model assumption and variance components
- Covariance (correlation) structure of observations: when two observations are independent or correlated?
- How to use ANOVA table (Table 25.5, 25.6) to test factor A, B, A:B interactions: obtain the correct F-test and F-distribution by comparing the E(MS), how to interpret the test results for the factor effects.

- Randomized Complete Block Design: random block

  - with random blocks and fixed treatment effect, as a special case the 2- factor mixed effect model.

  - How to use ANOVA table (Table 25.8) to test treatment effects.

- Three-factor studies with 3 random factors (Model II)
  - ANOVA model assumption and variance components
  - Covariance (correlation) structure of observations: when two observations are independent or correlated?
  - Hypothesis testing (ANOVA Table 25.9): how to construct the test-statistics for
    - the three-way interaction;
    - two-way interactions;
    - main effects (what is Approximate *F*-test, Pseudo *F*-test?)
- (**will not test)** three-factor ANOVA with mixed factor effects.

- (**will not test)** estimation and CIs of the fixed and random effects parameters in balanced or unbalanced one-factor, 2-factor and 3-factor studies (OK if you can use software to complete homework and understand example in lecture note)

Chapter 26-27 : Key concepts in other design topics (to be summarized in 12/5/19)

Chapter 26. Nested Designs
- **Difference of nested vs. cross factor**
- **Two factor nest design model (26.7) with factor A and factor B nested with A**
- (**will not test)** ANOVA table and estimation as long as you understand how to run the data analysis with nest designs in lecture note

Chapter 27. Repeated measured Design
- (**will not test in the final) It is OK if you can understand how the to mode and analyze data from repeated design treating subject as a random block factor and understand the split-plot design.**