

Name: _____ GWID: _____

1. It is important to check the model assumptions before we make inferences from the statistical model for a dataset. In a one-factor or two-factor study setting,

(1) (10 points) Please state the two assumptions for the ANOVA model.

a) constancy or homogeneity of the error variance across the factor or treatment levels

b) errors are independent, and normally distributed.

(2) (10 points) Please explain how you can examine whether these assumptions are appropriate.

(Answering either residuals plots or statistical test is OK to get all points) Use plots of residual from ANOVA model, or conduct statistical tests

(3) (10 points) If you find your data have a considerable departure from the ANOVA model assumptions, what would be proper remedial measures? (Please list at least two different approaches to study the relationship of the factor with the response in your study)

(Answering any two of the three approaches is OK to get all points) Use a proper transformation; use a nonparametric test; use weighted least square regression approach for the cases with normal errors but nonconstant variances.

2. Assume we have a balanced two-factor study with observation:

Y_{ijk} : subscript index i denotes the level of the factor A, j denotes the level of the factor B, k denotes the k th observation for treatment (i, j) , $k = 1$ to n , with $n \geq 1$:

Please write down the two-factor ANOVA models:

(1) (10 points) Cell means models _____ $Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$ _____

(2) (10 points) Factor Effects models _____ $Y_{ijk} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$ _____

3. A research laboratory was developing a new compound for the relief of severe cases of hay fever. In an experiment with 36 volunteers, the amounts of the two active ingredients (factors A and B) in the compound were varied at three levels each. Randomization was used in assigning four volunteers to each of the nine treatments. The data on hours of relief follow.

		Factor B (ingredient 2)		
Factor A (ingredient 1)		$j = 1$	$j = 2$	$j = 3$
		Low	Medium	High
$i = 1$ Low		2.4	4.6	4.8
	
		2.5	4.7	4.6
$i = 2$ Medium		5.8	8.9	9.1
	
		5.3	9.0	9.4
$i = 3$ High		6.1	9.9	13.5
	
		6.2	10.1	13.2

→→→ continue on next page

You have run a two-factor ANOVA analysis and obtain the following ANOVA table:

Source of variation	SS (sum of squares)	df	MS (mean squares)	Test statistic (F^*)
A (ingredient 1)	220	2	110	110
B (ingredient 2)	120	2	60	60
AB interaction	40	4	10	(F_{AB}^*) 10
Error	27	27	1	
Total	407	35		

- (20 points)** Please fill out the missing SS, df, MS and the test statistics in the above ANOVA table. Missing values are labels with question marks (?).
- (10 points)** What distribution does the test statistic F_{AB}^* for the AB interaction follow? $F(4, 27)$
- (10 points)** If the P -value for the test of interaction < 0.001 , what can you conclude from this result? **The interaction effect between factor A and B was significant.**
- (10 points)** A company conducted an study to evaluate whether the Pay (low or high) and Authority (Small or Great) impact the Productivity of Executives. They found the two factors, Pay and Authority, did not interact. Please choose a treatment means plot (from below) that was most likely to represent this study data.

