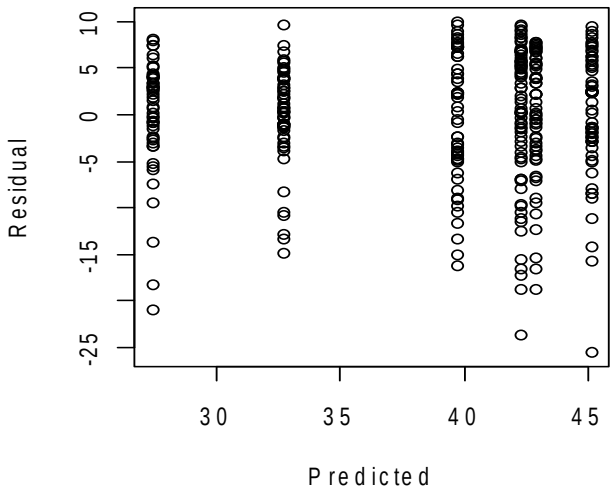
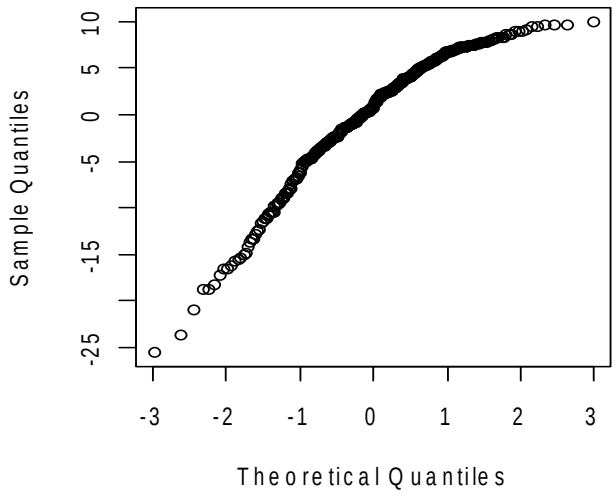
1. Diagnostics for the diet and longevity study

(a) Independence: look at the study design. Treatments were randomly assigned to individual mice; longevity was measured on individual mice. So, e.u. = mouse is the same as the o.u. = mouse. No obvious problem with lack of independence.

(b) Equal variances: a variety of diagnostic methods (e.g. residual vs predicted value plot, box plots of the observations in each group, or calculation of the within group s.d.) or tests (Levene's test) are acceptable. All groups have similar variabilities. No problem with equal variances.

(c) Normality: If you pool the residuals, then plot a histogram or quantile-quantile plot, you see some deviations from normality: there are some especially small values. If you look at each group separately, you see some groups that seem normal and others that do not. I think there is a problem with normality. Other well justified answers were accepted.





2) Marker effects in a genome-wide study

(a) 10 Markers: numbered 1-10

- Here are my computations:

marker	pvalue	adjP
7	0.000	0.004
9	0.000	0.013
1	0.001	0.046
5	0.002	0.046
3	0.002	0.040
6	0.003	0.053
2	0.004	0.050
8	0.004	0.051
4	0.005	0.052
10	0.005	0.047

63	0.008	0.071
73	0.009	0.072
24	0.010	0.073
17	0.017	0.119
64	0.027	0.179
46	0.040	0.249
62	0.078	0.460

Note that markers 6,2,8, and 4 are included because 10 has an adjusted p value < 0.05

(b) No markers. Here are my calculations:

marker	pvalue	adjP
83	0.032	3.225
96	0.046	2.275
64	0.051	1.710

The smallest adjusted P value is much larger than 0.05.

(c) You shouldn't be surprised. You are identifying markers in a region expected contain associations and no markers in a region expected to lack associations.

- Notice that without adjustment, you "identify" markers in the B data set.

3) Teaching methods.

(a) Treatment:	L+D	PT	PT+L	CI	CI+L
C1:	1/3	-1/2	1/3	-1/2	1/3
C2:	0	1/2	1/2	-1/2	1/2

C1 is more precisely estimated. The variance of C1 is $0.83\ \sigma^2 / n$. The variance of C2 is $1.00\ \sigma^2 / n$.

(b) The basic equation is $\delta = (t_{1-\alpha/2} + t_{1-\beta})$ s.e., Substituting $se = \sqrt{0.83\ \sigma^2 / n}$ and solving for n gives: $n = 0.83(t_{1-\alpha/2} + t_{1-\beta})^2\ \sigma^2 / \delta^2$. I give you $\sigma = 5$ and $\delta = 3$. If you start with quantiles of 2 and 0.85, you get n = 18.8, i.e. n = 19 per treatment. A CRD with 5 treatments and 19 subjects per treatment has 90df for the error. Those quantiles are 1.987 and 0.846. Evaluating n again gives you n = 18.5, i.e the same sample size: 19 subjects per treatment.

(c) You should use design D2, because it has the smaller s.e..

Notes: The variance of the contrast is $\sigma^2\ \Sigma\ c_i^2 / n_i$. You should choose the allocation of effort that gives the smaller $\Sigma\ c_i^2 / n_i$. That will have the smaller variance and s.e. for contrast C1. Design D1 gives you $\Sigma\ c_i^2 / n_i = 0.06944$; design D2 gives you $\Sigma\ c_i^2 / n_i = 0.06667$. Although the difference seems small, D1 is 1.041 times D2, so D2 is 4% more efficient. Of course, all this depends on C1 being the most important question. If all pairs of differences were the primary concern, D1 is much better.

4) Platelet data. My SAS code for all parts, without extra run commands, is:

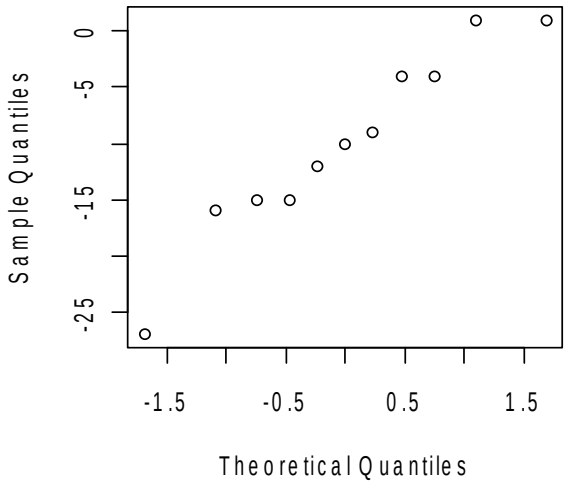
```
data platelet;
  infile 'platelet.txt';
  input before after;
  diff = before - after;
proc univariate cibasic;
  var diff before after;
data platelet2;
  infile 'platelet2.txt';
  input pair trt $ aggreg;
proc glm;
  class pair trt;
  model aggreg = pair trt;
  run;
```

(a) The estimated mean difference (before – after) is -10.0 with a 95% confidence interval of (-15.6, -4.4).

Note: the s.e./3 is 0.8, so numbers are reported to the tenth's digit. The 0.975 quantile of the 11 df T distribution is 2.200, so the 95% ci is -10 +/- 2.200*2.51.

(b) t = -3.98, with p-value 0.0026.

(c) The paired t-test is appropriate if the differences are approximately normal. A qqplot is reasonably close to a straight line, so yes, a paired t-test is appropriate.



(d) The variance for before is 243.7; that for after is 351.0. Those give you the variance of the difference, ignoring the pairing, = 54.1. Directly measuring the variance of the differences gives you 69.4, so the variance of the mean difference = 6.3. Yes, pairing is a very good idea for this sort of study.

(e) The F statistic for treatment (before or after) is 15.85, with p = 0.0026. Yes, you get the same p-value.

5) Fats in diets. My SAS code for all parts, without extra run commands, is:

```
data dietfat;
  infile 'dietfat.txt';
  input block fat $ lipid;
proc glm;
  class block fat;
  model lipid = block fat /clparm;
  estimate 'Extr. low - Fairly low' fat 1 -1;
  lsmeans fat /stderr;
  run;
```

(a) F = 273, p = < 0.0001. There is very strong evidence of at least one difference among lipid treatments.

Note: Various other wordings are reasonable, including: very strong evidence that not all the means are equal.

Claiming "all the means are different" is not correct. (All you know is that at least one is different; the rest may be exactly the same, but the F statistic will still be large.)

(b) The difference in mean lipid level between the Extremely low and Fairly low diets is estimated to be 0.12, with a 95% confidence interval of (0.05, 0.19). There is strong evidence that this difference is not zero (p=0.0053).

Note: no adjustment for multiple comparisons is needed. This is the primary question in the study and presumable was chosen before the data are analyzed.

(c) I=5, J=3, $MS_{blocks} = 0.3547$, $MS_{error} = 0.0024$ so $\sigma^2_{CR} = (4*0.3547 + 5*2*0.0024)/(14) = 0.103$.

Hence efficiency of the RCBD relative to the CRD = 0.103 / 0.0024 = 42.9.

(d) definitely. This is a huge efficiency!