Homework due November 3rd

Advanced Data Analysis 1, Stat 427/527 Name:

1. Guinea pigs: The data below are the survival times in hours of 72 guinea pigs after they were injected with a given dose of tubercule bacilli in a medical experiment. The data are from the article "Acquisition of resistance of guinea pigs injected with different doses of virulent tubercule bacilli," by T. Bjerkedal in the American Journal of Hygiene, (1960), pp. 130–148.

```
43 45 53 56 56 57 58 66 67 73 74
79 80 80 81 81 81 82 83 83 84 88
89 91 91 92 92 97 99 99 100 100 101
102 102 102 103 104 107 108 109 113 114 118
121 123 126 128 137 138 139 144 145 147 156
162 174 178 179 184 191 198 211 214 243 249
329 380 403 511 522 598
```

The data is located at

`http://math.unm.edu:/~james/guinea.csv`

(a) (10 pts) Obtain a 95% t-CI for the mean survival time.
(b) (10 pts) Repeat part (a) using a suitable nonparametric method.
(c) (10 pts) Take the log of the survival time and a 95% t-CI for the mean log survival time.
(d) (10 pts) Repeat part (c) using a suitable nonparametric method.
(e) (10 pts) Compare your 4 CIs, and contrast the nonparametric with the t-CIs. If they differ much, explain why they differ. Which analysis appears most appropriate? Explain.

2. Protoporphrin levels were determined for three groups of people—a control group of normal workers, a group of alcoholics with sideroblasts in their bone marrow, and a group of alcoholics without sideroblasts. The given data appeared in the paper "Erythocyte Coproporphyrin and Protoporphrin in Ethanol-Induced Sideroblastic Erythoporiesis" (Blood, 1974, p. 291295).

(a) (10 pts)Analyze the data using both a Kruskal-Wallis test and (b) t-test, assuming you are interested in comparing the typical protoporphrin level across groups.

(b) (10 pts) Make sure to clearly identify all population parameters, and assess the assumptions underlying both methods of analysis.
(c) (10 pts) Is it more appropriate to do the t-test on the original data or using a log transformation?
(d) (10 pts) Interpret the results of the KW and t-tests.
(e) (10 pts) In addition, do Bonferroni comparisons for pairwise differences using both Kruskal-Wallis and the t-test.

```
Normal  Alc_w_sb  Alc_wo_sb
22 78 37
27 172 28
47 286 38
30 82 45
38 453 47
78 513 29
28 174 34
58 915 20
72 84 68
56 153 12
30 780 37
39 NA 8
53 NA 76
50 NA 148
36 NA 11
```

The data is located at

`http://math.unm.edu:/~james/blood.txt`

3. (30 pts) The following data are from a study of a method for measuring the amount of chlorpheniramine maleate in tablets. Each tablet had a known quantity of 4.0 mg. Four labs were asked to report the amount measured in 10 tablets. The goal is to determine whether the labs have consistent measurements (in this case, the labs want the null hypothesis to be true).

Lab A	Lab B	Lab C	Lab D
4.00	4.02	4.02	4.00
4.02	3.95	3.86	4.02
4.01	3.91	4.02	4.03
4.01	3.89	3.97	4.04
4.04	3.92	3.91	4.00
3.99	4.01	3.82	3.81
4.03	3.89	3.98	3.91
3.97	3.89	3.99	3.96
3.98	3.97	3.99	4.05
3.98	3.90	4.00	4.06

I don't have a file with this data (except this pdf!). You can copy and paste the data or enter by hand to create the appropriate data set in R.

Select an appropriate method for comparing the means of the four labs. Use boxplots to informally assess the assumptions of the method you choose. Carefully describe the assumptions of the test you choose and interpret the results in the context of the problem.