Math 1710 Class 29

Type I/II Errors, Power, 2-Sample Prop., and Examples

Dr. Back

Nov. 4, 2009

Power, Type I/II Errors, α, and β

Given $H_0: \; p = p_0$, there are two ways an HT can report an inaccurate result:

Power, Type I/II Errors, α, and β

Given $H_0: \; p = p_0$, there are two ways an HT can report an inaccurate result:

	H_0 true	H_0 false
Retain H_0	*Good*	*Type II Error* probability $= \beta$ depends on value of p
Reject H_0	*Type I Error* probability $= \alpha$	*Good* probability $=$ 1-β $=$ *power*

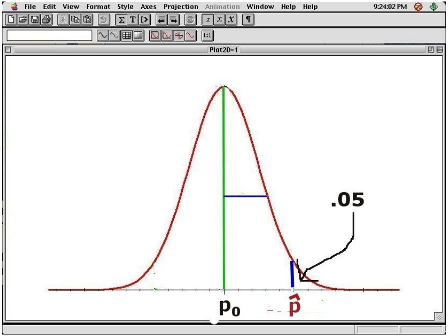
Power, Type I/II Errors, α, and β

Given $H_0: \ p = p_0$, there are two ways an HT can report an inaccurate result:

Why prob(Type I Error)$= \alpha$?

Borderline value of $\hat{p}$ when $\alpha = .05$

Remember this is the sampling distribution of $\hat{p}$ when $p = p_0$.



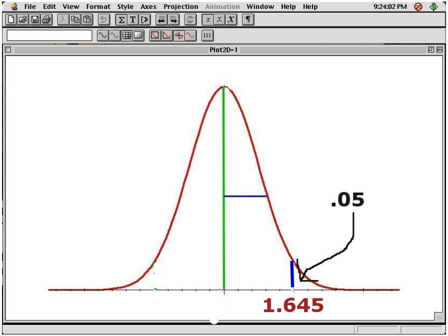
Power, Type I/II Errors, α, and β

Given $H_0: \ p = p_0$, there are two ways an HT can report an inaccurate result:

Why prob(Type I Error)= α?

Value of z^* when $\alpha = .05$

This is the borderline value of z in retaining/rejecting.



Power, Type I/II Errors, α, and β

Given $H_0: \; p = p_0$, there are two ways an HT can report an inaccurate result:

Type I Error Examples:

a: False Positive in a diagnosis; i.e. deciding a person is sick when they really are not. *(H_0 : The person is well.)*

b: Convicting an innocent person. *(H_0 : The person is innocent.)*

c: Producer Risk; the chance that a good good shipment erroneously fails a a test for quality. *(H_0 : A product meets a specification.)*

Power, Type I/II Errors, α, and β

Given $H_0: \; p = p_0$, there are two ways an HT can report an inaccurate result:

Type I Error Examples:

a: False Positive in a diagnosis; i.e. deciding a person is sick when they really are not. *(H_0 : The person is well.)*

b: Convicting an innocent person. *(H_0 : The person is innocent.)*

c: Producer Risk; the chance that a good good shipment erroneously fails a a test for quality. *(H_0 : A product meets a specification.)*

Type II Error Examples:

a: False Negative; missing a sick person.

b: Letting a guilty person go free.

c: Consumer Risk; the chance that a bad shipment erroneously passes a a test for quality.

Power Example

40% of employees are women.

Woman under-represented as executives?

What would it take in $\hat{p}$ for the company to prove that women are as well represented among executives?

Power Example

40% of employees are women.
Woman under-represented as executives?
What would it take in $\hat{p}$ for the company to prove that women are as well represented among executives?

$$H_0: \quad p = .4$$
$$H_a: \quad p > .4$$

Power Example

40% of employees are women.
Woman under-represented as executives?
What would it take in $\hat{p}$ for the company to prove that women are as well represented among executives?

$$H_0: \quad p = .4$$
$$H_a: \quad p > .4$$

One could also do a HT to see if a given $\hat{p}$ demonstrates women are under represented. Then H_a would be $p < .4$.

Power Example

$$H_0 : \quad p = .4$$
$$H_a : \quad p > .4$$

$n = 420, SD(\hat{p}) = .0239.$

$$z = \frac{\hat{p} - .4}{.0239}$$

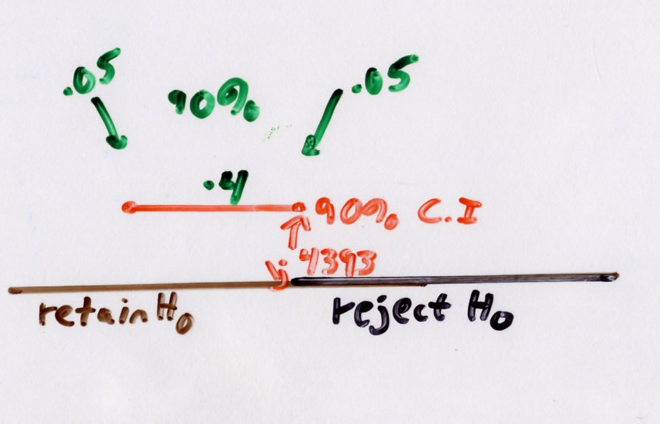
$z > z^*$ means $\hat{p} > .0239z^* + .4.$

$\alpha = .05 \Rightarrow z^* = 1.645$; rejection means $\hat{p} > .439.$

$\alpha = .01 \Rightarrow z^* = 2.326$; rejection means $\hat{p} > .456.$

Power Example

How a $\hat{p}$ will be dealt with in the HT



Power Example

How does the power depend on the effect size?

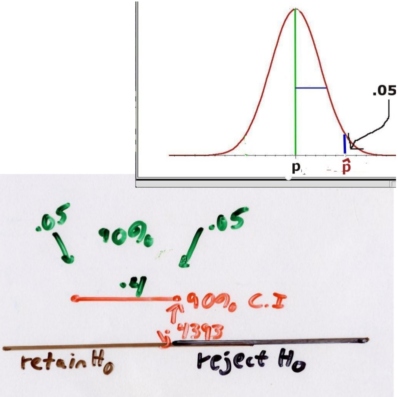
i.e. on the actual value of p

Power Example

How does the power depend on the effect size?

Sampling Dist of $\hat{p}$ when effect size is large.

Power nearly 1

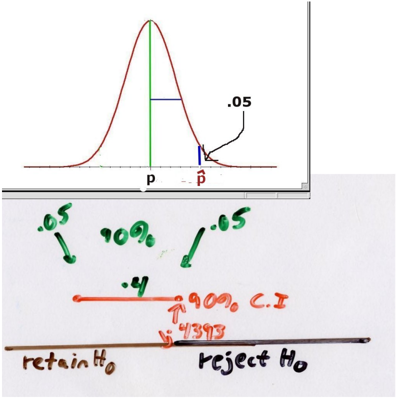


Power Example

How does the power depend on the effect size?

Sampling Dist of $\hat{p}$ when effect size is small.

Power nearly α

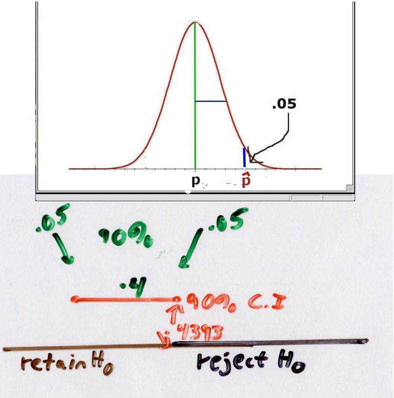


Power Example

How does the power depend on the effect size?

Sampling Dist of $\hat{p}$ when effect size is middle size.

Power in middle between 0 and 1; you could calulate it, but we won't ask you to.



Power Example

How does the power depend on the effect size?

$\alpha = .05$ case:

p	β	power
.4001	.95	.05
.41	.89	.11
.45	.33	.67
.5	.01	.99

Power Example

How does the power depend on the effect size?

$\alpha = .01$ case:

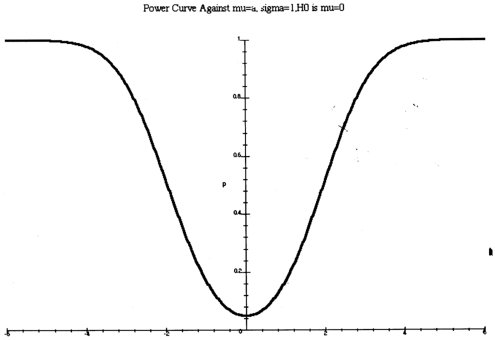
p	β	power
.4001	.99	.01
.41	.97	.03
.45	.59	.41
.5	.04	.96

Power Example

How does the power depend on the effect size?

Power vs. alternative value of p in 2-sided case.

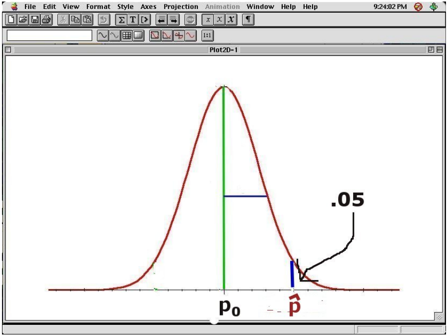
A Power Curve



Power Poperties

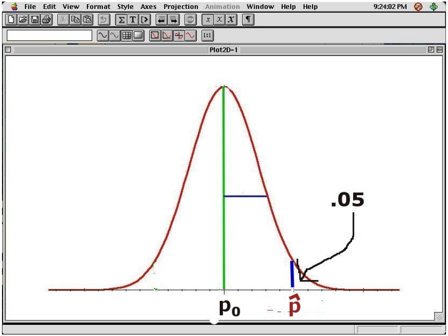
Prob of type I error $= \alpha$.

(H_0 1-sides, $\alpha = .05$ below)



Power Poperties

Power = 1 - β always.

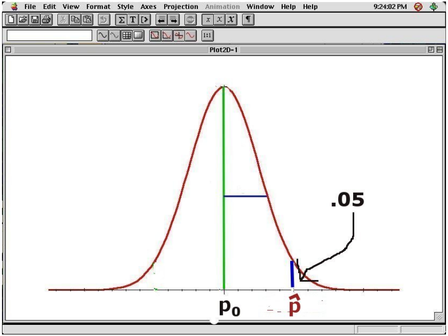


Power Poperties

Power depends on effect size. (i.e actual alternative value for p.)

Power Poperties

Power is approx. α when actual p is near p_0 but not exactly p_0.



Power Poperties

Power goes to 1 as effect size grows, assuming in the 1 sided case that the alternative value supports H_a.

Power Poperties

Power increases as sample size increases.

Power Poperties

Increasing α decreases β . (Easier to reject H_0.)

300 stdnts, 60% approve; 200 faclty, 65%

A significant difference? CI for difference in rates of approval?

300 stdnts, 60% approve; 200 faclty, 65%

A significant difference? CI for difference in rates of approval? Let p_1 and p_2 denote the true proportions of students and faculty that approve.

300 stdnts, 60% approve; 200 faclty, 65%

A significant difference? CI for difference in rates of approval?
2-sample inference based on the sampling distribution of
$\hat{p}_1 - \hat{p}_2$

300 stdnts, 60% approve; 200 faclty, 65%

A significant difference? CI for difference in rates of approval?
2-sample inference based on the sampling distribution of $\hat{p}_1 - \hat{p}_2$

$$\mu = p_1 - p_2$$

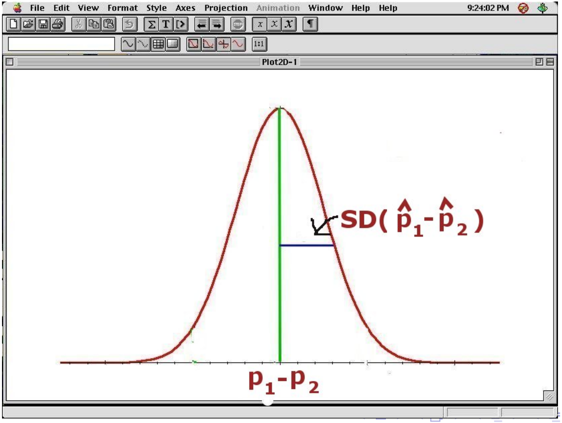
$$Var(\hat{p}_1 - \hat{p}_2) = Var(\hat{p}_1) + Var(\hat{p}_2) = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

$$SD(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

300 stdnts, 60% approve; 200 faclty, 65%

A significant difference? CI for difference in rates of approval?
2-sample inference based on the sampling distribution of $\hat{p}_1 - \hat{p}_2$

Samp. Dist. of $\hat{p}_1 - \hat{p}_2$: $N(p_1 - p_2, SD(\hat{p}_1 - \hat{p}_2))$



300 stdnts, 60% approve; 200 faclty, 65%

2-sample inference based on the sampling distribution of $\hat{p}_1 - \hat{p}_2$

Find a CI for $p_1 - p_2$:

300 stdnts, 60% approve; 200 faclty, 65%

2-sample inference based on the sampling distribution of $\hat{p}_1 - \hat{p}_2$

Find a CI for $p_1 - p_2$:

Since we don't know p_1 and p_2, we can't directly compute $SD(\hat{p}_1 - \hat{p}_2)$.

So we use $SE(\hat{p}_1 - \hat{p}_2)$ instead.

300 stdnts, 60% approve; 200 faclty, 65%

2-sample inference based on the sampling distribution of $\hat{p}_1 - \hat{p}_2$

Find a CI for $p_1 - p_2$:

$$SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

300 stdnts, 60% approve; 200 faclty, 65%

2-sample inference based on the sampling distribution of $\hat{p}_1 - \hat{p}_2$

Find a CI for $p_1 - p_2$:

$$SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Same argument as in the 1-sample case gives a CI for $p_1 - p_2$ of

$$\hat{p}_1 - \hat{p}_2 \pm z^* SE(\hat{p}_1 - \hat{p}_2).$$

300 stdnts, 60% approve; 200 faclty, 65%

2-sample inference based on the sampling distribution of $\hat{p}_1 - \hat{p}_2$

Find a CI for $p_1 - p_2$:

$$SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Same argument as in the 1-sample case gives a CI for $p_1 - p_2$ of

$$\hat{p}_1 - \hat{p}_2 \pm z^* SE(\hat{p}_1 - \hat{p}_2).$$

Here we have

$$SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{.6 \cdot .4}{300} + \frac{.65 \cdot .35}{200}} = .0440.$$

300 stdnts, 60% approve; 200 faclty, 65%

2-sample inference based on the sampling distribution of $\hat{p}_1 - \hat{p}_2$

Find a CI for $p_1 - p_2$:

Same argument as in the 1-sample case gives a CI for $p_1 - p_2$ of

$$\hat{p}_1 - \hat{p}_2 \pm z^* SE(\hat{p}_1 - \hat{p}_2).$$

Here we have

$$SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{.6 \cdot .4}{300} + \frac{.65 \cdot .35}{200}} = .0440.$$

A 95% CI for $p_1 - p_2$ is:

$$(.6 - .65) \pm 1.96 \cdot .0440 = -.05 \pm .0863 = (-.1363, .0363).$$

300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

Without the request for P-value, we could use the CI above.
But for the P-value we need to use "Method 1."

300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

Our hypotheses are:

- H_0: $p_1 = p_2$
- H_a: $p_1 \neq p_2$

300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

Our hypotheses are:

- H_0: $p_1 = p_2$
- H_a: $p_1 \neq p_2$

A twist enters. We are only interested in the reasonableness of our observed $\hat{p}_1 - \hat{p}_2$ with respect to the sampling dist if H_0 is true. There are many such distributions (since we don't know the common value of $p_1 = p_2$ to use.) In particular what we did with $SE(\hat{p}_1 - \hat{p}_2)$ above does not fit the $p_1 = p_2$ situation.

300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

Our hypotheses are:

- H_0: $p_1 = p_2$
- H_a: $p_1 \neq p_2$

We resolve this conflict by making our best estimate of the common value of p_1 and p_2, namely the weighted average

$$\hat{p}_{pooled} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$$

and then

$$SE_{pooled}(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\widehat{p_{pooled}}\,\widehat{q_{pooled}}}{n_1} + \frac{\widehat{p_{pooled}}\,\widehat{q_{pooled}}}{n_2}}.$$

300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

Here the weighted average is

$$\hat{p}_{pooled} = \frac{300 \cdot .60 + 200 \cdot .65}{200 + 300} = .6 \cdot 300 + .4 \cdot .65 = .63$$

and then

$$SE_{pooled}(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{.63 \cdot .37}{300} + \frac{.63 \cdot .37}{200}} = .0443.$$

300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

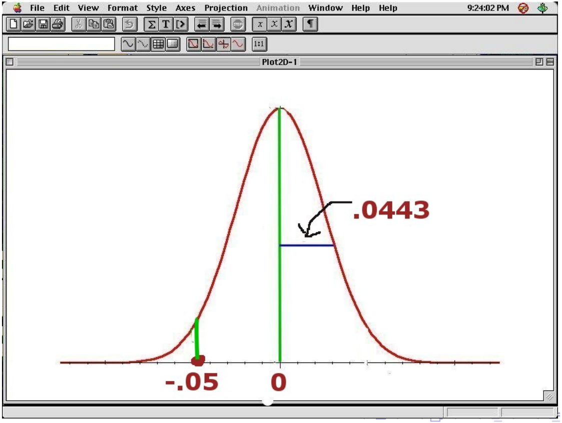
Our z-statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{SE_{pooled}(\hat{p}_1 - \hat{p}_2)} = \frac{-.05}{.0443} = -1.12.$$

300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

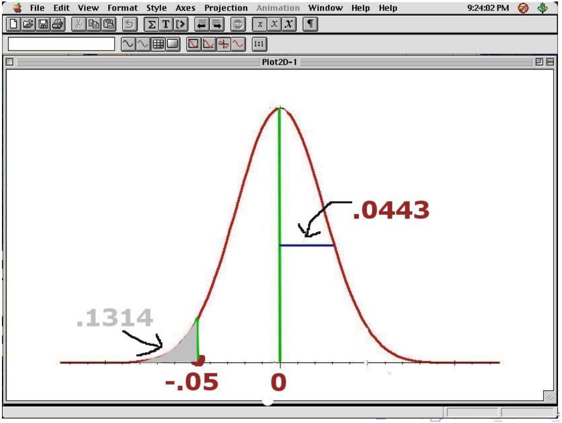
Approx Samp. Dist. of $\hat{p}_1 - \hat{p}_2$: $N(0, .0443)$ if H_0 is true.



300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

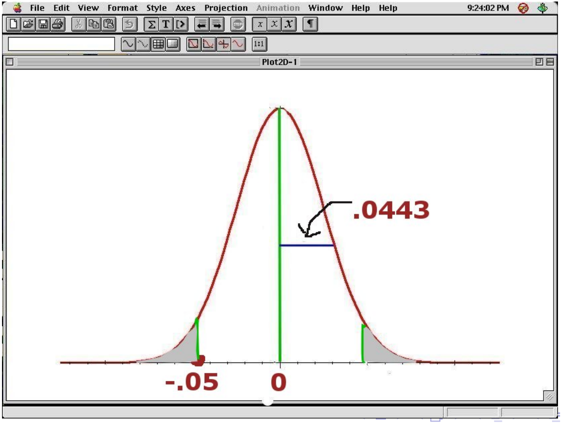
Tail Prob. is $P(Z < -1.12) = .1314$.



300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

P-value=2(Tail Prob.) $= 2(.1314) = .2628$



300 stdnts, 60% approve; 200 faclty, 65%

Carry out an HT at a sig level of $\alpha = .05$ of whether faculty and student approval rates are different. Calculate the P-value as well.

Our P-value is larger than $\alpha = .05$, so we retain H_0.