Name Jerry

Student Number

STA 302f 2015 Quiz 11

1. (6 points) For the general linear model with normal errors, we wish to test $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{t}$, where $\mathbf{C}$ is a $q \times (k+1)$ matrix of constants; the rows of $\mathbf{C}$ are linearly independent.

   (a) What is the distribution of $\mathbf{C}\widehat{\boldsymbol{\beta}}$?

   i. Write the answer in the space below.

   $$C\hat{\beta} \sim N(C\beta, \sigma^2 C(X'X)^{-1}C')$$

   ii. Copy down the *two* expressions from the formula sheet that justify your answer.

   $$\hat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1})$$

   $$Ay \sim N(A\mu, A\Sigma A')$$

   (b) What is the distribution of $(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})'(\mathbf{C}\sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}(\mathbf{C}\widehat{\boldsymbol{\beta}} - \mathbf{t})$ if H_0 is true?

   i. Write the answer in the space below.

   $$\chi^2(q)$$

   ii. Copy down the expression from the formula sheet that justifies your answer.

   $$W = (y-\mu)'\Sigma^{-1}(y-\mu) \sim \chi^2(p)$$

   (c) Write the test statistic in the space below, showing how σ^2 disappears.

   $$F = \frac{\frac{1}{\sigma^2}(C\hat{\beta} - t)'(C(X'X)^{-1}C')^{-1}(C\hat{\beta} - t)/q}{\frac{SSE}{\sigma^2}/(n-k-1)}$$

   (d) How do you know numerator and denominator are independent?

   Because $SSE = \hat{\varepsilon}'\hat{\varepsilon}$, $\hat{\beta}$ and $\hat{\varepsilon}$ are independent, and functions of independent random vectors are independent.

2. (4 points) In homework, you analyzed the extended `statclass` data.

(a) What is the predicted Final Exam score for a Female student from Race A, with a Quiz average of 6, a Computer average of 5, and a Midterm mark of 70%? The answer is a number. **Circle the number**. Show a little work.

$10.68 + 4.7 * 6 - 2.59 * 5 + 0.296 * 70$

$= 10.68 + 28.2 - 12.95 + 20.72$

$= 46.65$

Answer to full accuracy is 46.68 using R's predict.

(b) After correcting for the other independent variables in the model, mark on the midterm explains _____% of the remaining variation in final exam performance. The answer is a number. **Circle the number**. Show a little work.

$$P = \frac{gF}{gF + n - k - 1} = \frac{t^2}{t^2 + n - k - 1} = \frac{2.293^2}{2.293^2 + 51}$$

$= 0.0935$ or

9.35%

Either answer is okay

(c) Allowing for the other information we have about the students, are there any differences in the average performance of students from the different racial groups? Express your answer in plain, non-statistical language.

Race B does best.

Attach your *complete* R printout to your quiz. Make sure your name and student number are written clearly on the printout.