STA 2101/442 Assignment Eight[1]

Please bring your R printouts to the quiz. *Your printouts must* not *contain answers to the non-computer parts of this assignment.* The non-computer questions are just practice for the quiz, and will not be handed in. You may use R as a calculator, but please bring a real calculator to the quiz.

1. Suppose you have a random sample from a normal distribution, say $Y_1, \ldots, Y_n \stackrel{i.i.d.}{\sim} N(\mu, \sigma^2)$. If someone randomly sampled another observation from this population and asked you to guess what it was, there is no doubt you would say $\overline{Y}$. But what if you were asked for a prediction *interval*?

   Accordingly, suppose the normal model is reasonable and you observe a sample mean of $\overline{Y} = 7.5$ and a sample variance (with $n-1$ in the denominator) of $S^2 = 3.82$. The sample size is $n = 14$. Give a 95% prediction interval for the next observation. The answer is a pair of numbers. Be able to show your work. You can get the distribution result you need from the formula sheet, or you can re-derive it for this special case. Be able to do it both ways.

2. This suggests using Bonferroni-corrected deleted Studentized residuals for detecting outliers in simple independent random sampling, as long as a normal model seems reasonable. Try it with these data: `10.17 9.69 8.99 6.45 13.81 9.82 4.07 7.62 10.84 9.97`. Any outliers? Use R. Bring your printout to the quiz.

3. This question explores the practice of "centering" quantitative explanatory variables in a regression by subtracting off the mean.

   (a) Consider a simple experimental study with an experimental group, a control group and a single quantitative covariate. Independently for $i = 1, \ldots, n$ let

   $$Y_i = \beta_0 + \beta_1 x_i + \beta_2 d_i + \epsilon_i,$$

   where x_i is the covariate and d_i is an indicator dummy variable for the experimental group. If the covariate is "centered," the model can be written

   $$Y_i = \beta_0' + \beta_1'(x_i - \overline{x}) + \beta_2' d_i + \epsilon_i,$$

   where $\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$.

   i. Express the β' quantities in terms of the β quantities.
   ii. If the data are centered, what is $E(Y|x)$ for the experimental group compared to $E(Y|x)$ for the control group?
   iii. By the invariance principle (this takes you back all the way to Likelihood Part One), what is $\widehat{\beta}_0$ in terms of $\widehat{\beta}'$ quantities? Assume ϵ_i is normal.

[1] Copyright information is at the end of the last page.

(b) In this model, there are $p-1$ quantitative explanatory variables. The un-centered version is

$$Y_i = \beta_0 + \beta_1 x_{i,1} + \ldots + \beta_{p-1} x_{i,p-1} + \epsilon_i,$$

and the centered version is

$$Y_i = \beta_0' + \beta_1'(x_{i,1} - \overline{x}_1) + \ldots + \beta_{p-1}'(x_{i,p-1} - \overline{x}_{p-1}) + \epsilon_i,$$

where $\overline{x}_j = \frac{1}{n}\sum_{i=1}^n x_{i,j}$ for $j = 1, \ldots, p-1$.

i. What is β_0' in terms of the β quantities?

ii. What is β_j' in terms of the β quantities?

iii. By the invariance principle, what is $\widehat{\beta}_0$ in terms of the $\widehat{\beta}'$ quantities? Assume ϵ_i is normal.

iv. Using $\sum_{i=1}^n \widehat{Y}_i = \sum_{i=1}^n Y_i$, show that $\widehat{\beta}_0' = \overline{Y}$.

(c) Now consider again the study with an experimental group, a control group and a single covariate. This time the interaction is included.

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 d_i + \beta_3 x_i d_i + \epsilon_i$$

The centered version is

$$Y_i = \beta_0' + \beta_1'(x_i - \overline{x}) + \beta_2' d_i + \beta_3'(x_i - \overline{x}) d_i + \epsilon_i$$

i. Express the β' quantities from the centered model in terms of the β quantities from the un-centered model. Is the correspondence one to one?

ii. For the un-centered model, what is the difference between $E(Y|X = \overline{x})$ for the experimental group compared to $E(Y|X = \overline{x})$ for the control group?

iii. What is the difference between intercepts for the centered model? Compare this to your answer to Question 3(c)ii.

4. This question is about the union-intersection method for multiple comparisons, of which the Scheffé tests are the prime example. The discussion starts out in a general way. Especially at the beginning, it helps to think of the example of an initial test for difference among p treatment means, and a family of follow-up tests consisting of all pairwise comparisons.

(a) The model is $Y \sim P_\theta, \theta \in \Theta$. The parameter is θ and the parameter space is Θ. There is an initial size α test of $H_0 : \theta \in \Theta_0$, where $\Theta_0 \subset \Theta$. The initial test has critical region C_0, meaning H_0 is rejected if $Y \in C_0$. Size α means $\max_{\theta \in \Theta_0} P\{Y \in C_0\} \leq \alpha$. There is a family of follow-up tests indexed by γ, with γ belonging to an index set G. That is, there is one follow-up test for each $\gamma \in G$.

The follow-up tests have null hypothesis regions $\{\Theta_\gamma : \gamma \in G\}$ and critical regions $\{C_\gamma : \gamma \in G\}$. The idea is that if the initial null hypothesis is rejected, we carry out follow-up tests to reveal specific ways in which it is wrong.

i. Consider a particular follow-up test γ. We want truth of the initial null hypothesis to imply truth of the follow-up null hypothesis. Does this mean $\Theta_\gamma \subseteq \Theta_0$, or does it mean $\Theta_0 \subseteq \Theta_\gamma$?

ii. Let's make the truth of the initial null hypothesis equivalent to the truth of *all* the null hypotheses in the family. This means $\Theta_0 = \ldots$

iii. It would also be desirable for rejection of the follow-up null hypothesis to imply rejection of the initial null hypothesis. That way, if the initial test did not detect anything, the follow-ups could not detect anything either. Does this mean $C_\gamma \subseteq C_0$, or does it mean $C_0 \subseteq C_\gamma$?

iv. Suppose the property in the last item is enjoyed by *all* the follow-up tests in the family. Then $\cup_{\gamma \in G} C_\gamma \subseteq \ldots$

v. Thus, suppose the family of follow-up tests has these two properties:

- $\cap_{\gamma \in G} \Theta_\gamma = \Theta_0$
- $\cup_{\gamma \in G} C_\gamma \subseteq C_0$

Show that if all the null hypotheses in the family are true and these properties hold, the probability of rejecting one or more of them is less than or equal to α, the significance level of the initial test.

(b) Now consider a random sample from a normal distribution, and an initial test of $H_0 : \mu = \mu_0$ versus the alternative that $\mu \neq \mu_0$. The initial null hypothesis is rejected if $|T| > t_{\alpha/2}$.

There are two tests in the follow-up family, and they are both one-sided. The first follow-up test has null hypothesis is $\mu \leq \mu_0$, which is rejected if $T > t_{\alpha/2}$. The second follow-up test has null hypothesis is $\mu \geq \mu_0$, which is rejected if $T < -t_{\alpha/2}$. Please refer to the notation in the preceding question.

i. What is θ?

ii. What is Θ?

iii. What is Θ_0?

iv. What is Y?

v. What is C_0?

vi. How many elements are in the set G?

vii. Letting $\gamma = 1, 2$, what is Θ_1?

viii. What is C_1?

ix. What is Θ_2?

x. What is C_2?

xi. Show that the family of follow-up tests has the two properties given in Question 4(a)v, and thus is a union-intersection family.

The point of this is that you can draw directional conclusions from the one-sided follow-ups. In practice, this means just looking at the sign of the test statistic for any two-sided Z or t-test.

(c) Now we will specialize the discussion to regression, and obtain the Scheffé tests. Based on the usual linear model, the initial null hypothesis is $H_0 : \mathbf{L}_0\boldsymbol{\beta} = \mathbf{0}$, where as usual $\mathbf{L}_0$ is an $r \times p$ matrix with linearly independent rows. The null hypotheses of the follow-up tests are indexed by $\gamma \in G$, and the set G is big. For now, we will say that the null hypotheses of the follow-up tests are *all* the statements of the form $H_\gamma : \mathbf{L}_\gamma\boldsymbol{\beta} = \mathbf{0}$ that are logically implied by the null hypothesis of the initial test. In each case, the $s \times p$ matrix $\mathbf{L}_\gamma$ has linearly independent rows. Logical implication just means that each row of $\mathbf{L}_\gamma$ is a linear combination of the rows of $\mathbf{L}_0$.

  i. Verify that $\cap_{\gamma \in G}\Theta_\gamma = \Theta_0$.

  ii. Is the null hypothesis of the follow-up test a more restrictive version of the null hypothesis of the initial test, or is it a less restrictive version?

  iii. Consider the model where the null hypothesis of the initial test is true. This is the reduced model for the initial test. The model in which the follow-up null hypothesis is true is also a reduced model. Is it more reduced than the model of the initial test, or less reduced?

  iv. Regression sum of squares: Which is true? $SSR_0 \leq SSR_\gamma$ or $SSR_\gamma \leq SSR_0$?

  v. To actually carry out union-intersection follow-up F tests, it is very convenient to calculate an F statistic for the follow-up test in the usual way (using software, of course), and then compare it to a critical value that is adjusted to make the *joint* significance level $\leq \alpha$. Call this adjusted critical value f_a. All we need to satisfy the property $\cup_{\gamma \in G}C_\gamma \subseteq C_0$ is for $F_\gamma > f_a$ to imply $F_0 > f$, where f is the critical value of the initial test. Show that $f_a = \frac{r}{s}f$ works. That is, show $F_\gamma > \frac{r}{s}f \Rightarrow F_0 > f$. These are the classical Scheffé tests, although Scheffé obtained them through geometric considerations, with a strong focus on simultaneous confidence intervals for single linear combinations.

(d) Union-intersection families often have potentially more members than you think at first. For example, the Scheffé family in Question 4c could have been defined in terms of single linear combinations, but you see it can be expanded to contain multiple linear combinations too. Could it also contain one-sided tests like those of Question 4b? Briefly explain.

(e) Could the Scheffé family be further expanded to included non-linear null hypotheses? This won't be on the quiz or final exam.

(f) Can you give the rule for union-intersection follow-ups to a large-sample likelihood ratio test? This won't be on the quiz or final exam.

5. In the Chick Weights study, newly hatched chickens were randomly assigned to one of six different feed supplements, and their weight in grams after 6 weeks was recorded. The Chick Weights data are in an R dataset called `chickwts`. Type `chickwts` to see it.

   (a) Make sure a table of means, standard deviations and sample sizes for the 6 feed types is part of your output.

   (b) Test whether the six mean weights are different. Get the F statistic, degrees of freedom, p-value and proportion of explained variation.

   (c) You want to know which means are different from which other means. Carry out the multiple comparison procedure likely to be the most powerful in this situation. Base your conclusions on the usual $\alpha = 0.05$ *joint* significance level for the family of tests. Of course when you state your conclusions in plain language, you would not mention the significance level or joint significance level. But to be honest, stating the conclusions in plain language isn't easy. The pattern is complicated.

   (d) Test for differences among mean weights for the five feed types *excluding* horsebean.

      i. First, write the null hypothesis in terms of μ values.

      ii. Now obtain the F statistic, degrees of freedom and p-value. Do you reject H_0 at $\alpha = 0.05$?

      iii. What is the "reduced" (restricted) model for this problem? You should not do this test by fitting a full and a reduced model, but be aware that you could.

      iv. What proportion of the remaining variation does the effect explain? The answer is a number between zero and one that you could obtain from your printout with a calculator, though you may choose to do it with R and have it on your printout.

      v. Consider this test as a Scheffé follow-up to the initial test of difference among six feed types. What is the Scheffé critical value? Do you reject the null hypothesis with the follow-up test?

   (e) Obtain a 95% confidence interval for the difference between the expected weight for chicks fed horsebean, versus the average of the other expected values. Your answer is a pair of numbers: a lower limit and an upper limit.

   (f) Would you advise a chicken farmer to purchase the Horsebean feed supplement if she wanted big fat chickens?
