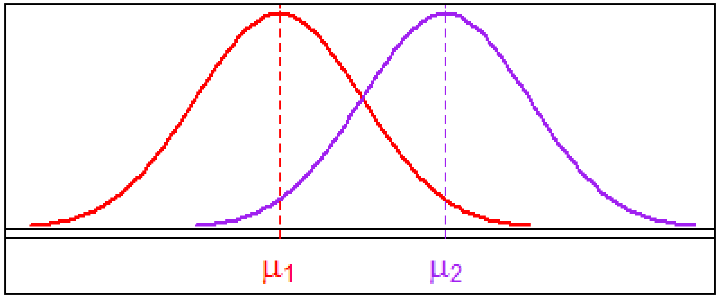
Analysis of Variance
Part 2: Single Factor Studies

STAT 705: Regression and Analysis of Variance

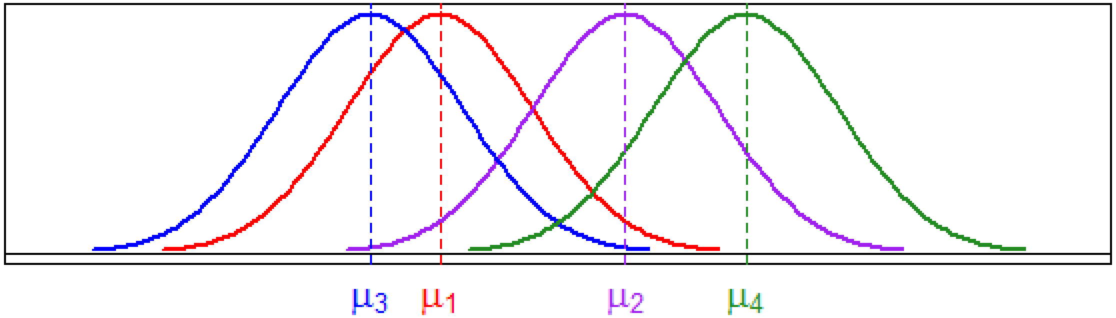
Introduction

- Recall the two-sample t test
  - Two groups (populations)
  - Y is some measured characteristic
  - For both groups, Y follows a normal distribution



Test H_0: $\mu_1 = \mu_2$ vs. H_a: $\mu_1 \neq \mu_2$

- What if there are more than two groups?



Are all of the means equal to each other?

Compare the Means of >2 Groups

- Suppose we have t groups, and we want to compare their means
- This is comparable to an experiment that has one factor with t levels (i.e., t treatments)
- We want to test if the mean response is the same for <u>all</u> the treatments

$$H_0: \mu_1 = \mu_2 = \mu_3 = . . . = \mu_t$$

vs. H_a: at least one mean is different

- This is an **Analysis of Variance (ANOVA)** hypothesis

Analysis of VARIANCE??

- We are comparing *means*
- How can we do this by comparing *variances*?
- We illustrate with an example

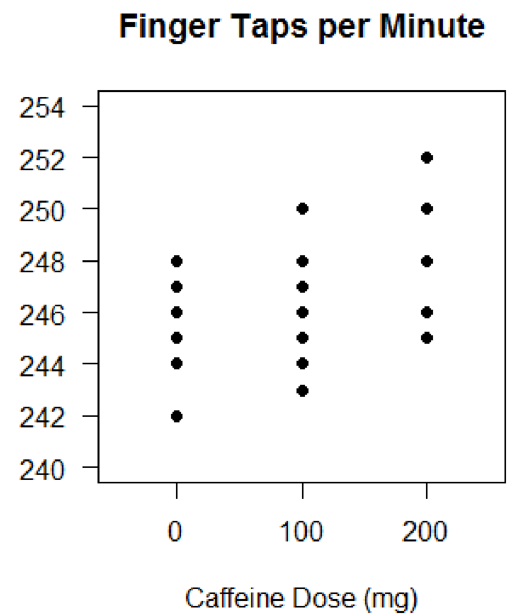
Example: Effect of Caffeine

In an effort to determine the effect of caffeine on muscle and nervous systems, the following experiment was conducted. Thirty male college students were recruited and randomly assigned to consume one of three dosages of caffeine: 0 mg, 100 mg or 200 mg. Two hours later, the student repeatedly taps his finger on the table, and the researcher records the number of taps per minute. The resulting data is shown in the table on the next slide.

DOES CAFFEINE AFFECT THE MEAN NUMBER OF TAPS?

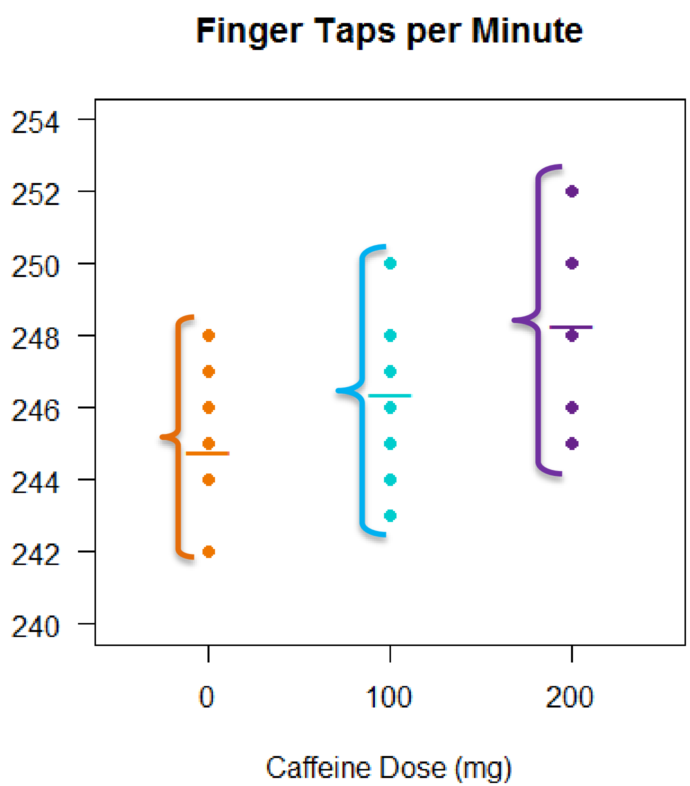
Caffeine Data and Scatterplot

0 mg	100 mg	200 mg
242	248	246
245	246	248
244	245	250
248	247	252
247	248	248
248	250	250
242	247	246
244	246	248
246	243	245
242	244	250
Source: Draper and Smith, 1981		



- Scatterplot appears 'stacked' because values for the predictor are fixed in advance.
- Some points are identical, so all are not visible on the graph.

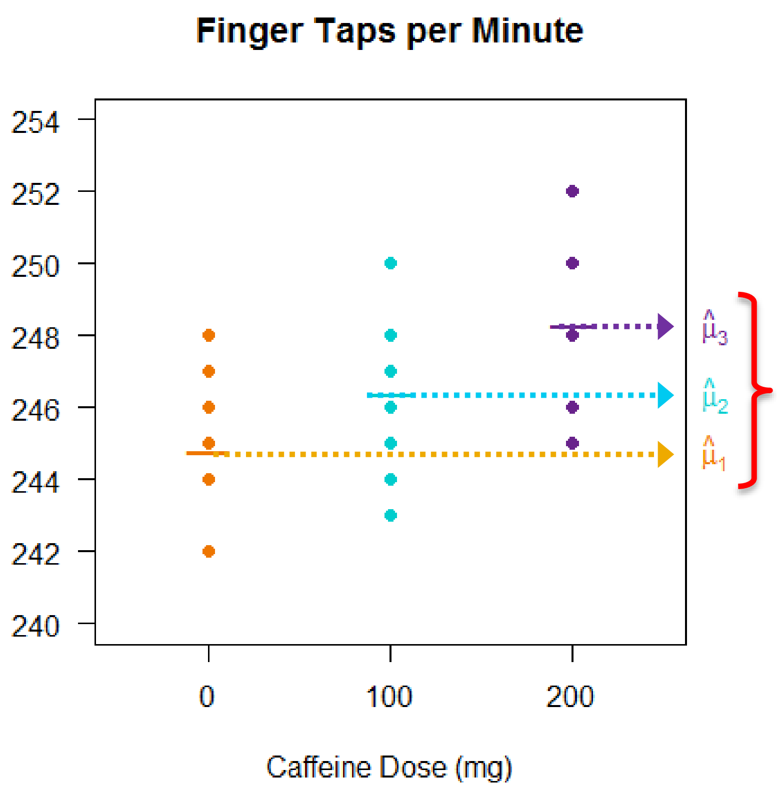
Within Group Variability



Each treatment group has a mean and variability around the mean.

If we combine the variability of these groups, we get the within group variability.

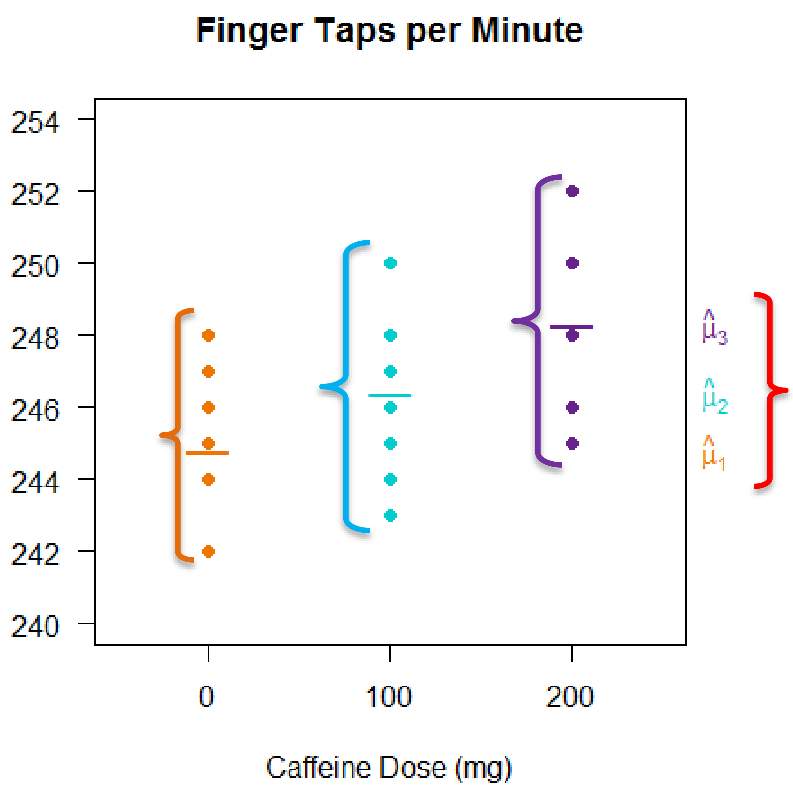
Between Group Variability



There is also variability among the means of the three groups.

This is the between group variability.

The Big Question



Is the variability in treatment means 'large' relative to the variability within treatments?

We answer this question with an F test, like we did with regression.

Intuitive Concept

- We do not expect the sample means to all be identically equal
  - There is always sampling variability associated with values calculated from a sample
- If the variability among the means (i.e. between groups) is much larger than the variability within groups, then
  - The sample means are relatively widespread
  - It is believable that the true population means are different
- However, if the variability between groups is about the same as the variability within groups, then
  - The sample means are NOT more widespread
  - It is believable that the true population means are all the same

Some Notation

n_i = sample size for i^{th} treatment

y_{ij} = observed value for the j^{th} subject in the i^{th} treatment

$\bar{y}_{i\bullet}$ = sample mean for i^{th} treatment

$\bar{y}_{\bullet\bullet}$ = "grand mean", i.e. the mean of all the observations

s_i^2 = sample variance for i^{th} treatment

N = total number of observations

Table layout:

Dose 0 mg	Y_{11}	Y_{12}	…	$Y_{1,10}$	$\bar{Y}_{1\bullet}$	s_1^2
Dose 100 mg	Y_{21}	Y_{21}	…	$Y_{2,10}$	$\bar{Y}_{2\bullet}$	s_2^2
Dose 200 mg	Y_{31}	Y_{32}	…	$Y_{3,10}$	$\bar{Y}_{3\bullet}$	s_3^2

observed values　　sample means & variances

Calculate SSE

For Each Treatment				
Dosage	Subscript	Sample Size	Mean	Variance
0 mg	1	10	244.8	5.73
100 mg	2	10	246.4	4.27
200 mg	3	10	248.3	4.90

$$\text{Dosage 0mg, } SS_1 = (n_1 - 1)s_1^2 = (9)5.73 = 51.57$$

$$\text{Dosage 100mg, } SS_2 = (n_2 - 1)s_2^2 = (9)4.27 = 38.43$$

$$\text{Dosage 200mg, } SS_3 = (n_3 - 1)s_3^2 = (9)4.90 = 44.10$$

$$\text{Total SS Within Treatments} = SS_1 + SS_2 + SS_3 = 134.1$$

SS Within = SSE = 134.1

Calculate SS Treatments

For Each Treatment				
Dosage	Subscript	Sample Size	Mean	Variance
0 mg	1	10	244.8	5.73
100 mg	2	10	246.4	4.27
200 mg	3	10	248.3	4.90

$$\text{grand mean} = \bar{Y}_{\bullet\bullet} = \frac{(10)244.8 + (10)246.4 + (10)248.3}{30} = 246.5$$

Total variation in the treatment means

$$= n_1(\bar{Y}_{1\bullet} - \bar{Y}_{\bullet\bullet})^2 + n_2(\bar{Y}_{2\bullet} - \bar{Y}_{\bullet\bullet})^2 + n_3(\bar{Y}_{3\bullet} - \bar{Y}_{\bullet\bullet})^2$$

$$= (10)\left(244.8 - 246.5\right)^2 + (10)\left(246.4 - 246.5\right)^2 + (10)\left(248.3 - 246.5\right)^2 = 61.4$$

SS between = SS Treatments = 61.4

Degrees of Freedom & Mean Squares

- Degrees of freedom
  - t = number of treatments
  - df Treatments = t – 1 = 3 – 1 = 2
  - df Error = N – t = 30 – 3 = 27
- Mean Square = SS / df
  - MS Treatments
    - » (SS Treatments) / (df Treatments) = 61.4 / 2 = 30.7
  - MS Error = MSE
    - » (SS Error) / (df Error) = 134.1 / 27 = 4.967

Replace 'Treatment' with 'Model' and this is exactly what we did with regression.

MS Treatments = 30.7 and MS Error = 4.967

Test Statistic for ANOVA F Test

$$F = \frac{\text{MS Between}}{\text{MS Within}} = \frac{\text{MS Treatments}}{\text{MS Error}} = \frac{\text{MSTrt}}{\text{MSE}}$$

- This follows an F distribution with numerator and denominator degrees of freedom (df Trt) and (df Error), respectively
- Large values indicate the means are NOT all equal
- For the caffeine data
  - F = 30.7 / 4.967 = 6.18 . . . is this “large”?
  - df = (2, 27)
  - From the F table, critical value is 3.35 (α=.05)

Decision and Conclusion

- We are testing

  H_0: $\mu_1 = \mu_2 = \mu_3$

  H_a: At least one mean is different
- Test statistic = F = 6.18; Critical value = 3.35
- Decision: Reject H_0 (F is larger than critical value)
- Conclusion:

  At significance level 0.05, the sample provides convincing evidence that the mean number of finger taps per minute is different for different caffeine doses.

ANOVA Table

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F	p-value
Treatment	t – 1	SS Treatment	MS Treatment	MS Trt / MSE	
Error	N – t	SS Error	MS Error		
Corrected Total	N – 1	SS Total			

This table should look familiar. It is the same ANOVA table we used with regression, except that

1) "Model" is now "Treatment", and
2) We are using the letter 't' (for number of treatments) instead of 'p' (for number of predictors)

The F statistic we just calculated 'by hand' and its associated p-value will be on the top row. As we did with regression, we will use software to get p-values.

What You Should Know

- Write the ANOVA hypotheses
- If you are given the sample sizes, means and variances for several treatments, you should be able to
  - calculate the ANOVA F statistic
  - perform the ANOVA hypothesis test
- Understand why the comparison of ‘within’ and ‘between’ group variability is valid to compare group means