Chapter 5: One-way ANOVA and Randomized Experiments - Part 3

Harsha Perera

Key Topics

- How Big is the Effect ? Confidence Intervals and Effect Sizes
- Multiple Comparisons and Fisher's Least Significant Difference (LSD)

Confidence Intervals and Effect Sizes

- ANOVA F-test attempts to address the question of whether there is a difference in means among different groups of the categorical predictor
  - i.e whether there is an association between continuous response and categorical predictor

- ANOVA F-test
  - does not tell us which differences in means led to the rejection of null hypothesis

- In order to address this we need
  - Interval estimates
  - Effect sizes

Confidence Interval for Group Means (μ_i)

- Point estimate for μ_i is $\bar{y}_i$
- Interval estimate for μ_i

$$\bar{y}_i \pm t^*_{df,\alpha} \cdot SD \cdot \sqrt{1/n_i}$$

- t^* depends on degrees of freedom (df) and significance level (α)
  - df = df of the error term
- $SD = \sqrt{MSE}$
- n_i: sample size of the corresponding group
- i: group index

Fruit fly lifetimes: 95% Confidence Intervals for Group Means

- Point estimates ($\bar{y}_i$)

```
##    Treatment Group_Means
## 1       none       63.56
## 2 1 pregnant       64.80
## 3 8 pregnant       63.36
## 4   1 virgin       56.76
## 5   8 virgin       38.72
```

- SD and df,

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## Treatment     4  11939  2984.8   13.61 3.52e-09 ***
## Residuals   120  26314   219.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- df $= 120$
- SD $= \sqrt{MSE} = \sqrt{219.3} = 14.81$
- When $\alpha = 0.05$, $t^*_{120} = 1.984$ (from t-table)
- $n_i = 25$

Fruit fly lifetimes: 95% Confidence Intervals for Group Means

- Margin of error $= t^*_{df,\alpha} \cdot SD \cdot \sqrt{1/n_i}$
  - $\text{MOE} = 1.984 \times 14.81 \times \sqrt{1/25} = 5.88$

- 95% CI for control group: $63.56 \pm 5.88 = [57.68, 69.44]$
  - For the male fruit flies living alone, we are 95% confident that the mean length of life is approximately between 57 days and 69 days

- 95% CI for '1 pregnant' group: $64.80 \pm 5.88 = [57.68, 69.44]$
- 95% CI for '8 pregnant' group: $63.36 \pm 5.88 = [57.48, 69.24]$
- 95% CI for '1 virgin' group: $56.76 \pm 5.88 = [50.88, 62.64]$
- 95% CI for '8 virgin' group: $38.72 \pm 5.88 = [32.84, 44.60]$

Fruit fly lifetimes: 95% Confidence Intervals for Group Means

- 95% CI of first four groups overlap
- 95% CI of '8 virgin' group do not overlap with rest of the groups
  - This could be the reason for rejection of null hypothesis of no mean difference in mean lifetime
- 95% CI of '8 virgin' group suggest that mean lifetime of fruit flies in this particular groups do vary between 32 days and 44 days
  - i.e maximum mean lifetime is 44 days, which is way smaller than the minimum mean lifetime of other groups
  - This observation suggests or justify that sexual activity of male fruit flies could shorten length of life

Confidence Interval for the Difference of Group Means $(\mu_i - \mu_j)$

- Point estimate for $\mu_i - \mu_j$ is $\bar{y}_i - \bar{y}_j$
- Interval estimate for $\mu_i - \mu_j$

$$(\bar{y}_i - \bar{y}_j) \pm t^*_{df,\alpha} \cdot SD \cdot \sqrt{1/n_i + 1/n_j}$$

- t^* depend on degrees of freedom (df) and significance level (α)
  - df = df of the error term
- $SD = \sqrt{MSE}$
- n_i, n_j: sample sizes of the corresponding groups

Fruit fly lifetimes: 95% Confidence Intervals for the Difference of Two Group Means

- Point estimates ($\bar{y}_i$)

```
##    Treatment Group_Means
## 1       none       63.56
## 2 1 pregnant       64.80
## 3 8 pregnant       63.36
## 4   1 virgin       56.76
## 5   8 virgin       38.72
```

- SD and df,

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## Treatment     4  11939  2984.8   13.61 3.52e-09 ***
## Residuals   120  26314   219.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- df $= 120$
- SD $= \sqrt{MSE} = \sqrt{219.3} = 14.81$
- When $\alpha = 0.05$, $t^*_{120} = 1.984$ (from t-table)
- $n_i = 25$

95% Confidence Intervals for the Difference of Two Group Means: Control vs. 1 Pregnant

$$(\bar{y}_1 - \bar{y}_2) \pm t^*_{120} \cdot SD \cdot \sqrt{1/n_1 + 1/n_2}$$

$$(63.56 - 64.80) \pm 1.984 \cdot 14.81 \cdot \sqrt{1/25 + 1/25}$$

$$-1.24 \pm 8.31$$

$$[-9.55, 7.07]$$

- We are 95% confident that difference of mean lifetime between 'control' and '1 pregnant' groups is between 9 days less or 7 days more
- Here, we are interested in the mean difference
  - $H_0 : \mu_1 = \mu_2$ or $H_0 : \mu_1 - \mu_2 = 0$
  - $H_1 : \mu_1 \neq \mu_2$ or $H_1 : \mu_1 - \mu_2 \neq 0$
- If confidence interval contains '0', this suggest that we fail to reject null hypothesis at a 0.05 level of significance

95% Confidence Intervals for the Difference of Two Group Means

- Control vs. 1 pregnant: [-9.55, 7.07]
- Control vs. 8 pregnant:[-8.09, 8.49]
- Control vs. 1 virgin: [-1.49, 15.09]
- **Control vs. 8 virgin:** $[16.55, 33.13]^*$
- 1 pregnant vs. 8 pregnant: [-6.85, 9.73]
- 1 pregnant vs. 1 virgin: [-0.25, 16.33]
- **1 pregnant vs. 8 virgin:** $[17.79, 34.37]^*$
- 8 pregnant vs. 1 virgin: [-1.69, 14.89]
- **8 pregnant vs. 8 virgin:** $[16.35, 32.93]^*$
- **1 virgin vs. 8 virgin:** $[9.75, 26.33]^*$

95% Confidence Intervals for the Difference of Two Group Means

- Confidence intervals for the difference of two group means enables us to perform a formal hypothesis test
- Confidence intervals for groups means do have the same interpretation, without any formal inference

Effect Size

- Measure of how much practical importance a numerical difference might make in real life
- It is the ratio of mean difference to SD
- Can be done for
  - Single group: $D_i = \frac{\bar{y}_i - \bar{y}}{SD}$
  - Pair of groups: $D_{ij} = \frac{\bar{y}_i - \bar{y}_j}{SD}$
  - SD $= \sqrt{MSE}$

Fruit Flies: Effect Sizes - Single Group

- Control: $\frac{\bar{y}_1 - \bar{y}}{SD} = \frac{63.56 - 57.44}{14.81} = 0.41$
  - Estimate that the increase in mean lifetime for a male fruit fly when living alone is about 41% of the SD in lifetime.
- 1 pregnant: $\frac{64.80 - 57.44}{14.81} = 0.50$ - Increase by 50%
- 8 pregnant: $\frac{63.36 - 57.44}{14.81} = 0.40$ - Increase by 40%
- 1 virgin: $\frac{56.76 - 57.44}{14.81} = -0.04$ - Decrease by 4%
- 8 virgin: $\frac{38.72 - 57.44}{14.81} = -1.26$ - Decrease by 126%

Fruit Flies: Effect Sizes - Pair of Groups

- Control vs 1 pregnant: $\frac{\bar{y}_1 - \bar{y}_2}{SD} = \frac{63.56 - 64.80}{14.81} = -0.08$
  - Effect size is 8%. Difference between environments hardly matters
- Control vs. 8 virgin: $\frac{63.56 - 38.72}{14.81} = 1.68$
  - Effect size is 168%. Difference between environment matters.
- 1 virgin vs. 8 virgin: $\frac{56.76 - 38.72}{14.81} = 1.22$
  - Effect size is 122%. Difference between environment matters.

Effect of Sample Size

- For large samples
  - value of the t statistic will be higher, which results in a smaller p-value
    - higher confidence in mean difference
  - Margin of error will be smaller
    - more precise interval estimates
  - Sample size has a lesser impact on effect size
    - Doesn't change much
- In One-way ANOVA we need to consider all three aspects
  - When sample size smaller, estimate of effect size becomes important

Summary

- With One-way ANOVA analysis we try to answer following questions

1. Is there an effect? or Is there a mean difference among groups? If 'YES'

- One-way ANOVA F-test

2. Which groups had different means from the rest and how precisely can we measure the group means?

- Confidence intervals for group means and mean difference

3. Is the effect size big enough to make an impact ?

- Effect sizes

Fisher's Least Significant Difference (LSD) / Multiple Comparisons

- An alternative method to find which mean differences led to the rejection of null hypothesis of no mean difference among groups
- Three step approach

1. Is the ANOVA F-test statistically significant: If 'YES'
2. Find the least significant difference (LSD):

$$LSD = t *_{df,\alpha} \cdot SE$$

$$LSD = t^*_{df,\alpha} \cdot \sqrt{MSE} \cdot \sqrt{1/n_i + 1/n_j}$$

3. If difference in group means $(\bar{y}_i - \bar{y}_j) > LSD$, there is a significant difference in group means

Fruit flies: Fisher's Least Significant Difference (LSD)

- Step 1: ANOVA F-test
  - p-value < 0.05

- Step 2: Calculate LSD

$$LSD = t^* \sqrt{MSE} \sqrt{1/n_i + 1/n_j} = 1.984 \times 14.81 \times 0.28 = 8.22$$

Fruit flies: Fisher's Least Significant Difference (LSD)

- Step 3: Compare LSD with differences in groups means

```
##    Treatment Group_Means
## 1       none       63.56
## 2 1 pregnant       64.80
## 3 8 pregnant       63.36
## 4   1 virgin       56.76
## 5   8 virgin       38.72
```

- 1 pregnant vs. 1 virgin: 64.80 - 56.76 $= 8.04 <$ LSD
  - No mean difference between the groups

- 1 virgin vs. 8 virgin: 56.76 - 38.72 $= 18.04 >$ LSD
  - Group means are different