Stat 1001 Spring 2009 — Some helpful ideas for Chapters 19–21

1. In a random sample of $n = 100$ adults who regularly drive the Twin Cities freeways, 45 report that they feel safer with the ramp meters turned on.

   The **population** is the collection of all people who regularly drive Twin Cities freeways. We would like to know about their views, and so we take a random sample from them. We let p be the **unknown** fraction of this population who feel safer with ramp meters on. The quantity p is a **parameter**, a quantity that we are willing to believe is fixed that is a characteristic of the population whose value we (typically) do not know, and will collect data to learn about.

   In the sample, we get a **sample estimate** $\hat{p} = 45/100 = .45$ or 45%. The sample estimate will be different every time we draw a sample, but it should be "close" to p. Just how close depends on the standard error (the book calls this a standard deviation, but I think the term standard error is clearer):

   $$\text{The standard error of } \hat{p} \text{ is estimated to be } \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{.45 \times .55}{100}} = 0.05$$

   95% of the time, the sample estimate $\hat{p}$ will be within $2 \times$ SD of the true value p. We call the interval

   $$\hat{p} - 2 \times \text{SD to } \hat{p} + 2 \times \text{SD} = 0.35 \text{ to } 0.55 \text{ or } 35\% \text{ to } 55\%$$

   a **95% confidence interval for** p. *If we repeat this process over and over, 95% of the time the true value of the parameter* p *will be included in the interval*. To change the confidence level, we need only change the multiplier from 2 to some other number. For example, from the normal table, 68% of the probability is in the interval ± 1 SD, and so $\hat{p} -$ SD to $\hat{p} +$ SD is a 68% confidence interval. Check in a normal table to verify that the multiplier 1.64 will give a 90% confidence interval.

2. If we obtained a sample of 100 by obtaining volunteers through advertisement in a newspaper and asked the same question as in #1, the sample proportion is not an estimate of the population value because this is not a random sample from the population.

3. You travel to work on the bus each day. For $n = 30$ days, your average travel time was 18 minutes with a SD of 4 minutes. What you you say about the **long-term average** travel time? (The words *mean* and *average* are synonyms.)

   The long-term average is also a parameter, as it is the mean travel time you would expect if you continued to take the same bus route essentially forever. What you do observe is a sample mean of 18 minutes, and you expect the population mean to be "close" to 18 minutes. How close? To answer this, compute the standard deviation of the sample mean, which is called its **standard error**

   $$\text{SEM} = \text{standard error of the mean} = \frac{\text{SD}}{\sqrt{n}} = \frac{4}{\sqrt{30}} = 0.7 \text{ minutes}$$

   Use the fact that if $n > 30$ or so, the **sample mean is approximately normally distributed** and so the interval

   $$(\text{Sample mean} - 2\text{SE}) \text{ to } (\text{Sample mean} + 2\text{SE}) = (18 - 1.4) \text{ to } (18 + 1.4) = 16.6 \text{ to } 19.4$$

   is a 95% confidence interval for the long-run average. If you repeat this process over and over, 95% of such intervals will include the true value.

4. If you collect more data, the confidence interval for the population mean or proportion (or any other quantity) will get narrower and narrower because of the "divide by the square root of the sample size" part of the standard error.

5. A common mistake is contained in the following statement: *Since the 95% confidence interval for the mean travel time is from 16.6 minutes to 19.4 minutes, then 95% of the time I can expect the actual travel time to be in this interval.* The confidence interval is for the long-term average, not for any particular travel-time. If travel times are normally distributed, then about 95% of the time the travel time will be within 2 SD of the population mean. We don't know the population mean or the population variance, but estimates of these are given ni #3. Thus a reasonable estimate would be that any individual travel time will be in the interval from $18 - 2 \times 4 = 10$ minutes to $18 + 2 \times 4 = 26$ minutes. Because we used estimates and not true values this interval is a bit too narrow and the actual coverage is somewhat less than 95%.

6. In a large study in in Utah of mothers with adequate prenatal care, the relative risk of a low birthweight baby for mothers between the ages of 13–17 to mothers between ages 20-24 was 1.4, and 95% confidence interval 1.3 to 1.5. While we have **not** learned how this interval is computed, we **do** know what it means: 95% of the time we compute such intervals, the true relative risk (in the population, not the sample) will be in this interval. Do you remember the definition of relative risk?

7. Most of the time we want to make **comparisons**: is the treatment different from the control; are roads safer with ramp meters off than on; and so on.

8. Continuing with #3 above, suppose during the time the ramp meters were off, your travel time had a sample average of 23 minutes and a SD of 8 minutes, for $n_2 = 40$ days. What can you say about the population difference in travel times between the two conditions?

   The populations of interest are the long-run travel time if ramp meters are kept off, and the long-run travel time if the ramp meters are on. The sample difference is $18 - 23 = -5$ minutes, or the sample difference is 5 minutes in favor of ramp meters on. This is the sample difference and we expect the population difference to be similar. Exactly how similar depends on the standard deviations, and the standard error of the difference has the following formula. Let the subscript "$_1$" refer to the first sample and "$_2$" refer to the second sample.

$$\text{Standard error of the difference} = \sqrt{\frac{\text{SD}_1^2}{n_1} + \frac{\text{SD}_2^2}{n_2}} = \sqrt{\frac{4^2}{30} + \frac{8^2}{40}} = 1.46 \approx 1.5$$

   A 95% confidence interval for the difference is $-5 - 2 \times 1.5$ to $-5 + 2 \times 1.5$ or from -7 minutes to -3 minutes. Since zero is not included in the interval, it is likely that turning of the ramp meters will slow our bus ride.

9. In a survey of 583 Twin Cities freeway drivers about the ramp meter experiment, 16% thought the freeways were extremely congested at the morning commute, and 25% thought that freeways were extremely congested at the afternoon commute. While in the **sample** it is true that more people thought the freeways were more congested in the afternoon, do we have enough information to say this in the population as well? To answer this, we assume that the answers to the two questions were independent, meaning that the way a driver answered the first question did not influence how they answered the second question.

   Let the subscript "$_1$" refer to the first sample and "$_2$" refer to the second sample. The sample difference is $\hat{p}_1 - \hat{p}_2 = .16 - .25 = -0.09$. The standard error of the difference is

$$\text{Standard error of the difference} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{.16 \times .84}{583} + \frac{.25 \times .75}{583}} = 0.02$$

   As before the 95% confidence interval for the true difference in rates is from $-0.09 - 2 \times .02$ to $0.09 + 2 \times .02$ or -0.13 to -0.05. 95% of such intervals will cover the true difference in proportions. Since the value of 0 (no difference) is well outside the interval, we believe that the difference in the population is not zero, and people view the afternoon as more congested.

10. O-rings are used to connect different parts of the space shuttle together, and failure of the O-rings can cause catastrophe. Prior to the ill-fated flight of the *Challenger*, 18 tests of O-rings were done with are temperature less than 60°F, and of these 4 failed (failure rate was 22%). At higher temperatures, 120 were tested, and 5 failed (rate = 4.1%). On the day of the launch of the *Challenger*, the air termperature was slightly above freezing. Should the launch have taken place?

    The confidence interval for the difference in failure rates is computed according to our rules to be from .16 to .18, suggesting that the failure rate is much higher at low temperature. A more meaningful statistic is the relative risk, which is $.22/.041 = 5.5$, with 95% confidence interval (we don't know how to compute this) from 3.17 to 13.61, meaning that we are 95% confident that the chance of failure of an O-ring as between 3.17 and 13.61 times as high at low temperature than at high temperature. If the leaders of the *Challenger* mission had done this calculation, do you think they would have launched on a very cold day?