One Sample Hypothesis Testing for Means: Raise What They Raise In Ireland

Ben Babcock
University of Minnesota

Hypothesis Testing for Means: The Idea

We are going to:

1. form null and alternative hypotheses and gather data.

2. calculate a mean difference (effect).

3. calculate a standard error (sd of the sampling distribution)

4. divide the ___________ by the ________________ in order to obtain the test statistic.

5. check the probability of the statistic from #3 occurring.

6. choose to reject or not reject the _________ based on this fact.

The Basic Form of Most Parametric Test Statistics

$$\frac{\textit{Mean Difference} \text{ or } \textit{Effect}}{\textit{Standard Error}}$$

I used a whole slide for this thing. You should probably remember it.

The Research Question

The one-sample mean question often arises when you have previous expectations or a “norm” for comparison.

IO Psychology:
Suppose you are consulting for a factory making gloves for the U. S. Army. You know that the number of flaw-free pairs of gloves an employee can sew in a week is approximately normally distributed with a mean of 800 and a standard deviation of 50 pairs.

Suppose that we want to try to increase productivity by boosting morale. How do we test if the morale-boosting improves or hurts ***mean individual productivity***?

Setting Up the Hypotheses and Alpha

Null:

Alternative (experimental):

Choose a critical alpha level:

The Central Limit Theorem

In order to answer the research question, we need to know the Central Limit Theorem (CLT).

The CLT states that, for almost any distribution:

If a distribution has a mean μ and a standard deviation of σ, the mean of the sampling distribution of the mean for sample size N will be μ with a standard deviation of $\sigma/\sqrt{N}$. As N becomes large, the sampling distribution of the mean will become normal.

The quantity $\sigma/\sqrt{N}$ is called the ______________________.

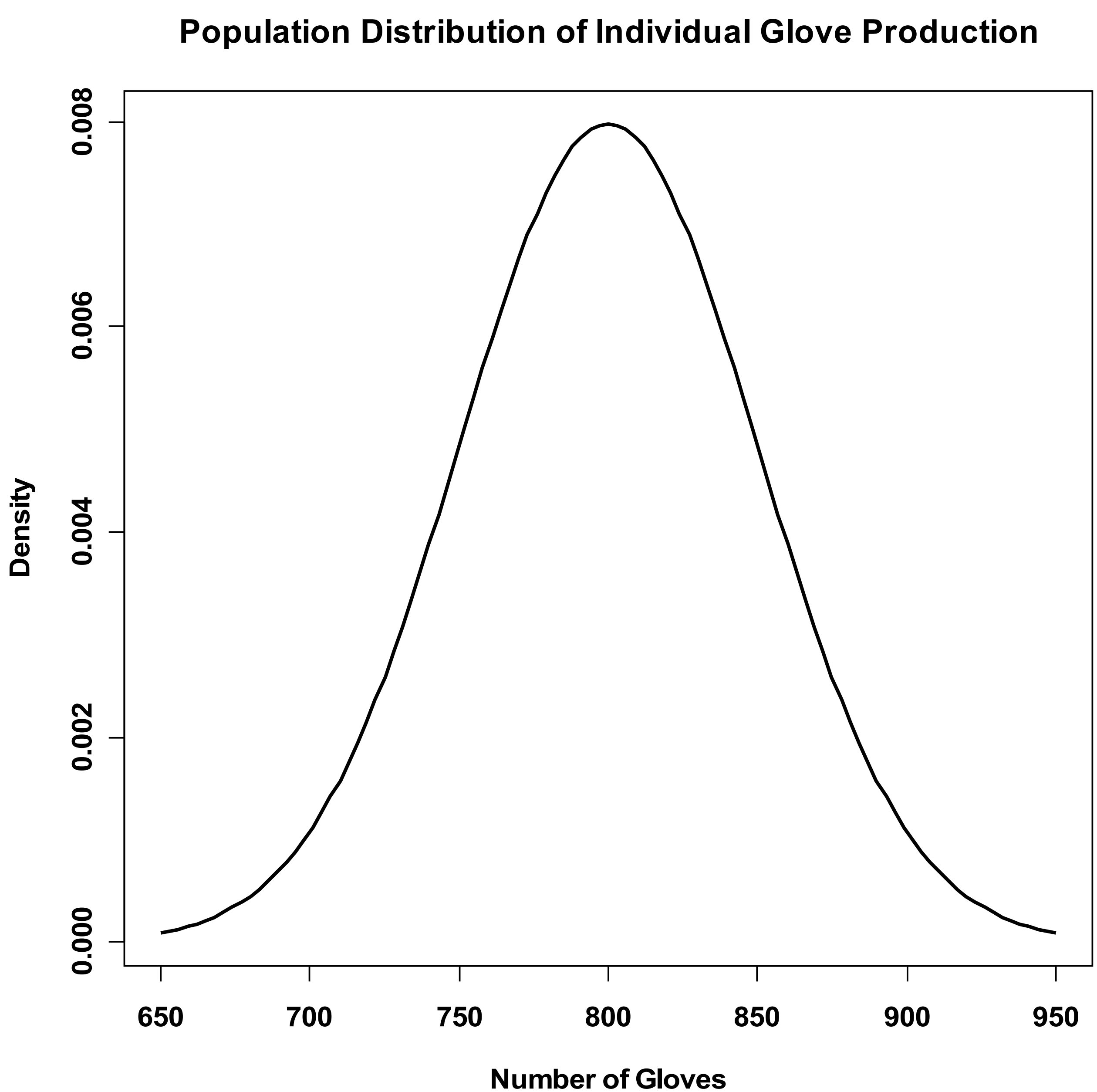
H, 281

Back to the Research Question

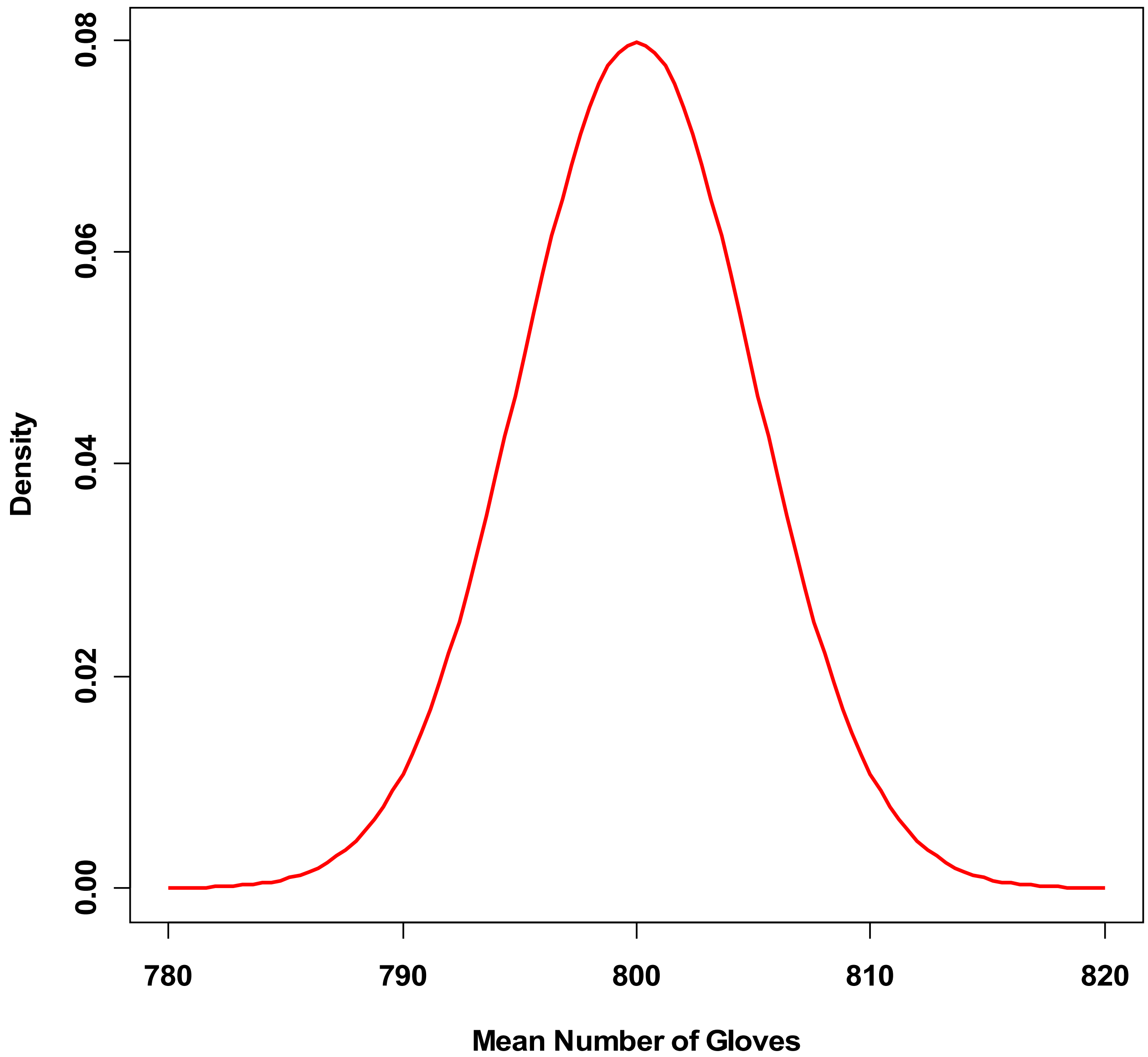
4 Things that we know:

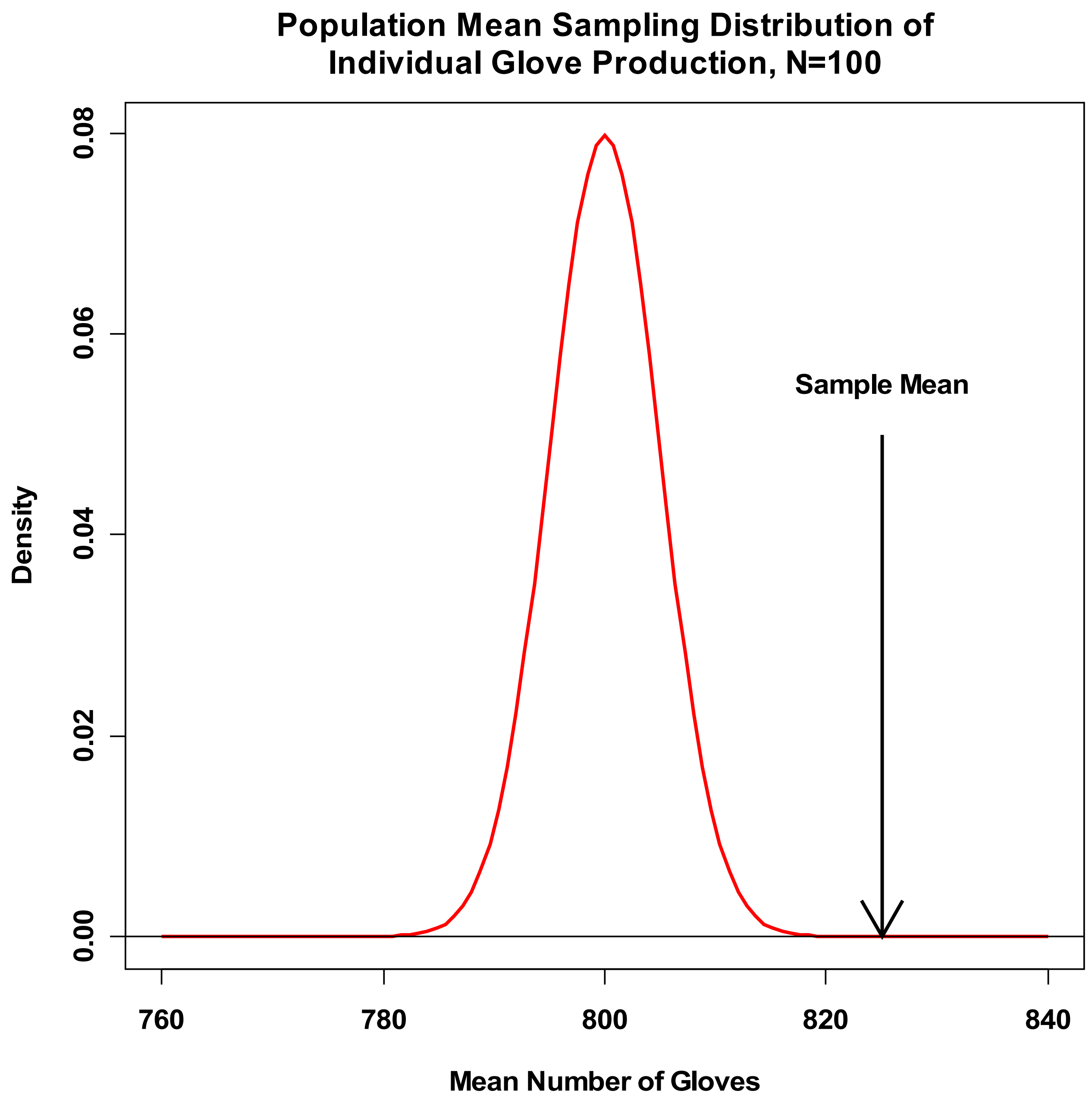
1. μ gloves: 800
2. σ gloves: 50
3. The Central Limit Theorem
4. We used the morale-boosting technique on 100 people. They have a mean of 825 pairs with a standard deviation of 100.

How do we test to see if our data are rare given the null (null=population distribution)?


Population Distribution of Individual Glove Production
Density
0.000
0.002
0.004
0.006
0.008
650
700
750
800
850
900
950
Number of Gloves

Population Mean Sampling Distribution of
Individual Glove Production, N=100




Population Mean Sampling Distribution of
Individual Glove Production, N=100
Density
0.00
0.02
0.04
0.06
0.08
Sample Mean
760
780
800
820
840
Mean Number of Gloves

Back to the Research Question

4 Things that we know:
1. μ gloves: 800
2. σ gloves: 50
3. The Central Limit Theorem
4. We used the morale-boosting technique on 100 people. They have a mean of 825 pairs with a standard deviation of 100.

How do we test to see if our data are rare given the null (population distribution)?

Remember the __-score...

It's basic form was:

The *z*-test

We have a population mean (800). We have a sample mean (825). We have a population standard error ($50/100^{1/2}$). We have everything we need to produce a *z*-score. H, 286

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{N}}$$

where $\bar{X}$ is the sample mean,
μ and σ are the known population mean and standard deviation and N is sample size.
Calculations here:

p-value (area "more extreme"):
Conclusion and interp. here:

One-Sample *t*-test: σ Unknown

When you *do not* know σ, the statistic is nearly the same. We just use the ____________________ in place of σ. We also need something called "degrees of freedom". H, 288 and 546

The one-sample *t* statistic (*t*) is

$$t(df) = \frac{\bar{X} - \mu}{s_X / \sqrt{N}}$$

where s_X is the sample sd and *df* is the degrees of freedom.

What does this equation mean? Remember the basic form:

$$\frac{\textit{Mean difference} \text{ or } \textit{effect}}{\textit{Standard Error}}$$

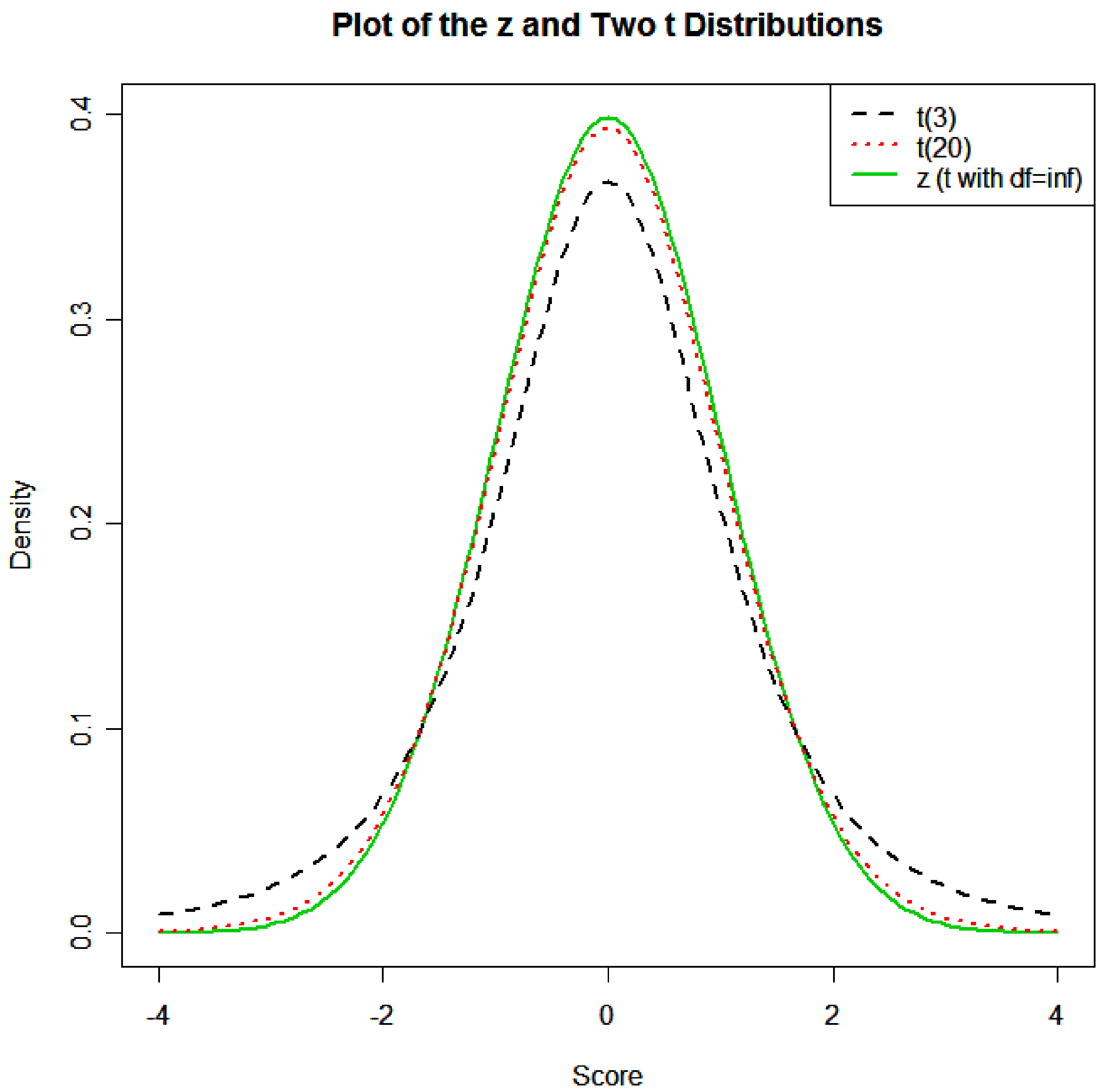
The *t*-distribution: Similar to the *z*

The *t*-distributions are like the *z*-distribution with ____________________. This makes *t*-tests more conservative for sampling error.

William Gossett published the *t* family of distributions while working for Guinness® Brewing Company.

When you are working with the *t*-distribution, you must know your degrees of freedom (df).

The *t*-distribution is not one but ***many*** distributions. They depend on your sample size. The larger your sample size (df), the more the *t*-distribution ____________________________________.


Plot of the z and Two t Distributions
t(3)
t(20)
z (t with df=inf)
0.0
0.1
0.2
0.3
0.4
Density
-4
-2
0
2
4
Score

Degrees of Freedom

The df is an indicator of "independent pieces" of information you have when estimating a parameter.

It is generally the number of people that you have (*N*) minus 1 or 2. For the one-sample *t*, it is _______________.

H, 290

Example of degrees of freedom:
Suppose that we know the mean of the following 5 numbers is 3. The numbers are 2, 3, 4, 5, and *x*. What *must x* be?

We have ____ degrees of freedom in the estimation of the mean.

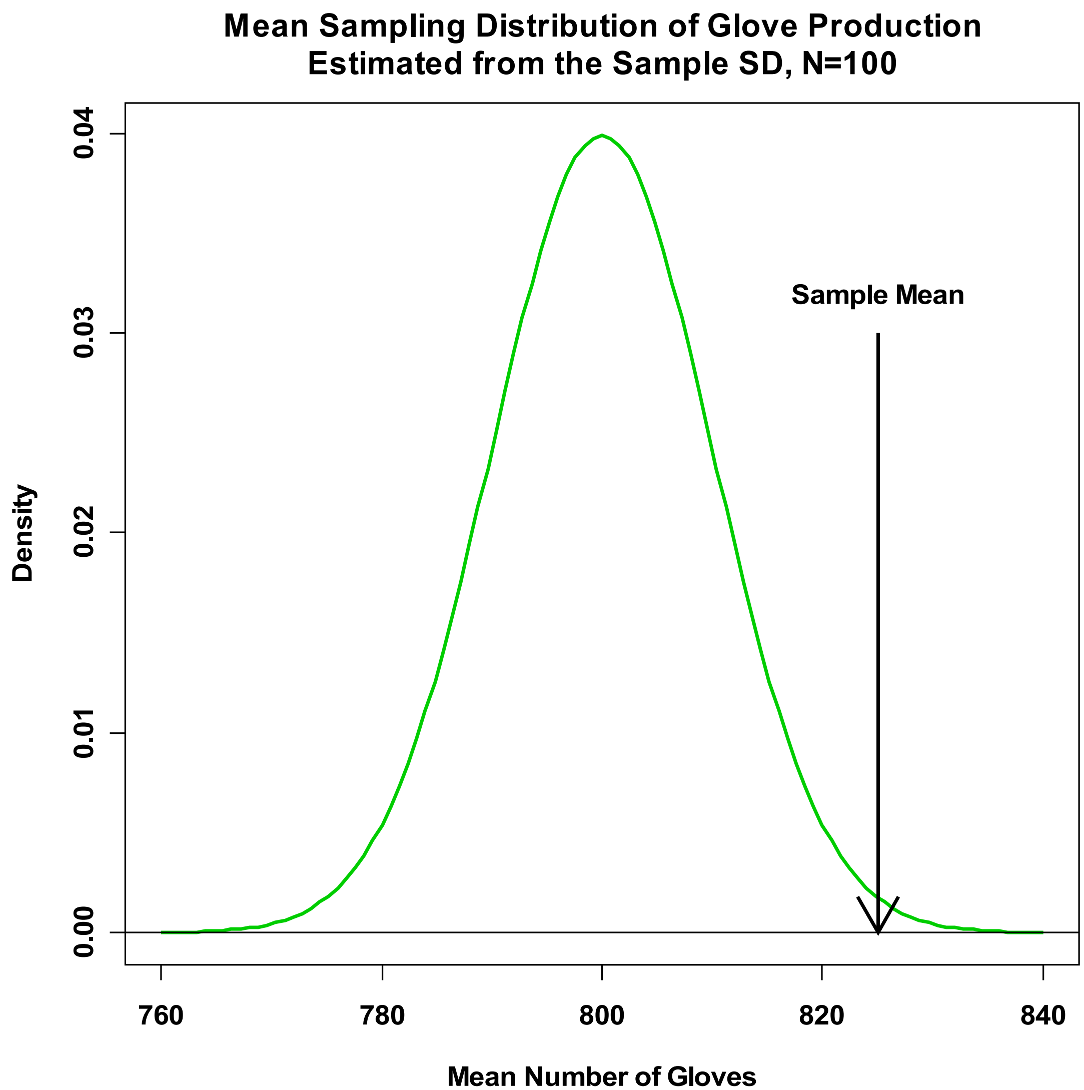
One-Sample *t*-test: σ Unknown

Previous question: You know that the number of pairs of flaw-free pairs of gloves an employee can sew in a week is approximately normally distributed with a mean of 800. **You do not know the population $\boldsymbol{\sigma}$**. You do your morale thing, and the 100 lucky people have a mean of 825 with $s_X = 100$.

$$t(df) = \frac{\bar{X} - \mu}{s_X / \sqrt{N}}$$

Calculations:

Conclusion:



The estimated sampling distribution of the mean is wider because the sample SD was greater than σ (when we knew it).

Effect Size

There are two types of effect sizes in the *t*-test situation: unstandardized and standardized.

Standardized: Divide the mean difference by the standard deviation you use. It's similar to a *z*-score.
H, 295

Unstandardized: Simply the ___________________________.

The unstandardized effect is good if your units of measurement are well known or meaningful in many contexts. What are some examples of well known units?

Confidence Intervals

We need a ***range*** that is a reasonable guess for the value of a parameter. H, 296-297. Howell gives us b.s. on pages 298-299. Basic form of all confidence intervals:

$$C.I. = statistic \pm crit.\ value \times standard\ error$$
$$\approx statistic \pm 2 \times standard\ error$$

95% Confidence Interval around a single mean using a *t*-value.
Sample Mean = 50 S. Dev = 12 N = 145 *df* =
Critical *t*-value = _______

Calculations go here:

Remember to always use the 2-tailed value. Confidence intervals are of VITAL IMPORTANCE to modern statistics!

Applets to Visualize Concepts

Confidence Intervals
http://www.ruf.rice.edu/~lane/stat_sim/conf_interval/index.html

Sampling Distributions of the Mean
http://wise.cgu.edu/sdmmod/sdm_applet.asp

Online *p*-value calculator
http://www.graphpad.com/quickcalcs/PValue1.cfm
or just do a Google search for “p-value calculator”

Critical values from *p*-values (good for confidence intervals)
http://www.graphpad.com/quickcalcs/Statratio1.cfm