STAT 500 — HW 3 — Solution

Grading: one point per part of questions 1-3, 5 points for question 4.

1. Cloud seeding experiment

   (a) eu = cloud, ou = cloud. No, don't expect a problem with non-independence.
   Note: If the eu and ou were different things, you would expect a problem.

   (b) No, the data are not consistent with the assumption of normally distributed errors.
   Note: You could use a variety of graphical assessments, e.g box plots, QQ plot, or a formal test. All that I've looked at give the same answer. Your explantion should include what you did and what in that suggested non-normality. For example, a QQ plot of the residuals or two QQ plots of the residuals for each group show clear departures from a straight line.

   (c) It seems that the assumption of equal variances is not reasonable. Both residual or abs(residual) plot show the same pattern: the seeded group of clouds is more variable.
   Note: other analyses are reasonable so long as the explanation was reasonable.

   (d) A two-sample t-test of abs(residual) gives p=0.017. There is evidence of unequal variances.

   (e) $\hat{\beta} = \frac{\log 651 - \log 278}{\log 442 - \log 164} = 0.86$. This suggests a log transformation.
   Note: you might suggest a transformation of $Y^{0.14}$. This would make sense only if prior knowledge of this type of data indicates that this makes sense.

   (f) Yes, log transformed data seem approximately normal. Again, I expect your explanation to include what you did and what you saw in that plot that indicated approximately normal. For example, a QQ plot of the pooled residuals shows a straight line.

   (g) Yes, the assumption of equal variances seems reasonable after log transformation. Again, there are many possible ways to reach this conclusion. Some likely possibilities are:
   A residual plot shows about equal spread.
   Levene's test gives a p-value of 1.0

   (h) The p-value from the Wilcoxon test using untransformed rainfall is 0.014. Using log transformed rainfall, you get the same p-value: 0.014. The ranks of the original observations are the same as the ranks of the log transformed values, so the test gives the same answer.
   Note: The appropriate p-value is that for the normal approximation. There is no theory to support a t approximation (p=0.017), but we didn't mark that wrong. SAS uses what is called a continuity correction, a detail that I didn't talk about. It improves the fit of normal distribution. If you turn that off, you get p=0.013. Both answers were accepted.

   My SAS code for all parts:

```
data rain;
  infile 'rainfall.txt' firstobs=2;
  input trt $ rain;
  lograin = log(rain);
  run;
```

```
proc glm;
  class trt;
  model rain=trt;
  output out=resids r=resid p=pred;
  run;

proc univariate;
  var resid;
  qqplot / normal;
  title 'Hi-res qq plot against normal distribution';
  run;

data resid2;
  set resids;
  absresid = abs(resid);
  run;

proc gplot;
  plot resid*pred;
  plot absresid*pred;
  title 'residual and abs(residual) vs predicted value plot';
  run;

proc ttest;
  class trt;
  var absresid;
  title "Levene's test, untransformed rainfall"
  run;

proc sort data=rain;
  by trt;
proc means mean std data=rain;
  by trt;    /* class trt also works and avoids sorting */
  var rain;
  title "means and sd's for each treatment";
  run;

/* not shown: graphical diagnostics for lograin */

proc npar1way wilcoxon;
  class trt;
  var rain;
  title 'Wilcoxon rank sum test';
  run;
```

2. Trauma data.

   (a) There many reasonable ways to look at these data. All give the same conclusion: the errors are not normally distributed. There is one unusually large observation.

   (b) With all observations, p = 0.75. There is no evidence of a difference between the two means.

   (c) Omitting the patient with thyroid disease gives very different conclusion: p = 0.0020. Very strong evidence of a difference.

   (d) If there were more patients with thyroid disease, you would omit them from the analysis too.

   Note: SAS code not shown because there is nothing not used in question 1.

3. Illustrative data.

   (a) Means: group 1: 17.3, group 2: 14.9. 95% ci for difference: (-12.4, 17.2)

   (b) There is one value in group 1 (with mean of 17.3) that is much larger than the other values.

   (c) Means: group 1: 10.1, group 2: 14.9. 95% ci for difference: (-7.0, -2.7). There is a very big change in the conclusions!

   Note: In this case, I corrected the data, instead of dropping the outlier.

4. Mercury in Fish
   Many different answers are possible. Preliminary inspection (of observations or residuals) indicates a couple of very large values, one in each group. These do not come from different populations (at least in any way that I know of, otherwise I would have told you). I would be very reluctant to remove these values. In particular, if you consider transforming the values, these two unusual points are no longer unusual if you look at log(Y).

   Your explanation was more important than the actual choice of analysis.

   Full credit for a reasonable justification of a single set of conclusions with test, point estimate and confidence interval. These can be on the transformed scale if decided to transform the data. No loss of points for not reporting a ci if used a non-parametric analysis (because that's hard to do and we haven't talked about it).
   -1 point if missing something requested (e.g. test, estimate or interval). (except for above exception)
   -2 points for reporting results from multiple analyses (e.g. multiple point estimates or multiple confidence intervals). Analyses done to choose an approach (e.g. using Levene's test do not count as multiple analyses).

   Notes: My own analysis was based on non-parametric methods because of one additional complication that I didn't tell you about. Some of the measurements (and most of the really small values) are reported as <X by the chemistry lab. Those are censored values. Non-parametric tests are much less affected by censored values.