STA 113 Spring 2005
Computing Assignment 3

Submitted by: I. H. Dinwoodie

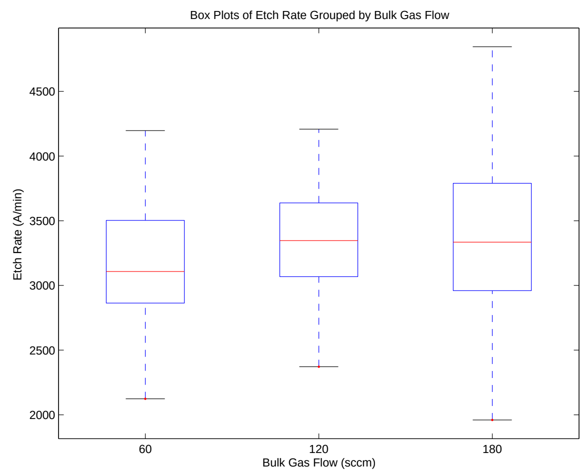
March 22, 2005

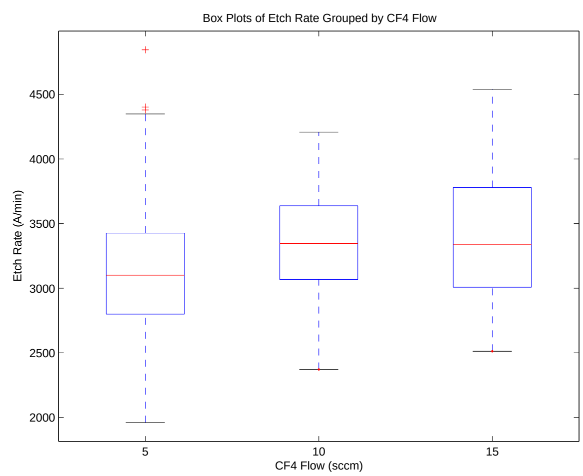
I have adhered to the Duke Community Standard in completing this assignment.

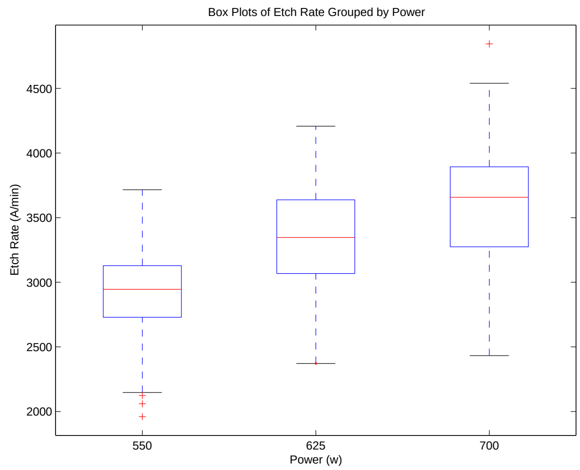
IAN

1. Multiple Regression

   (a) A few outliers show up in the boxplots. The boxplots also give a rough idea of how etch rate varies with each of the variables separately. In particular, etch rate and Bulk gas flow seem not to have a linear relationship.







(b) The fitted models are

$$e = 34.85 + 1.71b + 24.81c + 4.47p$$

for etch rate and

$$n = 15.56 + 0.0063b - 0.28c - 0.0053p$$

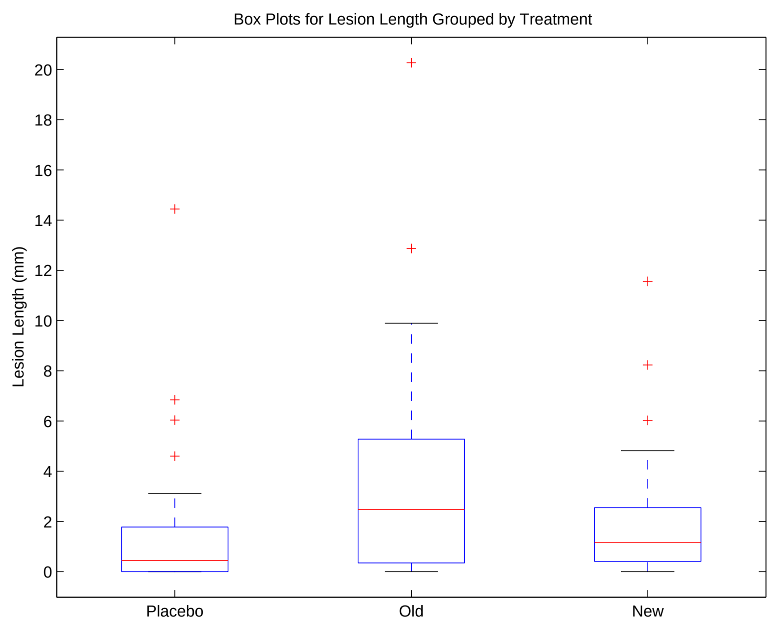
for nonuniformity.

(c) The combinations of Bulk gas flow (sccm) and Power (W) that give a mean response with etch rate at least 3100 and nonuniformity less than 9 are (60, 625), (60, 700), (120, 625), (120, 700), (180, 700).

2. ANOVA

   (a) One-way ANOVA gives a p-value of 0.03 on the test that all three means are equal, which would indicate that one or more pairs of means from μ_P, μ_O, μ_N are unequal. This methodology assumes that the variances of the observations in each treatment group are the same (p. 414), which is contradicted by the boxplots, so the conclusions are suspect.

   (b) The three two-sample t-tests with the methodology for unequal variances (note the very different values of sample standard deviations on different treatments) shows that the test of $H_0 : \mu_P = \mu_O$ is significant with a p-value of 0.0258, and a confidence interval for $\mu_P - \mu_O$ which does not contain 0.0. The boxplot confirms this comparison, and the unequal variances.



3. Binomial Confidence Intervals

   (a) Use `binornd(10,.1,1,10000);` in Matlab.

   (b) By counting the number of intervals that contain 0.1, it is found that the old method includes 0.1 64% of the time, and the new method includes 0.1 93% of the time.

   (c) The new method seems to have a coverage probability much closer to the correct value of 90%.

Matlab Diary

1.

```
tdfread('etchratedata.txt',' ')
  Name                 Size                    Bytes  Class

  Bulkgasflow       490x1                       3920  double array
  CF4flow           490x1                       3920  double array
  EtchRate          490x1                       3920  double array
  Power             490x1                       3920  double array
  Site              490x1                       3920  double array
  Wafer             490x1                       3920  double array

Grand total is 2940 elements using 23520 bytes

boxplot(EtchRate,Bulkgasflow)
boxplot(EtchRate,CF4flow)
boxplot(EtchRate,Power)



format long


beta=robustfit([Bulkgasflow,CF4flow,Power],EtchRate)

beta =

  34.85120774609542
   1.70898348630921
  24.80906596639595
   4.46972982255813


tdfread('nonuniformitydata.txt',' ');
  Name                   Size                    Bytes  Class

  B                    10x1                         80  double array
```

```
  CF4               10x1                       80  double array
  Nonuniformity     10x1                       80  double array
  P                 10x1                       80  double array
  R                 10x1                       80  double array

Grand total is 50 elements using 400 bytes

alpha=robustfit([B,CF4,P],Nonuniformity)

alpha =

  15.56237086621604
   0.00626884566927
  -0.28023347305300
  -0.00530971951742

B

B =

    60
   180
    60
   180
    60
   180
    60
   180
   120
   120

P

P =

   550
   550
   550
   550
   700
```

```
   700
   700
   700
   625
   625


m=[1,60,15,550;1,60,15,625; 1,60,15,700;
1,120,15,550; 1,120,15,625; 1,120,15,700;
1,180,15,550;  1,180,15,625; 1,180,15,700];


performance=m*[beta,alpha]

performance =

   1.0e+03 *

   2.96787760882756   0.00881465377600
   3.30310734551942   0.00841642481219
   3.63833708221128   0.00801819584838
   3.07041661800611   0.00919078451615
   3.40564635469797   0.00879255555235
   3.74087609138983   0.00839432658854
   3.17295562718467   0.00956691525631
   3.50818536387653   0.00916868629250
   3.84341510056839   0.00877045732870

save

Saving to: matlab.mat

quit
```

2.

```
[p,anova,stats]=anova1(LesionLength,Treatment)

p =
```

```
    0.0302


anova =

    'Source'    'SS'            'df'    'MS'          'F'         'Prob>F'
    'Groups'    [   84.4812]    [ 2]    [42.2406]     [3.6317]    [0.0302]
    'Error'     [1.1166e+03]    [96]    [11.6309]           []          []
    'Total'     [1.2011e+03]    [98]           []           []          []


stats =

    gnames: {3x1 cell}
         n: [33 32 34]
    source: 'anova1'
     means: [1.5456 3.7174 2.0218]
        df: 96
         s: 3.4104




crosstab(Treatment)

ans =

    33
    32
    34

Placebo=LesionLength(1:33);
Old=LesionLength(33+(1:32));
New=LesionLength(33+32+(1:34));
std(Old)

ans =
```

```
    4.5321

std(New)

ans =

    2.5287


[h,p,ci]=ttest2(Placebo,Old,.05,'both','unequal')
%% Note: we are not assuming equal variances.

h =

     1


p =

    0.0258


ci =

   -4.0713   -0.2723

[h,p,ci]=ttest2(Placebo,New,.05,'both','unequal')

h =

     0


p =

    0.4768


ci =
```

```
   -1.8056    0.8532

[h,p,ci]=ttest2(Old,New,.05,'both','unequal')

h =

     0


p =

    0.0688


ci =

   -0.1361    3.5274

boxplot(LesionLength,Treatment)
quit
```

3.

```
sim=binornd(10,.1,1,10000);

ciold=binciold(sim,10,90);
cinew=binci(sim,10,90);
ciold(1:2,1)

ans =

   -0.0560
    0.2560


fracold=sum(ciold(1,:)<0.1 & ciold(2,:)>0.1)/10000;
fracold

fracold =

    0.6371
```

```
fracnew=sum(cinew(1,:)<0.1 & cinew(2,:)>0.1)/10000;
fracnew

fracnew =

    0.9315

diary off
```