Lecture 8 practice

Stats 7 Summer Session II 2022

Fill in the blank

When doing an ANOVA, you observe large differences in means between groups. Within the ANOVA framework, this would most likely be interpreted as evidence strongly favoring the ______________ hypothesis.

Alternative

For ANOVA the hypotheses are always:

H_0: All population means are the same ($\mu_1 = \mu_2 = \ldots = \mu_K$)

H_A: At least one population mean differs

Chicken diet and weight

Previously we compared the effects of two types of feed at a time.

A better analysis would first consider all feed types at once: casein, horsebean, linseed, meat meal, soybean, and sunflower.

Conduct a hypothesis test to determine if the data provide convincing evidence that the average weight of chicks varies across some (or all) groups. Make sure to check relevant conditions.

Chicken diet and weight

First, what type of test is appropriate?

We want to compare a numerical variable, weight, for each of the 6 feed groups. ANOVA allows us to compare multiple means at once.

What are our hypotheses?

Let μ_i Represent the average weight of chicks raised on feed i

H_0: $(\mu_{casein} = \mu_{horsebean} = \mu_{linseed} = \mu_{meatmeal} = \mu_{soybean} = \mu_{subflower})$

H_A: The average weight varies across some (or all) feed groups

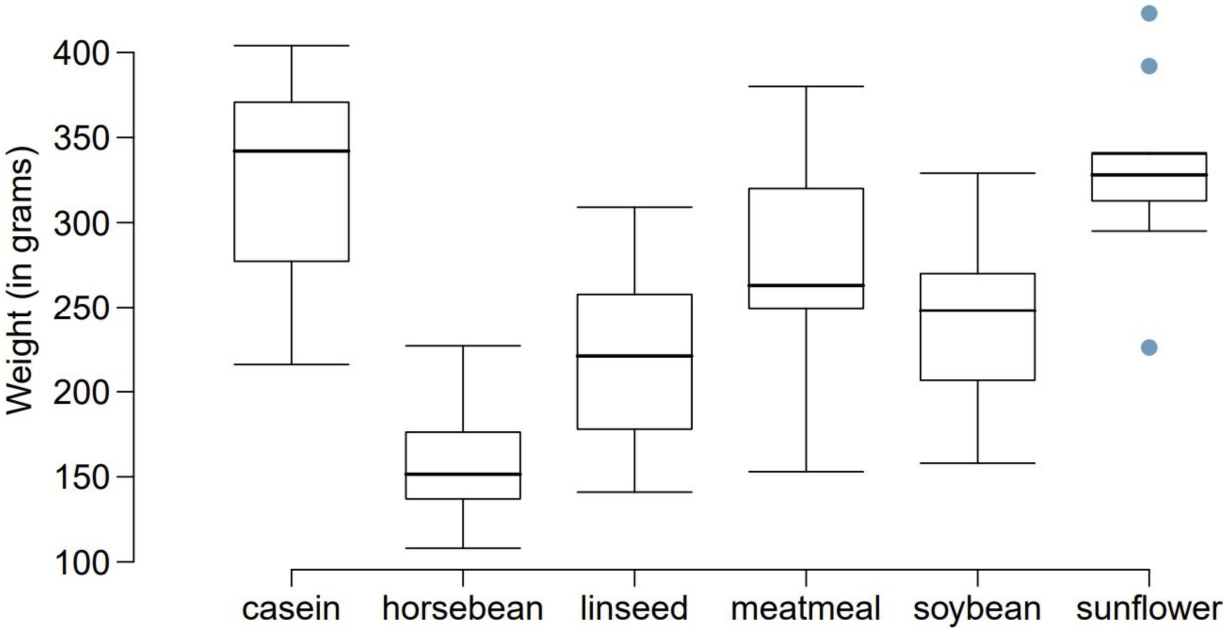
Chicken diet and weight

Now we need to check the conditions:

- Independence: There was random assignment of chicks to the groups so we have independence between groups and it seems reasonable to assume weights of chicks within a group are independent

Chicken diet and weight

Now we need to check the conditions:



- Normality: From the boxplot all we can tell is that the distribution of weights within each feed group appears to be fairly symmetric so normality seems to be an okay assumption

Chicken diet and weight

Now we need to check the conditions:

	Mean	SD	n
casein	323.58	64.43	12
horsebean	160.20	38.63	10
linseed	218.75	52.24	12
meatmeal	276.91	64.90	11
soybean	246.43	54.13	14
sunflower	328.92	48.84	12

- Similar standard deviations: The smallest standard deviation is 38.63 so 2 x 38.63 = 77.26. No group's standard deviation exceeds 2 times the minimum standard deviation so this assumption is okay

Chicken diet and weight

Since the conditions appear to be met we can proceed to the analysis

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
feed	5	231,129.16	46,225.83	15.36	0.0000
Residuals	65	195,556.02	3,008.55		

From the ANOVA table above identify the test statistic and the p-value.

Test statistic = 15.36

P-value = 0.0000

Chicken diet and weight

Make a conclusion in context with the p-value of 0.0000.

The p-value is less than any significance level so reject the null hypothesis in support of the alternative hypothesis.

The data provide convincing evidence (p-value = 0.0000 $< \alpha =$ 0.05) that the average weight of chicks varies across some (or all) feed supplement groups.

Coffee, depression, and physical activity

Caffeine is the world's most widely used stimulant, with approximately 80% consumed in the form of coffee.

Participants in a study investigating the relationship between coffee consumption and exercise were asked to report the number of hours they spent per week on moderate (e.g., brisk walking) and vigorous (e.g., strenuous sports and jogging) exercise.

Based on these data the researchers estimated the total hours of metabolic equivalent tasks (MET) per week, a value always greater than 0.

Coffee, depression, and physical activity

(a) Write the hypotheses for evaluating if the average physical activity level varies among the different levels of coffee consumption.

H_0: The population mean of MET for each group is equal to the others.

H_A: At least one pair of means is different.

Coffee, depression, and physical activity

The table below gives summary statistics of MET for women in this study based on the amount of coffee consumed.

	Caffeinated coffee consumption					
	$\leq$ 1 cup/week	2-6 cups/week	1 cup/day	2-3 cups/day	$\geq$ 4 cups/day	Total
Mean	18.7	19.6	19.3	18.9	17.5	
SD	21.1	25.5	22.5	22.0	22.0	
n	12,215	6,617	17,234	12,290	2,383	50,739

(b) Check conditions and describe any assumptions you must make to proceed with the test.

Independence: We don't have any information on how the data were collected, so we cannot assess independence. To proceed, we must assume the subjects in each group and between groups are independent. In practice, we would inquire for more details.

Coffee, depression, and physical activity

The table below gives summary statistics of MET for women in this study based on the amount of coffee consumed.

	Caffeinated coffee consumption					
	$\leq$ 1 cup/week	2-6 cups/week	1 cup/day	2-3 cups/day	$\geq$ 4 cups/day	Total
Mean	18.7	19.6	19.3	18.9	17.5	
SD	21.1	25.5	22.5	22.0	22.0	
n	12,215	6,617	17,234	12,290	2,383	50,739

(b) Check conditions and describe any assumptions you must make to proceed with the test.

Normality: The data are bound below by zero and the standard deviations are larger than the means, indicating very strong skew. However, since the sample sizes are extremely large, even extreme skew is acceptable.

Coffee, depression, and physical activity

The table below gives summary statistics of MET for women in this study based on the amount of coffee consumed.

	Caffeinated coffee consumption					
	$\leq$ 1 cup/week	2-6 cups/week	1 cup/day	2-3 cups/day	$\geq$ 4 cups/day	Total
Mean	18.7	19.6	19.3	18.9	17.5	
SD	21.1	25.5	22.5	22.0	22.0	
n	12,215	6,617	17,234	12,290	2,383	50,739

(b) Check conditions and describe any assumptions you must make to proceed with the test.

Constant variance: This condition is sufficiently met, as the standard deviations are reasonably consistent across groups.

Coffee, depression, and physical activity

Below is the output associated with this test.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Coffee	0	10508	2627	5.2	0.0003
Residuals	50734	25,564,819	504		

Identify the test statistic, p-value, and make a conclusion with α = 0.10.

Test statistic = 5.2, p-value = 0.0003 < significance level

Reject H_0 in support of H_A.

The data provide convincing evidence (p-value = 0.0003 < α = 0.10) that the average MET differs between at least one pair of groups.

True or false: ANOVA

Determine if the following statements are true or false in ANOVA, and explain your reasoning for statements you identify as false.

(a) As the number of groups increases, the modified significance level for pairwise tests increases as well.

False. As the number of groups increases, so does the number of comparisons and hence the modified significance level decreases.

True or false: ANOVA

Determine if the following statements are true or false in ANOVA, and explain your reasoning for statements you identify as false.

(b) As the total sample size increases, the degrees of freedom for the residuals increases as well.

True, the degrees of freedom for the residuals increases with sample size and decreases with the number of groups being compared.

True or false: ANOVA

Determine if the following statements are true or false in ANOVA, and explain your reasoning for statements you identify as false.

(c) The constant variance condition can be somewhat relaxed when the sample sizes are relatively consistent across groups.

True, the constant variance condition is especially necessary when the sample sizes differ.

True or false: ANOVA

Determine if the following statements are true or false in ANOVA, and explain your reasoning for statements you identify as false.

(d) The independence assumption can be relaxed when the total sample size is large.

False. As the number of groups increases, so does the number of comparisons and hence the modified significance level decreases.

Prison isolation experiment

Subjects from Central Prison in Raleigh, NC, volunteered for an experiment involving an “isolation” experience. The goal of the experiment was to find a treatment that reduces subjects’ psychopathic deviant T scores. This score measures a person’s need for control or their rebellion against control, and it is part of a commonly used mental health test called the Minnesota Multiphasic Personality Inventory (MMPI) test.

Prison isolation experiment

The experiment had three treatment groups:

(1) Four hours of sensory restriction plus a 15 minute “therapeutic” tape advising that professional help is available.

(2) Four hours of sensory restriction plus a 15 minute “emotionally neutral” tape on training hunting dogs.

(3) Four hours of sensory restriction but no taped message.

Forty-two subjects were randomly assigned to these treatment groups, and an MMPI test was administered before and after the treatment.

We want to compare the success of the treatments.

Prison isolation experiment

(a) What are the hypotheses?

H_0: ($\mu_1 = \mu_2 = \mu_3$) Average score difference is the same for all treatments.

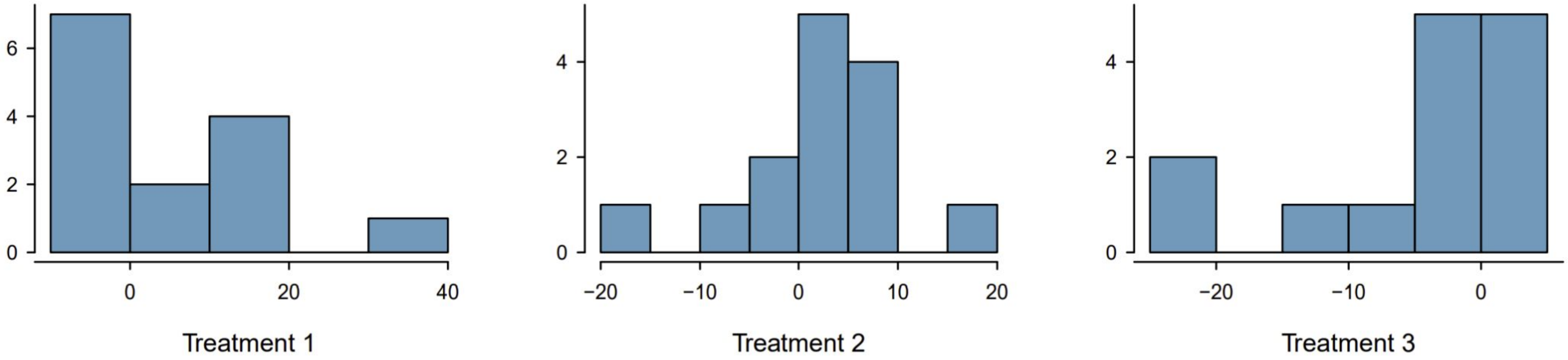
H_A: At least one pair of means are different.

(b) Complete the ANOVA. What is the conclusion of the test? Use a 5% significance level.

We have our hypotheses so the next step is to check conditions: independence, data normally distributed, similar standard deviations

Prison isolation experiment

Distributions of the differences between pre and post treatment scores (pre - post) are shown below, along with some sample statistics.



	Tr 1	Tr 2	Tr 3
Mean	6.21	2.86	-3.21
SD	12.3	7.94	8.57
n	14	14	14

Prison isolation experiment

(b) What is the conclusion of the test? Use a 5% significance level.

Conditions:

- Patients were randomized so independence is satisfied
- There are some minor concerns about skew, especially with the third group, though this may be acceptable.
- The standard deviations across the groups are reasonably similar, the max is not larger than double the minimum standard deviation

Prison isolation experiment

(b) What is the conclusion of the test? Use a 5% significance level.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
treatment	2	639.48	319.74	3.33	0.0461
Residuals	39	3740.43	95.91		

$s_{pooled} = 9.793$ on $df = 39$

The p-value is just less than the significance level so we reject the null.

The data provide convincing evidence of a difference between the average reduction in score among treatments.

Prison isolation experiment

(c) Conduct pairwise tests to determine which groups are different from each other.

We found evidence of at least one difference in means between the 3 groups. We can conduct individual pairwise (difference of two independent means) test to identify particularly which groups are different from each other.

When performing multiple comparisons/tests it is important to remember type I error increases. We will correct for this by modifying out 5% significance level accordingly.

Prison isolation experiment

(c) Conduct pairwise tests to determine which groups are different from each other.

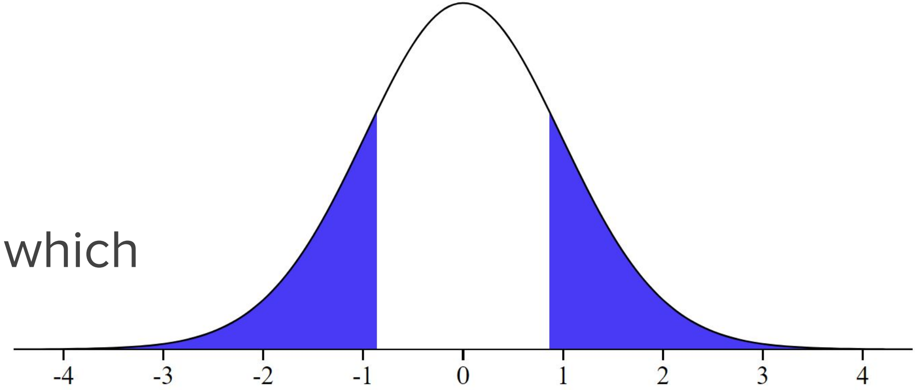
First we need to calculate the number of comparisons, then our modified significance level

$$K = \frac{k(k-1)}{2} = \frac{3(3-1)}{2} = 3$$

$$\alpha^* = \frac{\alpha}{K} = \frac{0.05}{3} = 0.01667$$

So we will be performing 3 tests in total, each with a significance level of 0.01667.

Prison isolation experiment



(c) Conduct pairwise tests to determine which groups are different from each other.

We will compare the means for treatments 1 and 2 first.

H_0: ($\mu_1 - \mu_2 = 0$) The two means are equal.

H_A: ($\mu_1 - \mu_2 \neq 0$) The two means are different.

$$\mathrm{T} = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(6.21 - 2.86) - 0}{\sqrt{\frac{12.3^2}{14} + \frac{7.94^2}{14}}} = 0.86$$

```
> 2 * pt(-0.86, df = 39)
[1] 0.395045
```

Prison isolation experiment

(c) Conduct pairwise tests to determine which groups are different from each other.

The p-value is 0.3950, which is fairly large.

We did not find strong evidence (p-value = 0.3950 > $\alpha^* = 0.01667$) of a difference between the average reduction in score among treatments 1 and 2.

Prison isolation experiment

(c) Conduct pairwise tests to determine which groups are different from each other.

The following are the results for all of the pairwise comparisons

- Treatment 1 and 2 p-value = 0.3950
- Treatment 2 and 3 p-value = 0.0619
- Treatment 1 and 3 p-value = 0.0239

The p-values are all larger than the modified significance level of $\alpha^* =$ 0.01667, so we do not have strong evidence to conclude that it is this particular pair of groups that are different.

That is, we cannot identify if which particular pair of groups are actually different, even though we've rejected the notion that they are all the same!

What we've learned for numerical variables

	1 Mean	1 Mean of paired differences	Difference of 2 means	ANOVA
Data scenario	1 sample of a numerical variable	2 samples of a numerical variable with paired data	2 samples of a numerical variable with independent data	multiple samples of a numerical variable with independent data
Distribution used	Normal (if σ is known) t (if σ is unknown)	Normal (if σ is known) t (if σ is unknown)	Normal (if σ is known) t (if σ is unknown)	F-distribution
Inference	Hypothesis test and Confidence interval	Hypothesis test and Confidence interval (Same as 1 mean inference)	Hypothesis test and Confidence interval	Hypothesis test

Conditions and hypotheses for each?

Credits

Examples adapted from OpenIntro Statistics (4th edition) by David Diez, Mine Cetinkaya-Rundel, and Christopher D Barr https://www.openintro.org/book/os/ protected under the Creative Commons License