



Who skips more class??

An Observational Study

by: Rob Cannon
John Trachsel

Study overview

- ▶ Two independent populations
 - Male Undergrads
 - Female Undergrads
- ▶ We are interested in the center or means of these two populations.

Hypothesis formation

- ▶ The null and alternative hypotheses for this observational study are:

$$H_0: \mathcal{M}_{\text{males}} = \mathcal{M}_{\text{females}}$$

- ▶ $H_a: \mathcal{M}_{\text{males}} - \mathcal{M}_{\text{females}} \neq 0$

Survey Strategy

- Collect data

- Format Data for SAS

- Use SAS to compare population means.

Study Assumptions



- ▶ Individuals answers are accurate
- ▶ Samples taken are a good representation of University student population
- ▶ Data was entered correctly

Analysis Method

- ▶ Two methods of analysis.
 - Calculate a confidence interval for the difference between the means (test H_a :)
 - Calculate a two sample t statistic (test H_0 :)

90% Confidence Interval

- ▶ SAS code for 90% Confidence interval of for males and females.

```
data skip;
```

- ▶ infile 'H:\documents\skip.txt';
- ▶ input mhours fhours;
- ▶ run ;
- ▶ proc means data = skip n mean stddev stderr clm alpha = .10 ;
- ▶ var fhours mhours ;
- ▶ run ;

90% Confidence interval SAS Output

Variable	N	Mean	Std Dev	Std Error	Lower 90% C level for Mean	Upper 90% C level for Mean
Female Hours	124	1.43548	1.57852	0.14175	1.2005	1.67042
Male Hours	137	2.13868	1.82379	0.15581	1.88063	2.39674

Female Hours Skipped: 1.44 +/- .24

Male Hours Skipped: 2.14 +/- .26

90% Confidence Interval

- ▶ Female Hours Skipped: 1.44 +/- .24
- ▶ Male Hours Skipped: 2.14 +/- .26
- ▶ Because the confidence intervals do not overlap we can confirm that $H_a: \mathcal{M}^{males} - \mathcal{M}^{females} \neq 0$ at the 90% confidence level.

Probability test for Mean Equality

- ▶ SAS Code:

```
proc univariate mu0 = 1.44 2.14 data = skip ;  
var mhours fhours ;  
run ;
```

- ▶ This is a test for the location of the mean.

Testing $H_0: \mathcal{M}_{males} = \mathcal{M}_{females}$

- ▶ Use SAS to compute location of means.

Probability Test Results

- ▶ SAS provided an indirect route to give P, the probability, of Ho: $\mathcal{M}_{males} = \mathcal{M}_{females}$ proven true
- ▶ SAS Output

Group	T Value	Probability
Male	4.48	<.0001
Female	-4.97	<.0001

Data Collection

- ▶ Created three question surveys
 - Questions:
 - What is your sex?
 - Are you taking at least 12 semester hours?
 - How many hours on average do you miss a week?
- ▶ Pooled 137 males and 124 females for the survey
- ▶ Samples collected at random from the University of Iowa Campus
- ▶ Threw out any students who were not full time or any that presented a ridiculous amount of hours missed

Conclusions

- ▶ Based on the data analysis girls on average miss less class than guys
- ▶ Future test could determine the effect of going to class and variability of hours missed between different majors

Questions

