AP Statistics – Chapter 8 Notes §8.1 Confidence Intervals: The Basics

A _________________________________ is a statistic that provides an estimate of a population parameter. The value of that statistic from a sample is called a _________________________. Ideally a ___________________________ is our "best guess" at the value of an unknown parameter.

Example:
In each of the following settings, determine the point estimator you would use and calculate the value of the point estimate.
a) The makers of a new golf ball want to estimate the median distance the new balls will travel when hit by mechanical driver. They select a random sample of 10 balls and measure the distance each ball travels after being hit by the mechanical driver.
Here are the distances (in yards): 285 286 284 285 282 284 287 290 288 285

(b) The golf ball manufacturer would also like to investigate the variability of the distance travelled by the golf balls by estimating the interquartile range.

(c) The math department wants to know what proportion of its students own a graphing calculator, so they take a random sample of 100 students and find that 28 own a graphing calculator.

| Confidence Interval – Margin of Error – Confidence Level |
| --- |
| A __________________________________ for a parameter has two parts:<br>• An interval calculated from the data, which has the form: estimate ± margin of error<br>• Margin of error tells how close the estimate tends to be to the unknown parameter in repeated random sampling.<br>• A confidence level C, which gives the overall success rate of the method for calculating the confidence interval. That is, in C% of all possible samples, the method would yield an interval that captures the true parameter value. |

Interpreting Confidence Level and Confidence Interval:

Confidence Level:

95% of all possible samples of a size ( ) from this population will result in an interval that captures the unknown parameter.

Confidence Interval:

We are 95% confident that the interval from ___ to ___ captures the actual value of the (population parameter in context).

Example: Interpret the confidence interval and the confidence level.

According to www.gallup.com, on August 13, 2010, the 95% confidence interval for the true proportion of Americans who approved of the job Barack Obama was doing as president was $0.44 \pm 0.03$.

Interpret the confidence interval and the confidence level.

CYU: Pg.476

| Calculating a Confidence Interval |
|---|
| statistic ± (critical value)(standard deviation of statistic)<br><br>statistic = point estimator of the parameter<br>(critical value)(standard deviation of statistic) = margin of error<br>critical value = depends on both the confidence level and the sampling distribution of the statistic<br><br>Conditions:<br>RANDOM: The data come from a well-designed random sample or randomized experiment.<br>NORMAL: The sampling distribution of the statistic is approximately Normal.<br>INDEPENDENT: Individual Observations are independent. When sampling without replacement, the sample size n should be no more than 10% of the population size N (the 10% condition) to use our formula for the standard deviation statistic. |

About Confidence Intervals:

- Standard deviation of the statistic depends on the sample size n: larger samples give more precise estimates, which means less variability in the statistic.
- Smaller margin of error = lower confidence
- Increasing sample size reduces the margin of error.
- Our method of calculation assumes that the data come from an SRS of size n from the population of interest.
- The margin of error in a confidence interval covers only chance variation due to random sampling or random assignment.

Constructing a Confidence Interval for p:

Use the following formula: statistic ± (critical value)(standard deviation of statistic)

- ✓ Check conditions: RANDOM, NORMAL, INDEPENDENT
- ✓ Find the critical value based on the desired confidence interval.
- ✓ Calculate the confidence interval

  **when the standard deviation of a statistic is estimated from data, the result is called the standard error of the statistic.

See Example "The Beads" (Pg. 486)
CYU: Pg.487

See Example "80% Confidence" (Pg. 488)
You Try: Find the critical value z* for a 96% confidence interval. Assume that the Normal condition is met. (Show work)

See Example "The Beads" (Pg.489)
CYU: Pg. 490

Example: Kissing the Right Way?

According to an article in the *San Gabriel Valley Tribune* (2-13-03), "Most people are kissing the 'right way'." That is, according to the study, the majority of couples tilt their heads to the right when kissing. In the study, a researcher observed a random sample 124 couples kissing in various public places and found that 83/124 (66.9%) of the couples tilted to the right. Construct and interpret a 95% confidence interval for the proportion of all couples who tilt their heads to the right when kissing.

Choosing Sample Size:

Formula:

Example:

Suppose that you wanted to estimate the $p$ = the true proportion of students at your school that have a tattoo with 95% confidence and a margin of error of no more than 0.10.

Determine how many students should be surveyed to estimate $p$ within 0.10 with 95% confidence.

CYU: Pg.494

AP Statistics – Chapter 8 Notes §8.3 Estimating a Population Mean

| Interval for Population Mean | |
|---|---|
| When population standard deviation is known | When population standard deviation is not known |
| Choosing sample size:<br>(It is the size of the sample that determines the margin of error) | t-distribution and degrees of freedom<br>(pg. 505) |
| | |

Conditions for Inference about a Population Mean: RANDOM, NORMAL, INDEPENDENT

CYU: Pg. 501

The ________________________________ of the sample mean ___ is , where $s_x$ is the sample standard deviation. It describes how far _____ will be from _____, on average, in repeated SRSs of size n

Example: Video Screen Tension
A manufacturer of a high-resolution video terminals must control the tension on the mesh of fine wires that lies behind the surface of the viewing screen. Too much tension will tear the mesh, and too little will allow wrinkles. The tension is measured by an electrical device with output readings in millivolts (mV). Some variation is inherent in the production process. Here are the tension readings from a random sample of 20 screens from a single day's production:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 269.5 | 297.0 | 269.6 | 283.3 | 304.8 | 280.4 | 233.5 | 257.4 | 317.5 | 327.4 |
| 264.7 | 307.7 | 310.0 | 343.3 | 328.1 | 342.6 | 338.8 | 340.1 | 374.6 | 336.1 |

Construct and interpret a 90% confidence interval for the mean tension μ of all the video terminals produced on this day.

**INFERENCE FOR PROPORTIONS USES Z
INFERENCE FOR MEANS USES T
When the actual df does not appear in table B, use the greatest df available that is less than your desired df.

CYU: Pg.511

An inference procedure is called _______________ if the probability calculations involved in that procedure remain fairly accurate when a condition for using the procedure is violated.
(t procedures are quite robust against non-Normality of the population except when outliers or strong skewness are present)

*The Normal Condition:

Sample size less than 15: Use t procedures if the data appear close to Normal. Graph it and check for roughly symmetric, single peak, no outliers. If the data are clearly skewed or if outliers are present, do not use t.

Sample size at least 15: The t procedures can be used except in the presence of outliers or strong skewness.

Large sample size: The t procedures can be used even for clearly skewed distributions when the _________ is large, roughly $n \geq 30$.

Example: How much homework?
The principal at a large high school claims that students spend at least 10 hours per week doing homework on average. To investigate this claim, an AP Statistics class selected a random sample of 250 students from their school and asked them how many long they spent doing homework during the last week. The sample mean was 10.2 hours and the sample standard deviation was 4.2 hours.
Problem:
(a) Construct and interpret a 95% confidence interval for the mean time spent doing homework in the last week for students at this school.
(b) Based on your interval in part (a), what can you conclude about the principal's claim?