Answers for Homework 6

1. The likehood for a sample $x_1, x_2, ..., x_n$ is

$f(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} a x_i^{a-1}$

$\log f(x_1, x_2, ..., x_n) = (a-1)\log(x_1 x_2 ... x_n) + n \log a$

Now, find the mode of the log-likelihood. Let

$\frac{\partial \log f(x_1, x_2, ..., x_n)}{\partial a} = \log(x_1 x_2 ... x_n) + \frac{n}{a} = 0$

Note: To make sure we don't get the minimum instead of maximum, let's look at the second derivative

$\frac{\partial^2 \log f(x_1, x_2, ..., x_n)}{\partial a^2} = -\frac{n}{a^2}$

This is always negative, which means we are getting the maximum.

Solving a out of the equation, we get

$a_{MLE} = -\frac{n}{\log(x_1 x_2 ... x_n)}$

In this instance, $\hat{a} = -\frac{4}{\log(0.200.900.800.90)} \approx 1.96$.

8.33 We should use small-sample confidence interval on page 374. We can do this without assuming equal variance. (Reason: fist paragraph below Figure 8.11, page 376).

$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{11 * 0.104^2 + 11 * 0.035^2}{22} = 0.006$;

$t_{0.025}$(on 12+12-2=22 degrees of freedom)= 2.074;

$\mu_2 - \mu_1 = (\bar{y}_2 - \bar{y}_1) \pm t_{0.025} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = -0.20 \pm 0.066$

We can conclude that, with 95% confidence, the Ca^{2+} uptake in fast muscles is more than that in slow muscles.

8.34 Since $n_1 = n_2 = 30$, we use large-sample approximation (on page 372), so a 90% confidence interval of the difference would be,

$\mu_{old} - \mu_{new} = (\bar{y}_1 - \bar{y}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} = \frac{s_2^2}{n_2}}$

$= (9.80 - 9.42) \pm z_{0.05} \sqrt{\frac{.2926}{30} + \frac{.2293}{30}} = 0.38 \pm 0.22$

Since 0 is not contained in the interval, there is strong evidence that $\mu_{old} > \mu_{new}$. So it may not be a good idea to establish the process locally.

8.40

```
MTB > erase c1, c2, c3
MTB > name c1='Joseph', c2='Iowa', c3='Diff'
MTB > set 'Joseph'
DATA> 782
DATA> 965
DATA> 948
DATA> 1181
DATA> 1414
DATA> 1633
DATA> 1852
DATA> end
MTB > set 'Iowa'
DATA> 593
DATA> 672
DATA> 750
DATA> 988
DATA> 1226
DATA> 1462
DATA> 1698
DATA> end
MTB > subtract 'Iowa' 'Joseph' 'Diff'
MTB > tinterval 95 'Diff'

             N      MEAN    STDEV  SE MEAN   95.0 PERCENT C.I.
Diff         7     198.0     44.5     16.8   (  239.2,  156.8)
```

So a 95% confidence interval for the differnce between St. Joseph and Iowa is 198.0 ± 41.2, which means the readings in St. Joseph is significantly greater than those in Iowa Great Lakes.

8.50 Since $\hat{p} = 21/35 = 0.60$, $n\hat{p} > 4$ and $n\hat{q} > 4$, we can refer to the large-sample confidence interval for the proportion on page 388.

$p = \hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.60 \pm 2.575\sqrt{\frac{0.60*0.40}{35}} = 0.60 \pm 0.21$

So, we are 99% confident that the proportion of patients preferring Caridex is between 39% and 81%.

8.72 Use the formulae on page 409 for the sample size for estimating the difference between to proportions:

$n = (\frac{z_{0.05}}{H})^2(0.1 * 0.9 + 0.1 * 0.9) \approx 195$

So at least 195 residences should be inspected.

8.80 We can do it in 2 ways: assume the variances are equal and use the small sample approximation on page 374 (easier); or use the second approximation on page 377 (harder).

I'll do it the first way.

(a.) $t_{0.005}$ with $(11 + 17 - 2 = 26)$ degrees of freedom $= 2.779$
$S_p^2 = \frac{10*24.4+16*19.9}{26} = 21.63$
Therefore, a 99% C.I. would be
$\mu_1 - \mu_2 = 43.7 - 48.5 \pm t_{0.005}\sqrt{S_p^2(1/11 + 1/17)} = -4.8 \pm 5.0$

(b.) 1) Independent sample;

2) Both populations are approximately normally distributed;

3) **The population variances are equal**.

(c.) We cannot conclude that the capacities for the two alloys are different, because 0 is contained in the confidence interval.

(d.) Refer to page 408, the sample size for estimating the difference between two population means:
$n = (\frac{z_{0.005}}{2})^2(\sigma_1^2 + \sigma_2^2) \approx 74$