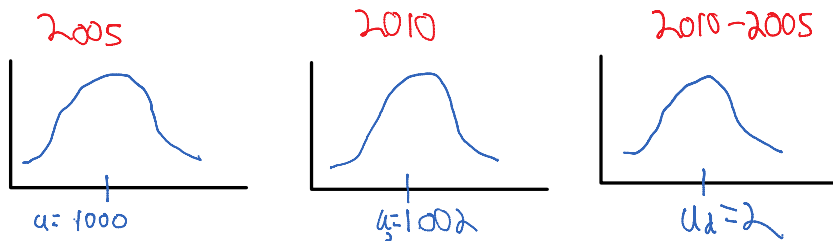


## Lecture 10, comparing groups 2

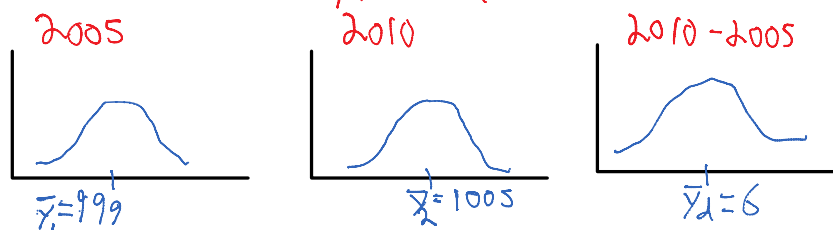
Saturday, October 20, 2012  
12:23 PM

Matched Data

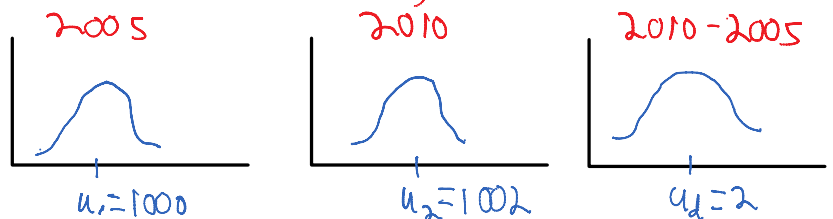
Population data



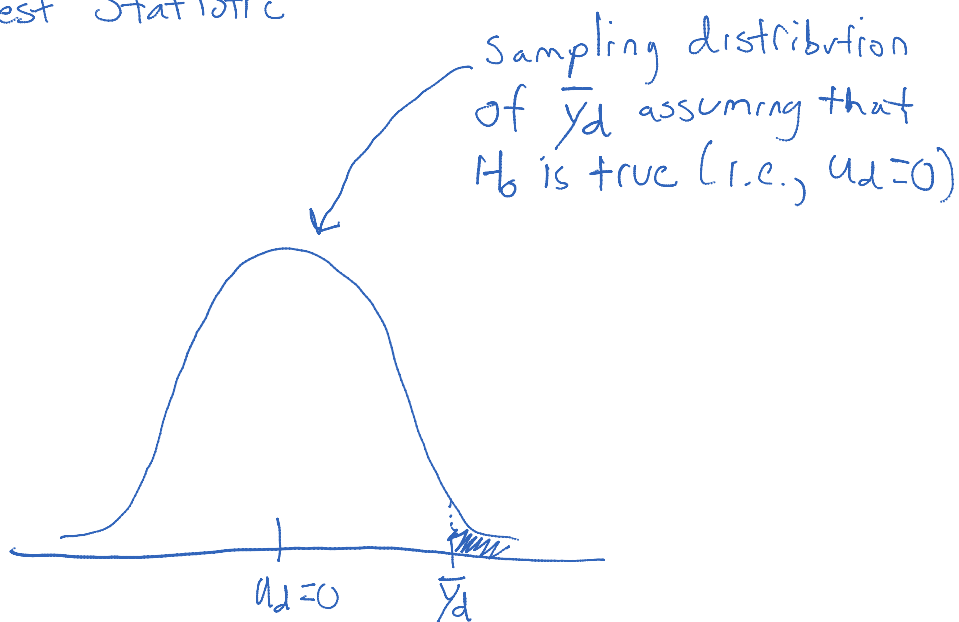
Sample data



Sampling dist

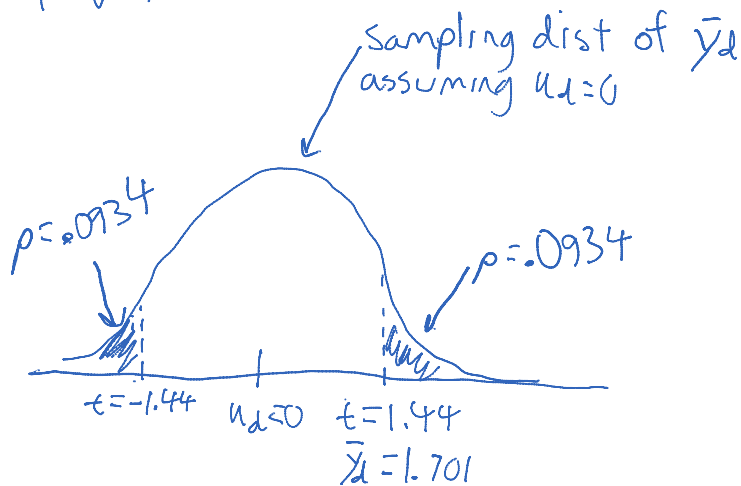


Test Statistic



Test statistic: How unlikely would it be to observe the point estimate we observed if the null hypothesis was true

P-value



$$p\text{-value} = \Pr(\text{obs} > 1.44) + \Pr(\text{obs} < -1.44) = .0934 \times 2 = .1868$$

`. ttest dsatp50==0`

One-sample t test

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
dsatp50	1321	1.701363	1.288401	46.82767	- .8261748 4.2289

mean = mean(dsatp50)

Ho: mean = 0

Ha: mean < 0  
Pr(T < t) = 0.9066

Ha: mean != 0  
Pr(|T| > |t|) = 0.1869

Ha: mean > 0  
Pr(T > t) = 0.0934

t = 1.3205

degrees of freedom = 1320

$H_0: u_d = 0$

$H_a: u_d \neq 0$

$H_a: u_d > 0$

p-value = .0934 x 2

t-test

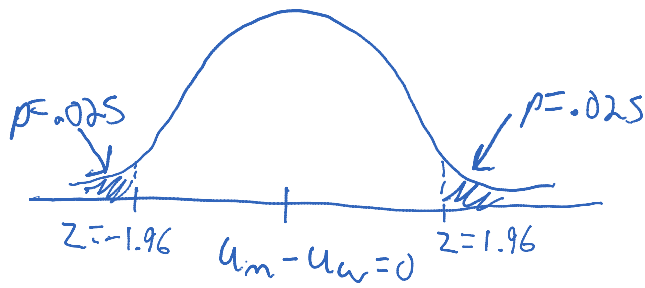
Type I error

$H_0: u_{\text{men}} - u_{\text{women}} = 0$

$H_a: u_{\text{men}} - u_{\text{women}} \neq 0, \alpha = .05$

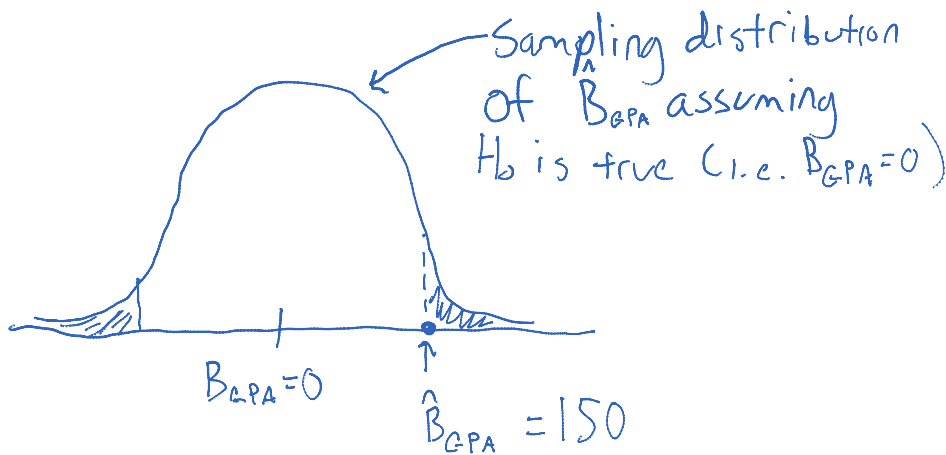
Show rejection region on sampling

distribution



What is probability of making a type I error?

## Significance test for Regression Coefficient



$y$	$\hat{\beta}$	$\hat{\beta}$	test statistic	p-value (two sided)	95% CI of $\beta$
b2004v2cat8	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
hsgpaw	.6141679	.0355743	17.26	0.000	.5444345 .6839013
pareduc	.1871284	.0147169	12.72	0.000	.1582802 .2159767
sat100	.4336535	.0109311	39.67	0.000	.4122261 .4550809
female	.1130442	.0354756	3.19	0.001	.0435044 .1825841
black	.5954396	.0591407	10.07	0.000	.4795111 .7113681
hispanic	.2082926	.06008	3.47	0.001	.0905228 .3260623
asian	.2600058	.0562783	4.62	0.000	.1496881 .3703234
native	.1333464	.1954941	0.68	0.495	-.2498644 .5165572

hispanic	.2082926	.06008	3.47	0.001	.0905228	.3260623
asian	.2600058	.0562783	4.62	0.000	.1496881	.3703234
native	.1333464	.1954941	0.68	0.495	-.2498644	.5165572
mixedoth	.1011369	.0846062	1.20	0.232	-.0647095	.2669833
_cons	-3.151401	.1124993	-28.01	0.000	-3.371924	-2.930877

$\hat{\beta}_{GPA}$        $se \hat{\beta}_{GPA}$        $t \hat{\beta}_{GPA}$

Sometimes regression output in journal articles looks like this:

```
. est store mod1
```

```
. est table mod1, star b(%9.3fc)
```

Variable	mod1
hsgpaw	0.614***
pareduc	0.187***
sat100	0.434***
female	0.113**
black	0.595***
hispanic	0.208***
asian	0.260***
native	0.133
mixedoth	0.101
_cons	-3.151***

$\hat{\beta}_{GPA} = 0.614$   
 $H_0: \beta_{GPA} = 0$   
 $H_a: \beta_{GPA} \neq 0$   
 $p\text{-value} < .001$

legend: \* p<0.05; \*\* p<0.01; \*\*\* p<0.001