STA 2101/442 Assignment 1 (Mostly Review)[1]

Except for Question 6, the questions are practice for the quiz on Friday Sept. 20th, and are not to be handed in. For the linear algebra part starting with Question 7, there is an excellent review in Chapter Two of Renscher and Schaalje' *Linear models in statistics*. The chapter has more material than you need for this course.

1. Let $Y_1, \ldots, Y_n$ be numbers, and $\overline{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$. Show that the sum of squares $Q_m = \sum_{i=1}^n (Y_i - m)^2$ is minimized when $m = \overline{Y}$.

2. Let $Y_1, \ldots, Y_n$ be a random sample from a distribution with mean μ and standard deviation σ.

   (a) Show that the sample variance $S^2 = \frac{\sum_{i=1}^n (Y_i - \overline{Y})^2}{n-1}$ is an unbiased estimator of σ^2.

   (b) Denote the sample standard deviation by $S = \sqrt{S^2}$. Assume that the data come from a continuous distribution, so it's easy to see that $Var(S) \neq 0$. Using this fact, show that S is a *biased* estimator of σ.

3. Let $Y_1, \ldots, Y_n$ be a random sample from a normal distribution with mean μ and variance σ^2, so that $T = \frac{\sqrt{n}(\overline{Y} - \mu)}{S} \sim t(n-1)$. This is something you don't need to prove, for now.

   (a) Derive a $(1-\alpha)100\%$ confidence interval for μ. "Derive" means show all the high school algebra. Use the symbol $t_{\alpha/2}$ for the number satisfying $Pr(T > t_{\alpha/2}) = \alpha/2$.

   (b) A random sample with $n = 23$ yields $\overline{Y} = 2.57$ and a sample variance of $S^2 = 5.85$. Using the critical value $t_{0.025} = 2.07$, give a 95% confidence interval for μ. The answer is a pair of numbers.

   (c) Test $H_0 : \mu = 3$ at $\alpha = 0.05$.

      i. Give the value of the T statistic. The answer is a number.

      ii. State whether you reject H_0, Yes or No.

      iii. Can you conclude that μ is different from 3? Answer Yes or No.

      iv. If the answer is Yes, state whether $\mu > 3$ or $\mu < 3$. Pick one.

   (d) Show that using a t-test, $H_0 : \mu = \mu_0$ is rejected at significance level α if and only the $(1-\alpha)100\%$ confidence interval for μ does not include μ_0. The problem is easier if you start by writing the set of T values for which H_0 is *not* rejected.

   (e) In Question 3b, does this mean $Pr\{1.53 < \mu < 3.61\} = 0.95$? Answer Yes or No and briefly explain.

4. For each of the following distributions, derive a general expression for the Maximum Likelihood Estimator (MLE); don't bother with the second derivative test. Then use the data to calculate a numerical estimate; you should bring a calculator to the quiz in case you have to do something like this.

   (a) $p(x) = \theta(1-\theta)^x$ for $x = 0, 1, \ldots$, where $0 < \theta < 1$. Data: `4, 0, 1, 0, 1, 3, 2, 16, 3, 0, 4, 3, 6, 16, 0, 0, 1, 1, 6, 10`. Answer: 0.2061856

   (b) $f(x) = \frac{\alpha}{x^{\alpha+1}}$ for $x > 1$, where $\alpha > 0$. Data: `1.37, 2.89, 1.52, 1.77, 1.04, 2.71, 1.19, 1.13, 15.66, 1.43` Answer: 1.469102

   (c) $f(x) = \frac{\tau}{\sqrt{2\pi}} e^{-\frac{\tau^2 x^2}{2}}$, for x real, where $\tau > 0$. Data: `1.45, 0.47, -3.33, 0.82, -1.59, -0.37, -1.56, -0.20` Answer: 0.6451059

   (d) $f(x) = \frac{1}{\theta} e^{-x/\theta}$ for $x > 0$, where $\theta > 0$. Data: `0.28, 1.72, 0.08, 1.22, 1.86, 0.62, 2.44, 2.48, 2.96` Answer: 1.517778

5. The random variable X has density $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$ for $0 < x < 1$, where $\alpha > 0$ and $\beta > 0$. Find $E(X)$; show your work. Hint: $f(x)$ is a density for *any* $\alpha > 0$ and $\beta > 0$.

6. In the United States, admission to university is based partly on high school marks and recommendations, and partly on applicants' performance on a standardized multiple choice test called the Scholastic Aptitude Test (SAT). The SAT has two sub-tests, Verbal and Math. A university administrator selected a random sample of 200 applicants, and recorded the Verbal SAT, the Math SAT and first-year university Grade Point Average (GPA) for each student. The data are given in the file `sat.data`. There is a link on the course web page in case the one in this document does not work.

   The university administrator knows that the Verbal and Math SAT tests have the same number of questions, and the maximum score on both is 800. But are they equally difficult on average for this population of students?

   Using R, do a reasonable analysis to answer the question. Bring your printout to the quiz; you may be asked to hand it in. Be ready to

   - State your model.
   - Justify your choice of model. Would you expect Verbal and Math scores from the same student to be independent?
   - State your null and alternative hypotheses, in symbols.
   - Express your conclusion (if any) in plain, non-statistical language that could be understood by someone who never had a Statistics course. Your answer is something about which test is more difficult for these students. Marks will be deducted for use of technical terms like null hypothesis, significance level, critical value, p-value, and so on even if what you say is correct.

Remember, the computer assignments in this course are *not group projects*. You are expected to do the work yourself. There is more than one correct answer. I did the analysis several different ways, and I consider all of them correct. I can think of about five more acceptable ways that I did not try. The number of bad ways to analyze the data is virtually unlimited.

7. Which statement is true? (Quantities in boldface are matrices of constants.)
   (a) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$
   (b) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{BA}+\mathbf{CA}$
   (c) Both a and b
   (d) Neither a nor b

8. Which statement is true?
   (a) $a(\mathbf{B}+\mathbf{C}) = a\mathbf{B}+a\mathbf{C}$
   (b) $a(\mathbf{B}+\mathbf{C}) = \mathbf{B}a+\mathbf{C}a$
   (c) Both a and b
   (d) Neither a nor b

9. Which statement is true?
   (a) $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{AB}+\mathbf{AC}$
   (b) $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$
   (c) Both a and b
   (d) Neither a nor b

10. Which statement is true?
    (a) $(\mathbf{AB})' = \mathbf{A}'\mathbf{B}'$
    (b) $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$
    (c) Both a and b
    (d) Neither a nor b

11. Which statement is true?
    (a) $\mathbf{A}'' = \mathbf{A}$
    (b) $\mathbf{A}''' = \mathbf{A}'$
    (c) Both a and b
    (d) Neither a nor b

12. Suppose that the square matrices $\mathbf{A}$ and $\mathbf{B}$ both have inverses and are the same size. Which statement is true?
    (a) $(\mathbf{AB})^{-1} = \mathbf{A}^{-1}\mathbf{B}^{-1}$
    (b) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
    (c) Both a and b
    (d) Neither a nor b
13. Which statement is true?
    (a) $(\mathbf{A}+\mathbf{B})' = \mathbf{A}' + \mathbf{B}'$
    (b) $(\mathbf{A}+\mathbf{B})' = \mathbf{B}' + \mathbf{A}'$
    (c) $(\mathbf{A}+\mathbf{B})' = (\mathbf{B}+\mathbf{A})'$
    (d) All of the above
    (e) None of the above
14. Which statement is true?
    (a) $(a+b)\mathbf{C} = a\mathbf{C} + b\mathbf{C}$
    (b) $(a+b)\mathbf{C} = \mathbf{C}a + \mathbf{C}b$
    (c) $(a+b)\mathbf{C} = \mathbf{C}(a+b)$
    (d) All of the above
    (e) None of the above
15. Let $\mathbf{A}$ be a square matrix with the determinant of $\mathbf{A}$ (denoted $|\mathbf{A}|$) equal to zero. What does this tell you about $\mathbf{A}^{-1}$? No proof is required here.
16. Recall that an inverse of the matrix $\mathbf{A}$ (denoted $\mathbf{A}^{-1}$) is defined by two properties: $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ and $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Prove that inverses are unique, as follows. Let $\mathbf{B}$ and $\mathbf{C}$ both be inverses of $\mathbf{A}$. Show that $\mathbf{B} = \mathbf{C}$.
17. Let $\mathbf{X}$ be an n by p matrix with $n \neq p$. Why is it incorrect to say that $(\mathbf{X}'\mathbf{X})^{-1} = \mathbf{X}^{-1}\mathbf{X}'^{-1}$?
18. Suppose that the square matrices $\mathbf{A}$ and $\mathbf{B}$ both have inverses. Prove that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. You have two things to show.
19. Let $\mathbf{A}$ be a non-singular square matrix. Prove $(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}$.
20. Using Question 19, prove that the if the inverse of a symmetric matrix exists, it is also symmetric.
21. Let $\mathbf{a}$ be an $n \times 1$ matrix of real constants. How do you know $\mathbf{a}'\mathbf{a} \geq 0$?

22. Recall the *spectral decomposition* of a square symmetric matrix (for example, a variance-covariance matrix). Any such matrix $\boldsymbol{\Sigma}$ can be written as $\boldsymbol{\Sigma} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'$, where $\mathbf{P}$ is a matrix whose columns are the (orthonormal) eigenvectors of $\boldsymbol{\Sigma}$, $\boldsymbol{\Lambda}$ is a diagonal matrix of the corresponding (non-negative) eigenvalues, and $\mathbf{P}'\mathbf{P} = \mathbf{P}\mathbf{P}' = \mathbf{I}$.

    (a) Let $\boldsymbol{\Sigma}$ be a square symmetric matrix with eigenvalues that are all strictly positive.

        i. What is $\boldsymbol{\Lambda}^{-1}$?

        ii. Show $\boldsymbol{\Sigma}^{-1} = \mathbf{P}\boldsymbol{\Lambda}^{-1}\mathbf{P}'$

    (b) Let $\boldsymbol{\Sigma}$ be a square symmetric matrix, and this time some of the eigenvalues might be zero.

        i. What do you think $\boldsymbol{\Lambda}^{1/2}$ might be?

        ii. Define $\boldsymbol{\Sigma}^{1/2}$ as $\mathbf{P}\boldsymbol{\Lambda}^{1/2}\mathbf{P}'$. Show $\boldsymbol{\Sigma}^{1/2}$ is symmetric.

        iii. Show $\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2} = \boldsymbol{\Sigma}$.

    (c) Now return to the situation where the eigenvalues of the square symmetric matrix $\boldsymbol{\Sigma}$ are all strictly positive. Define $\boldsymbol{\Sigma}^{-1/2}$ as $\mathbf{P}\boldsymbol{\Lambda}^{-1/2}\mathbf{P}'$, where the elements of the diagonal matrix $\boldsymbol{\Lambda}^{-1/2}$ are the reciprocals of the corresponding elements of $\boldsymbol{\Lambda}^{1/2}$.

        i. Show that the inverse of $\boldsymbol{\Sigma}^{1/2}$ is $\boldsymbol{\Sigma}^{-1/2}$, justifying the notation.

        ii. Show $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Sigma}^{-1}$.

    (d) The (square) matrix $\boldsymbol{\Sigma}$ is said to be *positive definite* if $\mathbf{v}'\boldsymbol{\Sigma}\mathbf{v} > 0$ for all vectors $\mathbf{v} \neq \mathbf{0}$. Show that the eigenvalues of a positive definite matrix are all strictly positive. Hint: start with the definition of an eigenvalue and the corresponding eigenvalue: $\boldsymbol{\Sigma}\mathbf{v} = \lambda\mathbf{v}$.

    (e) Let $\boldsymbol{\Sigma}$ be a symmetric, positive definite matrix. Putting together a couple of results you have proved above, establish that $\boldsymbol{\Sigma}^{-1}$ exists.

23. Let $\mathbf{X}$ be an $n \times p$ matrix of constants. The idea is that $\mathbf{X}$ is the "design matrix" in the linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, so this problem is really about linear regression.

    (a) Recall that $\mathbf{A}$ symmetric means $\mathbf{A} = \mathbf{A}'$. Let $\mathbf{X}$ be an n by p matrix. Show that $\mathbf{X}'\mathbf{X}$ is symmetric.

    (b) Recall the definition of linear independence. The columns of $\mathbf{A}$ are said to be *linearly dependent* if there exists a column vector $\mathbf{v} \neq \mathbf{0}$ with $\mathbf{A}\mathbf{v} = \mathbf{0}$. We will say that the columns of $\mathbf{A}$ are linearly *independent* if $\mathbf{A}\mathbf{v} = \mathbf{0}$ implies $\mathbf{v} = \mathbf{0}$.

        Show that if the columns of $\mathbf{X}$ are linearly independent, then the columns of $\mathbf{X}'\mathbf{X}$ are also linearly independent. Use Problem 21 and the definition of linear independence.

    (c) Show that if the columns of $\mathbf{X}'\mathbf{X}$ are linearly independent, then the columns of $\mathbf{X}$ are linearly independent.

    (d) Show that if the columns of $\mathbf{X}'\mathbf{X}$ are linearly independent, then $(\mathbf{X}'\mathbf{X})^{-1}$ exists.

    (e) Show that if $(\mathbf{X}'\mathbf{X})^{-1}$ exists, then the columns of $\mathbf{X}'\mathbf{X}$ are linearly independent.