# ANOVA and Results from YES

Feb 19, 2014

Qu Jin

# Outline

- **Why** do we want to use ANOVA?
- **What** is ANOVA? **What** are the assumptions?
- **How** do we interpret ANOVA table?
- **What** are post-hoc analyses? **Which** method should we use?
- Results from YES

# **Why** do we want to use ANOVA?

- Example: Innovation Orientation of students from four groups
- Business > Engineering > Other > STM
- Are the four groups different from each other? How can we test the significance?
- One option is t-tests: E-B, E-S, E-O, B-S, B-O, and S-O
- The other option is ANOVA

| | Innovation Orientation | |
|---|---|---|
| Disciplines | Mean | SD |
| Engineering (N = 518) | 2.93 | 1.06 |
| Business (N = 471) | 3.54 | 1.01 |
| Science, Technology, and Mathematics (STM) (N = 668) | 2.55 | 1.09 |
| Other (N = 1230) | 2.87 | 1.14 |

# **What** is ANOVA?
# **What** are the assumptions?

- Primary question: does discipline help explain Innovation Orientation?
- $H_0$: $\mu_E = \mu_B = \mu_S = \mu_O$
- $H_1$: at least one $\mu$ is different
- Assumptions:
  - All observations are independent
  - All observations are normally distributed
    - Means may depend on the levels of the factors
    - Constant variance

# **How** do we interpret ANOVA table?

| | Sum of Squares | df | Mean Square | F | P-value |
|---|---|---|---|---|---|
| Model | 18.82 | 3 | 6.27 | 11.51 | .000 |
| Error | 1550.64 | 2845 | 0.55 | | |

Levene's Test of Equality of Error Variance

The null hypothesis is that the error variance is equal across groups

| F | df1 | df2 | P-value |
|---|---|---|---|
| 0.351 | 7 | 2845 | .93 |

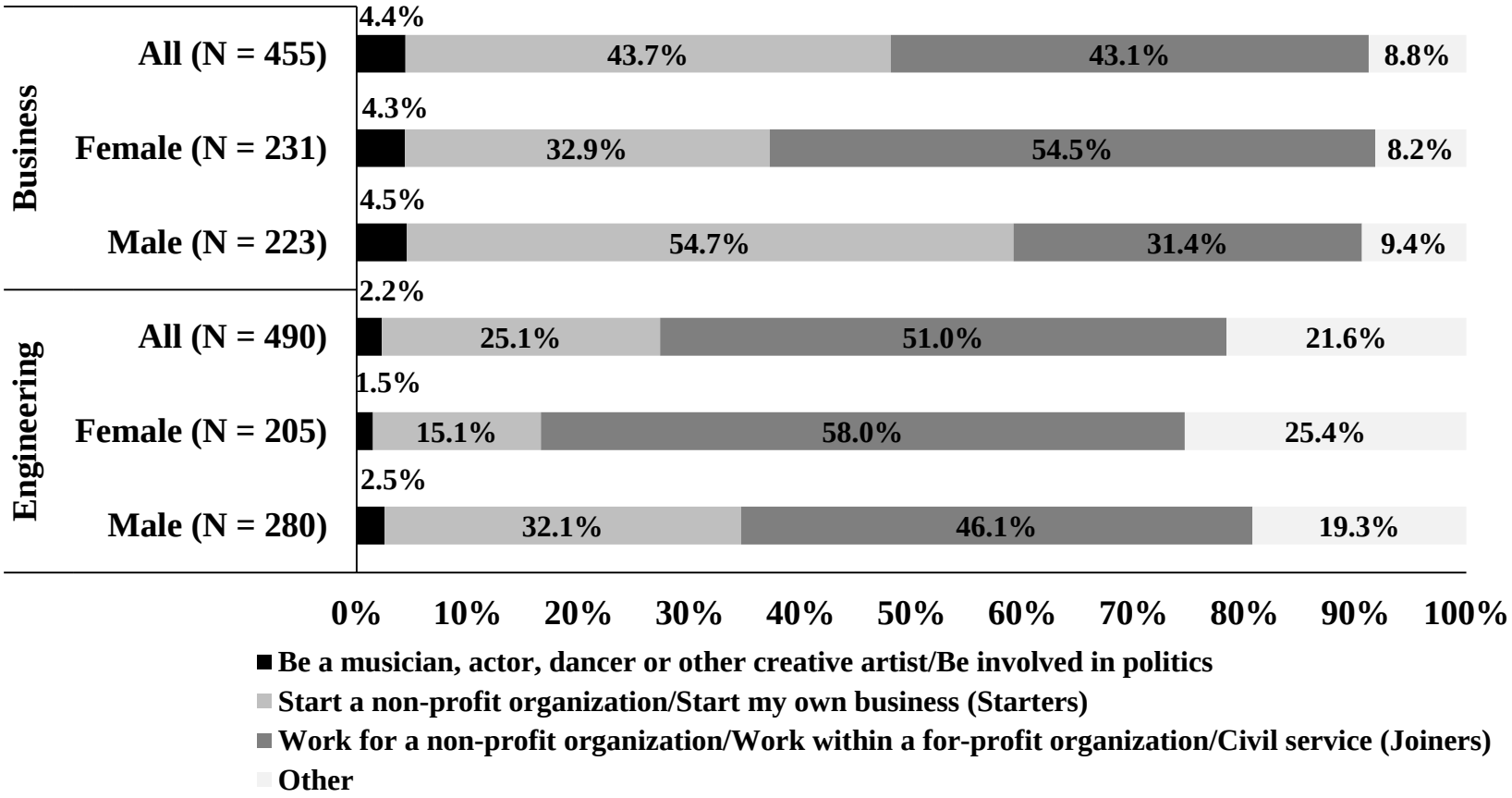
| Source of Variation | df | SS |
|---|---|---|
| Model | $r-1$ | $\sum_{i=1}^{r} n_i(\overline{Y}_{i.}-\overline{Y}_{..})^2$ |
| Error | $n_T-r$ | $\sum_{i=1}^{r}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_{i.})^2$ |
| Total | $n_T-1$ | $\sum_{i=1}^{r}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})^2$ |

$$\overline{Y}_{..}=\sum_{i=1}^{r}\sum_{j=1}^{n_i}Y_{ij}/n_T \qquad \overline{Y}_{i.}=\sum_{j=1}^{n_i}Y_{ij}/n_i$$

# **What** are post-hoc analyses?
# **Which** method should we use?

- Least Significant Difference (LSD): least conservative, likely to have Type I error
- Tukey: more conservative than LSD, often used
- Bonferroni: more conservative than Tukey
- Scheffe: most conservative
- Conservative: strong control of overall Type I error (incorrect rejection of a true null hypothesis)– avoids false positives
- Powerful: able to pick up differences that exist – avoids false negatives

# Results from YES



# Results from YES

| | Starters | | | | Joiners | | | | ANOVA Post-Hoc Analysis[1] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Engineering (ES) | | Business (BS) | | Engineering (EJ) | | Business (BJ) | | | | | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | ES | BS | EJ | BJ |
| Entrepreneurial Intent | 3.90 | .62 | 4.17 | .67 | 2.62 | .99 | 2.97 | .95 | b | a | d | c |
| Career Value-Challenge | 4.11 | .52 | 4.10 | .57 | 3.81 | .53 | 3.82 | .58 | a | a | b | b |
| Innovation Orientation | 3.56 | .65 | 3.71 | .74 | 3.21 | .74 | 3.28 | .73 | a | a | b | b |
| Goal Selection-Novel | 3.97 | .71 | 3.99 | .74 | 3.68 | .69 | 3.60 | .69 | a | a | b | b |
| Goal Selection-Challenge | 4.09 | .68 | 4.05 | .67 | 3.81 | .71 | 3.71 | .73 | a | a | b | b |
| Sense of Self-Movers & Shakers | 3.55 | .59 | 3.69 | .55 | 3.32 | .60 | 3.48 | .63 | a b | a | c | b c |

[1] Groups with the same label are not significantly different; Tukey-Kramer method

- In four measures, starters were different from joiners
- Business starters had the highest scores in entrepreneurial intent and Sense of Self-Movers & Shakers