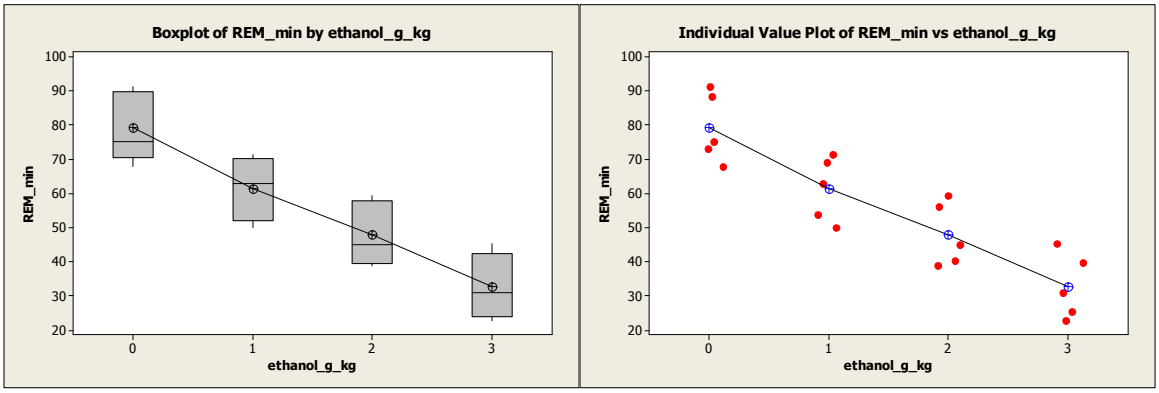
Homework 7, due Thursday October 27 in lab ---- **Solutions**

(assignment for these solutions can be found on the last page)

a) *Make boxplots and dotplots of the data to compare the groups. Compute, in Minitab, the sample means and standard deviations for the 4 treatment groups.*




```
                                Individual 95% CIs For Mean Based on
                                Pooled StDev
Level   N    Mean   StDev  --------+---------+---------+---------+-
0       5  79.280  10.181                            (----*---)
1       5  61.540   9.344                  (----*---)
2       5  47.920   9.461         (----*----)
3       5  32.760   9.559  (---*----)
                           --------+---------+---------+---------+-
                                  40        60        80       100
```

b) *Looking at the graphical 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.*

There are certainly differences. The amount of REM is inversely proportional to the amount of ethanol they are subjected to, in almost a linear way (nearly a straight line). The amount of REM is centered around 79 minutes for no ethanol, while those receiving 3g/kg ethanol only experience 33 minutes, each with only 10 minutes standard deviation.

c) *Do a one-way Analysis of Variance to compare the mean REM levels for the 4 groups. Are the groups significantly different at the 5% level?*

The ANOVA table below tests
H_0: $\mu_0=\mu_1=\mu_2=\mu_3$ that all means are equal, versus
H_A: At least one mean is different from the others.
There are three required assumptions: (1) Independent random samples from each population, (2) the population frequency curves are normal, and (3) the populations have equal standard deviations, $\sigma_0=\sigma_1=\sigma_2=\sigma_3$.

Checking the assumptions: We'll assume (1) is satified. If you perform normality tests for these 4 groups, normality (2) seems reasonable (small plots below). Based on the standard deviations above, assumption (3) appears to be reasonable.

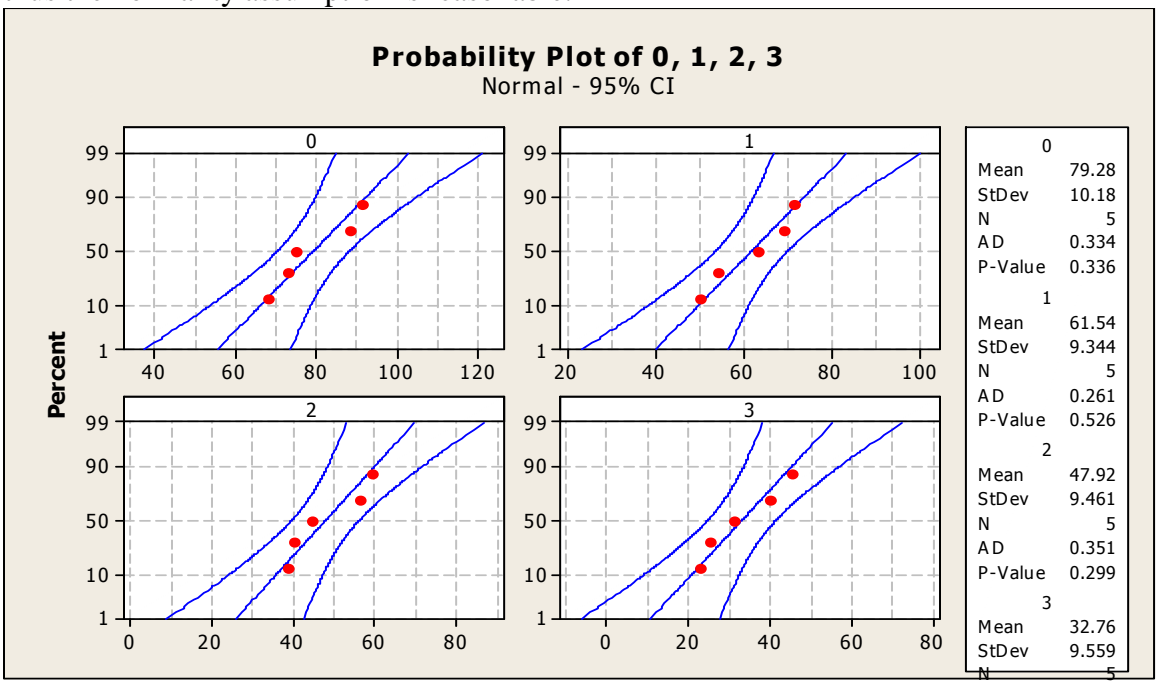
The conclusion of the test: Because the *p*-value below is nearly 0 (and less than $\alpha=0.05$), we reject H_0 in favor of H_A, concluding that at least one mean is different from the others.

One-way ANOVA: REM_min versus ethanol_g_kg

Source	DF	SS	MS	F	P
ethanol_g_kg	3	5882.4	1960.8	21.09	**0.000**
Error	16	1487.4	93.0		
Total	19	7369.8			

S = 9.642 R-Sq = 79.82% R-Sq(adj) = 76.03%

Plots confirming that the four samples do not indicate that the populations are not normal, thus the normality assumption is reasonable.



d) Compare all possible pairs of groups using Fisher's LSD method, and summarize the results of the multiple comparison. Repeat for Tukey's HSD method, and using Bonferroni comparisons. Do the three methods find different groupings? If so, what accounts for that?

Below are the results for the three multiple comparison methods. They find a couple different groupings. These intervals are listed from widest to narrowest: Bonferroni, Tukey's HSD, Fisher's LSD. Thus, the Bonferroni method result in the least differences detected, and Fisher's LSD result in the most differeces detected.

Fisher's LSD: The first plot shows that groups 1, 2, and 3 are different from group 0 (since confidence intervals do not include 0). The second plot shows that groups 2 and 3 are different from group 1 (2 is almost not different from 1). The third plot shows that group 3 is different from group 2.

```
Fisher 95% Individual Confidence Intervals
All Pairwise Comparisons among Levels of ethanol_g_kg

Simultaneous confidence level = 81.11%


ethanol_g_kg = 0 subtracted from:
ethanol_g_kg    Lower   Center    Upper    +---------+---------+---------+---------
1             -30.667  -17.740   -4.813              (---*---)
2             -44.287  -31.360  -18.433         (----*---)
3             -59.447  -46.520  -33.593    (---*----)
                                          +---------+---------+---------+---------
                                        -60       -30         0        30

ethanol_g_kg = 1 subtracted from:
ethanol_g_kg    Lower   Center    Upper    +---------+---------+---------+---------
2             -26.547  -13.620   -0.693              (---*----)
3             -41.707  -28.780  -15.853          (---*----)
                                          +---------+---------+---------+---------
                                        -60       -30         0        30

ethanol_g_kg = 2 subtracted from:
ethanol_g_kg    Lower   Center   Upper    +---------+---------+---------+---------
3             -28.087  -15.160  -2.233              (---*---)
                                         +---------+---------+---------+---------
                                       -60       -30         0        30
```

Tukey's HSD: The first plot shows that groups 1, 2, and 3 are different from group 0 (since confidence intervals do not include 0). The second plot shows that only group 3 is different from group 1 (2 is not different from 1 since the interval includes 0). The third plot shows that group 3 is not different from group 2.

```
Tukey 95% Simultaneous Confidence Intervals
All Pairwise Comparisons among Levels of ethanol_g_kg

Individual confidence level = 98.87%

ethanol_g_kg = 0 subtracted from:
ethanol_g_kg    Lower   Center    Upper   -+---------+---------+---------+--------
1             -35.203  -17.740   -0.277           (-----*-----)
2             -48.823  -31.360  -13.897        (-----*----)
3             -63.983  -46.520  -29.057   (----*-----)
                                          -+---------+---------+---------+--------
                                        -60       -30         0        30
```

```
ethanol_g_kg = 1 subtracted from:
ethanol_g_kg    Lower   Center    Upper   -+---------+---------+---------+--------
2             -31.083  -13.620    3.843             (----*-----)
3             -46.243  -28.780  -11.317         (----*-----)
                                           -+---------+---------+---------+--------
                                          -60       -30         0        30

ethanol_g_kg = 2 subtracted from:
ethanol_g_kg    Lower   Center  Upper   -+---------+---------+---------+--------
3             -32.623  -15.160  2.303            (-----*-----)
                                          -+---------+---------+---------+--------
                                         -60       -30         0        30
```

Bonferroni: First use the Fisher's LSD but adjust the significance level by dividing the individual significance level by the number of comparisons (with k groups, we have k*9k-1)/2, we have 4 groups) 4*3/2=6. Thus, we have .05/6=0.0083333. We enter this significance level for the Fisher test.

The first plot shows that groups 2 and 3 are different from group 0 (since confidence intervals do not include 0), but 1 is not different from 0. The second plot shows that group 3 is different from group 1, but 2 is not different from. The third plot shows that group 3 is not different from group 2.

```
Fisher 99.1667% Individual Confidence Intervals
All Pairwise Comparisons among Levels of ethanol_g_kg

Simultaneous confidence level = 96.24%

ethanol_g_kg = 0 subtracted from:
ethanol_g_kg    Lower   Center    Upper  --+---------+---------+---------+-------
1             -36.085  -17.740    0.605           (-----*-----)
2             -49.705  -31.360  -13.015       (------*-----)
3             -64.865  -46.520  -28.175  (-----*------)
                                          --+---------+---------+---------+-------
                                          -60       -30         0        30

ethanol_g_kg = 1 subtracted from:
ethanol_g_kg    Lower   Center    Upper  --+---------+---------+---------+-------
2             -31.965  -13.620    4.725            (-----*------)
3             -47.125  -28.780  -10.435        (-----*------)
                                          --+---------+---------+---------+-------
                                          -60       -30         0        30

ethanol_g_kg = 2 subtracted from:
ethanol_g_kg    Lower   Center  Upper  --+---------+---------+---------+-------
3             -33.505  -15.160  3.185            (-----*-----)
                                         --+---------+---------+---------+-------
                                         -60       -30         0        30
```

e) 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.

In part (c) we rejected H_0 in favor of H_A, concluding that at least one mean is different from the others. Each of the three multiple comparison methods we used in part (d) indicates at least one difference between the group population means. Thus, they are consistent.

f) Looking at the numerical and graphical summaries does it appear that the distributions of REM levels are reasonably normal,and have constant variance across groups? Discuss.

Also in (c), the assumptions of normality and constant variance were reasonably satisfied.

g) Comment on the Levene's and Bartlett's formal tests of equal variances in light of what you see when you look at the data.

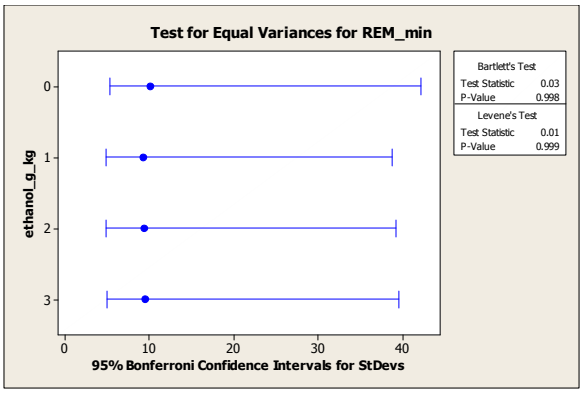
Below, both Bartlett's and Levene's test for equal variances indicate that there is not enough evidence (virtually no evidence with .99 *p*-values) that the variances are different. The confidence intervals in the plot are virtually the same.

Test for Equal Variances: REM_min versus ethanol_g_kg

```
95% Bonferroni confidence intervals for standard deviations
ethanol_g_kg  N    Lower    StDev    Upper
           0  5  5.37464  10.1810  42.2383
           1  5  4.93288   9.3441  38.7666
           2  5  4.99461   9.4611  39.2517
           3  5  5.04626   9.5589  39.6576

Bartlett's Test (normal distribution)
Test statistic = 0.03, p-value = 0.998

Levene's Test (any continuous distribution)
Test statistic = 0.01, p-value = 0.999
```



Stat 538 Biostat I

Homework 7 due Thursday October 27 in lab

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. 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

We are interested in comparing the typical amount of REM sleep across treatments.

a) Make boxplots and dotplots of the data to compare the groups. Compute, in Minitab, the sample means and standard deviations for the 4 treatment groups.

b) Looking at the graphical 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) Do a one-way Analysis of Variance to compare the mean REM levels for the 4 groups. Are the groups significantly different at the 5% level?

d) Compare all possible pairs of groups using Fisher's LSD method, and summarize the results of the multiple comparison. Repeat for Tukey's HSD method, and using Bonferroni comparisons. Do the three methods find different groupings? If so, what accounts for that?

e) 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) Looking at the numerical and graphical summaries does it appear that the distributions of REM levels are reasonably normal,and have constant variance across groups? Discuss.

g) Comment on the Levene's and Bartlett's formal tests of equal variances in light of what you see when you look at the data.

Summarize all your results carefully in a report. Make certain you define all populations and parameters under consideration, and that you make clear with each step which parameter(s) you are analyzing. I am particularly concerned that you keep straight the distinction between sample and population means, and sample and population variances.