**8.3 Estimating a Population Mean**

When estimating the Population Mean ($μ)$, there are two different possibilities:

1. **When** $σ$ **is known - One Sample z-Interval for a Population Mean:**



**Conditions for Interference About A Population Mean:**

1. Random
2. Normal – sample size is greater than 30
3. Independent – check 10% condition if necessary

\*We rarely, if ever, use z procedures to calculate a confidence interval for a population mean. Why?

**Choosing Sample Size for a Desired Margin of Error When Estimating :**

1. **When** **is Unknown - The *t* distribution**

Draw an SRS of size *n* from a large population that has a Normal distribution with mean  and standard deviation . The statistic is given by the following formula. This is similar to the z-statistic.

$$t= \frac{\overbar{x}-μ}{\frac{s\_{x}}{\sqrt{n}}}$$

* Has the t distribution with degrees of freedom df = n – 1. This statistic will have approximately a  distribution as long as the sampling distribution of  is close to Normal.
* Using s to estimate  introduces additional variability, resulting in a statistic whose distribution is more spread out than in the z distribution.



Normal distribution is in blue and t distribution is in red

What is a *t* distribution? Describe the shape, center, and spread of the *t* distributions.

* *Shape:*
* *Center:*
* *Spread:*

**One-Sample t-Interval for a Population Mean:**

 

**Conditions for Using One-Sample t Procedures:**

 1. Random

1. Normal-
	1. *Sample size less than 15:* Use t procedures if the data appear close to Normal. If the data are clearly skewed or if outliers are present, do not use t. (Use a normal probability plot to check Normality)
	2. *Sample size at least 15:* The t procedures can be used except in the presence of outliers or strong skewness
	3. *Large samples:* The t procedures can be used even for clearly skewed distributions when the sample size is greater than 30
2. Independent – check 10% condition if necessary

For most purposes, you can safely use the one-sample t-procedures when n $\geq $ 15 unless an outlier or strong skewness is present. Except, in the case of small samples, the condition that the data come from a random sample or randomized experiment is more important than the Normal condition.

**Robust Procedures –** An inference procedure is called robust if the probability calculations involved in that procedure remain fairly accurate when a condition for using the procedure is violated

* The results of one-sample t procedures are exactly correct when the population is Normal. Real populations are never exactly Normal.
* The usefulness of the t procedures in practice therefore depend on how strongly they are affected by non-normality.
* Fortunately, the t procedures are quite robust against nonnormality of the population except in the case of outliers or strong skewness.

**Example**

Below are the graphs of a standard Normal distribution and a *t*-distribution with 3 degrees of freedom.



(a) Indicated which graph is which and explain how you know.

Dotted graph =

Solid graph =

(b) On the same figure sketch a graph of a *t*-distribution with 1 degree of freedom.

**Example**

Find the critical *t*\* value for each of the following confidence intervals:

1. 95% confidence interval with 8 degrees of freedom.

(b) 80% confidence interval when *n* = 20

**Example**

National Fuelsaver Corporation manufactures the Platinum Gasaver, a device they claim “may increase gas mileage by 30%.” Here are the percent changes in gas mileage for 15 identical, randomly-selected vehicles, as presented in one of the company’s advertisements:



(a) The sample mean is $\overbar{x}$ = 29.43 and the sample standard deviation is *s =* 16.23. Calculate and interpret the standard error of the mean for these data.

(b) Construct and interpret a 90% confidence interval to estimate the mean change (in percent) in gas mileage. Does the data support the company’s claim? Use the four-step process.

**Example**

You want to estimate the mean fuel efficiency of Ford Focus automobiles with 99% confidence and a margin of error of no more than 1 mile per gallon. Preliminary data suggests that $σ$ = 2.4 miles per gallon is a reasonable estimate of the standard deviation for all cars of this make and model. How large a sample do you need?