Math 2311

Bekki George – bekki@math.uh.edu

Office Hours: MW 11am to 12:45pm in 639 PGH

Online Thursdays 4-5:30pm

And by appointment

Class webpage: http://www.math.uh.edu/~bekki/Math2311.html

Ch 7 - Confidence Intervals

statistic ± margin of error

$\uparrow$ $\bar{x}$ or $\hat{p}$

$\uparrow$ z^* or t^* times st. dev.

$\bar{x} \pm z^*(\sigma/\sqrt{n})$ or $\bar{x} \pm t^*(s/\sqrt{n})$

$n-1 = df$

$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Popper 19

1. What is the z* critical value for a 95% confidence interval?
   a. 1.96
   b. 1.645
   c. 2.33
   d. 2.06
   e. none of these

8.1 - Inference for the Mean of a Population

In chapter 7, we discussed one type of inference about a population – the confidence interval. In this chapter we will begin discussion the **significance test**.

H_0 : is the **null hypothesis.** The **null hypothesis** states that there is no effect or change in the population. It is the statement being tested in a test of significance.

$\mu = \boxed{} \leftarrow$ stated population mean

H_a : is the **alternate hypothesis**. The **alternative hypothesis** describes the effect we suspect is true, in other words, it is the alternative to the "no effect" of the null hypothesis.

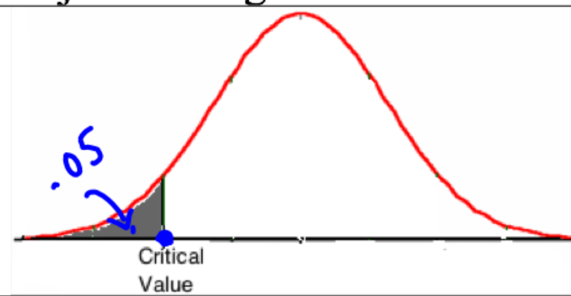
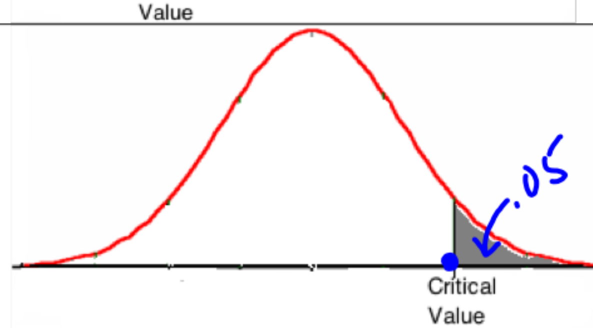
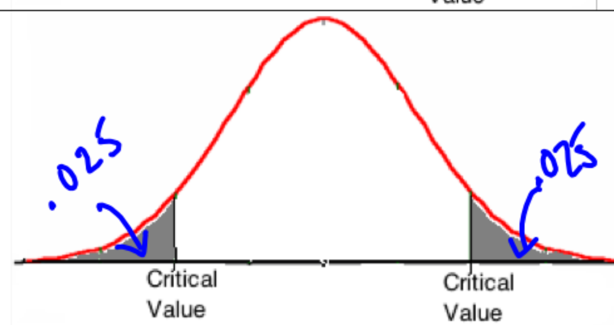
changing $=$ to: $<$, $>$, or $\neq$

Since there are only two hypotheses, there are only two possible decisions: *reject the null hypothesis in favor of the alternative* or *don't reject the null hypothesis*. We will never say that we accept the null hypothesis.

For inference about a population mean:

claimed

$H_0 : \mu = \mu_0$ where μ_0 represents the given population mean.

Alternate Hypothesis	Rejection Region
$H_a : \mu < \mu_0$	
$H_a : \mu > \mu_0$	
$H_a : \mu \neq \mu_0$	

If I want to look at a critical value

$\alpha = .05$

one tailed tests

← two tailed test

The probability, computed assuming that the null hypothesis is true, that the test statistic would take a value as extreme or more extreme than that actually observed is called the ***p*-value** of the test. A result with a small *p*-value is called **statistically significant**. This means that chance alone would rarely produce so extreme a result. We say that a value is **statistically significant** when the *p*-value is as small as, or smaller than, the given significance level, α. If we are not given α, we can interpret the results like this:

- If the *p*-value is less than 1%, we say that there is overwhelming evidence to infer that the alternative hypothesis is true. (We also say that the test is highly significant)
- If the *p*-value is between 1% and 5%, we say that there is strong evidence to infer that the alternative hypothesis is true. (We also say that the test is significant)
- If the *p*-value is between 5% and 10%, we say that there is weak evidence to infer that the alternative hypothesis is true. (We also say that the test not statistically significant)
- If the *p*-value is exceeds 10%, we say that there is no evidence to infer that the alternative hypothesis is true.

if no α given, use $\alpha = .05$

When performing a significance test, we follow these steps:

1. Check assumptions.
2. State the null and alternate hypotheses.
3. Graph the rejection region, labeling the critical values.
4. Calculate the test statistic.
5. Find the p-value. If this answer is less than the significance level, α, we can reject the null hypothesis in favor of the alternate.
6. Give your conclusion using the context of the problem. When stating the conclusion you can give results with a confidence of $(1-\alpha)(100)\%$.

(Means)

z – test

Assumptions:

1. An SRS of size n from the population.
2. Known population standard deviation, σ.
3. Either a normal population or a large sample ($n \geq 30$).

To compute the z – test statistic, we use the formula: $\boxed{z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}.}$

test statistic

t – test

Assumptions:

1. An SRS of size n from the population.
2. Unknown population standard deviation. have s
3. Either a normal population or large sample ($n \geq 30$).

To compute the t – test statistic, we use the formula: $\boxed{t = \dfrac{\bar{x} - \mu_0}{s / \sqrt{n}},}$ where s is the sample standard deviation. The t – test will use $n - 1$ degrees of freedom.

Popper 19

2. A *t* test is used instead of a *z* test when
   a. The population mean is unknown
   b. The population size is unknown
   c. The population variance is unknown
   d. None of these

σ^2

Examples:

10. Mr. Murphy is an avid golfer. Suppose he has been using the same golf clubs for quite some time. Based on this experience, he knows that his average distance when hitting a ball with his current driver (the longest-hitting club) under ideal conditions is 200 yards with a standard deviation of 9. After some preliminary swings with a new driver, he obtained the following sample of driving distances:

$\sigma = 9$ claiming $\mu = 200$

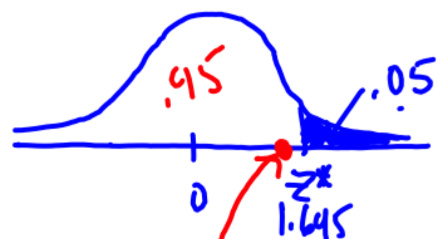
205	198	220	210	194	201	213	191	211	203

$n = 10$ Assume pop. is normal

He feels that the new club does a better job. Do you agree?

$H_0: \mu = 200$ $\bar{x} = 204.6$

$H_a: \mu > 200$ use $\alpha = .05$



invNorm(.95)

$z^* = 1.645$

test statistic: $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{204.6 - 200}{9/\sqrt{10}} = \boxed{1.616}$

pvalue: $P(z > 1.616) = \text{normalcdf}(1.616, \text{big\#}, 0, 1)$

$= .053 > \alpha = .05$

Fail to reject the null hypothesis.

if test statistic is in rejection region means that the pvalue $< \alpha$

$\Rightarrow$ Reject H_0 in favor of H_a

if t.s. not in rejection region and pvalue $> \alpha$

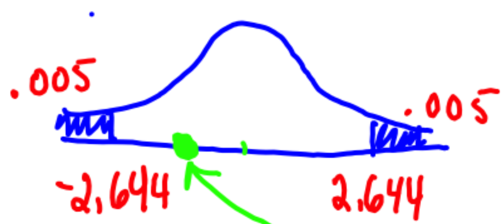
$\Rightarrow$ fail to reject H_0

13. An association of college bookstores reported that the average amount of money spent by students on textbooks for the Fall 2010 semester was $325.16. A random sample of 75 students at the local campus of the state university indicated an average bill for textbooks for the semester in question to be $312.34 with a standard deviation of $76.42. Do these data provide significant evidence that the actual average bill is different from the $325.16 reported? Test at the 1% significance level.

$n = 75 \quad \bar{x} = 312.34 \quad s = 76.42 \quad \alpha = .01$

$H_0: \mu = 325.16$

$H_a: \mu \neq 325.16$



InvT(.005, 74)

$t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} = \dfrac{(312.34 - 325.16)}{(76.42/\sqrt{75})} = -1.4528$

If $H_a \neq$ then pvalue use $<$ if test stat neg $>$ if pos and mult. by 2

pvalue: $2p(t < -1.4528) = \boxed{.1505} > .01$

$2 \cdot pt(-1.4528, 74)$ or $2 \cdot tcdf(\text{small}, -1.4528, 74)$

Fail to reject the null hyp

Popper 19

3. Suppose a test of the hypothesis in a question gives a p-value of 0.02. The correct action based on $\alpha = 0.05$ would be to

   a. Reject H_0
   ~~b.~~ Reject H_a
   ~~c.~~ Accept H_0
   d. Fail to reject H_0
   ~~e.~~ None of these

Matched pairs is a special test when we are comparing corresponding values in data. This test is used only when our data samples are DEPENDENT upon one another (like before and after results).

Matched pairs t – test assumptions:

1. Each sample is an SRS of size n from the same population.
2. The test is conducted on paired data (the samples are NOT independent).
3. Unknown population standard deviation.
4. Either a normal population or large samples ($n \geq 30$).

Example:

15. A new law has been passed giving city police greater powers in apprehending suspected criminals. For six neighborhoods, the numbers of reported crimes one year before and one year after the new law are shown. Does this indicate that the number of reported crimes have dropped?

Neighborhood	1	2	3	4	5	6
Before	18	35	44	28	22	37
After	21	23	30	19	24	29

assume normal population

Before-After -3 12 14 9 -2 8 ← find averge of these for $\bar{x}_D = 6.333$

H_0: no change : $\mu_D = 0$

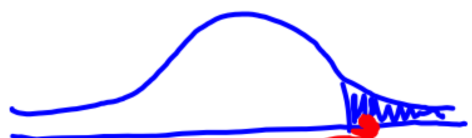
H_a: $\mu_D > 0$

$\alpha = .05$

$df = 5$

```
> crime_diff=c(-3,12,14,9,-2,8)
> mean(crime_diff)
[1] 6.333333
> sd(crime_diff)
[1] 7.174027
```

S



$t^* = 2.015$

invT(.95, 5)

$$t = \frac{6.333 - 0}{7.174/\sqrt{6}} = \boxed{2.162}$$

1-pt(2.162, 5) $< \alpha$

p value: $p(t > 2.162) = $ tcdf(2.162, 1000000, 5) $= .0415 < .05$

Reject the null in favor of the alt. hypothesis

Popper 19

4. A significance test was performed to test $H_0: \mu = 23$ versus $H_a: \mu < 23$. The test statistic is $z = -1.70$. What is the p-value for this test?

$p(z < -1.70)$

a. 0.09
b. 0.04
c. 0.02
d. 0.96
e. none of these

5. A significance test was performed to test $H_0: \mu = 23$ versus $H_a: \mu \neq 23$. The test statistic is $z = -1.70$. What is the p-value for this test?

$2 \cdot p(z < -1.70)$

a. 0.09
b. 0.04
c. 0.02
d. 0.96
e. none of these