12.10

We must find the sample size, n, such that

$$1.645\sqrt{\frac{(0.7)(0.3)}{n}} \leq 0.04$$

$$\frac{(0.7)(0.3)}{\left(\frac{0.04}{1.645}\right)^2} \leq n$$

So the desired sample size is 356.

If 50% responds favorably, the margin of error of the 90% confidence interval is

$$1.645\sqrt{\frac{(0.5)(0.5)}{356}} = 0.0429$$

12.12

We must find the sample size, n, such that

$$1.96\sqrt{\frac{(0.75)(0.25)}{n}} \leq 0.04$$

$$\frac{(0.75)(0.25)}{\left(\frac{0.04}{1.96}\right)^2} \leq n$$

So the desired sample size is 451.

12.14

We will test whether p, the proportion of patients who take this medication, is fewer than 10%. That is, we test

$$H_0: p = 0.1 \quad \text{vs.} \quad H_A: p < 0.1$$

Note that $n\hat{p} = 23 > 10$, and $n\hat{q} = 417 > 10$, so our sample size is large enough. Also it is reasonable to assume that 440 < 0.1(All people who may take this medication.) Assume the patients were randomly assigned to the treatment group.

The TI-84 1-PropZTest gives: z = -3.337, p-value = 0.000423.

This very low p-value shows strong statistical evidence against H_0. That is we have strong evidence to support the claim that the proportion of patients taking this medication is less than 10%.

12.18

(a) $$2.5758\sqrt{\frac{(0.2)(0.8)}{n}} \le 0.015$$

$$n \ge 4718.148$$

So the required sample size is 4719.

(b) The TI-84 99% C.I. is (0.08877, 0.11127), so the margin of error is 0.01127, which is well within the required margin of error from part (a).

12.19

(a) We will test whether *p*, the proportion of free throws Shaq makes is greater than 0.533. That is,

$$H_0: p = 0.533$$
$$H_A: p > 0.533$$

Consider the 39 free throws in the most recent game an independent SRS of Shaq's free throws. Shaq has thrown way more than 390 free throws.
$n\hat{p} = 26 > 10$ and $n\hat{q} = 13 > 10$ So we can proceed.

The TI-84 1-PropZTest gives $z = 1.673$, *p*-value = 0.0471. This low *p*-value (<5%) is moderately strong evidence against H_0, in support of the claim that Shaq's free throw percentage is greater than 0.533.

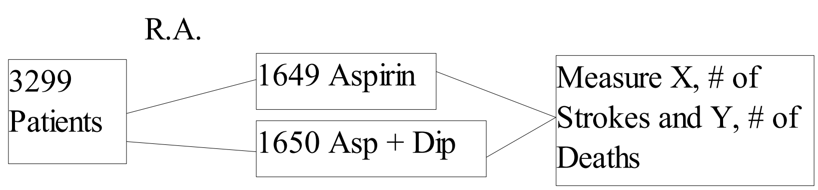
(b) Type I error means that we claim that Shaq's free throw percentage has gone up when it has not. Type II error means that we would claim Shaq's free throw percentage did not go up when in fact it did.

(c) Note that we will reject at the 5% level H_0 if $\hat{p} > 0.66440654$. So

$$P(\hat{p} > 0.66440654 \text{ if } p = 0.6) = \text{normalcdf}\,(0.66440654, 1\text{E}99, 0.6, \sqrt{(0.6)(0.4)/39}) = 0.2058$$

(d) P(Type I) = 0.05, and P(Type II) = 1 – 0.2058 = 0.7942

12.26



(b) We will test whether p_1, the proportion of strokes among those who take only aspirin is different from p_2, the proportion of strokes among those to take aspirin + dipyridamole. We were told the experiment was randomized, it is reasonable to assume that there are more than 32990 people who have had strokes and take these medications, and since
$n_1\hat{p}_1 = 206,\ n_1\hat{q}_1 = 1443,\ n_2\hat{p}_2 = 157,\ n_2\hat{q}_2 = 1493,\ \text{all} > 5$, we proceed.

Under the null hypothesis, our pooled estimate of $\hat{p} = \frac{206+157}{1649+1650}$.

$$P(\hat{p}_1 - \hat{p}_2 < -0.0297 \text{ or } \hat{p}_1 - \hat{p}_2 > 0.0297) = 2 \cdot \text{normalcdf}\,(-1\text{E}99, -0.0297, 0, \sqrt{\hat{p}\hat{q}(1/n_1 + 1/n_2)}) = 0.006$$

This very low *p*-value gives strong statistical evidence against H_0. That is we have strong evidence to support the claim that the proportion of strokes among those who take only aspirin is different from the proportion of strokes among those who take aspirin and dipyridamole.

(c) We will test whether p_1, the proportion of deaths among those who take only aspirin is different from p_2, the proportion of deaths among those to take aspirin + dipyridamole. We were told the experiment was randomized, it is reasonable to assume that there are more than 32990 people who have had strokes and take these medications, and since $n_1\hat{p}_1 = 182,\ n_1\hat{q}_1 = 1467,\ n_2\hat{p}_2 = 185,\ n_2\hat{q}_2 = 1465$, all > 5, we proceed.

Under the null hypothesis, our pooled estimate of $\hat{p} = \frac{182+185}{1649+1650}$.

$$P(\hat{p}_1 - \hat{p}_2 < -0.00175 \text{ or } \hat{p}_1 - \hat{p}_2 > 0.00175) = 2 \cdot \text{normalcdf}\left(-1\text{E}99, -0.00175, 0, \sqrt{\hat{p}\,\hat{q}\,(1/n_1 + 1/n_2)}\right) = 0.8729$$

This very high *p*-value does not give strong statistical evidence against H_0. That is we do not have strong evidence to support the claim that the proportion of deaths among those who take only aspirin is different from the proportion of deaths among those who take aspirin and dipyridamole.

12.30

(a) We will test whether p_1, the proportion of urban/suburban students who succeed in the chemical engineering course is different from p_2, the proportion of rural students who succeed.

$H_0 : p_1 = p_2$

$H_A : p_1 \neq p_2$

- We are not told that our sample was random, and if it was not, we cannot trust the results of this test.
- It is reasonable to assume that 65<0.1(All urban/suburban students) and that 52<0.1(All rural students).
- 52 >5, 13>5, 30>5, and 25>5, so all counts of successes and failures are greater than 5.

The TI-84 2-PropZTest: $z = 2.986$, *p*-value = 0.0028. This low p-value gives strong evidence against the null hypothesis.

There is strong statistical evidence that the proportion of urban/suburban students who succeed in this chemical engineering course is different from the proportion of rural students who succeed.

(b) The 90% confidence interval for $p_1 - p_2$ is (0.11723, 0.39186). We are 90% confident that the difference between the proportion of successful urban/suburban students and the proportion of successful rural students is in this interval.

12.31

We will test whether there is a difference between p_1, the proportion of women who succeed in the course, and p_2 the proportion of men who succeed. That is, we test

$H_0 : p_1 = p_2$ vs. $H_A : p_1 \neq p_2$

The TI-84 2-PropZTest gives $z = 0.0244$, *p*-value = 0.98.

This very high *p*-value indicates no statistical evidence against H_0. That is, there is no evidence to support the claim that there is a difference between the proportion of women and the proportion of men who succeed in the course.

12.32

(a) We will test whether p_1 the proportion of high-tech companies that offer stock options, is higher than p_2, the proportion of non-high-tech companies that offer stock options. That is we test the hypotheses...

$$H_0: p_1 = p_2$$
$$H_A: p_1 > p_2$$

We are told to treat the samples as SRS. It is reasonable to assume that there are more than 2000 companies out there. We note that
$n_1 \hat{p}_1 = 73,\ n_1 \hat{q}_1 = 18,\ n_2 \hat{p}_2 = 75,\ n_2 \hat{q}_2 = 34$ are all greater than 5, so we proceed.

The TI-84 2-PropZTest gives $z = 1.832$, *p*-value = 0.033.

This low p-value, indicates strong statistical evidence against H_0. That is, there is strong evidence to support the claim that the proportion of high-tech companies that offer stock options is higher than the proportion of non-high-tech companies that offer stock options.

The 95% confidence interval for the difference in the proportions of the two types of companies that offer stock options (high – non) is (-.0053, 0.23355). We are 95% confident that the difference in these proportions is in this interval.

12.35

(a) We will construct a 95% confidence interval for *p*, the proportion of vehicles going faster than 65 mph when no radar is present.

We will assume that 12931 is < 1/10 (Number of all cars). We will consider the sample an SRS, although in fact it is not. Since $n\hat{p} = 5690$ and $n\hat{q} = 7241 \gg 10$ we will proceed.

The TI-84 1-PropZInt = (0.43147, .44858). We are 95% confident that the proportion of vehicles traveling over 65 mph when no radar is present is in this interval.

(b) We will construct a 95% confidence interval for the difference between p_1, the proportion of vehicles going over 65 mph without radar, and p_2, the proportion of those going over 65 mph with radar. Note that $n_2 \hat{p}_2 = 1051,\ n_2 \hat{q}_2 = 2207$ are both greater than 5. So we proceed.

The TI-84 2-PropZInt = (.10199, .13819). We are 95% confident that the difference in the proportions described above is in this interval.

(c) We require the members (trials) of our sample to be **independent** of one another so that our probability calculations (which require independence) will be valid. The clusters described here makes the speeds of the cars dependent on the cars around them.

12.37

(a) We will construct a 95% confidence interval for *p*, the proportion of businesses of this type that fail within three years.

We were told the sample was an SRS and we will assume that there are more than 1480 businesses of this type in Indiana. Also $n\hat{p}=22,\ n\hat{q}=126$ are bigger than 10.

The TI-84 1-PropZInt = (0.09134, 0.20596). We are 95% confident that the proportion of businesses of this type that fail after three years is in this interval.

(b) We require

$$1.9599\sqrt{\frac{(.14865)(.85135)}{n}}\le 0.04$$
$$n\ge 303.823$$

So we would need a sample size of n = 304.

(c) There are two problems with this study. The first is the high non-response rate. The second is that the sample was only taken in Indiana. The findings of this study might apply to the population of a businesses of this type in Indiana who take the time to respond to surveys.

12.38

We will test whether there is a difference between p_1, the proportion of food and drink businesses in central Indiana headed by men that fail, and p_2, the proportion of these businesses headed by women that fail.

H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 \neq 0$

Assuming all conditions held from the previous exercise and noting that

$$n_1\hat{p}_1=15>5$$
$$n_1\hat{q}_1=91>5$$
$$n_2\hat{p}_2=7>5$$
$$n_2\hat{q}_q=35>5$$

we proceed with our 2-prop z test: z = -0.387, p-val = 0.698.

This very high p-value indicates that we do not have significant evidence against H_0. That is, we do not have statistically significant evidence to support the claim that there is a difference in the proportion of food and drink businesses headed by men and the proportion of those headed by women that fail.

12.42

Since the station did not obtain an SRS, any inference procedures or conclusions drawn from them are completely invalid.

12.39

(a) The TI-84 2-PropZTest gives a p-value of 0.698.

(b) The sample proportions are still $\hat{p}_1=0.1415$ and $\hat{p}_2=0.1666$, the same as in the previous study. But the TI-84 2-PropZTest gives a p-value of 0.033, which is significant at the 5% level.

(c) Larger sample decrease the margin of error in our estimates.

12.42
The conclusion is NOT justified because conditions necessary for carrying out the inferencr procedure are violated. Namely, the sample is not an SRS.

12.45