Solutions to the Review Problems

1.a. $\overline{x} \pm z\frac{\sigma}{\sqrt{n}} = 7.3 \pm 1.96\frac{1.9}{\sqrt{90}} = 7.3 \pm .3925$.Could use t^* instead of z^* with 89 df (need to use 80). t=1.99

b. $\widetilde{p} = \frac{18+2}{90+4} = .2128$. $\widetilde{p} \pm z\frac{\sqrt{\widetilde{p}(1-\widetilde{p})}}{\sqrt{n+4}} = .2128 \pm 1.96\sqrt{.2128 * .7872}/\sqrt{94} = .2128 \pm .0827$

c. $H_0 : \mu \leq 7 \; vs \; H_a : \mu \; > 7.$ $t = \frac{7.3-7}{1.9/\sqrt{90}} = 1.5$ Since t is less than the table value of 1.664(80 d.f) retain the null hypothesis that $\mu \leq 7$ (could use z).

d. $H_0 : p \leq .15 \, vs \, H_a : p > .15$ $z = \frac{.2-.15}{\sqrt{.15*.85}/\sqrt{90}} = 1.328$. Since z is less than the table value of 1.645 retain the null hypothesis that $p \leq .15$

e. $n = \frac{z^2*.25}{e^2} = \frac{1.96^2*.25}{.04^2} = 600.25 \, or \, 601.If \;\; p \leq .3$ then .25 is replaced by $.3 * .7 = .21$.Sample size becomes 504.21 or 505. Note that we can subtract 4 from n if we use $\widetilde{p}$

f. We assumed that the 90 observations that we have is a random sample of the population. But 60 individuals did not respond. It is entirely possible that those who did not respond are very different from those who did respond.

2. a. i) $H_0 : p \leq .6 \; vs \; H_a : \; p > .6$

ii) $n = \frac{z^2*.25}{e^2} = \frac{1.96^2*.25}{.08^2} = 150.0625$ or 151 Note that we can subtract 4 from n if we use $\widetilde{p}$

b. z needs to be 1.645. $z = \frac{\widehat{p}-.6}{\sqrt{.6*.4}/\sqrt{400}}$Solving for $\widehat{p}$ yields $\widehat{p} = .6403$. Number correct is 400(.6403) or at least 257 correct.

c. The closer that p is to .6 from above the harder it is to distinguish it from .6, so the greater error occurs when p=.63. This is a Type II error which in this context is failing to conclude that ESP satisfies the criterion of $p > .6$

3. ai) $H_0 : \; \mu_d \geq 0 \; vs. \; H_a : \mu_d < 0.$ $t = \frac{-41.2-0}{\sqrt{3077.767}/\sqrt{25}} = -3.7133$ Since $t \leq -t_1 = -1.711$(24 d.f.) Reject the null hypothesis and conclude that the review course works.

ii) $H_0 : \; \mu_d \geq -20 \; vs. \; H_a : \mu_d < -20.$ $t = \frac{-41.2-(-20.0)}{\sqrt{3077.767}/\sqrt{25}} = -1.911$ Since $t \leq -t_1 = -1.711$(24 d.f.) Reject the null hypothesis and conclude that the review course works.

b. $H_0 : \; p \leq .5 \; vs. \; H_a : p > \; .5$ $\widehat{p} = \frac{19}{25} = .76$ $z = \frac{.76-.5}{\sqrt{.5*.5}/\sqrt{25}} = 2.6$ Since $z \geq z_1$ =1.645 Conclude that the review course does raise a student's score over 50% of the time.

4.ai) $H_0 : \mu_{1994} \geq 3 \, vs \, H_a : \mu_{1994} < 3.$ $t = \frac{2.5-3}{2/\sqrt{150}} = -3.062$ Since this is less than -1.66 we conclude that $\mu < 3$

aii) Limit of the acceptance region is $3 - 1.645\frac{2}{\sqrt{150}} = 2.7314$ Probability of an error is $P\left(\overline{X} \geq 2.7314|\mu = 2.5\right) = P(Z \geq \frac{2.7314-2.5}{2/\sqrt{150}} = 1.42) = 1 - .9222 = .0778$

b. $n = \frac{(z_\alpha+z_\beta)^2\sigma^2}{(\mu_a-\mu_b)^2} = \frac{(1.645+1.645)^2 2^2}{(2.5-3)^2} = 173.1856 \, or 174$

c. $H_0 : p_{1994} \leq .8 \, vs \, H_1 : p_{1994} > .8$ $\widehat{p} = \frac{125}{150} = .8333$ $z = \frac{.8333-.8}{\sqrt{.8*.2}/\sqrt{150}} = 1.0205 < 1.645$.Retain the possibility that $p \leq .8$

5.a. $H_0 : \mu_d \leq 0 \, vs \, H_a : \mu_d > 0$ $t = \frac{6.25}{9.9535/\sqrt{8}} = 1.776 < 1.895$ $(7d.f.)$ Retain the possibility that it does not increase the percentage who goes to college.

b. Since $\overline{d} = 6.25$ The sum is 50 so if we eliminate the 29 we get a $\overline{d} = \frac{50-29}{7} = 3$ Hence s^2 also changes to 17. $t = \frac{3}{\sqrt{17}/\sqrt{7}} = 1.9251 < 1.943$ $(6d.f.)$ Retain the possibility that it does not increase the percentage who goes to college.

c. There is no difference in the conclusions to a and b

6.a. $H_{0:}\, \mu \geq 30$ $H_{a:}\, \mu < 30$ $t = \frac{20-30}{20/\sqrt{15}} = -1.9365$. Looking in the t- table (14 df) puts this value between the columns of .025 and .05. Hence the P-value must be in that range.

b. Since the standard deviation is large relative to the mean and the numbers must be positive the times are not normally distributed. That makes the test in part a. suspect.

c. Limit for acceptance region is: $30 - 1.645\frac{20}{\sqrt{15}} = 21.5053$.If the mean is 20 then the average is normal with a mean of 20 and a standard deviation of $\frac{20}{\sqrt{15}}$ $P\left(\overline{X} \leq 21.5053\right) = P\left(Z \leq \frac{21.5053-20}{20/\sqrt{15}} = .29\right) = .6141$

d. $H_0 : p \geq .5 \, vs \, H_a : p < .5$ $\widehat{p} = \frac{3}{15} = .2$ $z = \frac{.2-.5}{\sqrt{.5*.5}/\sqrt{15}} = -2.3238 < -1.645$ This shows that the median is below 30 minutes.

e. Since the t-value is not large enough (e.g. at least 1.645) retain the possibility that the average for airline 1 is no worse than that for airline 2.

f. Test in part d is based on the median which does not require that the data be normal or an outlier would not have much effect. The new average would be $\overline{x} = \frac{300-90}{14} = 15$.This turns out to give the same average as that for airline 2.

7 a. Retain if rating is 2,3 4 or 5 and reject if rating is 1. If the null hypothesis is true the probability of a type I error is the probability of getting a rating of 1 with a benign tumor which equals .05. If the alternative hypothesis is true the probability of a type II error is the probability of getting a rating of 2,3,4, or 5 with a malignant tumor which is .2+.3+.2+.1=.8

b.

c	Acceptance Region	Rejection Region	Alpha	Beta	Alpha+Beta
1	{1,2,3,4,5}	Never	0	1	1
2	{2,3,4,5}	{1}	.05	.8	.85
3	{3,4,5}	{1,2}	.1	.6	.7
4	{4,5}	{1,2,3}	.3	.3	.6
5	{5}	{1,2,3,4}	.6	.1	.7
>5	Never	{1,2,3,4,5}	1	0	1

It is best when c=4.

8. a. For each additional hour spent on school related functions the average GPA increases by an estimated amount of .0342 GPA points.

b. $\widehat{GPA}$ = 1.76447+(.0342453)20= 2.45. Since RMSE=.402178 the 95% prediction interval is 2.45 $\pm$1.96(.402178) = 2.45 $\pm$ .79

c. i) The P-value is 1 in ten thousand. This says the probability that observing a coefficient for school hours that is at least as far from zero as the observed value of .0342 when there is no relationship between school hours and GPA happens only 1 in ten thousand times. We therefore would reject the null hypothesis in favor of H_a that says the slope is not zero. In this context, it is best to do a one-sided test as it does not make sense (we hope) that working more hours lowers one's anticipated GPA.

ii) The 95% confidence interval is of the form: Estimate $\pm$ Margin of error where Margin of error is 1.96(Standard Error). In this case, it is $.0342 \pm 1.96(.00183) = .0306\,to\,$ .0378 We are 95% certain that the range of (.0306,.0378) includes the true average increase in GPA for each additional hour spent on school related functions.