11.1E: Analysis of Variance (Exercises)

1. What are the three pieces of variance analyzed in ANOVA?

Answer:

Variance between groups (SSB), variance within groups (SSW) and total variance (SST).

2. What does rejecting the null hypothesis in ANOVA tell us? What does it not tell us?
3. What is the purpose of post hoc tests?

Answer:

Post hoc tests are run if we reject the null hypothesis in ANOVA; they tell us which specific group differences are significant.

4. Based on the ANOVA table below, do you reject or fail to reject the null hypothesis? What is the effect size?

Source	SS	df	MS	F
Between	60.72	3	20.24	3.88
Within	213.61	41	5.21	
Total	274.33	44		

5. Finish filling out the following ANOVA tables:

a. $K = 4$

Source	SS	df	MS	F
Between	87.40			
Within				
Total	199.22	33		

b. $N = 14$

Source	SS	df	MS	F
Between		2	14.10	
Within				
Total	64.65			

c.

Source	SS	df	MS	F
Between		2		42.36
Within		54	2.48	
Total				

Answer:

a. $K = 4$

Source	SS	df	MS	F
Between	87.40	3	29.13	7.81
Within	111.82	30	3.73	
Total	199.22	33		

b. $N = 14$

Source	SS	df	MS	F
Between	28.20	2	14.10	4.26
Within	36.45	11	3.31	
Total	64.65	13		

c.

Source	SS	df	MS	F
Between	210.10	2	105.05	42.36
Within	133.92	54	2.48	
Total	344.02			

6. You know that stores tend to charge different prices for similar or identical products, and you want to test whether or not these differences are, on average, statistically significantly different. You go online and collect data from 3 different stores, gathering information on 15 products at each store. You find that the average prices at each store are: Store 1 xbar = $27.82, Store 2 xbar = $38.96, and Store 3 xbar = $24.53. Based on the overall variability in the products and the variability within each store, you find the following values for the Sums of Squares: SST = 683.22, SSW = 441.19. Complete the ANOVA table and use the 4 step hypothesis testing procedure to see if there are systematic price differences between the stores.
7. You and your friend are debating which type of candy is the best. You find data on the average rating for hard candy (e.g. jolly ranchers, $\overline{X}$= 3.60), chewable candy (e.g. starburst, $\overline{X}$ = 4.20), and chocolate (e.g. snickers, $\overline{X}$= 4.40); each type of candy was rated by 30 people. Test for differences in average candy rating using SSB = 16.18 and SSW = 28.74.

Answer:

Step 1: $H_0 : \mu_1 = \mu_2 = \mu_3$ "There is no difference in average rating of candy quality", H_A: "At least one mean is different."

Step 2: 3 groups and 90 total observations yields $df_{num} = 2$ and $df_{den} = 87$, $\alpha = 0.05$, $F^* = 3.11$.

Step 3: based on the given SSB and SSW and the computed df from step 2, is:

Source	SS	df	MS	F
Between	16.18	2	8.09	24.52
Within	28.74	87	0.33	
Total	44.92	89		

Step 4: $F > F^*$, reject H_0. Based on the data in our 3 groups, we can say that there is a statistically significant difference in the quality of different types of candy, $F(2, 87) = 24.52, p < .05$ Since the result is significant, we need an effect size: $\eta^2 = 16.18/44.92 = .36$which is a large effect

8. Administrators at a university want to know if students in different majors are more or less extroverted than others. They provide you with data they have for English majors ($\overline{X}$ = 3.78, n = 45), History majors ($\overline{X}$= 2.23, n = 40), Psychology majors ($\overline{X}$= 4.41, n = 51), and Math majors ($\overline{X}$ = 1.15, n = 28). You find the $SSB = 75.80$ and $SSW = 47.40$ and test at $\alpha = 0.05$.

9. You are assigned to run a study comparing a new medication ($\overline{X}$= 17.47, n = 19), an existing medication ($\overline{X}$= 17.94, n = 18), and a placebo ($\overline{X}$= 13.70, n = 20), with higher scores reflecting better outcomes. Use $SSB = 210.10$ and $SSW = 133.90$ to test for differences.

Answer:

Step 1: $H_0 : \mu_1 = \mu_2 = \mu_3$ "There is no difference in average outcome based on treatment", H_A: "At least one mean is different."

Step 2: 3 groups and 57 total participants yields $df_{num} = 2$ and $df_{den} = 54$, $\alpha = 0.05, F^* = 3.18$.

Step 3: based on the given SSB and SSW and the computed df from step 2, is:

Source	SS	df	MS	F
Between	210.10	2	105.02	42.36
Within	133.90	54	2.48	
Total	344.00	56		

Step 4: $F > F^*$, reject H_0. Based on the data in our 3 groups, we can say that there is a statistically significant difference in the effectiveness of the treatments, $F(2, 54) = 42.36, p < .05$ Since the result is significant, we need an effect size: $\eta^2 = 210.10/344.00 = .61$, which is a large effect.

10. You are in charge of assessing different training methods for effectiveness. You have data on 4 methods: Method 1 ($\overline{X}$ = 87, n = 12), Method 2 ($\overline{X}$= 92, n = 14), Method 3 ($\overline{X}$ = 88, n = 15), and Method 4 ($\overline{X}$= 75, n = 11). Test for differences among these means, assuming $SSB = 64.81$ and $SST = 399.45$.
