The process of *statistical inference* refers to *estimation, hypothesis testing, and prediction*.

Estimation is the process of inferring or estimating a population parameter (e.g.: mean or standard deviation) from the corresponding statistic of a sample drawn from the population.

Hypothesis testing is accomplished by first making an assumption with regard to an unknown population characteristic. We then take a random sample from the population, and on the basis of the corresponding sample characteristic, we either accept or reject the hypothesis with a particular degree of confidence.

Estimation – we can construct a point estimate or an interval estimate

Eg want to estimate true pop mean (mu)

"Parameter" = population value (eg: mean mu, standard deviation sigma)
"Statistic" = corresponding sample value (eg: sample mean x-bar, sample standard deviation s)

- point estimate of mu (true pop mean) is x-bar (sample mean)

- interval estimate would be a confidence interval

?? which is more useful?
?? Which uses more info?

Interval estimates use more info, therefore they will reveal more about the data.

?? How do we build a confidence interval for the pop mean?

--- Depends on what info we have.

Specifically, do we know the true pop standard deviation (sigma) or not?

In most cases, if we don't know the mean, we're not going to know the std dev either

1. Know sigma: 95% CI for mu is: x-bar $\pm$ $Z\alpha/2$ $\bullet$ (sigma/root n)

2. Don't know sigma: 95% CI for μ is: x-bar $\pm$ t (α/2, n-1 degrees of freedom) $\bullet$ (s/root n)
x-bar $\pm$ t (α/2, n-1 degrees of freedom) $\bullet$ ($S_{x\text{-}bar}$)

Ok what's all this stuff?

95% - is the confidence level = the probability that the interval estimate will contain the true population parameter.
Mu = pop mean = what we're trying to estimate
x-bar = sample mean
sigma = pop standard deviation
s = sample standard deviation
n = sample size
$S_{x\text{-bar}}$ = the <u>standard error of the mean.</u>

? What is the standard error of the mean?

→ Any time you calculate a statistic this means we're dealing with a single sample from the population. If we had taken a different sample, we would have gotten a different value for the statistic. So imagine we had taken a bunch of different samples. We would have a bunch of different estimates of the mean.

We could calculate the mean of these means, and similarly we could calculate the standard deviation of these means.

This latter measure (the standard deviation of the sample means) is known as the standard error of the mean.

We interpret this standard error the same as any standard deviation – it indicates approximately how far the observed value of the statistic (the sample mean we just calculated) is from the mean of all the sample means (which the central limit theorem tells us will equal the true population mean).

So the <u>standard error of the mean</u> tells us approximately the standard deviation we would find if we took a bunch of samples, calculated their means, and worked with all those means as a data set.

**** standard error of the mean = $S_{x\text{-bar}}$ = how far the sample average (x-bar) is from the population mean μ.**

* Note the difference between s (the standard deviation) and $S_{x\text{-bar}}$: The standard deviation tells us how far individuals in our sample are from the average, while the standard error of the mean tells us how far the average from our sample is from the true population average.

$S_{x\text{-bar}}$ = S/root n

Z ?
t?
Alpha?
Degrees of freedom?

Z and t are "critical values" in each of their respective distributions (marking points)
Ref: table E in appendix C for Z and table F for t.

Zexamples… 95% confidence: Z = 1.96, 99% confidence: Z = 2.58

These come from the central limit theorem, which states that with large sample sizes, 95% of the sample means will fall within 1.96 standard errors of the true pop mean, or that the interval built1.96 standard errors around a sample mean will contain the true population mean with 95% probability.

- ⇨ mu +- 1.96 (sigma over root n) - will contain 95% of sample means
- ⇨ x-bar +- 1.96 (sigma over root n) - will contain true pop mean with 95% probability.

The number of degrees of freedom in a particular statistic is the number of independent observations in the sample (sample size) minus the number of population parameters we've estimated with sample info. = The number of data points that are free to vary

here, statistic we're using = std dev = s = $\sqrt{\sum_{i=1}^{n} (x_i - \text{x-bar})^2 / (n-1)}$

where x_i = are the individual obs and n = sample size, and x-bar = sample mean

since we've calculated the mean (we know the mean) then only n-1 of the values are free to vary

eg: pick row – I know avg GPA for people in this row so, if I know 5 of the actual GPA's then I automatically know the last using the avg.

n-1 are freely chosen, but since I know the mean I also know the last.

Alpha = (1- the conf level) = the area in both tails of the normal distribution.

How do we read or interpret a 95% CI?

If we took a bunch of samples, 95% of the means would be in this CI - or –

"we are 95% confident that the true mean is in this range"

<u>The confidence interval statement</u> (assuming $\alpha = 0.05$):

"We are 95% sure that the true population mean is somewhere between our estimate of the mean (that is the sample mean) minus *t* standard errors and the estimate plus t standard errors. That is, we are 95% sure that the true population mean μ is somewhere between:

x-bar – $Z_{\alpha/2} \bullet$ (sigma/root n) and x-bar + $Z_{\alpha/2} \bullet$ (sigma/root n)

x-bar – t ($S_{\text{x-bar}}$) and x-bar + t ($S_{\text{x-bar}}$) –or-- **x-bar ± t ($S_{\text{x-bar}}$)**

So, $Z_{\alpha/2} \bullet$ (sigma/root n) and t ($S_{\text{x-bar}}$) will be numbers.

Eg: x-bar = 70

You could construct a CI such that LB = 67 and UB = 73 or 70 ± 3

3 is known as the maximum error of the estimate = E = the maximum difference between the point estimate and the true population parameter.

- If you have an idea of an acceptable E, you can use it to calculate the sample size, n, necessary for a CI.

$n = [(Z_{\alpha/2} \bullet \text{sigma}) / E]^2$

Examples:

1. **CI with Z (sigma known)** **8-13: p 296**

Know sigma: 95% CI for mu is:

x-bar ± Zα/2 • (sigma/root n)
12.6 ± 1.645 • 2.5/ (root 40)
12.6 ± 1.645 • 2.5/ 6.325
12.6 ± 1.645 • 0.395
12.6 ± 0.650
LB:11.95 and UB:13.25

2. CI with t (sigma unknown) **8-40 p 302**

Don't know sigma: 95% CI for μ is:

x-bar ± t (α/2, n-1 degrees of freedom) • (s/root n)
126 ± 2.262 • (4/root 10)
126 ± 2.262 • 1.26
126 ± 2.86
LB: 123.14 and UB: 128.86

3. sample size using E **8-25: p 297**

$n = [(Z_{\alpha/2} \bullet \text{sigma}) / E]^2$
$n = [(1.645 \bullet 8) / 6]^2$
$n = [13.16 / 6]^2$
$n = 2.19^2$
n = 4.81 = 5

you do: 8-15, 8-24, 8-39

Hypothesis testing is accomplished by first making an assumption with regard to an unknown population characteristic. We then take a random sample from the population, and on the basis of the corresponding sample characteristic, we either accept or reject the hypothesis with a particular degree of confidence.

Eg: **testing a hypothesis about the population mean**
Suppose the Dean of the B-School made an assertion that the average GPA of students in the business school was 3.25.

? How could we test to se if this assertion is true?

→ We could find the average of every single student in the business school, and calculate the true population mean, or we could sample a portion of the students, and use *Inferential Statistics* to make inferences (predictions, conclusions, and decisions) about the population data set based on information contained in a sample.

Hypothesis testing -- first make an assumption with regard to an unknown population parameter (eg: mu) = the null hypothesis = the "status quo" condition we wish to test, and also specify an alternative hypothesis = that which is true of the null hypothesis is wrong.

-- then take a random sample from the population, and on the basis of the sample (test) statistic corresponding to the pop parameter in Ho (mu), we either accept or reject the hypothesis with a particular degree of confidence.

? Recall what we mean by a "random sample"?

→ Each member of the population has an equal chance of being sampled

So, we can say that if the test statistic falls in a certain range of its distribution, then its very likely that ho is true. Or we can say, if the test statistic is in a certain range then Ho is likely false.

Will we always be correct?

Two types of *errors* can occur in testing a hypothesis about the population based on sample data:

1. We can reject a hypothesis that is in fact true = "*type I error*" , or
2. We could accept a hypothesis that is actually false = "*type II error*".

* What we typically do is pick the probability of making a type I error (this probability is represented by the symbol α).

* The accepted probability of type I error, α, is known as **the *level of significance*** of the hypothesis test, and 1 - α is termed the *level of confidence* of the test.

? What do we give-up by achieving a lower chance of type I error?

$\rightarrow$ we accept a greater probability of making a type II error, (β). The accepted probability of type I error, α, is known as the *level of significance* of the hypothesis test, and 1 - α is termed the *level of confidence* of the test.

Formal steps in testing hypotheses about the population mean:

1. Assume that μ equals some hypothetical value μ_0 (in the gpa example, $\mu_0 = 3.25$).

   We write this assumption as: H_0: $\mu = \mu_0$, which is called the ***null hypothesis***.

2. The alternative hypothesis is then: H_1: $\mu \neq \mu_0$, H_1: $\mu > \mu_0$, or H_1: $\mu < \mu_0$, depending on how the problem is stated.

Our E.g.: Ho: $\mu = 3.25$
H_1: $\mu \neq 3.25$

-- We haven't specified whether we believe true GPA is greater than or less than 3.25 – we only want to test if it is exactly equal to 3.25 or not. So a GPA > or < 3.25 would cause us to reject the null hypothesis.

3. Pick a level of significance for the test, and define the acceptance region and rejection region for the test using the appropriate distribution.

Recall that a lower value for the probability of type I error, α, results in a higher probability of type II error.

It is common to specify $\alpha = 0.05$ or $\alpha = 0.01$, hence the level of confidence will be 95% or 99%.

4. Take a random sample from the population and

a) compute the test statistic based on the appropriate sample statistic

b) If test stat falls in the acceptance region, accept H_0, otherwise reject H_0 in favor of H_1.

?how do we do it? How do we use the sample statistic?

We use the sample statistic to construct a test statistic about which we know something – specifically, we know its distribution...

Construct a **test statistic** and compare the value of the statistic to some critical value that separates an AR and RR at some value for level of significance.

? What's a test statistic?

A value determined from sample information used to reject or not reject the null hypothesis.

Here we'll use the t statistic to express the difference between x-bar and μ in units of the standard deviation:

We know the distribution of a mean divided by its standard error, so we can use that knowledge to determine whether the observed difference bt the sample statistic (x-bar) and the pop parameter (mu) is large enough to be statistical evidence of a true difference bt the pop parameter and the hypothesized value.

General formula for Z-statistic (use Z **whenf pop std deviation is known)**

Z stat = $\dfrac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = Z_{calc}$

General formula for t-statistic (use t when pop std dev is not known):

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{\bar{x} - \mu_0}{S_{\bar{x}}} = t_{stat} = t_{calc}$$

Continue with our example… suppose n = 62, x-bar = 3.10, and we know that sigma = 0.59

$$Z = \frac{3.10 - 3.25}{\frac{0.59}{\sqrt{62}}} = -0.15 / 0.075 = -2.00$$

? what do we do with this value?

→ compare it to a "critical value".

? where does the critical value come from?

Z comes from Z table

T comes from the t – table.

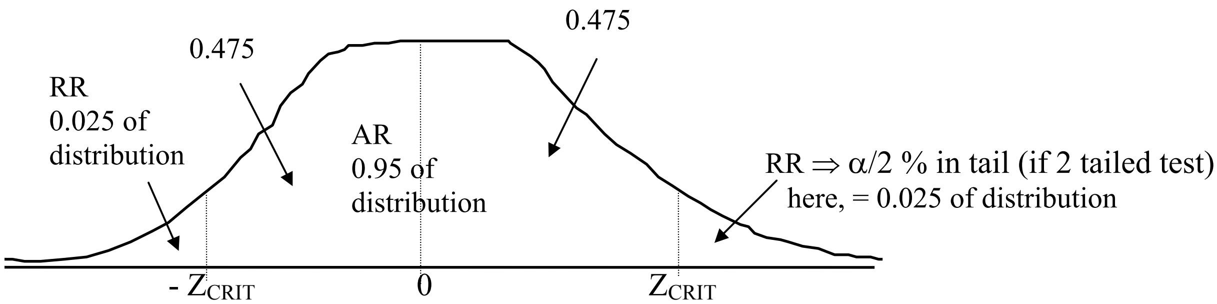
In fact, they will be the same Z and t values we used in constructing our CI's.

Critical Z = Z crit = $Z_{\alpha/2}$

(note again that this is a *two-tailed test* because the rejection region is in both tails of the distribution).

2-tailed test ⇒ want RR to be split in 2 tails of normal distribution

w/ $\alpha = 0.05$ ⇒ $0.05/2 = .025$ in each tail



Look up 0.475 in t table ⇒ $Z_{CRIT} = 1.96$
Since the calculated Z value falls in the rejection region, we reject H_0: that the Dean's claim is true, at the 5% significance level.

? do you think that we would reject this same hypothesis at the 99% level?
Not likely

w/ $\alpha = 0.01$ ⇒ $0.01/2 = .005$ in each tail ⇒ 0.495 in center

Zcrit = 2.575

So we reject at $\alpha = 0.05$ but we accept at $\alpha = 0.01$

"p-value" is the lowest alpha value at which you would reject.
Here its somewhere close to 0.05

We can also use the 95% CI to test the null hypothesis that the true population mean = 3.25.

See if the value we are testing for (in the null hypothesis) 3.25 is in our 95% CI

If not ⇒ we reject the null hypothesis.
If yes ⇒ we accept the null hypothesis

- if the value to be tested in the null μ_0 is <u>not in the CI</u>, then we reject Ho in favor of H1

→ the sample average x-bar is statistically different from the hypothesized value, and this difference could not be reasonably due to random chance alone.

- if the value to be tested in the null μ_0 <u>is in the CI</u>, then we accept Ho

→ the sample average x-bar is <u>not</u> statistically different from the hypothesized value, and the observed difference is likely due to random chance alone.

what about a one-sided test of the statement "our students study <u>at least</u> 25 hours per week"

null hypothesis $H_0: \mu \geq \mu_0$
alternative hypothesis is then $H_1: \mu < \mu_0$

$$t = \frac{21.82 - 25.00}{5.14} = -0.62$$

t crit = 1.697

⇒ abs value of the test statistic < t critical

⇒ Do not reject Ho, although the sample mean is smaller than the hypothesized value, the difference is small enough to be considered reasonably due to chance.

Do: 9-16, 9-26, 9-29 (c, d, e), 9-49, 9-54