Homework Assignment 7
Solutions

1. Please see the code, and make sure that 'dose' is a factor!
   
   (a) The simplest way is to use the function `stripchart`. Not particularly nice, but it does the job. Note that we jittered the response a bit.
   
   (b) We carry out a one-way ANOVA, and conclude that there are differences between the doses.
   
   (c) We can use the function `ci.bonf()` from `lab` and the code to the lecture notes. We conclude that the response to the 200mg dose is larger than that of the 0mg dose, but that we cannot detect any significant differences between the 0mg and 100mg doses or the 100mg and 200mg doses.
   
   (d) We use the function `TukeyHSD()` and the output from `aov()` to get Tukey's HSD confidence intervals. Our conclusions are the same as for part (c).
   
[ 4 points ]

2. Please see the code for details. This is a one-way ANOVA situation. We do the ANOVA and then study the diagnostic plots. The variance is not constant, and the residuals are not normally distributed. After a logarithmic data transformation the results look a lot better! We go ahead and calculate the ANOVA table:

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>4</td>
<td>1.15</td>
<td>4.29</td>
<td>0.0041</td>
</tr>
<tr>
<td>Residuals</td>
<td>59</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We get a P-value of 0.004, indicating strong evidence for differences among the cancer types.

[ 4 points ]

3. Please see the code for more details.

   (a) A stripchart for example is a good way of showing the data.
   
   (b) The null hypothesis is that all group means are the same, \( H_0 : \mu_1 = \cdots = \mu_5 \), versus the alternative that they are not all the same. A one-way ANOVA yields a F-statistic of 3.76, and a p-value of 0.009 (4 and 60 degrees of freedom), so we conclude that indeed there are differences in the stem lengths of daffodils from the different sites.
   
   (c) Both approaches (assuring a 5% family-wise error rate) detect significant differences between the open area and the areas north and south of the building, but otherwise, do not indicate differences at this significance level.
   
   (d) The Kruskal-Wallace test yields a p-value of 0.005, and thus, the differences between sites are more significant than those from the parametric one-way ANOVA.
(e) Both tests are still significant, with the one-way ANOVA becoming way more significant (0.0009 versus 0.009 before removing the smallest observation). This appears to be largely due to the within-group variance being much smaller after removing the smallest observation (2192.4 versus 3034.3). The Kruskal-Wallace test yields about the same significance (0.003 versus 0.005 before).

(f) We do not detect any significant differences between any of the four sites around the building (assuring a 5% family-wise error rate via Bonferroni), but do now detect significant differences between all of the sites around the building when comparing the stem lengths to those in the open area nearby (Tukey’s HSD yield the same info, btw).

(g) Overall, it appears that the stem lengths do not differ much when comparing the daffodils from the four sides of the building, but the stem lengths appear to be longer compared to the daffodils from the open area. Maybe they daffodils have to “stretch” some more around the building to catch the sunlight? There was one very short daffodil collected on the west side of the building that had a fair amount of influence on the statistical inference: the differences between the stem lengths from the open area and those from the west and east side of the building were not statistically significant (though close!) after multiple comparisons correction when including that flower.

4. Please see the code for the details of the simulation. When the means are 3, 4, 5 and 6 respectively, the estimate I get is 81.4% power. Since 8,140 out of the 10,000 iterations yielded a p-value less than 0.05, the 95% confidence interval for the power can be calculated with `binom.test()`, which gives (0.806;0.821). Note that a rough 95% confidence interval for the power can be calculated as 81.4% ± 1% (i.e., $\hat{p} \pm 1/\sqrt{n}$). When the means are 3, 3, 6 and 6 respectively, we expect more power because the squared treatment effects ($\alpha_i$) are larger, and therefore the expected means squares for the treatment effects are larger. The simulation shows 98% power.

5. Please see the code for details. We have two strains (pHH and pHL), and the litters are nested within the strains. The strains were selected for high and low blood pH, and thus, are fixed effects. The litters were randomly chosen, and thus, are random effects. We can therefore analyze the data using a nested ANOVA with a fixed group and random subgroup effect. In other words,

$$Y_{ijk} = \mu + \alpha_i + B_{ij} + \epsilon_{ijk},$$

where $Y_{ijk}$ is the $k^{th}$ reading of litter $j$ within strain $i$. We assume $B_{ij} \sim N(0, \sigma^2_B)$ and independent, and $\epsilon_{ijk} \sim N(0, \sigma^2)$, independent. We use the function `nested.anova()` from the lab to carry out the nested ANOVA, and get the following ANOVA table:

```
                      Df  Sum Sq  Mean Sq   F value  Pr(>F)
  x1                   1 0.006645 0.006645 1.26380  0.2829
  x1:x2                12 0.063093 0.005258 2.22211  0.0282 *
Residuals             42 0.099375 0.002366
```

Looking at some diagnostic plots, we do not see any model violations. So we conclude that there are no significant differences in the blood pH readings between the two strains. We also conclude that there is an appreciable variation in readings between litters. The mean squares from the residuals are our estimator for $\sigma^2$, i.e. $\hat{\sigma}^2 = 0.0024$. We estimate the litter variance component as $\hat{\sigma}^2_B = (0.0053 - 0.0024)/4 = 0.0007$, which corresponds to $0.0007/(0.0024 + 0.0007) \approx 23.4\%$ of the total variability.

[5 points]