Practical Guide to Analysis of Variance (ANOVA)

Core Knowledge Seminar
Wednesday, 30 March 2022

Taehoon Ha

Contents

Example Data: Diet and Ovariectomy (Ovx)

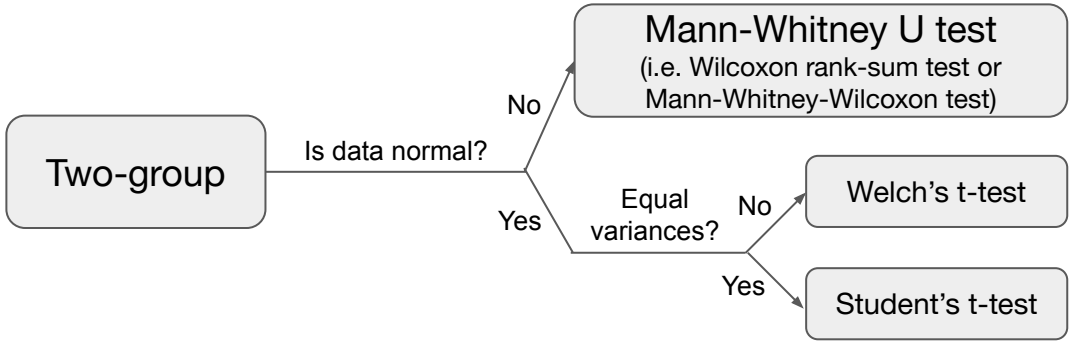
- **Study Design:**
  - Each mouse is an experimental unit
  - Balanced and complete randomized factorial design
  - **Treatment**: **OVX** (shamOVX vs. OVX) and **Diet** (low fat vs. high fat) – **4 groups**
    - Combine the levels of the factors into one categorical variable
  - **Outcome**: mouse weight (continuous variable)

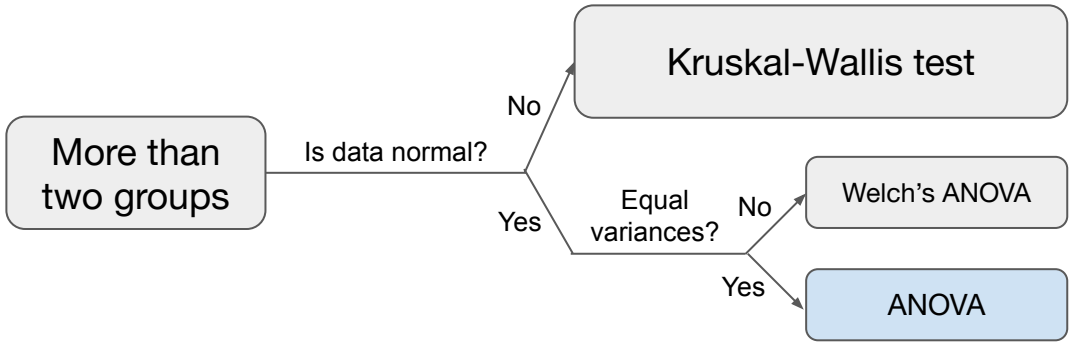
	LowFat	LowFat.OVX	HighFat	HighFat.OVX
N	10	10	10	10

  - **Goal**: To identify whether there is a group(s) that has a significantly different mean weight.

One-way ANOVA

Where We Stand: to compare continuous data in multiple independent groups





- **Assumptions**:
  - Errors should be random and independent
  - Normality
  - Homogeneity of variance
- If assumptions violated,
  - Transform your data and see if they meet assumptions
  - If still violated, try non-parametric approach **(Kruskal-Wallis test)**

Fisher's Solution: ANOVA

- **Idea:** Instead of doing multiple pairs of comparisons, why don't we do a single test?
  - This test will tell us whether there is difference in any of the means.
  - We do multiple comparisons between pairs **only after** we know there is difference in means across the groups.

- **Hypotheses**:
  - H_0: All group means are the same. (H_0: $\mu_1 = \mu_2 = \ldots = \mu_p$)
  - H_a: At least one group mean is different.

- **Process:**
  - ($p>\alpha$) fail to reject H_0 → all group means are the same → No further investigation

  - ($p<\alpha$) reject H_0 → At least one group mean is different → Post-hoc analysis (i.e., pairwise comparison) to identify which group(s) mean(s) are significantly different.

Step by Step of One-way ANOVA

1. Combine the levels of the factors into one categorical variable (**Diet** & **OVX**)
2. Linear regression fitting to check group means
3. One-way ANOVA
4. Post-hoc analysis to assess hypotheses of interest
5. Model assumption assessment
6. Analysis with additional methods to improve the model

One-way ANOVA: Cell Means Model in R

```
cellmeans_model <- lm(MouseWt ~ GroupName - 1, data = dat.work)
summary(cellmeans_model)
```

```
## Call:
## lm(formula = MouseWt ~ GroupName - 1, data = dat.work)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -5.220 -1.123 -0.080  1.298  6.310
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## GroupNameLF     21.0800     0.7139   29.53   <2e-16 ***
## GroupNameLF.OVX 24.2200     0.7139   33.93   <2e-16 ***
## GroupNameHF     26.9900     0.7139   37.81   <2e-16 ***
## GroupNameHF.OVX 33.0300     0.7139   46.27   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.257 on 36 degrees of freedom
## Multiple R-squared:  0.9936, Adjusted R-squared:  0.9929
## F-statistic:  1398 on 4 and 36 DF,  p-value: < 2.2e-16
```

- In R, 1 means Intercept.
- Hence, -1 means intercept-free model, which is "**cell means model**".

Each coefficient indicates each group mean directly

One-way ANOVA

```
anova(cellmeans_model)
```

```
## Analysis of Variance Table
##
## Response: MouseWt
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## GroupName  4 28504.2  7126.0  1398.3 < 2.2e-16 ***
## Residuals 36   183.5     5.1
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
```

As p<0.001, we can conclude that at leat one group mean is different.

Now, we are wondering which group mean(s) are different → **Post-hoc analysis**

One-way ANOVA: Post-hoc Analysis

```
library(multcomp)
aov.cellmeans <- aov(MouseWt ~ GroupName -1, data = dat.work)
summary(glht(aov.cellmeans, linfct=mcp(GroupName="Tukey")), test = adjusted("none"))
```

```
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = MouseWt ~ GroupName - 1, data = dat.w
##
## Linear Hypotheses:
##                                Estimate Std. Error t value Pr(>|t|)
## HighFat.OVX - HighFat == 0        6.04       1.01   5.983 7.33e-07 ***
## LowFat - HighFat == 0            -5.91       1.01  -5.854 1.09e-06 ***
## LowFat.OVX - HighFat == 0        -2.77       1.01  -2.744  0.00941 **
## LowFat - HighFat.OVX == 0       -11.95       1.01 -11.837 5.73e-14 ***
## LowFat.OVX - HighFat.OVX == 0    -8.81       1.01  -8.726 2.07e-10 ***
## LowFat.OVX - LowFat == 0          3.14       1.01   3.110  0.00365 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- none method)
```

Depending on your hypotheses:

- Compare all possible pairs: “Tukey”
- Compare control to each treatment group: “Dunnett”
- Or Customize the contrast matrix K

One-way ANOVA: Contrast Matrix

```
K <- rbind("OVX effect in LF"      = c(-1, 1, 0, 0),
           "OVX effect in HF"      = c(0, 0, -1, 1),
           "HF effect in shamOVX" = c(-1, 0, 1, 0),
           "HF effect in OVX"      = c(0, -1, 0, 1),
           "OVX effect"            = c(-1, 1, -1, 1),
           "HF effect"             = c(-1, -1, 1, 1),
           "OVX HF Interation"     = c(1, -1, -1, 1))
summary(glht(aov.cellmeans, linfct=mcp(GroupName=K)), test = adjusted(type="none"))
```

If you are interested in a specific treatment effect, we can identify the effect of interest by designing and inputting a contrast matrix

One-way ANOVA: Contrast Matrix (cont'd)

```
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: User-defined Contrasts
##
## Fit: aov(formula = MouseWt ~ GroupName - 1, data = dat.work)
##
## Linear Hypotheses:
##                             Estimate Std. Error t value Pr(>|t|)
## OVX effect in LF == 0          3.140      1.010   3.110  0.00365 **
## OVX effect in HF == 0          6.040      1.010   5.983 7.33e-07 ***
## HF effect in shamOVX == 0      5.910      1.010   5.854 1.09e-06 ***
## HF effect in OVX == 0          8.810      1.010   8.726 2.07e-10 ***
## OVX effect == 0                9.180      1.428   6.430 1.86e-07 ***
## HF effect == 0                14.720      1.428  10.310 2.73e-12 ***
## OVX HF Interation == 0         2.900      1.428   2.031  0.04967
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- none method)
```

Significance Level and Multiple Comparisons

- **Family-wise error rate (FWER)**: Probability of having at least one false positives (i.e.,Type I error) in multiple comparisons
  - When comparing more than 2 group means, using significance level of α, what is the probability of making at least one wrong decisions?

- FWER for different number of comparisons given different significance levels:

	1	**3**	**6**	**10**	**15**	**21**	**28**	**36**	**45**
0.05	0.05	0.14	0.26	0.4	0.54	0.66	0.76	0.84	0.90
0.01	0.01	0.03	0.06	0.1	0.14	0.19	0.25	0.30	0.36

One-way ANOVA: Post-hoc Analysis

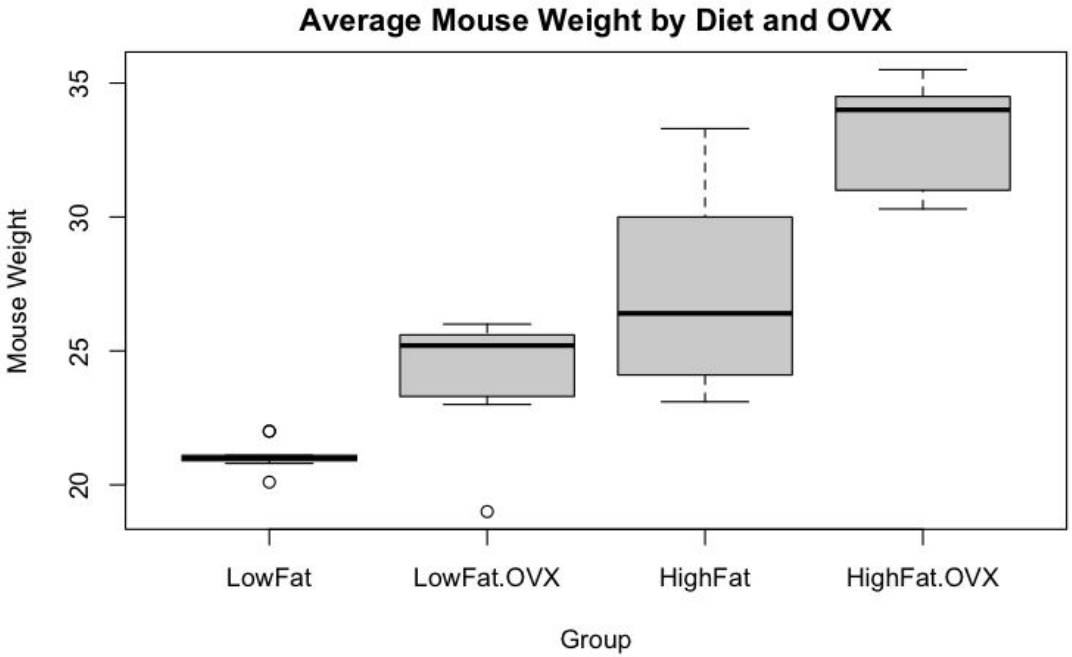
```
library(multcomp)
aov.cellmeans <- aov(MouseWt ~ GroupName -1, data = dat.work)
summary(glht(aov.cellmeans, linfct=mcp(GroupName="Tukey")), test = adjusted("none"))
```

In case you need p-value adjustment due to the multiple comparison, here we can select p-value adjustment method. For more, check out **`?multcomp::adjusted`** in R.

```
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = MouseWt ~ GroupName - 1, data = dat.work)
##
## Linear Hypotheses:
##                                Estimate Std. Error t value Pr(>|t|)
## HighFat.OVX - HighFat == 0        6.04       1.01   5.983 7.33e-07 ***
## LowFat - HighFat == 0            -5.91       1.01  -5.854 1.09e-06 ***
## LowFat.OVX - HighFat == 0        -2.77       1.01  -2.744  0.00941 **
## LowFat - HighFat.OVX == 0       -11.95       1.01 -11.837 5.73e-14 ***
## LowFat.OVX - HighFat.OVX == 0    -8.81       1.01  -8.726 2.07e-10 ***
## LowFat.OVX - LowFat == 0          3.14       1.01   3.110  0.00365 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- none method)
```

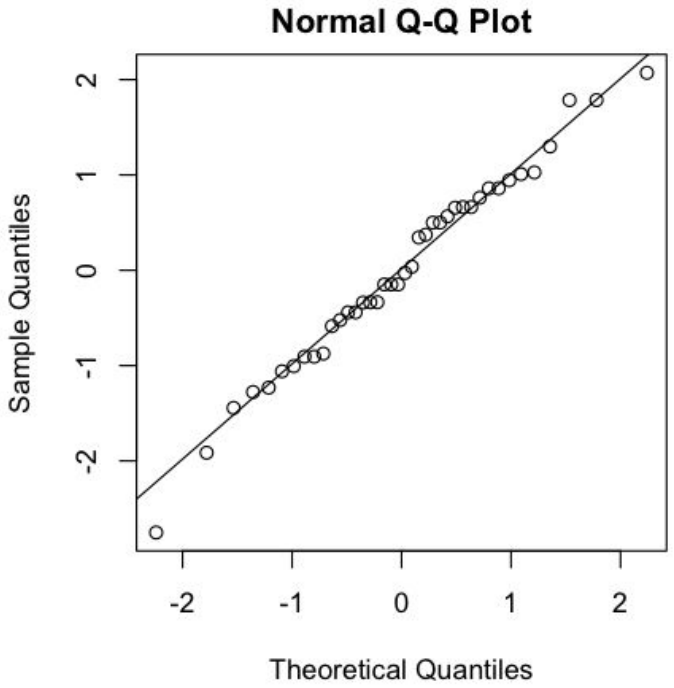
One-way ANOVA: Boxplot

```
boxplot(MouseWt ~ GroupName, data = dat.work,
        xlab = "Group", ylab = "Mouse Weight",
        main = "Average Mouse Weight by Diet and OVX")
```

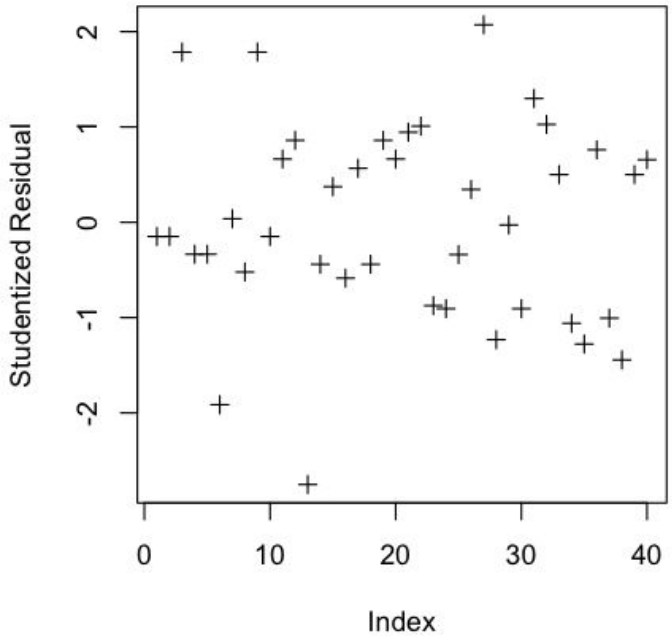


Model Diagnostics - Normality & Equal Variances

```
qqnorm(rstudent(cellmeans_model))
qqline(rstudent(cellmeans_model))
```




```
plot(1:nrow(dat.work), rstudent(cellmeans_model), pch=3,
     xlab="Index", ylab="Studentized Residual")
```



Two-way ANOVA

Two-way ANOVA: Basic Idea

- **Idea:** An extension of one-way ANOVA to the two factor setting
- **Process:**
  - Partition the total variation into 3-4 parts
    - Variation due to treatment factor1 (Diet)
    - Variation due to treatment factor2 (OVX)
    - Variation due to interaction between factor1 and factor2 (Diet and OVX) - optional
    - Variation due to random error
  - Compare each of the first three parts with the fourth part
- Two-way ANOVA does NOT have strong rationale with FWER control.
  - Allows control for FWER for each factor separately.

Step by step of two-way ANOVA

1. Linear regression fitting to check group means
2. Two-way ANOVA for assessing main effects and interaction effects (optional)
3. Post-hoc analysis to assess hypotheses of interest
4. Model assumption assessment
5. Analysis with additional methods to improve the model

Two-way ANOVA: effects model in R

```
effects_model <- lm(MouseWt ~ Diet * OVX, data = dat.work)
summary(effects_model)
```

```
Call:
lm(formula = MouseWt ~ Diet * OVX, data = dat.work)

Residuals:
   Min     1Q Median     3Q    Max
-5.220 -1.123 -0.080  1.298  6.310

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        21.0800     0.7139  29.529  < 2e-16 ***
DietHighFat         5.9100     1.0096   5.854 1.09e-06 ***
OVXOVX              3.1400     1.0096   3.110  0.00365 **
DietHighFat:OVXOVX  2.9000     1.4277   2.031  0.04967 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.257 on 36 degrees of freedom
Multiple R-squared:  0.8083, Adjusted R-squared:  0.7923
F-statistic: 50.59 on 3 and 36 DF,  p-value: 5.42e-13
```

This time, intercept was included, which is “**effects model**.”

In effects model,

- the intercept indicates the **reference group’s mean**.
- Here, LowFat and shamOVX group.
- Other coefficients show the mean difference between each of other groups and LowFat and shamOVX group.

Two-way ANOVA

```
anova(effects_model)
```

```
## Analysis of Variance Table
##
## Response: MouseWt
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## Diet       1 541.70  541.70 106.2948 2.731e-12 ***
## OVX        1 210.68  210.68  41.3411 1.859e-07 ***
## Diet:OVX   1  21.02   21.02   4.1256   0.04967 *
## Residuals 36 183.46    5.10
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
```

- Interaction term is significant
- Conclusion: there is a significant association between **Diet** and **Ovx**.

Two-way ANOVA: Contrast Matrix

```
K2 <- rbind("LF.OVX - LF"                = c(0,0,1,0),
            "HF.OVX - HF"                = c(0,0,1,1),
            "HF - LF"                    = c(0,1,0,0),
            "HF.OVX - LF.OVX"            = c(0,1,0,1),
            "LF.OVX + HF.OVX - LF - HF" = c(0,0,2,1),
            "HF + HF.OVX - LF - LF.OVX" = c(0,2,0,1),
            "HF.OVX - HF - LF.OVX + LF" = c(0, 0, 0, 1))

summary(glht(effects_model, linfct=K2), test=adjusted(type="none"))
```

Likewise, we can input a contrast matrix depending on your hypotheses.

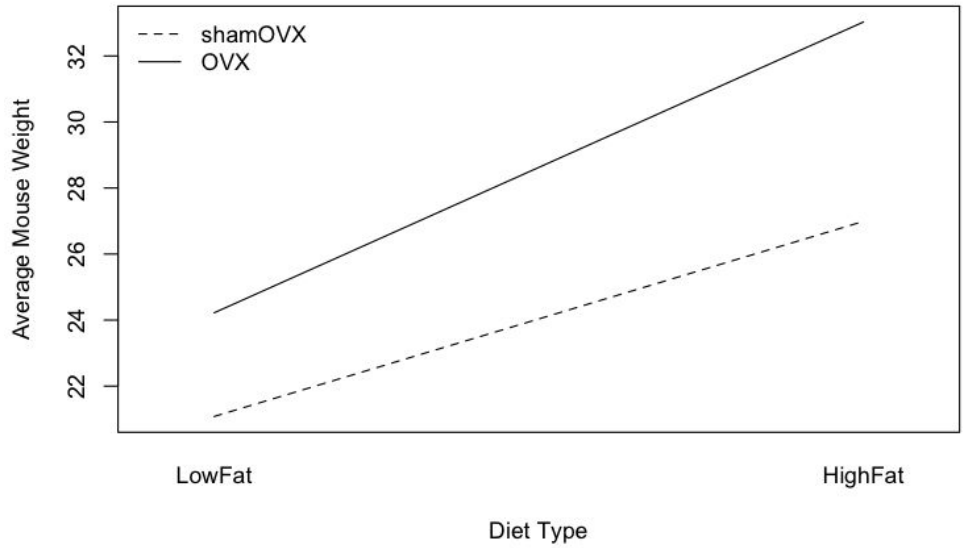
However, this time the contrast matrix is different as we used effects model instead of cell means model.

Two-way ANOVA: Contrast Matrix

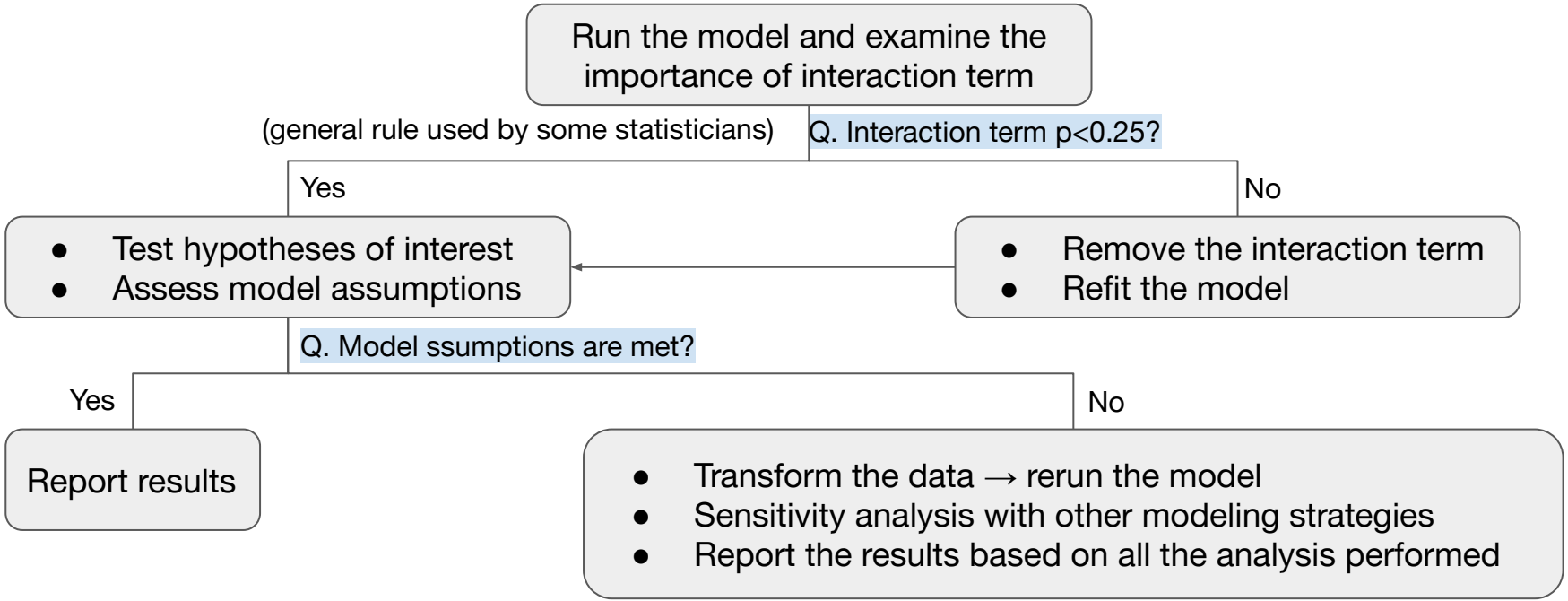
```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Fit: lm(formula = MouseWt ~ Diet * OVX, data = dat.work)
##
## Linear Hypotheses:
##                                  Estimate Std. Error t value Pr(>|t|)
## LF.OVX - LF == 0                    3.140      1.010   3.110  0.00365 **
## HF.OVX - HF == 0                    6.040      1.010   5.983 7.33e-07 ***
## HF - LF == 0                        5.910      1.010   5.854 1.09e-06 ***
## HF.OVX - LF.OVX == 0                8.810      1.010   8.726 2.07e-10 ***
## LF.OVX + HF.OVX - LF - HF == 0      9.180      1.428   6.430 1.86e-07 ***
## HF + HF.OVX - LF - LF.OVX == 0     14.720      1.428  10.310 2.73e-12 ***
## HF.OVX - HF - LF.OVX + LF == 0      2.900      1.428   2.031  0.04967 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- none method)
```

Two-way ANOVA: Interaction Plot

```
interaction.plot(dat.work$Diet, dat.work$OVX, dat.work$MouseWt,
                 xlab="Diet Type", ylab="Average Mouse Weight",
                 legend=F, lty=2:1)
legend("topleft", legend=levels(dat.work$OVX), lty=2:1, bty="n")
```



Interaction Term Analysis Process



- Keeping interaction term when there is no interaction → reduced efficiency in estimation
- Dropping interaction term when there is interaction → biased main treatment effect estimation

Partial F-test

Partial F-test: Basic Idea

- **When to Use:** To compare model fitting performances between a complex model and a simpler subset model.
  - e.g. $Y \sim b_0 + b_1X_1 + b_2X_2 + b_3X_3$ vs. $Y \sim b_0 + b_1X_1 + b_2X_2$
  - **Nested models**: A complex model should include all predictors that a simpler model has.
  - Cannot use to compare non-nested models → Use AIC, BIC, or Vuong's test instead.

- **Hypotheses**:
  - H_0: Simpler (subset) model with predictors p_{simple} is better.
  - H_a: Complex model with predictors $p_{complex}$ is better.

- **Process:**
  - ($p>\alpha$) fail to reject H_0 → Simpler (i.e., subset or reduced) model is better.
  - ($p<\alpha$) reject H_0 → Complex (i.e., full) model is better

Partial F-test in R

```
reduced_model <- lm(MouseWt ~ Diet, data = dat.work)
full_model <- lm(MouseWt ~ Diet + OVX + Diet * OVX, data = dat.work)
anova(reduced_model,full_model)
```

```
## Analysis of Variance Table
##
## Model 1: MouseWt ~ Diet
## Model 2: MouseWt ~ Diet + OVX + Diet * OVX
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     38 415.17
## 2     36 183.46  2    231.71 22.733 4.129e-07 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’
```

As p<0.05, there is enough evidence that either **OVX** or **Diet * OVX interaction term** are statistically significant.

In other words, complex (full) model is better than the simpler (reduced) model.

Repeated Measures and Mixed Effects Model

Repeated Measures ANOVA and Mixed-Effects Model

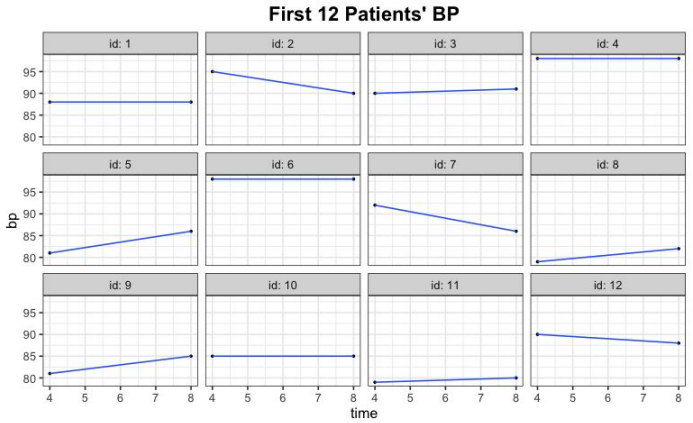
- **Idea:** to compare means across one or more variables that are based on repeated observations.

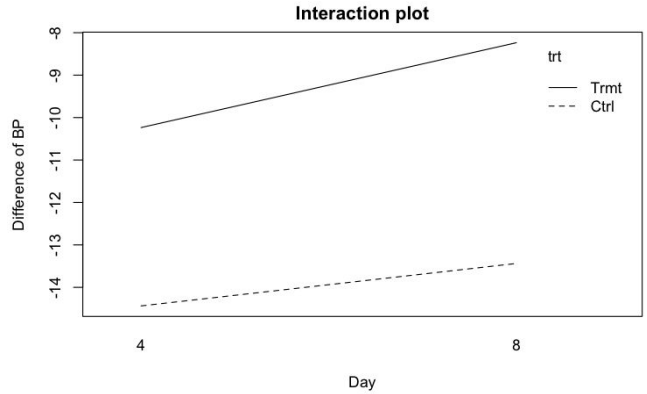
- **Common issues:**
  - **NA values**: If there is a missing value, you'd need to ignore all data for that sample → Solution: Imputation
  - What if we cannot do imputation? Any alternative approach?

- **Alternative: Mixed Effects Model**
  - Mixed effects model consists of two parts:
    - Fixed effect: when you test for variation among the means of the particular groups
    - Random effect: individual sample effects (animals, participants, rounds …) are considered random

Example Data: Blood Pressure

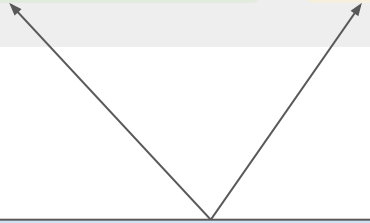
- **Study Design:**
  - Total 66 patients blood pressure measured over time
  - **Factors**: **Drug** (control vs. treatment) and **Time** (Day4 vs. Day8)
  - **Outcome**: blood pressure (continuous)
  - **Issue**: Patients' bp change over time differs from sample to sample
  - **Solution**: Treat drug and time effects fixed and randomize sample effects





Mixed Effects Model in R

```
blood_fit <- lmer(d ~ trt + time + trt * time + (1|id), data = blood_data)
summary(blood_fit)
```



Green: fixed effects (treatment and time effect and the interaction of between treatment and time)

Yellow: random effects (individaul sample effect randomized)

Mixed Effects Model in R (cont'd)

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [lmerModLmerTest]
## Formula: d ~ trt + time + trt * time + (1 | id)
##    Data: blood_data
## REML criterion at convergence: 862
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.92294 -0.44315  0.05264  0.36653  2.14310
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept) 58.00    7.616
##  Residual             14.87    3.857
## Number of obs: 132, groups:  id, 66
##
## Fixed effects:
##               Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)   -14.4375     1.5091  78.3627  -9.567 8.24e-15 ***
## trtTrmt         4.2022     2.1026  78.3627   1.999   0.0491 *
## time8           1.0000     0.9642  64.0000   1.037   0.3036
## trtTrmt:time8   1.0000     1.3434  64.0000   0.744   0.4594
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtTrm time8
## trtTrmt     -0.718
## time8       -0.319  0.229
## trtTrmt:tm8  0.229 -0.319 -0.718
```

Between sample variation is much bigger than within sample variation (almot 4 times)

- BP treatment effect was significant (P=0.049)
- Time effect was not significant (P=0.303)
- Interaction between treatment and time was not significant (P=0.459), thereby no significant effect of a new drug over time.

Biostatistics Support

Cold Spring Harbor Laboratory

Biostatistics Services Provided

1. Office Hours: Thursday 2-4pm - calendly.com/cshlbiostat
2. Research Collaborations - Model Development
3. Research Data Analysis
4. Biostatistical Support Letters
5. Study Design and Power Calculations
6. Review/Writing of Methods Sections

Questions?