MATH324 - Final Exam Review

Introduction

These notes are to be used as an additional resource in studying for the final exam for MATH324 for Fall 2018. The notes are organized into three separate sections: "Study Modules", "General Review Questions" and "Specific Review Questions". The "Study Modules" section is meant to summarize the key points of the course into distinct "modules" so that one may study by module as opposed to the course as a whole. The "General Review Questions" will complement the "Study Modules" section by stating general questions related to the course material but will not contain any specific calculation type problems. The "Specific Review Questions" section will contain specific problems that require the application of tools covered in the other two sections.

Study Modules

This section attempts to summarize the course into distinct modules that may be studied indvidually. While there is overlap between the sections, this is meant as a guide as to how to structure studying for the course.

- **Module 1 - The Normal Distribution**
  - Distributions related to the Normal Distribution:
    + Sample Average of Normals is Normal
    + Sum of Squares of Standardized Normals is Chi-Squared
    + Sample Variance Standardized by Population Variance is Chi-Squared
    + A Standard Normal Divided by a Square-Root of a Standardized Chi-Squared Distribution is t-distributed
    + The ratio of Standardized Chi-Squared Distributions is F-distributed
  - Using the Z-table:
    + Calculate the value of the probability from an observed quantity
    + Calculate the observed quantity that yields the specific probability value
  - The Central Limit Theorem:
    + Statement of the Theorem
    + Application of Theorem to calculate probability statements related to the sample mean
  - The Normal Approximation to the Binomial Distribution:
    + Why can the approximation be made?
    + Continuity Correction
    + Re-express Binomial as Normal and calculate probability

- **Module 2 - Basic Estimator Properties**
  - Estimator:
    + Bias
    + Mean Square Error
    + The Sample Average
    + Error of Estimation
- **Module 3 - Confidence Intervals**
  - Language:
    + One Tail vs. Two Tail Confidence Intervals
    + Lower & Upper Confidence Limits
    + The Pivotal Quantity
    + Interpretation of the Interval
  - Large-Sample Confidence Intervals:
    + Relation to Central Limit Theorem
    + Use of the Z-table
    + Types of Target Parameters: μ, p, $\mu_1 - \mu_2$, $p_1 - p_2$
    + Selecting a Sample Size
  - Small-Sample Confidence Intervals:
    + Required Assumptions to form interval
    + Use of the t-table
    + Notion of a Pooled Sample Variance
    + Types of Target Parameters: μ, $\mu_1 - \mu_2$
  - Confidence Interval for σ^2:
    + Required Assumptions to form interval
    + Use of the Chi-Squared table
- **Module 4 - Advanced Estimator Properties**
  - Variance and Convergence in Probability:
    + Relative Efficiency
    + Consistency
    + Consistency of Functions of Consistent Estimators
    + Slutsky's Lemma
  - Sufficiency and Related Ideas:
    + Definition of Sufficiency
    + The Factorization Theorem
    + Definition of Minimal Sufficiency
    + Lehmann-Scheffé Lemma
    + Definition of Minimum-Variance Unbiased Estimator (MVUE)
    + The Rao-Blackwell Theorem
  - Estimation Techniques:
    + Method of Moments (M.O.M)
    + Method of Maximum Likelihood (M.L.E)
    + M.L.E. Invariance, Consistency, Asymptotic Normality
    + Fisher Information Matrix

- **Module 5 - Hypothesis Testing**
  - Language:
    + Null Hypothesis
    + Alternative Hypothesis
    + Test Statistic
    + Rejection Region
    + Type I Error
    + Type II Error
  - Simple Hypotheses:
    + Definition
    + Neyman-Pearson Lemma
    + Definition of Randomized Test
  - Composite Hypotheses:
    + Definition
    + Likelihood Ratio
  - More Common Hypothesis Testing Problems:
    + Population mean (μ)
    + Population variance (σ^2)
    + Population proportion (p)
    + Difference in population means ($\mu_1 - \mu_2$)
    + Difference in population proportions ($p_1 - p_2$)
  - The p-value:
    + Definition
    + Calculating the p-value
    + Interpretation
- **Module 6 - Linear Models**
  - Simple Linear Regression:
    + The Model
    + Least Squares
    + Simple Linear Regression Parameter Estimators
    + Properties of Simple Linear Regression Parameter Estimators
    + Hypothesis Testing for Regression Parameter Estimators

General Review Questions

Module 1

1. Let $Y_1, Y_2, ..., Y_n$ be a random sample from a Normal distribution with mean μ and variance σ^2. What is the exact distribution of $\bar{Y}$? What is the exact distribution of $\sum_{i=1}^{n}((Y_i - \mu)/\sigma)^2$?

2. Let $Y_1, Y_2, ..., Y_n$ be a random sample from a Normal distribution with mean μ and variance σ^2. What is the exact distribution of $\frac{(n-1)S^2}{\sigma^2}$?

3. Let Y be Normally distributed with mean μ and variance σ^2 and let W be an independently χ^2 distributed random variable with v degrees of freedom. What is the distribution of $\frac{\frac{Y-\mu}{\sigma}}{\sqrt{\frac{W}{v}}}$?

4. If the Z-table (Table 4 in the textbook) only lists probability values for one-tail of the Normal distribution, why is it possible to calculate probabilities which use both tails? Is it possible to use Table 4 for a Normally distributed random variable with arbitrary mean μ and variance σ^2?

5. State the Central Limit Theorem

6. Given a random sample $X_1, X_2, ..., X_n$, is it required that the individual X_i are Normally distributed in order to apply the Central Limit Theorem?

7. What does it mean that a random variable, such as $\bar{Y}$ is *asympotically Normally distributed*?

8. If the Central Limit Theorem is only specific to the sample size tending to ∞, why can we use it when calculating specific probabilties? Will our calculated answer be exact or only approximate?

9. When approximating a Binomially distributed random variable with a Normally distributed random variable, why is the *continuity correction* suggested?

10. Given a Binomially distributed random variable, X, with parameters n and p, what are the expected value and variance of X?

Module 2

1. Give the mathematical definition of *bias* for an estimator an unknown parameter θ.

2. What does it mean for an estimator to *unbiased*? Express your answer to this question using your answer to the previous question.

3. Give the mathematical definition of *mean squared error* for an estimator of an unknown parameter θ. Express your formula in terms of the variance of the estimator and the bias of the estimator.

4. Give the mathematical definition of *error of estimation.*

5. Rewrite Table 8.1 on page 397 of the textbook.

Module 3

1. How do one-tail confidence intervals differ from two-tail confidence intervals?
2. What is a pivotal quantity? Does the pivotal quantity usually depend on the unknown parameter? What is special about the distribution of a pivotal quantity?
3. Suppose we are interested in whether a parameter is equal to 0 and we found a two-tailed confidence interval of $(-3, 4)$. What does this *suggest* about the unknown parameter?
4. In creating a large-sample confidence interval, is it necessary that the data are Normally distributed?
5. How is the Central Limit Theorem used for making large-sample confidence intervals?
6. For target parameters, μ, p, $\mu_1 - \mu_2$ and $p_1 - p_2$, write out the forms of their respective two-tailed 95% confidence intervals. For 99% confidence intervals, what quantity changes in your previous formulas?
7. How can the large sample confidence intervals in the previous question be used to calculate the sample size? Is the pivotal quantity required for this type of calculation?
8. What assumptions are required in constructing a small-sample confidence interval?
9. When forming a small-sample confidence interval, why is it necessary to refer to a t-table rather than the z-table?
10. What assumptions are required when using the pooled sample variance?
11. When forming a confidence interval for the variance parameter σ^2, what table must you refer to in order to obtain the critical values? Why do we refer to this table and not say the z-table instead?

Module 4

1. State the mathematical formula for the *efficiency* of one estimator relative to another. What can be concluded if the relative efficiency is > 1, $= 1$ or < 1?
2. Give the definition of *consistency.*
3. State Tchebysheff's Theorem.
4. How does Tchebysheff's Theorem relate to proving that an unbiased estimator is consistent? (Hint: Refer to Theorem 9.1)
5. If $\hat{\theta}$ is consistent, what conditions on the function g are required so that $g(\hat{\theta})$ is also consistent for $g(\theta)$?
6. Is it true that sums, differences, products and division of consistent estimators will still remain consistent? Is there any additional assumption required for the ratio of two consistent estimators?
7. State Slutsky's theorem.
8. Give the definition of a *sufficient* estimator.

9. Is the entire sample $(X_1, X_2, ..., X_n)$ always a sufficient statistic by definition?
10. For a random sample of i.i.d. data with density $f(x;\theta)$, state the form of the likelihood.
11. How can the likelihood be used in the factorization theorem to obtain a sufficient statistic?
12. What does it mean for a statistic to be *minimally sufficient*?
13. Given a random sample of i.i.d. data, how would one find the minimally sufficient statistic using the likelihood function? What is the name of this lemma?
14. Give the definition of MVUE. Explain how the Rao-Blackwell Theorem can be used to find an MVUE.
15. Suppose that T is a sufficient statistic for θ, and that $g(T)$ is an unbiased estimator for θ, for some function g. What can be said about the estimator $g(T)$? How does this result follow from the Rao-Blackwell theorem?
16. When deriving a method of moments estimator, explain how the sample moments are used to estimate the population moments.
17. When using the method of moments estimation approach in a problem where there are 5 unknown parameters, at least how many method of moments equations will be required to derive a method of moments estimator?
18. How is the likelihood used when determining the maximum likelihood estimator?
19. Is the maximum likelihood estimator obtained when maximizing the likelihood the same as when the log-likelihood is maximized?
20. Suppose $\hat{\theta}$ is the maximum likelihood estimator and g is a function, then what conclusion can be made about $g(\hat{\theta})$?
21. Are maximum likelihood estimators consistent? Are they asymptotically Normally distributed?
22. State a form of the Fisher Information Matrix. Under certain regularity conditions, it can be shown that the Fisher Information Matrix has an alternate form. What is it? How does the Fisher Information matrix relate to the asymptotic distribution of the maximum likelihood estimator?

Module 5

1. Give the definitions of the following terms: Null Hypothesis, Alternative Hypothesis, Test Statistic, Rejection Region, Type I Error and Type II Error.
2. How does Type I Error differ from Type II Error?
3. Given a null hypothesis, an alternative hypothesis, a test statistic and a rejection region, what statement can be made if the test statistic falls in the rejection region?
4. What are simple hypotheses?
5. State the Neyman-Pearson Lemma. What is the conclusion of the theorem?
6. What is a randomized test? How does it differ from testing for simple hypotheses?
7. What is the likelihood ratio statistic? How can it be used for simple hypothesis testing?

8. Rewrite the Large-Sample α-Level Hypothesis Tests box on page 500 of the textbook. Relate this box to table 8.1 on page 397 of the texbtook.

9. For the parameters μ and $\mu_1 - \mu_2$, how do the large sample tests change when the sample size is small? (Refer to the highlighted box in page 10.8)

10. What is a p-value? If the significance level $\alpha = 0.05$ and the p-value is 0.01, what statement can be made regarding the null and alternative hypotheses? What if the p-value was 0.08?

Module 6

1. Give the definition of a linear statistical model.

2. A simple linear model contains only an intercept and a slope parameter. Give the mathematical definition of this type of model.

3. State the sum of squares expression for the simple linear model between the observed response and the fitted model.

4. What are the Least Squares estimators for the simple linear regression model?

5. State the expected value of the slope parameter estimator for the simple linear regression model.

6. State the variance of the intercept parameter and the slope parameter for the simple linear regression model.

7. Summarize the "Properties of the Least-Squares Estimators; Simple Linear Regression" box on page 582 of the textbook. Make reference to the previous text to explain why each item is true.

8. Give the Test of Hypothesis for β_1, the Test statistic and the Rejection region.

9. Give a $100(1 - \alpha)\%$ Confidence Interval for β_1.

Specific Review Questions

Module 1

1. (**Normal Distribution Properties**) Suppose that $X_1, X_2, ..., X_n$ are a random sample from a Normal distribution with mean μ_1 and variance σ_1^2 and $Y_1, Y_2, ..., Y_n$ are a random sample from Normal distribution with mean μ_2 and variance σ_2^2.
   (a) If $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, what is the distribution of $\bar{X}$?
   (b) What is the distribution of $\bar{Y}$?
   (c) What is the distribution of $\bar{X} + \bar{Y}$?

2. (**Using the Normal Table**) For each of the problems listed below, use Table 4 in the Appendix of the textbook.
   (a) Let Z be a random variable with a standard Normal distribution (i.e. $N(0, 1)$). Find the value of z_0 such that $\mathbb{P}(Z \geq z_0) = 0.10$
   (b) Let Z be a random variable with a standard Normal distribution (i.e. $N(0, 1)$). Find the value of z_0 such that $\mathbb{P}(-z_0 \leq Z \leq z_0) = 0.95$
   (c) Let X be a random variable with a Normal distribution. The mean parameter is 27 and the standard deviation parameter is 3. What is $\mathbb{P}(X < 20)$?
   (d) Let X be a random variable with a Normal distribution. The mean parameter is 550 and the standard deviation parameter is 100. Find the value of x_0 such that $\mathbb{P}(X \leq x_0) = 0.90$

3. (**Approximating the Binomial with a Normal**) Let X be Binomially distributed with parameters $n = 200$ and $p = 0.06$. Using the Normal approximation (with the 0.5 continuity correction), determine the following two quantities in parts (a) and (b):
   (a) $\mathbb{P}(X \geq 20)$
   (b) $\mathbb{P}(X \leq 18)$.
   (c) Why is it necessary to use the continuity correction of 0.5 in order to calculate the required probabilities?

4. (**Using the Central Limit Theorem**) Let $Y_1, ..., Y_n$ be independent and identically distributed Gamma random variables with parameters α and β. Let $\overline{Y}$ be the sample mean of the Gamma random variables.
   (a) Derive the exact distribution of $\overline{Y}$.
   (b) What is the approximate distribution of $\overline{Y}$ for large n? Be sure to clearly state the values of the parameters for the approximate distribution.
   (c) Was it necessary to invoke the Central Limit Theorem in part (a)? Was a distributional approximation used in part (a)? Was an approximation used in part (b)? What conditions are required to make an approximate distributional statement for the sample mean?

5. (**Relationships between Distributions**) - Let $Y_1, Y_2, ..., Y_n$ be a sample of size n from a $N(\mu, \sigma^2)$ distribution, and let $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ and $S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$. Let $X_1, X_2, ..., X_n$ be a sample of size n from a $N(\mu, \sigma^2)$ distribution, and let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$. Name the distribution of each of the following:
   (a) $\left(\frac{\overline{Y}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$

(b) $\sum_{i=1}^{n} \frac{(Y_i-\mu)^2}{\sigma^2}$

(c) $\sum_{i=1}^{n} \frac{Y_i^2+X_i^2-2\mu(X_i+Y_i)+2\mu^2}{\sigma^2}$

(d) $\sum_{i=1}^{n} \frac{(Y_i-\overline{Y})^2}{\sigma^2}$

(e) $\frac{\overline{Y}-\mu}{\frac{S_Y}{\sqrt{n}}}$

(f) $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sum_{i=1}^{n}(Y_i-\mu)^2}$

Module 2

1. (**Bias and Mean Squared Error**) Suppose Y has a Binomial distribution with parameters n and p. Let $\hat{p}_1 = \frac{Y}{n}$ and $\hat{p}_2 = \frac{Y+1}{n+2}$ be two different estimators for the unknown parameter p.

   (a) Show that $\mathbb{E}(\hat{p}_1) = p$. What term do we give to an estimator wihen it has this property?

   (b) Derive the bias of $\hat{p}_2$. Is $\hat{p}_2$ unbiased?

   (c) Derive $MSE(\hat{p}_1) = \mathbb{E}[(\hat{p}_1 - p)^2]$

   (d) Derive $MSE(\hat{p}_2) = \mathbb{E}[(\hat{p}_2 - p)^2]$

2. (**Order Statistics**) Give the derivations for the cdf and pdf of the minimum and maximum order statistics for a collection i.i.d. random variables.

3. (**Comparison of Estimators**) Review Problem 8.8 (page 394) from Assignment 2.

Module 3

1. (**Pivotal Quantities**) Let $Y_1, Y_2, ..., Y_n$ denote a random sample of size n from a population with a uniform distribution on the interval $(0, \theta)$. Let $Y_{(n)} = \max(Y_1, Y_2, ..., Y_n)$ and $U = (\frac{1}{\theta})Y_{(n)}$

   (a) Show that U has distribution function

$$F_U(u) = \begin{cases} 0 & u < 0 \\ u^n & 0 \leq u \leq 1 \\ 1 & u > 1 \end{cases}$$

   (b) Because the distribution of U does not depend on θ, U is a pivotal quantity. Find a 95% lower confidence bound for θ.

2. (**Large Sample Confidence Intervals**) Given the assumption of a large sample size, give the theoretical formula for the 95% confidence interval if the parameters of interest are μ, p, $\mu_1 - \mu_2$ or $p_1 - p_2$. Are any assumptions required on the distribution of the underlying random sample? Which theorem do you need to use to form the approximate confidence interval?

3. (**Small Sample Confidence Intervals**) Given the assumption of a small sample size, give the theoretical formula for the 99% confidence interval if the parameters of interest are μ and $\mu_1 - \mu_2$. Are any assumptions required on the distribution of the underlying random sample?

4. (**The Pooled Sample Variance**) Give the formula of the pooled sample variance. What assumptions are required for the Normal distributions of the random samples when the pooled sample variance is used?

5. (**Confidence Interval for σ^2**) Give the formula for the 95% confidence interval if the parameter of interest is σ^2. What distribution do the critical values come from? Give the upper and low crtiical values for a 95% confidence interval.

Module 4

1. (**Relative Efficiency**) Review Problem 9.2 (page 447) from Assignment 2.

2. (**Consistency**) Suppose $X_1, X_2, ..., X_n$ is a random sample of size n from a Bernoulli(p) distribution. Let $Y = \sum_{i=1}^{n} X_i$.

   (a) Find $\mathbb{E}(\frac{Y}{n})$ and $Var(\frac{Y}{n})$

   (b) Using Chebyshev's inequality, show that for any positive number $\epsilon > 0$, that
   $$\lim_{n\to\infty} \mathbb{P}(|\frac{Y}{n} - p| < \epsilon) = 1$$

   (c) Using Theorem 9.1, show that $\frac{Y}{n}$ is consistent for p.

   (d) Explain how Chebyshev's Inequality is used to prove Theorem 9.1. (Hint: Generalize your solution to part (b))

   (e) Determine a consistent estimator for $p^3 + 3p^2 - sin(p) + p$. Let $g(p)$ be some function of p. Under what condition(s) will $g(\hat{p})$ be a consistent estimator for $g(p)$?

3. (**Slutsky's Theorem**) Suppose $Y_1, Y_2, ..., Y_n$ is a random sample of size n from a distribution with $\mathbb{E}(Y_i) = \mu$ and $\mathbb{V}(Y_i) = \sigma^2$. Define $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ where $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$. Show that the distribution function of $\sqrt{n}\left(\frac{\bar{Y}-\mu}{S_n}\right)$ converges to a standard Normal distribution function.

4. (**Sufficiency, Minimal Sufficiency and MVUE**) Suppose $X_1, X_2, ..., X_n$ are a random sample from an Exponential distribution with unknown parameter λ.

   (a) Is the entire sample $(X_1, X_2, ..., X_n)$ a sufficient statistic for λ?

   (b) Is the entire sample $(X_1, X_2, ..., X_n)$ always a sufficient statistic when given a random sample?

   (c) Determine a sufficient statistic for λ which is not the entire sample.

   (d) Is your statistic unique? If not, determine another sufficient statistic for λ.

   (e) Determine the minimal sufficient statistic for λ. Show your alternative sufficient statistic from the previous part can be written as a function of your minimal sufficient statistic.

   (f) Using your answer from part (e), determine the MVUE of λ.

   (g) Determine the MVUE of λ^2.

5. (**Sufficiency, Minimal Sufficiency and MVUE**) Suppose $X_1, X_2, ..., X_n$ are a random sample from a $Uniform(-\theta, 0)$ for $\theta > 0$.

   (a) Determine a non-trivial sufficient statistic.

   (b) What is the distribution of the sufficient statistic found in part (a)?

(c) Using the distribution calculated in part (b), what is the expectation of your sufficient statistic?

(d) Is it unbiased? If no, how can you alter the statistic to make it unbiased?

(e) Given a function g such that $g(T)$ is an unbiased estimator for the unknown parameter θ where T is a sufficient statistic, what conclusion can be made about the statistic $g(T)$? What theorem is needed to make this conclusion?

(f) Repeat parts (a)-(e) when $X_1, ..., X_n$ are a random sample from a $Uniform(-5\theta, 0)$.

6. (**Method of Moments**) Suppose that $Y_1, Y_2, ..., Y_n$ constitutes a random sample from a Poisson distribution with mean λ. Find the method-of-moments estimator of λ.

7. (**Method of Moments**) Suppose that $Y_1, Y_2, ..., Y_n$ constitutes a random sample from the Normal distribution with mean μ and variance σ^2, find the method-of-moments estimator of σ^2.

8. (**Method of Moments**) Suppose that $Y_1, Y_2, ..., Y_n$ is a random sample from a Gamma distribution with unknown parameters α and β, find the method-of-moments estimator for the pair (α, β).

9. (**Method of Maximum Likelihood**) Suppose that $Y_1, Y_2, ..., Y_n$ denote a random sample from the Poisson distribution with mean λ.

   (a) Find the MLE $\hat{\lambda}$ for λ.

   (b) Find the expected value and variance of $\hat{\lambda}$

   (c) Show that the estimator of part (a) is consistent for λ.

   (d) What is the MLE for $\mathbb{P}(Y = 0) = e^{-\lambda}$?

   (e) Calculate the Fisher Information matrix.

   (f) What is the asymptotic distribution of $\sqrt{n}(\hat{\lambda} - \lambda)$ as $n \to \infty$?

Module 5

1. (**Neyman-Pearson Lemma**) Complete Problems 1-3 from Tutorial 13.

2. (**Likelihood Ratio Test**) Complete Problem 4 from Tutorial 13.

3. (**Common Hypothesis Tests**) Complete Problem 5 from Tutorial 13.

Module 6

1. Refer to the problems for Chapter 11 posted on MyCourses.