Solutions for HW #2
STAT 554, Spring 2008

1) Minitab (or other appropriate software, or calculator and tables) can be used to obtain the confidence intervals for parts (a) through (c).
(a) **(61.2, 65.1)**
(b) **(61.6, 64.7)**
(c) **(60.1, 66.2)**
(d) Using

$$\sqrt{n^*}\left(\frac{\overline{X}^* - \mu_X}{S_0}\right)$$

as a pivot, one can arrive at (using steps very similar to those given on p. 55 of the course notes)

$$(\overline{x}^* \pm t_{n_0-1,0.025} s_0/\sqrt{n^*})$$

as a 95% confidence interval for μ_x. For the half-width to be no greater than 0.5, we need

$$t_{n_0-1,0.025} s_0/\sqrt{n^*} \le 0.5,$$

or equivalently,

$$n^* \ge (t_{n_0-1,0.025} s_0/0.5)^2.$$

Plugging in $t_{n_0-1,0.025} \doteq 2.57058$ and $s_0 \doteq 1.85221$, we have that the lower bound for n^* is about 90.68, and so the smallest value that will work for n^* is 91. This means that the fewest number of additional observations that will work is $91 - 6 = \mathbf{85}$. If the sample mean of the 85 additional observations is 61.38, the sum of the 85 additional observations is 85(61.38), and so the sample mean of all 91 observations is

$$(62.7 + 66.3 + 60.6 + 63.0 + 62.7 + 63.7 + 85(61.38))/91 \doteq 61.4978,$$

and so the resulting confidence interval is about

$$(61.4978 \pm 2.57058(1.85221)/\sqrt{91}) \doteq (61.4978 \pm 0.4991),$$

or **(61.0, 62.0)**.

2) The rejection region for a two-sided size 0.05 test is the union of the rejections regions for a lower-tailed size 0.025 test and an upper-tailed size 0.025 test. The power function for a two-sided size 0.05 test is the sum of the power functions for a lower-tailed size 0.025 test and an upper-tailed size 0.025 test. For a lower-tailed size 0.025 z test, the power function is

$$\beta(\mu) = P_\mu(\sqrt{n}(\overline{X}-\mu_0)/\sigma \le -z_{0.025}) = P_\mu(\sqrt{n}(\overline{X}-\mu)/\sigma \le \sqrt{n}(\mu_0-\mu)/\sigma - z_{0.025}) = \Phi(\sqrt{n}(\mu_0-\mu)/\sigma - z_{0.025}).$$

To obtain the power function for a two-sided size 0.05 test, we can add to this the power function for an upper-tailed size 0.025 test, which can be obtained from p. 62 of the class notes. The resulting power function (for the two-sided test) is

$$\beta(\mu) = \Phi(\sqrt{n}(\mu_0 - \mu)/\sigma - z_{0.025}) + \Phi(\sqrt{n}(\mu - \mu_0)/\sigma - z_{0.025}).$$

(a) Plugging in $n = 25$, $\sigma = \sqrt{400} = 20$, $\mu_0 = 200$, and $\mu = 195$, we get

$$\beta(195) = \Phi(\sqrt{25}(200 - 195)/20 - z_{0.025}) + \Phi(\sqrt{25}(195 - 200)/20 - z_{0.025}),$$

which is approximately equal to

$$\Phi(1.25 - 1.95996) + \Phi(-1.25 - 1.95996) \doteq \Phi(-0.7100) + \Phi(-3.2100) \doteq 0.23886 + 0.00066,$$

or about **0.240**.

(b) To have
$$\beta(195) = \Phi(\sqrt{n}(200 - 195)/20 - z_{0.025}) + \Phi(\sqrt{n}(195 - 200)/20 - z_{0.025})$$
equal to 0.80, we need to find the value of n for which
$$\Phi(\sqrt{n}/4 - z_{0.025}) + \Phi(-\sqrt{n}/4 - z_{0.025}) = 0.80.$$
Since a sample of size 25 results in a power of about 0.24, to increase the power to 0.80 we're going to have to increase n. For the increased sample size, the contribution from the upper-tail portion of the rejection region will be even smaller than the negligible contribution which results from using $n = 25$. So we can ignore the upper-tail contribution, and seek n to make $\Phi(\sqrt{n}/4 - z_{0.025}) = 0.80$. So we need
$$\sqrt{n}/4 - z_{0.025} = \Phi^{-1}(0.80) = z_{0.2},$$
or equivalently,
$$\sqrt{n}/4 = z_{0.025} + z_{0.2},$$
and so
$$n = (4(z_{0.025} + z_{0.2}))^2 \doteq 125.58.$$
Since n must be an integer, one can use **126** for n. (125 is also okay.)

3)
(a) Using pmf values (obtainable from Minitab), it can be seen that the exact p-value is the sum of all of the pmf probabilities that are less than or equal to $p(11) \doteq 0.0282$. So, letting Y be a binomial (15, 9/19) random variable, the p-value is $P(Y \leq 3) + P(Y \geq 11) = P(Y \leq 3) + 1 - P(Y \leq 10) \doteq$ **0.067**.
(b) For the approximate p-value, we need to double the approximation of $P(Y \geq 11)$, and so the approximate p-value is
$$2(1 - \Phi((10.5 - 15(9/19))/\sqrt{15(9/19)(10/19)})) \doteq 2(1 - \Phi(1.75547)),$$
which is about **0.079**.

I'll use my Minitab macro to obtain the confidence intervals for parts (c) through (g).
(c) **(0.51, 0.96)**
(d) **(0.48, 0.99)**
(e) **(0.48, 0.90)**
(f) **(0.45, 0.92)**
(g) **(0.45, 0.92)**

4) I'll use my Minitab macro to obtain the confidence intervals.
(a) **(0.477, 0.648)**
(b) **(0.473, 0.652)**
(c) **(0.476, 0.645)**
(d) **(0.472, 0.649)**
(e) **(0.472, 0.650)**

5) We have $r = 25$, $n_0 = 113$, and $n_1 = 15$. The null hypothesis expected value and variance are about 27.484375 and 5.3144704, respectively. The normal approximation results in an approximate p-value of
$$\Phi((25.5 - 27.484375)/\sqrt{5.3144704}) \doteq 2\Phi(-0.8608),$$
which is about **0.39**.

6) 14% of 250 is 35, and so the p-value is $P(W \geq 35)$, where W is a hypergeometric random variable. (W is the number of successes when a simple random sample of size 250 is drawn from a population consisting of 100 successes and 900 failures (where here a success is a machine that *required* service).) One can use the normal approximation (with continuity correction) here to obtain a decent approximation of the exact p-value. We have
$$P(W \geq 35) = 1 - P(W \leq 34) \simeq 1 - \Phi((34.5 - 250(0.1))/\sqrt{(750/999)(250)(0.1)(0.9)}),$$

which is about $1 - \Phi(2.31145) \doteq 0.0104$, and so we can say that the p-value is about **0.010**. (Note: The exact p-value obtained from the hypergeometric distribution is about 0.012.)

7) The desired confidence interval is

$$(\bar{x} \pm t_{14,0.01} s/\sqrt{15}).$$

Minitab can be used (as shown below) to determine that the confidence interval is about **(83.0, 89.0)**.