Variance Estimates and the F Ratio

ERSH 8310

Lecture 3

September 2, 2009

Today's Class

- Completing the analysis (the ANOVA table)
- Evaluating the F ratio
- Errors in hypothesis testing
- A complete numerical example

The ANOVA Table

COMPLETING THE ANALYSIS

Completing the Analysis

- Table 3-1 presents an ANOVA table – look at the entries
- Note that variance is SS/df

Table 3.1: Summary of the analysis of variance

Source	Bracket Term	SS	df	MS	F
A	$[A] = \frac{\sum A_j^2}{n}$	$[A] - [T]$	$a - 1$	$\frac{SS_A}{df_A}$	$\frac{MS_A}{MS_{S/A}}$
S/A	$[Y] = \sum Y_{ij}^2$	$[Y] - [A]$	$a(n-1)$	$\frac{SS_{S/A}}{df_{S/A}}$	
Total	$[T] = \frac{T^2}{an}$	$[Y] - [T]$	$an - 1$		

Running Example

- To provide context, let's take the example shown in Chapter 2, Table 2.2 (p. 22)

Control	Drug A	Drug B
16	4	2
18	7	10
10	8	9
12	10	13
19	1	11

Do You Remember?

- Now, let's compute a few things from this table:
  - What are the treatment sums A_j for each group j?

$$A_j = \sum_{i=1}^{n_j} Y_{ij}$$

  - Using the treatment sums, what is the mean for each group?

$$\bar{Y}_j = \frac{A_j}{n_j} = \sum_{i=1}^{n_j} \frac{Y_{ij}}{n_j}$$

  - What is the grand sum?

$$T = \sum Y_{ij} = \sum A_j$$

Building Bracket Terms

- Recall how to build the bracket terms:
  1. Square all quantities in a given set (Y_{ij}, A_j, and T).
  2. Sum these squared quantities (if more than one are present)
  3. Divide this sum by the number of scores that went into each component
- So:

$$[Y] = \sum Y_{ij}^2$$

$$[A] = \frac{\sum A_j^2}{n}$$

$$[T] = \frac{T^2}{an}$$

Sums of Squares

- Recall also that we used the bracket terms to construct the Sums of Squares:
  - $SS_T = [Y] - [T]$
  - $SS_A = [A] - [T]$
  - $SS_{S/A} = [Y] - [A]$
- What are these for Table 2.2?

Degrees of Freedom

- Degree of freedom (df):
  - The number of independent scores in a Sum of Squares

- The general rule for computing the df of any sum of squares is:
  - df = (number of independent observations) -(number of restraints)

-or-

  - df = (number of independent observation) -(number of population estimates)

Degrees of Freedom

- The total df comes from the sum of the within and between df

  - $df_T = df_A + df_{S/A} = N - 1$

    -or-

  - $(a)(n)-1 = (a-1)+a(n-1)$.

DF Example

- Table 2.2 (p. 22) is replicated to the right

Control	Drug A	Drug B
16	4	2
18	7	10
10	8	9
12	10	13
19	1	11

- What are the total df for this table?

- Bonus:
  - What is df_A?
  - What is $df_{S/A}$?

Mean Squares

- The mean squares are given by: MS = SS/df.
- So far, we have two mean squares:

$$MS_A = \frac{SS_A}{df_A} \qquad MS_{S/A} = \frac{SS_{S/A}}{df_{S/A}}$$

- Compute the Mean Squares for Table 2.2

The F Ratio

- The formula of the F ratio is $F = MS_A / MS_{S/A}$

- The degrees of freedom are:
  - $df_A = df_1 = a-1$: (numerator df)
  - $df_{S/A} = df_2 = a(n-1)$: (denominator df)

- The F ratio is approximately 1.0 when the null hypothesis is true and is greater than 1.0 when the null hypothesis is false

- What is the F ratio for Table 2.2?

EVALUATING THE F RATIO

Sampling Distributions

- Recall from your previous statistics course the idea of a sampling distribution
- A frequency distribution of a statistic is called a sampling distribution of the statistic
- It comes from replications of the same experiment over and over, with replacement for the sample
  - Each time, you record the mean.

Sampling Distribution Example

- To demonstrate sampling distributions, I would like your help and the help of our TAs

1. Please write down your height in inches on a small square of paper
2. Ask 5 different (and random) people for their heights
3. Compute the mean and median of your sample
4. Write it next to your height on your sheet
5. Hand that to the back where our TAs are
6. Take a 10 minute break

Survey Says…

- Using your data, we computed the following:

  - The true mean height (in inches) for our overall class

  - The distribution of each of your sample means

Sampling Distributions

- The distribution of our sample means is called a sampling distribution

- Its shape follows well known statistical properties
  - Here, we have the central limit theorem in play

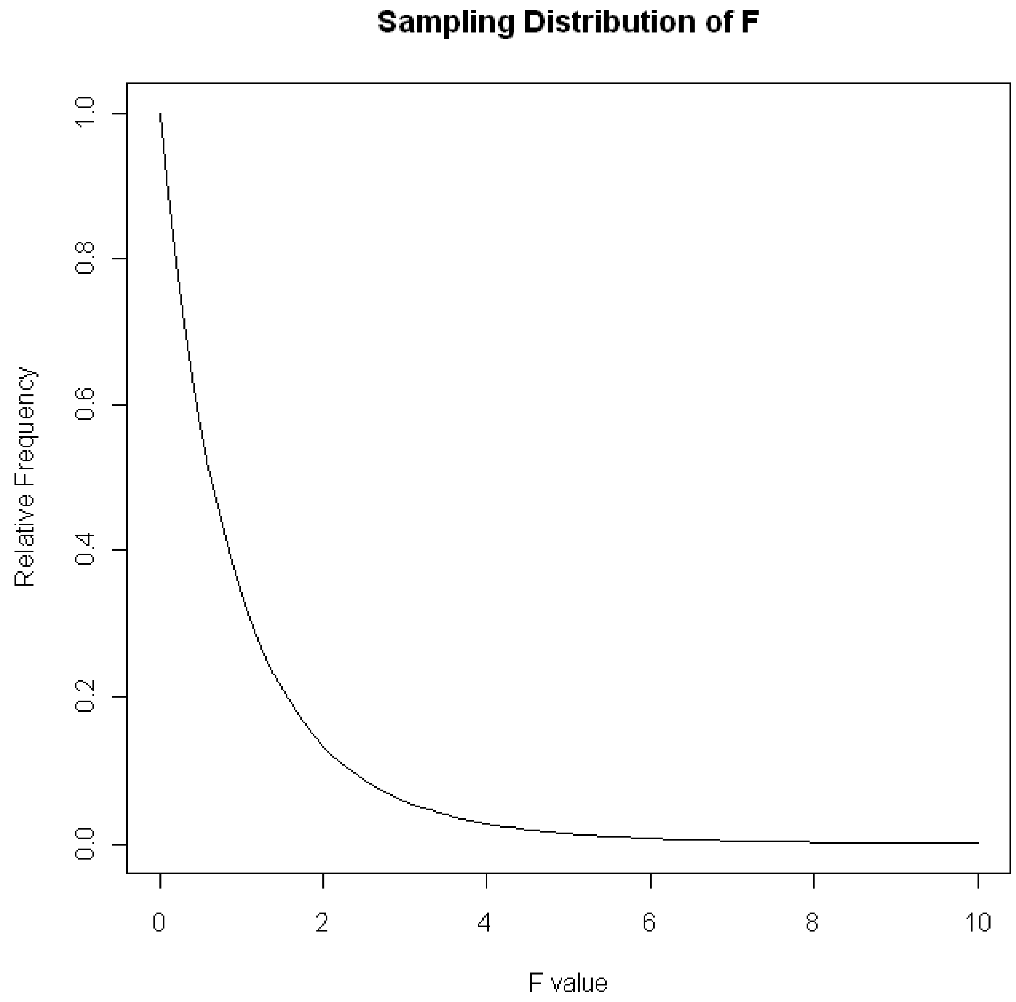
- Any statistic can have a sampling distribution
  - Here is the sampling distribution for the median

- What we will construct is a sampling distribution for the F-ratio when the null hypothesis is true

The Sampling Distribution of F

- Suppose that for an experiment a = 3, n = 5 (or N = 15), and the null hypothesis is true so that $\mu_1 = \mu_2 = \mu_3$
- Assume that we draw a very large number of such experiments each consisting of three groups of 5 scores, and that we compute the value of F for each case
- We can construct a graph relating F and frequency of occurrence [see Figure 3-1 for F(2,12)]

Theoretical Distribution



The F Table

- The shape of the F distribution is determined by the degrees of freedom for the ratio
- We will use $F(df_{num.}, df_{denom.})$ or $F_{df1,df2}$
- An F table is found in Appendix A.1 (see pp. 571-575)
- Also, you can compute exact p-values in Excel
  - Type "=fdist(F,df1, df2)" and you will get a p-value

Using the F Table

- A particular value of F in this table is specified by three factors:
  - (1) the numerator df
  - (2) the denominator df
  - (3) the value of α
    - Where α refers to the proportion of area to the right of an ordinate drawn at F_a
    - The α levels, α = .10, .05, .025, .01, and .001, are ones most commonly encountered, and these are listed in the table

The Distribution of F When the Null Hypothesis Is False

- The theoretical distribution when H_0 is false is called the noncentral F (i.e., F') distribution

- This is used frequently to calculate power
  - Used to figure out how big of a sample you need to detect a difference for a given effect size

- Very useful in planning research
  - Always consult a statistician first!

Testing the Null Hypothesis

- The two hypotheses are:

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

H_1: not all μ's are equal.

- Assume that we:
  - Conducted an experiment
  - Computed the value of F

- Suppose we could agree on a dividing line for any F distribution, where values of F falling above the line are considered to be unlikely (i.e., incompatible with H_0) and values of F falling below the line as considered to be likely (i.e., compatible with H_0)

Decision Rules

- One decision rule is to reject the null hypothesis when the observed F falls within the region of incompatibility
- In practice, there is fairly common agreement on a probability of $\alpha = .05$ to define the region of incompatibility for the F distribution
  - Doesn't always have to be this level, though

Decision Rules

- This probability may be called the significance level. The rejection rule is:
  - Reject H_0 when $F_{observed} \geq F_{(\alpha)}(df_{num.}, df_{denom.})$; retain H_0 otherwise

- It used to be that any reports would say that an F is significant at a given level of significance (e.g., $\alpha = .01$)

- In modern applications a researcher reports the exact probability of $F_{observed}$, the p-value

Decision Rules, Continued

- The p-value refers to the proportion of the sampling distribution of the F statistic fall at or above the F found in an experiment

- We do not need to consult an F table to determine significance,
  - Simply compare the exact probability with the chosen significance level
  - Reject H_0 if the exact probability is smaller than the chosen significance level

- For example, provided that $\alpha = .05$, reject H_0 if $p \leq .05$ and retain H_0 otherwise

Avoiding Common Misuses of Hypothesis Testing (p. 45-46)

- Never lose site of your data.
- Remember that not all null hypotheses represent interesting or plausible outcomes of the experiment and that their rejection is not inevitably informative
- Remember that the null-hypothesis test only gives information about how likely the sampling operations are to have produced your effect
- When you find a significant effect, ask what it means
- Always interpret your results in the context of other studies in the field

ERRORS IN HYPOTHESIS TESTING

Errors in Hypothesis Testing

- If we reject H_0 when it is true, then we make a type I error
- If we retain H_0 when it is false, then we make a type II error
- Note that α is the probability of type I error, and that β is the probability of type II error (see Table 3-3 on page 47)
- Power refers to the probability of rejecting the null hypothesis when an alternative hypothesis is true (i.e., power = $1-\beta$)

A COMPLETE NUMERICAL EXAMPLE

Vigilance Task While Sleep Deprived

- There are a = 4 conditions, namely, 4, 12, 20, and 28 hours without sleep
- There are n = 4 subjects randomly assigned to each of the different treatment conditions
- The vigilance task score represents the number of failures to spot objects on a radar screen during a 30-minute test period

Data (computation p. 51)

Hours without sleep			
4 hr	12 hr	20 hr	28 hr
a_1	a_2	a_3	A_4
37	36	43	76
22	45	75	66
22	47	66	43
25	23	46	62

Construct the ANOVA Table

Source	SS	Df	MS	F	P-value
A	3314.25	3	1104.75	7.34	0.005
S/A	1805.50	12	150.458		
Total	5119.75	15			

Do You Remember?

- Now, let's compute a few things from this table:

  - What are the treatment sums A_j for each group j?

$$A_j = \sum_{i=1}^{n_j} Y_{ij}$$

  - Using the treatment sums, what is the mean for each group?

$$\bar{Y}_j = \frac{A_j}{n_j} = \sum_{i=1}^{n_j} \frac{Y_{ij}}{n_j}$$

  - What is the grand sum?

$$T = \sum Y_{ij} = \sum A_j$$

Building Bracket Terms

- Recall how to build the bracket terms:
  1. Square all quantities in a given set (Y_{ij}, A_j, and T).
  2. Sum these squared quantities (if more than one are present)
  3. Divide this sum by the number of scores that went into each component
- So:

$$[Y] = \sum Y_{ij}$$

$$[A] = \frac{\sum A_j^2}{n}$$

$$[T] = \frac{T^2}{an}$$

Sums of Squares

- Recall also that we used the bracket terms to construct the Sums of Squares:
    - $SS_T = [Y] - [T]$
    - $SS_A = [A] - [T]$
    - $SS_{S/A} = [Y] - [A]$

Degrees of Freedom

- The total df comes from the sum of the within and between df

  - $df_T = df_A + df_{S/A}$

    -or-

  - $(a)(n)-1 = (a-1)+a(n-1).$

Mean Squares

- The mean squares are given by: MS = SS/df.

- So far, we have two mean squares:

$$MS_A = \frac{SS_A}{df_A} \quad MS_{S/A} = \frac{SS_{S/A}}{df_{S/A}}$$

The F Ratio

- The formula of the F ratio is $F = MS_A / MS_{S/A}$

- The degrees of freedom are:
  - $df_A = df_1 = a-1$: (numerator df)
  - $df_{S/A} = df_2 = a(n-1)$: (denominator df)

But Now What…

- You have computed the ANOVA
  - And received a p-value
- What were your hypotheses?
  - Null:
  - Alternative
- What is your decision?
- What does that mean?

Wrapping Up…

- The F-Ratio is the vehicle to test the null hypothesis (es) of the experiment
- The F-Ratio is composed of two variances: between groups and within groups.
- The ratio follows the idea of variance partitioning that will be present throughout the class
- Don’t get too wrapped up into the formulae presented today: most every calculation will be done using a computer

Up Next

- Lab:
  - Sampling distributions of F
  - How to calculate ANOVA models in SPSS
  - Good clean fun
- Homework
  - Assigned tonight – due at the start of class on Wednesday
- Next week:
  - Read Chapters 4 and 5