PBIO 504
Introductory Biostatistics

Analysis of Variance (ANOVA)

Textbook chapter 12

Hypothesis Testing t-test

- When we want to compare means from two different groups, we use the Student t test (also called the 2 sample t test).
- Null hypothesis: ***H_0: $\mu_1 = \mu_2$***.
- The test determines whether the difference between the observed means $\bar{X}_1 - \bar{X}_2$ is too large to have occurred by chance.
- How large is large?

$$T = |\bar{X}_1 - \bar{X}_2| / S_p, \text{ and } \mathrm{P}_{H_0}[|\mathrm{T}| \geq \mathrm{t}] = 0.05$$

$$\text{When n is large, df=n}_1\text{+n}_2\text{-2, } \mathrm{P}_{H_0}[|\mathrm{T}| \geq z] = 0.05$$

t Tests for more than Two Groups

- Suppose we had three different groups

- We could test the following pairs of groups using a t-test for each pair:

- ***$mean_1$ vs $mean_2$***
- ***$mean_2$ vs $mean_3$***
- ***$mean_1$ vs $mean_3$***

  » However

Recall the Alpha Level

- We <u>reject</u> the null hypothesis if the p-value from the t test is less than ***α,*** the level of chance (probability) that we are willing to accept for wrongly rejecting the null hypothesis.

- If ***α*** is set at 0.05, it will have this value for <u>each and every</u> t-test we perform.

- So if there are 3 groups being compared in our study, there are 3 t tests, each carrying its own ***α*** of 0.05, and so the study wide ***α*** may become as large as 3 x 0.05 = 0.15

- ***Are we willing to accept such a level? No, see later.***

ANOVA

- ANOVA will address this problem by giving us a single p-value to test the hypothesis:

H_0: $mean_1 = mean_2 = mean_3 = \ldots = mean_k$

H_A: at least one of these means is different from the others

where **k** is the number of groups.

Some Features of ANOVA

- **ANOVA** (analysis of variance) gets its name from its ability to determine whether the variability within groups is greater or lesser than between groups. Clearly the latter outcome would be desirable.

- ANOVA incorporates multiple variables: the outcome variable is the dependent variable (which must be a continuous variable), and the explanatory variables are the independent variables. This total set of variables is called the model.

- If there are just 2 groups, the ANOVA p-value will be the same as the p-value of the 2-sample t test if there are no other variables in the model.

The ANOVA Model

- The general form:

  dependent variable (y) =
  =independent variables $(x_1 + x_2 + x_3 + \ldots + x_k)$

- An example would be:

  blood pressure = agegroup

- Blood pressure is being modeled as a function of age group, where group 1: 20-40 years old, group 2: 41-60 years old, and group 3: 61+ years old. This is an example of a one-way ANOVA because there is only one independent variable agegroup. Var agegroup has 3 levels.

One- Way ANOVA Table

Source	df	SS	MS	F	P
Model (m) (*between groups*)	df_m= K-1	SS_m	MS_m = SS_m / df_m	MS_m / MS_e	P-value from F distri-bution
Error (e) (*within groups*)	df_e= N-K	SS_e	MS_e= SS_e / df_e		
Total (*overall*)	df= N-1	SS_m + SS_e			

Note: N is the total sample size and K is the number of groups of the indep var. (in the previous ex age had 3 groups)

ANOVA Table

Source	df	SS	MS	F	P
Model (m) (*between groups*)	K-1	SS_m	SS_m / df_m	MS_m / MS_e	from F distri-bution with
Error (e) (*within groups*)	N-K	SS_e	SS_e / df_e		
Total (*overall*)	N-1	SS_m + SS_e			

Review Sums of Squares (SS)

- To calculate the total SS, we take the deviation of each observed data point from the mean of all points, square it, and sum all of these squared deviations.
- To find the F statistic we divide the mean square for the model MS_m to the mean square error Ms_e

- See Textbook page 289

The F Distribution

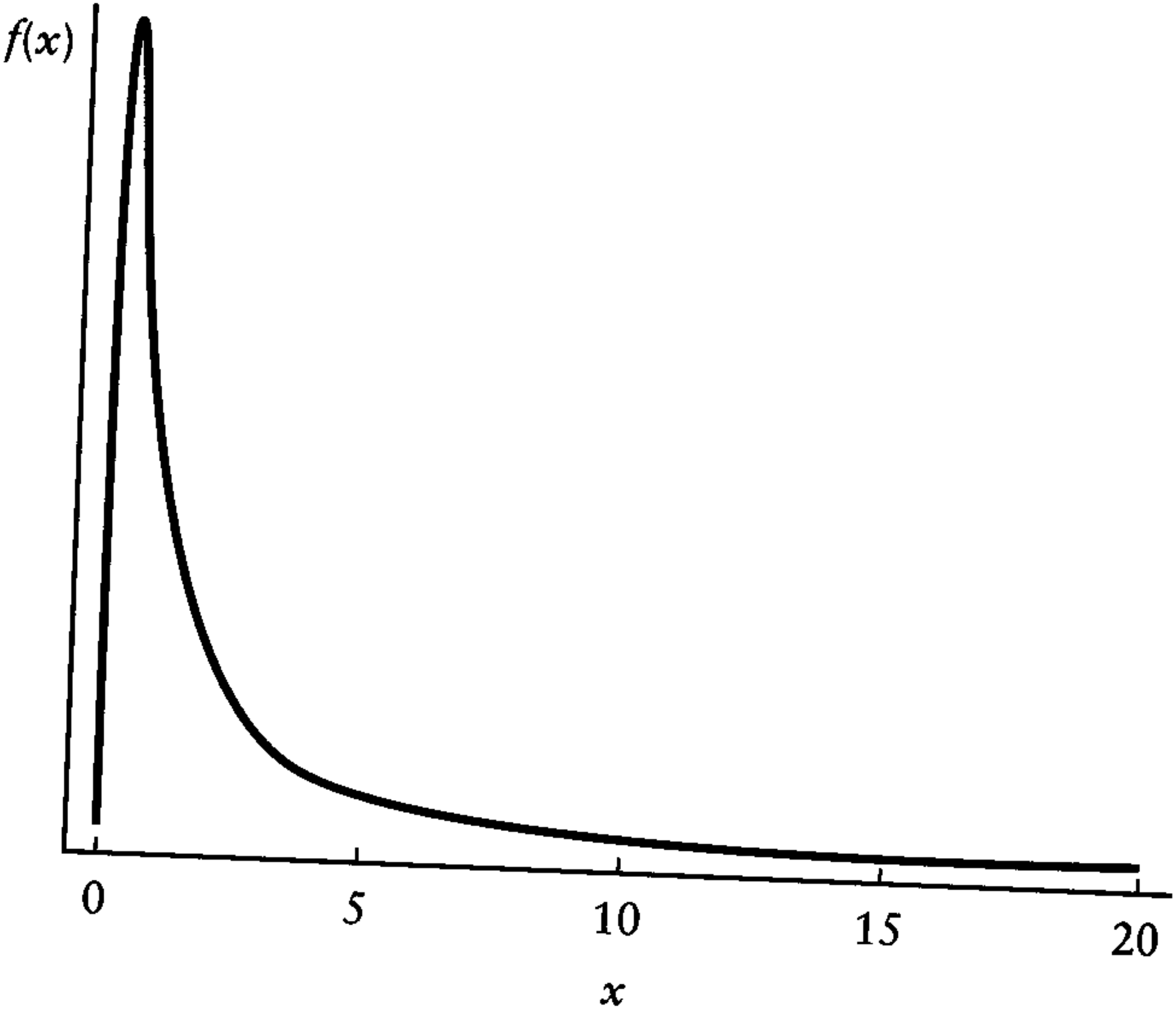
- F comes from the ratio of the between- groups variability and the within-groups variability (this is where ANOVA gets its name from, i.e. the statistical analysis of the variability in a study).

- Somewhat like the t distribution, there are many possible F distributions, but it is never negative.

- It is skewed to the right (see next figure).

- It depends on the df_m of the numerator and the df_e of the denominator.

The F Distribution for 4 and 2 df



ANOVA Assumptions

- Independent observations
- Normal distribution of the continuous dependent variable for each level of the categorical independent variable
- Homogeneity of variance across groups

One-Way ANOVA Example

- A new drug to treat acne is given to 35 subjects. It is tested against 2 drugs already on the market.
- Each person takes one of the drugs for 4 weeks, and the number of lesions is counted at the end of the study as the dependent variable.
- Model: lesions = type of drug
- Here lesions is the outcome and there is only one predictor variable type of drug (with three groups).

One-Way ANOVA Results

Source	df	SS	MS	F	P
Model	2	2133.66	1066.83	262.12	0.001
Error	32	130.30	4.07		
Total	34	2263.96			

This test statistic has an F distribution with 2 and 32 degrees of freedom. We obtain the p-value from this F distrib to be 0.001 and it means the probability of obtaining values **over** 262.12 assuming that Ho true.

P-val is signif and we can conclude that not all means are eq, or at least one differs from the other two.

2-Way and 3-Way ANOVA

- In a 2-way ANOVA model, there are two independent variables, and in a 3-way ANOVA model there are three independent variables

- 2-way ANOVA example:

  Blood pressure = agegroup + gender

- 3-way ANOVA example:

  Blood pressure = agegroup + gender + diabetes

Bonferroni Adjustment Method

- This test addresses the problem of multiple comparisons among the groups.

- If the overall model F test is statistically significant (e.g. $P<0.05$), then we reject the null hypothesis that all the groups' means are equal.

- However, it is often of interest to determine just which groups are different from the others.

- We don't want to inflate the alpha level, so we need a way to adjust the "pair-wise alpha" so that it keeps the "experiment-wide alpha" at 0.05

Bonferroni's Method

- First decide on the experiment-wide alpha, which is usually α = 0.05.
- Calculate the pair-wise alpha (a') that will be used for each comparison of means.
- $a' = 0.05 / c$ where ***0.05*** is the experiment-wide alpha, and c is the number of pairs of means that will be compared.

Bonferonni Example

- Suppose we are evaluating a study of mice where 3 different diets (A, B, and C) were given to independent groups of animals, and their weight was measured at the end of the experiment.

- The ANOVA model: weight = diet

- There are three possible diet comparisons
  (Number of pairs c is n(n-1)/2=3)
- Hence for each pair the new alpha is ***a*´= 0.05/3 =0.016**

- We'll reject the null hypothesis only if
  the p-value < 0.016

Another Bonferroni Example

- In this example, we will see how to apply the t -test to each pair of compared means, and using the Bonferroni Method to decide which actual mean differences are significant.

- A study was performed to determine which of 5 water temperature settings for the machinery would be the best for removing impurities from water at a waste water treatment plant.

- ANOVA Model: % impurities removed = temp. setting

- Alpha was set at 0.10 for this experiment.

Water Data

Temperature level				
I (low)	II (l-m)	III (med)	IV (m-h)	V (high)
40%	36%	49%	47%	55%
45%	42%	51%	49%	60%
42%	38%	53%	51%	62%
48%	39%	53%	52%	63%
50%	37%	52%	50%	59%
51%	40%	50%	51%	61%

Mean: **46.0** **38.7** **51.3** **50.0** **60.0**

ANOVA Table for this Example

Source	df	SS	MS	F	P
Model	4	1458	365.00	47.8	<0.01
Error	25	191	7.63		
Total	29	1649			

Overall reject Ho so at least one mean % impurities removed is diff from the others

Multiple Comparisons

- Note that there are 5 temperature settings, which will yield 10 comparisons.
- Bonferroni a' = α/c = 0.10/10 = 0.01
- What we do now is to calculate the p-values for all ten tests and use 0.01 to decide if they are statistically significant or not
- For example say that we obtain the following p-values: 0.001, 0.002, 0.005, 0.02, 0.05,…, 0.99
- Then only for the first three we can claim that the results are statistically significant and reject the null hypothesis.

One Last ANOVA Concept

- The **coefficient of variability** R^2 describes how much of the variability in the continuous dependent variable was accounted for by the categorical independent variable.

- R^2 ranges from 0 (no variability explained) to 1 (100% of the variability explained).

- The more variables in the model, larger the R^2

- We will not review how it is calculated; it is available in the software output.

Example for STATA Demo

- Dataset includes pre-test scores and final exam scores for sophomores and juniors in a college course.

- The variables used for ANOVA are final exam score (continuous dependent variable) and year (categorical independent variable with two levels sophomores and juniors).

STATA Program Code

- Generate descriptive statistics for final grade:
  - sum final

```
. sum final

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       final |        19    85.73684    8.818176         72         99
```

Program Code cont.

- One-way ANOVA for effect of year on final grade:
  - anova final year

Number of obs =	19		R-squared =	0.1344	
Root MSE =	8.44193		Adj R-squared =	0.0835	

Source	Partial SS	df	MS	F	Prob > F
Model	188.160401	1	188.160401	2.64	0.1226
year	188.160401	1	188.160401	2.64	0.1226
Residual	1211.52381	17	71.2661064		
Total	1399.68421	18	77.7602339		

Note that One-Way Anova with two groups is the same as the t-test (same p-value).