Statistics 410, Fall 2011
Solutions to homework 8, due 11/16/11

11.11 (a) This corresponds to being 1 standard deviation above the mean. So .15866.
(b) For the sample total, the standard deviation is $\sqrt{n}\sigma = 2{\cdot}5 = 10$. This is then two standard deviations over the mean, so the probability is .02275.
(c) This will be the probability from the first part raised to the fourth power. $(.15866)^4 = 0.00063368$.

11.12 (a) The mean will be $10{\cdot}40 = 400$ minutes. The standard deviation is $\sqrt{10} \cdot 15 = 47.43$.
(b) Let X be the repair time for 10 items repaired sequentially. 6 hours is 360 minutes. $40 = .8433\sigma$.

$$P(X \leq 360) = P(X \leq 400 - .8433\sigma) = \Phi(-.8433) = .199$$

11.19 $E(\overline{X}) = 1$. $Var(\overline{X}) = 4/9$. About 95% of the population lies between -3 and 5. The sample mean will lie between 1/9 and 17/9.
There is a 99% chance that the sample mean $\overline{x}$ will satisfy

$$-.144889 = 1 - 2.576(4/9) \leq \overline{x} \leq 1 + 2.576(4/9) = 2.14489.$$

11.20 $E(\overline{X}) = 60$. $SD(\overline{X}) = 5/4$.
About 95% of the population is between 50 and 70.
About 95% of the sample means lie between 57.5 and 62.5.
About 99% of the sample means lie between $60 - 2.576{\cdot}5/4 = 56.78$ and $60 + 2.576 \cdot 5/4 = 63.22$.

11.27 This corresponds to levels one standard deviation above the mean. So the probability is .15866.
Now the standard deviation is $30/\sqrt{10}$. So a mean at 200 or above is $\sqrt{10}$ standard deviations above the mean. The probability of this is .00079.

11.38 (a) For $n = 1$, there is obviously a .58 chance of a correct prediction.
(b) For $n = 5$, the prediction would be correct if there were 3, 4 or 5 "successes". This is

$$\sum_{i=3}^{5} \binom{5}{i} (.58)^i (.42)^{n-i} = .64746.$$

(c) For $n = 25$, we ask about having at least 13 successes. Using the continuity correction, we measure

$$P(K \geq 12.5) = P(z \geq \frac{12.5 - 25(.52)}{\sqrt{25(.52)(.48)}}) \approx 1 - \Phi(\frac{-1}{2.498}) = 1 - \Phi(-.40032) = .65542.$$

11.44 We have a dichotomous variable which is 1 if the income is over \$60000 and 0 otherwise. $\pi = .2$ by assumption. Let K be the number of “successes” in our sample of 400. 19% of 400 is 76. So we are asking for

$$P(K < 76) = P(K < 75.5) = P(z \le \frac{75.5 - 80}{\sqrt{\frac{50{,}000-400}{49{,}999}}\sqrt{400(.2)(.8)}}) = P(z \le \frac{-4.5}{.996 \cdot 5}) = \Phi(-.903614) = .1$$

The expected value of the total is $400 \cdot 45,000 = \$19,000,000$. The standard deviation of the total is $\sqrt{400}25,000 = \$500,000$.

So we are asking what the probability is of the total being two standard deviations or more above the expected value. This is .02275.

12.1 We are asking (if $\mu = 31,000$) what the probability of rejecting H_0 is.

To reject H_0 at the 1% significance level, we need a value for $\overline{z}$ which is in the top 1%, so above 2.33.

The population standard deviation is 7500. So the standard deviation for the sample means is $7500/\sqrt{400} = 375$. So we are asking for the probability that the sample mean is greater than $30,000 + 2.33 \cdot 375 = 30,873.8$.

$$P(\overline{X} \ge 30873.8) = P(\frac{\overline{X} - 31000}{375} \ge \frac{30873.8 - 31000}{375}) = P(z \ge -.336533) = .63307.$$

12.2 With $n = 100$, the standard deviation for the sample means is 750. If $\mu = 31000$ (a 1000 page improvement), then our sample mean is above 31000 50% of the time.

On the other hand, the test we'll be doing is asking what is

$$P(\overline{X} \ge 30,000 + z_{1-\alpha}750).$$

For this to be .5, we need $z_{1-\alpha} = 1.33$. So $1 - \alpha = .90824$, thus $\alpha = .09176$.

12.5 (a) $H_0 : \mu = 612$, that is this year's entering students had an average SAT verbal score of 612.
(b) $H_a : \mu > 612$.
(c) Significance level $\alpha = .05$.
(d) Our test statistics is 630, the average verbal SAT score of our sample. We use 80 (the sample standard deviation) to approximation σ (the population standard deviation) since $n = 100$. Then

$$\overline{z} = \frac{630 - 612}{80/\sqrt{100}} = \frac{18}{8} = 2.25.$$

(e) Our P-value is $P(z \ge 2.25) = .01222$. This is less than α, so we reject H_0, and conclude that SAT verbal scores have risen this year.

12.12 (a) $H_0 : \mu = 500$.
(b) $H_a : \mu \ne 500$.
(c) $\alpha = .05$.

(d) Our test statistic is the mean 477 from our sample of size 91. This gives

$$\overline{z} = \frac{477 - 500}{120/\sqrt{91}} = \frac{23}{12.58} = 1.83.$$

(e) The P-value is the probability that z is at least as extreme as $\overline{z}$, so is $P(|z| \geq 1.83) = 2 \times .03362 = .06724$.

(f) The P-value is greater than α, so we retain H_0. There is insufficient evidence to conclude the population mean is not equal to 500.

Since n is again large, we use s in place of σ.

12.13 (a) $H_0 : \mu = 1.5$.

(b) $H_a : \mu \neq 1.5$.

(c) $\alpha = .05$.

(d) Our test statistic is 1.504 (the mean of our sample of 400 bolts). This gives

$$\overline{z} = \frac{1.504 - 1.5}{.075/\sqrt{400}} = \frac{.004}{.00375} = 1.067.$$

We use s in place of σ since n is large.

(e) Our P-value is $P(|z| \geq 1.067) = 2 \cdot .14231 = .28462$.

(f) The P-value is greater than α so we retain H_0 and don't have evidence that the machine needs adjustment.

12.15 (a) $H_0 : \mu = 15$.

(b) $H_a : \mu > 15$.

(c) $\alpha = .05$.

(d) Our test statistic is $\overline{x} = 16.3$. This gives

$$\overline{t} = \frac{16.3 - 15}{3.32/\sqrt{19}} = \frac{1.3}{.7825} = 1.66$$

The P-value is $P(t \geq 1.66)$ for 17 df. But $P(t \geq 1.7396) = .05$, so our P-value must be greater than .05.

Another way to do this is to note that the critical value, $t_{.95} = 1.7396$, and $\overline{t} < 1.7396$.

In either analysis, we retain H_0, and conclude there is too little evidence to reject H_0.

12.16 (a) $H_0 : \mu = 60$ where μ is the actual mean drying time.

(b) $H_a : \mu > 60$. (Note: the problem is worded in such a way that you might end up doing a two-sided test here. I did a one-sided test because it is in the manufacturer's interest to advertise the lowest drying time possible, so I want to test if the speed of drying is being exaggerated.)

(c) $\alpha = .05$.

(d) Our test statistic is 66.3. This corresponds to

$$\bar{t} = \frac{66.3 - 60}{8.4/\sqrt{12}} = \frac{6.3}{2.425} = 2.60.$$

(e) The critical value $t_{.95} = 1.7959$ using the table at the back of the book for 11 df. Since 2.6 ¿ 1.7959, we reject H_0 and conclude that the paint takes longer than 60 minutes to dry.
Note that if we'd done a two-sided test, we would have asked the question: is $|2.6| \geq |t_{.975}| = 2.201$. So we still would have rejected H_0 in that case.

12.18 (a) $H_0 : \mu = 101$ (degrees Centigrade), and $H_a : \mu > 101$. Then reject H_0 if the probability of the test statistic given H_0 is under .05.

(b) The test statistic is $\bar{t} = \frac{101.2-101}{.3\sqrt{6}} = 1.633$, a t-statistic with 5 df.
$t_{.95} = 2.015$. So we retain H_0 since $1.633 < 2.015$. That is, we conclude that the temperature rise over the extra 1 degree that is allowed is such that there is a greater than 5% probability it occurred by chance.

12.21 Normalizing, this statistic becomes

$$\bar{z} = \frac{3.5 - 2.85}{.45/\sqrt{160}} = 18.27$$

This is incredibly unlikely, so is not reasonable to explain by sampling variability.

There are two likely explanations. The survey relies on self-reporting by students. If all students reported, some may have been exaggerating their GPA. If not all studnts responded, it is possible that those with good GPAs were more likely to respond.

13.3 Because we have 100 passengers, we assume that s is a good approximation to σ.

(a) About 95% of the sample of customers are between two s.d. below and two above. That is between 120 and 280 pounds.

(b)

$$192.16 = 200 - 1.96 \cdot 40/\sqrt{100} \leq \mu \leq 200 + 1.96 \cdot 40/\sqrt{100} = 207.84.$$

(c) Let X be the total weight of 50 randomly chosen passengers. We want to know $P(X \geq 10,900)$. If we assume that our sample is representative, and that the mean weight is 200 pounds, and standard deviation is 40 pounds, then $E(X) = 10,000$ with standard deviation $\sqrt{50}40 = 282.84$.

$$10,900 = 10,000 + 3.18 \cdot 282.84.$$

So $P(X \geq 10,900) = P(z \geq 3.18) = .00074$.

13.4 (a) Between $20 - 2 \cdot 5$ and $20 + 2 \cdot 5$, that is between 10 and 30.

(b) Between $20 - 2 \cdot 5/\sqrt{225}$ and $20 + 2 \cdot 5/\sqrt{225}$, that is, between 19.33 and 20.67.

(c)

$$19.14 = 20 - 2.5758 \cdot 1/3 \leq \mu \leq 20 + 2.5758 \cdot 1/3 = 20.86.$$

13.8 (a) $t_{.975}$ with 49 df is about 2.0086. So our confidence interval is

$$195,230 - 2.0086 \cdot 63,125/\sqrt{50} = \$177299 \leq \mu \leq 195,230 + 2.0086 \cdot 63,125/\sqrt{50} = \$213161.$$

(b) $t_{.95}$ with 49 df is about 1.6759. So we get

$$\mu \geq 195230 - 1.6759 \cdot 63125/\sqrt{50} = 180269.$$