Chapter 23 Summary
Inferences about Means

What have we learned?

- Statistical inference for means relies on the same concepts as for proportions—only the mechanics and the model have changed.
- The reasoning of inference, the need to verify that the appropriate assumptions are met, and the proper interpretation of confidence intervals and P-values all remain the same regardless of whether we are investigating means or proportions.

Getting Started

- Now that we know how to create confidence intervals and test hypotheses about proportions, it'd be nice to be able to do the same for means.
- Just as we did before, we will base both our confidence interval and our hypothesis test on the sampling distribution model.
- The Central Limit Theorem told us that the sampling distribution model for means is Normal with mean μ and standard deviation $SD(\bar{y}) = \frac{\sigma}{\sqrt{n}}$
- All we need is a random sample of quantitative data.
- And the true population standard deviation, σ.
    - Well, that's a problem…
- Proportions have a link between the proportion value and the standard deviation of the sample proportion.
- This is not the case with means—knowing the sample mean tells us nothing about $SD(\bar{y})$
- We'll do the best we can: estimate the population parameter σ with the sample statistic s.
- Our resulting standard error is $SE(\bar{y}) = \frac{s}{\sqrt{n}}$
- We now have extra variation in our standard error from s, the sample standard deviation.
    - We need to allow for the extra variation so that it does not mess up the margin of error and P-value, especially for a small sample.
- And, the *shape* of the sampling model changes—the model is no longer Normal. So, what is the sampling model?

Gosset's *t*

- William S. Gosset, an employee of the Guinness Brewery in Dublin, Ireland, worked long and hard to find out what the sampling model was.
- The sampling model that Gosset found has been known as Student's t.
- The Student's t-models form a whole *family* of related distributions that depend on a parameter known as degrees of freedom.
- We often denote degrees of freedom as df, and the model as t_{df}.

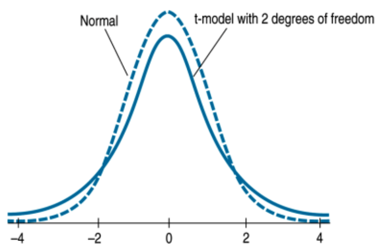
What Does This Mean for Means?

- **A practical sampling distribution model for means**
    - When the conditions are met, the standardized sample mean $t = \frac{\bar{y} - \mu}{SE(\bar{y})}$

      follows a Student's t-model with $n - 1$ degrees of freedom.
- We estimate the standard error with $SE(\bar{y}) = \frac{s}{\sqrt{n}}$

What Does This Mean for Means? (cont.)

- When Gosset corrected the model for the extra uncertainty, the margin of error got bigger.
  - Your confidence intervals will be just a bit wider and your P-values just a bit larger than they were with the Normal model.
- By using the *t*-model, you've compensated for the extra variability in precisely the right way.
- Student's *t*-models are unimodal, symmetric, and bell shaped, just like the Normal.
- But *t*-models with only a few degrees of freedom have much fatter tails than the Normal.

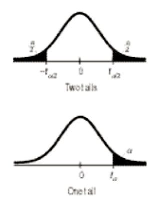


- As the degrees of freedom increase, the *t*-models look more and more like the Normal.
- In fact, the *t*-model with infinite degrees of freedom is exactly Normal.

Finding *t*-Values By Hand

- The Student's *t*-model is different for each value of degrees of freedom.
- Because of this, Statistics books usually have one table of *t*-model critical values for selected confidence levels.
- Alternatively, we could use technology to find *t* critical values for any number of degrees of freedom and any confidence level you need.
- What technology could we use?
  - The Appendix of *ActivStats* on the CD
  - Any graphing calculator or statistics program

Two tail probability		0.20	0.10	0.05
One tail probability		0.10	0.05	0.025
Table T	df			
Values of t_α	1	3.078	6.314	12.706
	2	1.886	2.920	4.303
	3	1.638	2.353	3.182
	4	1.533	2.132	2.776
	5	1.476	2.015	2.571
	6	1.440	1.943	2.447
	7	1.415	1.895	2.365
	8	1.397	1.860	2.306
	9	1.383	1.833	2.262
	10	1.372	1.812	2.228
	11	1.363	1.796	2.201
	12	1.356	1.782	2.179
	13	1.350	1.771	2.160
	14	1.345	1.761	2.145
	15	1.341	1.753	2.131
	16	1.337	1.746	2.120
	17	1.333	1.740	2.110
	18	1.330	1.734	2.101
	19	1.328	1.729	2.093



Part of Table T.

Assumptions and Conditions

- Gosset found the *t*-model by simulation.
- Years later, when Sir Ronald A. Fisher showed mathematically that Gosset was right, he needed to make some assumptions to make the proof work.
- We will use these assumptions when working with Student's *t*.
- Independence Assumption:
  - Randomization Condition: The data arise from a random sample or suitably randomized experiment. Randomly sampled data (particularly from an SRS) are ideal.
  - 10% Condition: When a sample is drawn without replacement, the sample should be no more than 10% of the population.

Assumptions and Conditions (cont.)

- Normal Population Assumption:
    - We can never be certain that the data are from a population that follows a Normal model, but we can check the
    - Nearly Normal Condition: The data come from a distribution that is unimodal and symmetric.
        - Check this condition by making a histogram or Normal probability plot.
- Nearly Normal Condition:
    - The smaller the sample size (n < 15 or so), the more closely the data should follow a Normal model.
    - For moderate sample sizes (n between 15 and 40 or so), the t works well as long as the data are unimodal and reasonably symmetric.
    - For larger sample sizes, the t methods are safe to use even if the data are skewed.

One-Sample t-Interval

- When the conditions are met, we are ready to find the confidence interval for the population mean, μ.
- The confidence interval is $\bar{y} \pm t^*_{n-1} \times SE(\bar{y})$
  where the standard error of the mean is $SE(\bar{y}) = \frac{s}{\sqrt{n}}$
- The critical value t^*_{n-1} depends on the particular confidence level, C, that you specify and on the number of degrees of freedom, $n - 1$, which we get from the sample size.

More Cautions About Interpreting Confidence Intervals

- Remember that interpretation of your confidence interval is key.
- What NOT to say:
    - "*90% of all the vehicles* on Triphammer Road drive at a speed between 29.5 and 32.5 mph."
        - The confidence interval is about the *mean* not the individual values.
    - "We are 90% confident that *a randomly selected vehicle* will have a speed between 29.5 and 32.5 mph."
        - Again, the confidence interval is about the *mean* not the individual values.
- What NOT to say:
    - "The mean speed of the vehicles is 31.0 mph *90% of the time.*"
        - The true mean does not vary—it's the confidence interval that would be different had we gotten a different sample.
    - "*90% of all samples* will have mean speeds between 29.5 and 32.5 mph."
        - The interval we calculate does not set a standard for every other interval—it is no more (or less) likely to be correct than any other interval.

Make a Picture, Make a Picture, Make a Picture

- Pictures tell us far more about our data set than a list of the data ever could.
- The only reasonable way to check the Nearly Normal Condition is with graphs of the data.
    - Make a histogram of the data and verify that its distribution is unimodal and symmetric with no outliers.
    - You may also want to make a Normal probability plot to see that it's reasonably straight.

One-Sample *t*-test for the Mean

- The conditions for the one-sample *t*-test for the mean are the same as for the one-sample *t*-interval.
- We test the hypothesis H_0: $\mu = \mu_0$ using the statistic $t_{n-1} = \dfrac{\bar{y} - \mu_0}{SE(\bar{y})}$
- The standard error of the sample mean is $SE(\bar{y}) = \dfrac{s}{\sqrt{n}}$
- When the conditions are met and the null hypothesis is true, this statistic follows a Student's *t* model with $n - 1$ *df*. We use that model to obtain a P-value.

Significance and Importance

- Remember that "statistically significant" does not mean "actually important" or "meaningful."
  - Because of this, it's always a good idea when we test a hypothesis to check the confidence interval and think about likely values for the mean.

Intervals and Tests

- Confidence intervals and hypothesis tests are built from the same calculations.
  - In fact, they are complementary ways of looking at the same question.
  - The confidence interval contains all the null hypothesis values we can't reject with these data.
- More precisely, a level *C* confidence interval contains *all* of the possible null hypothesis values that would not be rejected by a two-sided hypothesis text at alpha level $1 - C$.
  - So a 95% confidence interval matches a 0.05 level test for these data.
- Confidence intervals are naturally two-sided, so they match exactly with two-sided hypothesis tests.
  - When the hypothesis is one sided, the corresponding alpha level is $(1 - C)/2$.

Sample Size

- To find the sample size needed for a particular confidence level with a particular margin of error (*ME*), solve this equation for *n*: $ME = t^*_{n-1} \dfrac{s}{\sqrt{n}}$
- The problem with using the equation above is that we don't know most of the values. We can overcome this:
  - We can use *s* from a small pilot study.
  - We can use z^* in place of the necessary *t* value.
- Sample size calculations are *never* exact.
  - The margin of error you find *after* collecting the data won't match exactly the one you used to find *n*.
- The sample size formula depends on quantities you won't have until you collect the data, but using it is an important first step.
- Before you collect data, it's always a good idea to know whether the sample size is large enough to give you a good chance of being able to tell you what you want to know.

*The Sign Test—Back to Yes and No

- We could turn our quantitative data into a set of yes/no values (Bernoulli trials).
- We can test a median by counting the number of values above and below that value—this is called a sign test.
    - The sign test is a *distribution-free* method, since there are no *distributional* assumptions or conditions on the data.
        - Because we no longer have quantitative data, we don't require the Nearly Normal Condition.

What Can Go Wrong?

- Ways to Not Be Normal:
    - Beware multimodality.
        - The Nearly Normal Condition clearly fails if a histogram of the data has two or more modes.
    - Beware skewed data.
        - If the data are very skewed, try re-expressing the variable.
    - Set outliers aside—but remember to report on these outliers individually.

…And of Course:

- Watch out for bias—we can never overcome the problems of a biased sample.
- Make sure data are independent.
    - Check for random sampling and the 10% Condition.
- Make sure that data are from an appropriately randomized sample.
- Interpret your confidence interval correctly.
    - Many statements that sound tempting are, in fact, misinterpretations of a confidence interval for a mean.
    - A confidence interval is about the mean of the population, not about the means of samples, individuals in samples, or individuals in the population.