# Assignment 2 — due February 4, 2019

Note: For hypothesis testing, provide $H_0$, $H_A$, the test statistic, p-value, and a conclusion/interpretation based on the p-value, even when the question does not specifically ask for all of these elements of testing. Use language suggested by the following. You will be expected to interpret all p-values in future assignments and exams using this approach.

| p-value | Strength of evidence against the null hypothesis |
|---|---|
| Greater than 0.1 (or so) | No evidence |
| Between 0.05 (or so) and 0.1 | Weak evidence |
| Between 0.01 (or so) and 0.05 (or so) | Moderate evidence |
| Between 0.001 (or so) and 0.01 (or so) | Strong evidence |
| Less than 0.001 (or so) | Very strong evidence |

Table 1: Interpretation of p-values in terms of strength of evidence against the null hypothesis.

1. (a) Two independent samples of observations were taken; each sample is assumed to have come from a normal distribution. The sample mean for the first sample was 8.92, the sample variance was 1.40, and there were 9 observations in the first sample. For the second sample, there were 11 observations, the sample mean was 7.41, and the sample variance was 1.30.

    Some people use the following procedure for performing an $\alpha = 0.05$ independent (unpaired) sample test. "For each of the two samples, construct a one-sample 95% CI. If the two CIs overlap, conclude that the null hypothesis of equal means for the two groups should be accepted; if the CIs do not overlap, the null hypothesis should be rejected."

    Conduct such a "test" for the data given: is the null hypothesis of equal means accepted or rejected? (There is no test statistics to give and no p-value to give with this procedure). In addition, conduct the correct test (assuming equal variances).

    Why might these two procedures lead to different conclusions? What would you say to someone who recommends this CI overlap approach to two independent sample hypothesis testing?

    (b) Suppose the two samples come from normal populations with known variance $\sigma^2$, and that both samples are of size $n$. Find the probability of a Type I Error if you use the overlapping confidence interval rule.

2. The data set `donut.txt` provides an example of a balanced one-way experiment. The experiment went as follows: 24 batches of donuts were prepared, and each batch was randomly assigned to one of four treatment groups. The treatments consisted of 4 different fats which were used to fry the donuts. Of concern is whether the amount of fat absorbed by the donuts was the same for each type of fat. Variables in the data set consist of `gfa` — the grams of fat absorbed; `fat` — the particular type of fat used to fry the donuts; x1, x2, x3, x4, z1, z2, z3 — coded variables.

    (a) Use R to perform the overall $F$ test for one-way ANOVA. Plot the raw residuals versus predicted values. Comment on this plot. Repeat the plot, but this time use externally studentized residuals. Comment on the difference in plots, if any. Explain why you observe this difference (or lack thereof).

    (b) The model $Y_{ij} = \mu_i + \varepsilon_{ij}$ is often written as $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$. In essence, this amounts to changing from a problem with 4 model parameters $(\mu_1, \mu_2, \mu_3, \mu_4)$ to a problem with 5 parameters: $(\mu, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$. The problem then becomes overparametrized, which creates some difficulties. To see why, consider the model gfa $= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$ where $x_1, \ldots, x_4$ are defined in the file `donut.txt`. First, show that this is the same as the model $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$. What happens in R if you try to fit gfa $= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$?

(c) By using R, fit the following models. Note that some require you to change the fitting order.

     i. $\text{gfa} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$

     ii. $\text{gfa} = \beta_0 + \beta_3 x_3 + \beta_2 x_2 + \beta_1 x_1 + \varepsilon$

     iii. $\text{gfa} = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \varepsilon$

     iv. $\text{gfa} = \beta_0 + \beta_3 z_3 + \beta_2 z_2 + \beta_1 z_1 + \varepsilon$

Focus on the overall $F$ test for $H_0 : \mu_1 = \cdots = \mu_4$. For each of the models fit in R, how can you use the output from the **anova** function to construct that $F$ test? Does the fitting order matter? How does the R output change when you change the fitting order?

3. An experiment was conducted to determine reactions to a pain stimulus. (A related experiment can be found in Simone et al., 1989, *Pain*, 38:99-107.) There were 13 human subjects in total; each was assigned to have a dose of capsaicin injected intradermally, i.e. into the skin. (Capsaicin is the compound that makes peppers hot.) A typical reaction after this injection is that the skin near the injection site is painful to the touch for some period of time. For each subject, this amount of time was measured. Seven subjects received a dose of 100 $\mu l$, four subjects received a dose of 25 $\mu l$, and two subjects received a dose of 15 $\mu l$. The data are summarized below:

| Dose | | | Time (min.) | | | | |
|---|---|---|---|---|---|---|---|
| 15 $\mu l$ | 67 | 63 | | | | | |
| 25 $\mu l$ | 62 | 71 | 62 | 67 | | | |
| 100 $\mu l$ | 121 | 114 | 79 | 151 | 72 | 180 | 72 |

(a) Analyze these data using a one-way ANOVA based on the usual assumptions. Be sure to make a residual plot.

(b) Using Levene's test with the Brown-Forsythe modification, test for equality of variances. To do this in R, you may want to use the **leveneTest** function from the **car** package, and use the option **center=median**.

(c) Use a rank transformation of the data, and perform a nonparametric ANOVA based on those data. In R, you can use the **rank** function to get the rank of the "time" response. Alternatively, you may use the **kruskal.test** function.

(d) Would it be better to analyze the data by first transforming to square root or log?

(e) Summarize your findings for this question.