STA 303 | 1002

Note Title 1/12/2012

Next Thursday: SAS on Cquest demo

Spock Conspiracy Trial continued:

Is there evidence of a difference in means in % women on venires between Spock's judge and the other judges?

Last class:

$H_0: \mu_{spock} = \mu_{other}$

$H_a: \mu_{spock} \neq \mu_{other}$

Two-sample t-test

	test stat	p-value
Assuming equal variance (pooled)	5.67	<.0001
Satterthwaite approximation	7.16	<.0001

So there is strong evidence of a difference in the means.

A Linear Model Approach

Let $I_{spock} = \begin{cases} 1 & \text{if judge = Spock's} \\ 0 & \text{otherwise} \end{cases}$

Fit model $Y_i = \beta_0 + \beta_1 I_{spock,i} + e_i \quad , \quad E(e_i) = 0$

where Y_i = % women on i^{th} venire

$i = 1, \ldots, N$ where $N = n_{spock} + n_{other}$

$$E(Y_i) = \begin{cases} \beta_0 + \beta_1 & \text{if judge is Spock's} \\ \beta_0 & \text{if judge is another judge} \end{cases}$$

<u>Least squares estimates of β_0 and β_1:</u>

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$b_1 = \frac{\sum_{i=1}^{N} x_i y_i - N\bar{x}\bar{y}}{\sum x_i^2 - N\bar{x}^2}$$

$x_i = I_{spock,\, i}$

$\sum_{i=1}^{N} x_i = n_{spock}$

$\bar{x}$ = proportion of venires that are Spock's judge's $= n_{spock}/N$

$$\frac{\sum_{i=1}^{N} x_i y_i}{n_{spock}} = \frac{\sum_{(spock)} y_i}{n_{spock}} = \bar{y}_{spock}$$

$$\sum_{i=1}^{N} x_i^2 = n_{spock}$$

$$b_1 = \frac{n_{spock}\,\bar{y}_{spock} - n_{spock}\,\bar{y}}{n_{spock} - \frac{n_{spock}^2}{N}}$$

$$= \frac{N\bar{y}_{spock} - N\bar{y}}{N - n_{spock}}$$

$$= \frac{(n_{spock} + n_{other})\,\bar{y}_{spock} - \sum_{i=1}^{N} y_i}{n_{other}}$$

$$= \frac{\sum_{(spock)} y_i + n_{other}\,\bar{y}_{spock} - \sum y_i}{n_{other}}$$

$$= \frac{n_{other}\,\bar{y}_{spock} - \sum_{(other)} y_i}{n_{other}}$$

$$= \bar{y}_{spock} - \bar{y}_{other}$$

<u>Exercise</u>: Show: $b_0 = \bar{y}_{other}$

b_0 is the average of observations (y_i) for other judges

b_1 is the difference in average y_i's between Spock's and other judges

$$b_0 + b_1 = \bar{y}_{spock}$$

General: $Y_i = \beta_0 + \beta_1 I_{group,i} + e_i$

$$I_{group,i} = \begin{cases} 1 & \text{if } i^{th} \text{ observation is in one group} \\ 0 & \text{'' '' '' '' the other group} \end{cases}$$

Example of a ONE-WAY CLASSIFICATION
or "One Way Analysis of Variance"

If mean of the 2 groups are the same then $\beta_1 = 0$

Test with test statistic

$H_0: \beta_1 = 0$
vs $H_a: \beta_1 \neq 0$

$$t_{obs} = \frac{b_1}{\text{s.e. of } b_1}$$ (as in Simple Linear Regression)

Assuming model assumptions are true,

Assumptions: Form of model is correct

Gauss - Markov conditions
1. $E(e_i) = 0$
2. $Var(e_i) = \sigma^2$ const.
3. e_i's uncorr.

$e_i \sim$ Normal

and H_0 is true, then

t_{obs} is an observation from a t-distribution with $df = N - 2$ $(= n_{spoke} + n_{other} - 2)$