Lecture 31: more Chapter 11, Section 3
Categorical & Quantitative Variable
More About ANOVA

- ANOVA: Hypotheses, Table, Test Stat, *P*-value
- 1st Step in Practice: Displays, Summaries
- ANOVA Output
- Guidelines for Use of ANOVA

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production (discussed in Lectures 1-4)
  - Displaying and Summarizing (Lectures 5-12)
  - Probability (discussed in Lectures 13-20)
  - Statistical Inference
    - 1 categorical (discussed in Lectures 21-23)
    - 1 quantitative (discussed in Lectures 24-27)
    - cat and quan: paired, 2-sample, several-sample
    - 2 categorical
    - 2 quantitative

ANOVA Null and Alternative Hypotheses

H_0 : explanatory C & response Q *not* related

- Equivalently, $H_o : \mu_1 = \mu_2 = \cdots = \mu_I$

(difference among sample means just chance)

H_a : explanatory C & response Q *are* related

- Equivalently, H_a: not all the μ_i are equal

(difference too extreme to be due to chance)

Depending on formulation, the word "not" appears in Ho or Ha.

Example: *How to Refute a Claim about "All"*

- **Background**: Reader asked medical advice columnist: "Dear Doctor, does everyone with Parkinson's disease shake?" and doctor replied: *All patients with Parkinson's disease do not shake.*
- **Question:** Is this what the doctor meant to say?
- **Response:**

Example: *ANOVA Alternative Hypothesis*

- **Background**: Null hypothesis to test for relationship between race (3 groups) and earnings:

$$H_o: \mu_1 = \mu_2 = \mu_3$$

- **Question:** Is this the correct alternative?

$$H_a: \mu_1 \neq \mu_2 \neq \mu_3$$

- **Response:**

Words are better: say "   ".

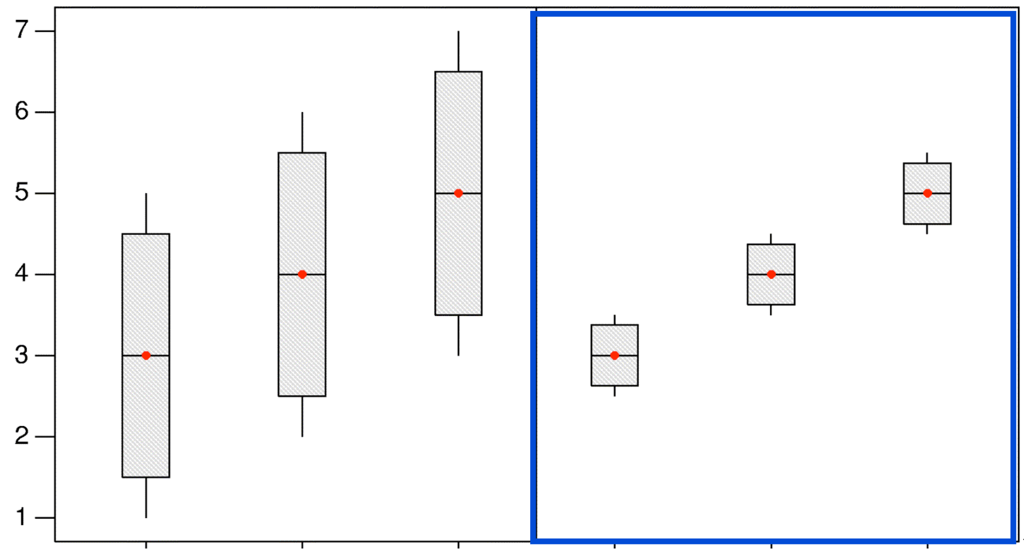
The F Statistic *(Review)*

$$F = \frac{\left[n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + \cdots + n_I(\bar{x}_I - \bar{x})^2\right] / (I-1)}{\left[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_I - 1)s_I^2\right] / (N - I)}$$

- **Numerator:** variation among groups
  - How different are $\bar{x}_1, \cdots, \bar{x}_I$ from one another?
- **Denominator:** variation within groups
  - How spread out are samples? (sds $s_1, \cdots, s_I$)

Role of Variations on Conclusion *(Review)*

Boxplots with same variation *among* groups (3, 4, 5) but different variation *within*: sds large (left) or small (right)



Scenario on right: smaller s.d.s → larger $F = \frac{var\ among}{var\ within}$ →smaller *P*-value→ likelier to reject H_0 →conclude pop means differ

ANOVA Table

Source	Degrees of Freedom	Sum of Squares	Mean Sum of Squares	F	P
Factor	$DFG = I - 1$	SSG	$MSG = SSG/DFG$	$F = \frac{MSG}{MSE}$	p-value
Error	$DFE = N - I$	SSE	$MSE = SSE/DFE$		
Total	$N - 1$	SST			

- Organizes calculations
  - "Source" refers to source of variation
  - DF: use I = no. of groups, N = total sample size
  - SSG measures overall variation among groups
  - SSE measures overall variation within groups
  - Mean Sums: Divide Sums by DFs
  - F: Take quotient of MSG and MSE
  - P-value: Found with software or tables

Example: *Key ANOVA Values*

- **Background**: Compare mileages for 8 sedans, 8 minivans, 12 SUVs; find SSG=42.0, SSE=181.4.
- **Question:** What are the following values for table:
  - **DFG? DFE? MSG? MSE? *F*?**
- **Response:**
  - **DFG** = 3 – 1 =
  - **DFE** = $N - I$ = (8+8+12) – 3 =
  - **MSG =** SSG/DFG = 42/2 =
  - **MSE**= SSE/DFE = 181.4/25 =
  - ***F***= MSG/MSE = 21/7.256 =

Practice: 11.29e-g p.561

Example: *Completing ANOVA Table*

- **Background**: Found these values for ANOVA:
  - **DFG**=3-1= 2
  - **DFE**=*N*-*I*=(8+8+12)-3= 25
  - **MSG**=SSG/DFG=42/2= 21
  - **MSE**=SSE/DFE=181.4/25= 7.256
  - ***F***=MSG/MSE=21/7.256= 2.89
- **Question:** Complete ANOVA table?
- **Response:** Software→*P*-val=0.0743→

Source	DF	SS	MS	F	P
Factor		42			
Error		181.4			

Practice: 11.32 p.563

ANOVA *F* Statistic and *P*-Value

- Sample means very different→
  F large→
  P-value small→
  Reject claim of equal population means.
- Sample means relatively close→
  F not large→
  P-value *not* small→
  Believe claim of equal population means.

How Large is "Large" *F*

Particular *F* distribution determined by DFG, DFE

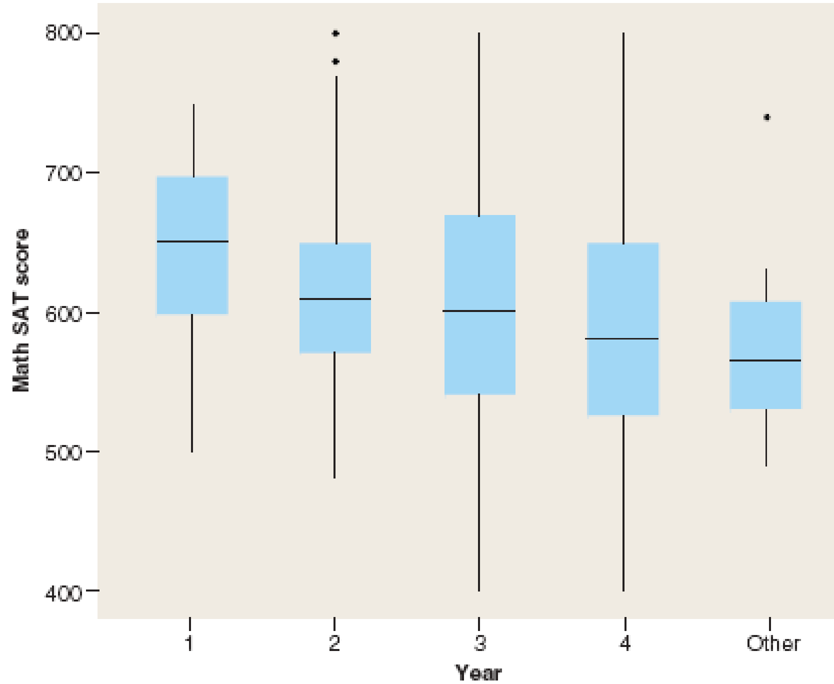
(these determined by sample size, number of groups)

P-value in software output lets us know if *F* is large.

Note: P-value is "bottom line" of test; "top line" is examination of display and summaries.

Example: *Examining Boxplots*

- **Background**: For all students at a university, are Math SATs related to what year they're in?



- **Question:** What do the boxplots suggest?
- **Response:** As year goes up, mean .
(Suggests students scored better in Math.)

Practice: 5.8b-c p.146

Example: *Examining Summaries*

- **Background**: For all students at a university, are Math SATs related to what year they're in?

Level	N	Mean	StDev
1	32	643.75	63.69
2	233	613.91	61.00
3	87	601.84	89.79
4	28	581.79	89.73
other	10	578.00	72.08

- **Question:** What do the summaries suggest?
- **Response:** Means decrease by about ______ points for each successive year 1 to 4. Standard deviations are around ______ and sample sizes are ______.

Practice: 11.29a-b p.561

Example: *ANOVA Output*

- **Background**: For all students at a university, are Math SATs related to what year they're in?

Analysis of Variance for Math

Source	DF	SS	MS	F	P
Year	4	78254	19563	3.87	0.004
Error	385	1946372	5056		
Total	389	2024626			

- **Question:** What does the output suggest?
- **Response:** Test H_o :

  P-value=0.004. Small? ____ Reject H_0? ____

  Conclude all 5 population means may be equal?

  Year and Math SAT related in population?

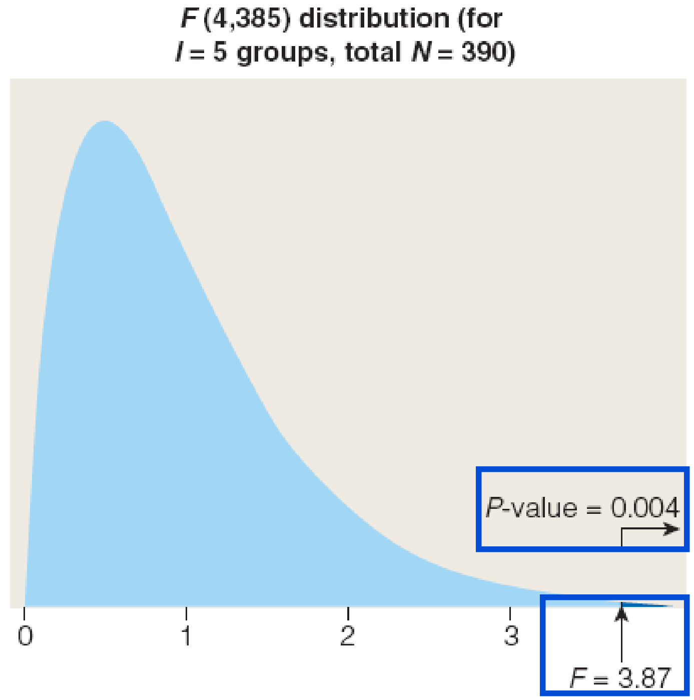
How Large is "Large" *F (Review)*

Particular *F* dist determined by DFG, DFE

(these determined by sample size, number of groups)

P-value in software output lets us know if *F* is large.

P-value = 0.004 → *F* = 3.87 is large (in given situation)



Example: *ANOVA Output*

- **Background**: A test for a relationship between Math SAT and year of study, based on data from a large sample of intro stats students at a university, produced a large *F* and a small *P*-value.
- **Question:** What issues should be considered before we use these results to draw conclusions about the relationship between year of study and Math SAT for all students at that university?
- **Response:**

Practice: 11.34 p.563

Guidelines for Use of ANOVA Procedure

- Need random samples taken independently from several populations.
- Confounding variables should be separated out.
- Sample sizes must be large enough to offset non-normality of distributions.
- Need populations at least 10 times sample sizes.
- Population variances must be equal.

Pooled Two-Sample *t* Procedure *(Review)*

If we can assume $\sigma_1 = \sigma_2$, standardized difference between sample means follows a pooled t distribution.

- Some apply Rule of Thumb: use pooled *t* if larger sample s.d. not more than twice smaller.

The F distribution is in a sense "pooled": our standardized statistic follows the F distribution only if population variances are equal (same as equal s.d.s)

Example: *Checking Standard Deviations*

- **Background**: For all students at a university, are Math SATs related to what year they're in?

Level	N	Mean	StDev
1	32	643.75	63.69
2	233	613.91	61.00
3	87	601.84	89.79
4	28	581.79	89.73
other	10	578.00	72.08

- **Question:** Is it safe to assume equal population variances?
- **Response:**

Largest s.d.=______ > 2(smallest s.d.) ?

Assumption of equal variances OK?

Example: *Reviewing ANOVA*

- **Background**: For all students at a university, are Verbal SATs related to what year they're in?

Level	N	Mean	StDev
1	32	596.25	86.91
2	234	592.76	65.87
3	86	596.51	77.26
4	29	579.83	79.47
other	10	551.00	124.32

Source	DF	SS	MS	F	P
Year	4	23559	5890	1.10	0.357

- **Questions:** Are conditions met? Do the data provide evidence of a relationship?
- **Response:** n_i large and 124.32 not > 2(65.87)→ *P*-val=0.357 small?____ Evidence of a relationship?

Guidelines for Use of ANOVA *(Review)*

- Need random samples taken independently from several populations
- Confounding variables should be separated out
- Sample sizes must be large enough to offset non-normality of distributions
- Need populations at least 10 times sample sizes
- Population variances must be equal.

Example: *Considering Data Production*

- **Background**: *F* test found evidence of relationship between Math SAT and year (*P*-value 0.004), but not Verbal SAT and year (*P*-value 0.357).
- **Question:** Keeping in mind that the sample consisted of students in various years taking an introductory statistics class, are there concerns about bias/confounding variables?
- **Response:** For Math, For Verbal,

Practice: 11.31 p.562

Lecture Summary

(Inference for Cat→Quan; More About ANOVA)

- ANOVA for several-sample inference
  - Formulating hypotheses correctly
  - ANOVA table
  - *F* statistic and *P*-value
- 1st step in practice: displays and summaries
  - Side-by-side boxplots
  - Compare means, look at sds and sample sizes
- ANOVA output
- Guidelines for use of ANOVA