

1. (35 pts) REM with ethanol treatment: An experiment was designed to investigate whether rapid eye movement (REM) sleep time depends on the amount of ethanol given in an injection. Four injection concentrations were considered: 0, 1, 2, and 4 grams per kilogram of body weight. Twenty rats were then chosen and randomly divided into four equal sized groups. Each group was given a different treatment (concentration of ethanol). The REM sleep time during the subsequent 24 hour period was recorded for each rat. We are interested in comparing the typical amount of REM sleep across treatments. The data are as follows

Treatment	REM sleep in minutes				
0 g/kg	88.6	73.2	91.4	68.0	75.2
1 g/kg	63.0	53.9	69.2	50.1	71.5
2 g/kg	44.9	59.5	40.2	56.3	38.7
4 g/kg	31.0	39.6	45.3	25.2	22.7

Summarize all your results carefully. Make certain you define all populations and parameters under consideration, and that you make clear with each step which parameter(s) you are analyzing. Here is a series of steps to get the table above into R in wide, then long, format.

```
# read table as it appears above, simply copy/paste
dat1 <- read.table(text="
0 g/kg      88.6 73.2 91.4 68.0 75.2
1 g/kg      63.0 53.9 69.2 50.1 71.5
2 g/kg      44.9 59.5 40.2 56.3 38.7
4 g/kg      31.0 39.6 45.3 25.2 22.7
", header=FALSE)
# extract only columns with data values
dat2 <- dat1[, 3:7]
# transpose the matrix so each treatment row is now in its own column
dat3 <- t(dat2)
# change matrix to a data.frame
rem <- as.data.frame(dat3)
# name the columns based on dosage
colnames(rem) <- c("0", "1", "2", "4")
# Convert to long format
library(reshape2)
rem.long <- melt(rem, variable.name = "dose", value.name = "minutes")
# R outputs: No id variables; using all as measure variables
# This warning is ok.
```

Questions begin on the next page.

- (a) (5 pts) Use boxplot of the data to compare the means and spread of the four groups. Compute, the sample means and sample standard deviation for the 4 treatment groups. Also, compute the standard errors for the 4 treatment group means.
- (b) (5 pts) Looking at the boxplot and descriptive summaries from part (a), do there appear to be differences in the typical REM levels for the 4 groups? Describe the differences you see that appear to be most pronounced.
- (c) 5 (pts) Do a one-way Analysis of Variance to compare the mean REM levels for the 4 groups.
- (c1) Construct the ANOVA table.
- (c2) Are the group means significantly different at the 5% level? Derive F statistic by using $F = MSM/MSE$ and compare it to the critical value with specified degrees of freedom. Then perform the test. (Use `qf(0.95, df1, df2)` to find the critical value).
- (c3) Use p-value reported in ANOVA table to do the test.
- (d) (5 pts) Compare all possible pairs of groups using Fishers LSD method, and summarize the results of the multiple comparison. Repeat for Tukeys HSD method, and using Bonferroni comparisons. Do the three methods find different groupings? If so, what accounts for that?
- (e) (5 pts) Are the results of the F -test in part (c), and the multiple comparisons in part (d) consistent with what you described in part (b)? Briefly discuss.
- (f) (5pts) Comment on the Levenes and Bartlett's formal tests of equal variances in light of what you see when you look at the data.
- (g) (5pts) Obtain the residuals (or centered values) and generate a normal probability plot based on them. This is an overall check on normality. Is the normality assumption reasonable here? Perform a formal test of normality on the residuals, report the p-value and comment. If the test says we did not sample from normally distributed populations, why does it appear to be saying that? Do you think there is a real difficulty here with applying the normal theory methods?