BMEG 802 – Advanced Biomedical Experimental Design and Analysis

One Way (Between) Analysis of Variance (ANOVA)

Joshua G. A. Cashaback, PhD

Recap

- One-Way (Between) ANOVA
    - Linear Model Approach
    - Test normality and sphericity assumption
    - Multiple mean comparisons

Today

- Two-Way ANOVA
  - linear model approach
  - interpret main effects and interactions
  - follow up mean comparisons
- n-Way ANOVA
  - general concepts
  - limitations
- Kruskal-Wallis
  - 1-way between levels
    - nonparametric version of 1-way ANOVA

Two Factor Design

In the example above we have two factors

- Factor A (e.g., Drug) with 2 levels (e.g., drug vs. no drug)
- Factor B (e.g., Biofeedback) with 2 levels (e.g., biofeedback vs. no biofeedback)

Fully crossed design

- every level of factor A is tested with every level of factor B
- total # groups (cells) is a x b

We will see how to formulate in terms of model comparisons:

- main effect of A
- main effect of B
- interaction effect A x B

2-Way ANOVA

Same approach as before

1. write the equation for the full and restricted models
2. derive the equations for model error $E_{restricted}$ and E_{full}
3. derive the expressions for degrees of freedom $df_{restricted}$ and df_{full}
4. end up with an equation for the F ratio

The Full Model

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha \cdot \beta)_{jk} + \epsilon_{ijk}$$

- Y_{ijk} is an individual score in the jth level of factor A and the kth level of factor B (i indexes subjects within each (j,k) cell)
- μ is the overall mean of all cells
- α_j is the effect of the jth level of factor A
- β_k is the effect of the kth level of factor B
- $(\alpha \cdot \beta)_{jk}$ is the interaction effect of level j of A and level k of B

Hypothesis testing using Restricted Models

Two-Factor (A x B) design: 3 null hypotheses to be tested:

- main effect of A
- main effect of B
- interaction effect of A x B

We will formulate a separate restricted model for each hypothesis test

- each test will involve the same full model
- we will use the usual F test

$$F = \frac{(E_{restricted} - E_{full})/(df_{restricted} - df_{full})}{(E_{full}/df_{full})}$$

Main Effect of A

Full Model: $Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha \cdot \beta)_{jk} + \epsilon_{ijk}$

null hypothesis is that A main effect is zero.

- $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_n = 0$

Restricted Model: $Y_{ijk} = \mu + \beta_k + (\alpha \cdot \beta)_{jk} + \epsilon_{ijk}$

F-Statistic for Main Effect A

$$E_{full} = \sum_{j=1}^{a} \sum_{k=1}^{b} \sum_{i=1}^{n} (Y_{ijk} - \bar{Y}_{jk})^2$$

$$df_{full = a \cdot b(n-1)}$$

$$E_{restricted} - E_{full} = nb \sum_{j=1}^{a} (\bar{Y}_j - \bar{Y})^2$$

$$df_{restricted} - df_{full} = a - 1$$

see Maxwell Delaney, Kelley (Chapter 7) for derivations

Now we can do our F-test!

$F = \frac{(E_{restricted} - E_{full})/(df_{restricted} - df_{full})}{(E_{full}/df_{full})}$

Main Effect of B

Full Model: $Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha \cdot \beta)_{jk} + \epsilon_{ijk}$

null hypothesis is that B main effect is zero.

- $H_0 : \beta_1 = \beta_2 = ... = \beta_n = 0$

Restricted Model: $Y_{ijk} = \mu + \alpha_j + (\alpha \cdot \beta)_{jk} + \epsilon_{ijk}$

F-Statistic for Main Effect B

$$E_{full} = \sum_{j=1}^{a} \sum_{k=1}^{b} \sum_{i=1}^{n} (Y_{ijk} - \bar{Y}_{jk})^2$$

$$df_{full=a \cdot b(n-1)}$$

$$E_{restricted} - E_{full} = nb \sum_{k=1}^{b} (\bar{Y}_k - \bar{Y})^2$$

$$df_{restricted} - df_{full} = a - 1$$

Now we can do our F-test!

$F = \frac{(E_{restricted} - E_{full})/(df_{restricted} - df_{full})}{(E_{full}/df_{full})}$

- note: denominator of F-test is the same as mean-square within from ANOVA table.

Interaction Effect of A x B

Full Model: $Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha \cdot \beta)_{jk} + \epsilon_{ijk}$

Restricted Model: $Y_{ijk} = \mu + \alpha_j + \beta_k + \epsilon_{ijk}$

F-Statistic for A x B Interaction

$$E_{full} = \sum_{j=1}^{a} \sum_{k=1}^{b} \sum_{i=1}^{n} (Y_{ijk} - \bar{Y}_{jk})^2$$

$$df_{full=a \cdot b(n-1)}$$

$$E_{restricted} - E_{full} = n \sum_{j=1}^{a} \sum_{k=1}^{b} (\bar{Y}_{jk} - \bar{Y}_j - \bar{Y}_k + \bar{Y})^2$$

$$df_{restricted} - df_{full} = (a-1)(b-1)$$

Now we can do our F-test!

$F = \frac{(E_{restricted} - E_{full})/(df_{restricted} - df_{full})}{(E_{full}/df_{full})}$

- don't worry, I won't make you do this by hand...

2-Way ANOVA Example

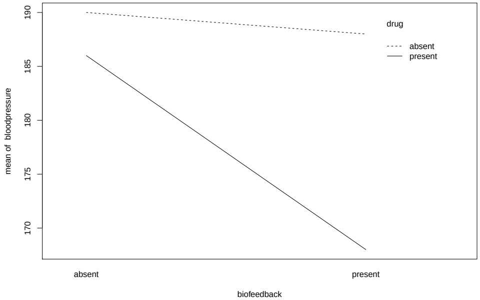
Hypothetical study: explore effects of biofeedback and drug therapy on blood pressure

- two independent variables: drug therapy and biofeedback

Biofeedback + Drug	Biofeedback, no Drug	no Biofeedback, Drug	no Biofeedback, no Drug
158	188	186	185
163	183	191	190
173	198	196	195
178	178	181	200
168	193	176	180
mean = 168	mean = 188	mean = 186	mean = 190
sd = 7.91	sd = 7.91	sd = 7.91	sd = 7.91

2-Way ANOVA Example

```r
bloodpressure <- c(158,163,173,178,168,188,183,198,178,193,
    186,191,196,181,176,185,190,195,200,180)
biofeedback <- factor(c(rep("present",10),rep("absent",10)))
drug <- factor(rep(c(rep("present",5),rep("absent",5)),2))
bpdata <- data.frame(bloodpressure, biofeedback, drug)
interaction.plot(biofeedback, drug, bloodpressure)
```



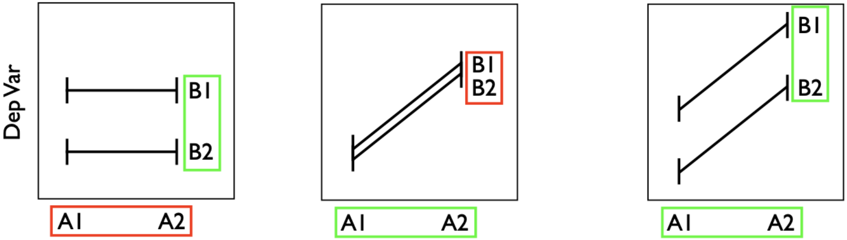
2-Way ANOVA Example

```r
myanova <- aov(bloodpressure ~ biofeedback*drug)
summary(myanova)
```

```
##                  Df Sum Sq Mean Sq F value  Pr(>F)
## biofeedback       1    500   500.0    8.00 0.01211 *
## drug              1    720   720.0   11.52 0.00371 **
## biofeedback:drug  1    320   320.0    5.12 0.03792 *
## Residuals        16   1000    62.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Significant interaction of drug, biofeedback, main effect of drug, main effect of biofeedback! But how do we interpret and perform followup mean comparisons?

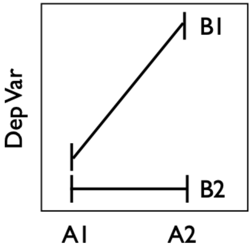
Main Effects



Main effect of B **Main effect of A** **Main effects of A and B**

in all 3 cases: **no A x B interaction effect**

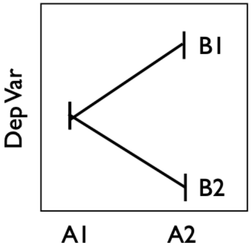
Interactions



A:

B:

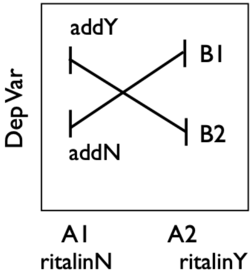
AxB:



A:

B:

AxB:

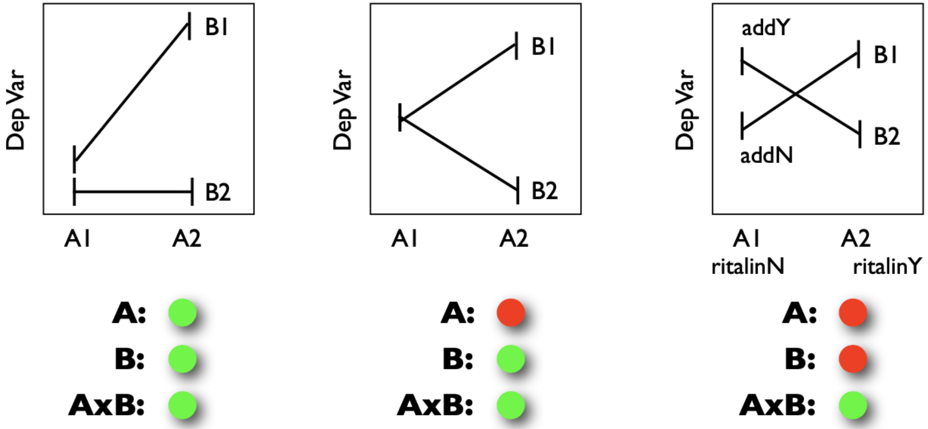


A:

B:

AxB:

Interactions



Rule of Thumb: parallel lines = main effect, non-parallel lines = interaction effect

Next Week

- Factorial (2-way, 3-way, etc.) ANOVA
- Kruskal Wallis