

Statistics 571: Statistical Methods

Summer 2003 Final Exam

Ramón V. León

Name: _____

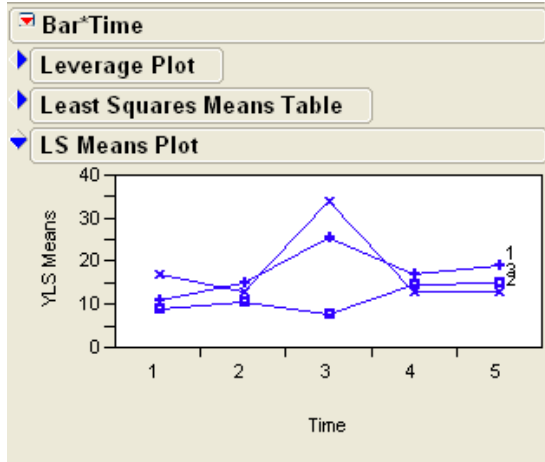
This exam is closed-book and closed-notes. However, you can use up to twenty pages of personal notes as an aid in answering your questions. Please be concise in your answers

1. Suppose you have a sample of 25 measurements and you wish to estimate the standard error of the sample median. How would you use the bootstrap procedure to do this?
2. Under what circumstances would one prefer to use the Spearman's or Kendall's rank correlation coefficient over the Pearson correlation coefficient?
3. For what type of statistical problem would one use the Kruskal-Wallis test? How about the Friedman test?
4. One can test that the median of distribution has a given value using either the sign test or the Wilcoxon signed rank test. What added assumption about the distribution of the observations is needed to use the Wilcoxon signed rank test?

5. An experiment was conducted to investigate the strength of a weld in a steel bar. Two factors were varied: the gauge bar setting (the distance the well die travels during the automatic weld cycle), which had three settings, and the welding time (total time of the automatic weld cycle), which was fixed at five levels. The weld strength data are shown in the JMP table below:

Ex13_7		Bar	Time	Y
	1	1	1	10
	2	1	1	12
	3	1	2	13
	4	1	2	17
	5	1	3	21
	6	1	3	30
	7	1	4	18
	8	1	4	16
	9	1	5	17
	10	1	5	21
	11	2	1	15
	12	2	1	19
	13	2	2	14
	14	2	2	12
	15	2	3	30
	16	2	3	38
	17	2	4	15
	18	2	4	11
	19	2	5	14
	20	2	5	12
	21	3	1	10
	22	3	1	8
	23	3	2	12
	24	3	2	9
	25	3	3	10
	26	3	3	5
	27	3	4	14
	28	3	4	15
	29	3	5	19
	30	3	5	11

Interpret the following JMP output: (Comment next to the JMP outputs)



Summary of Fit

RSquare 0.884928

RSquare Adj 0.777528

Root Mean Square Error 3.306559

Mean of Response 15.6

Observations (or Sum Wgts) 30

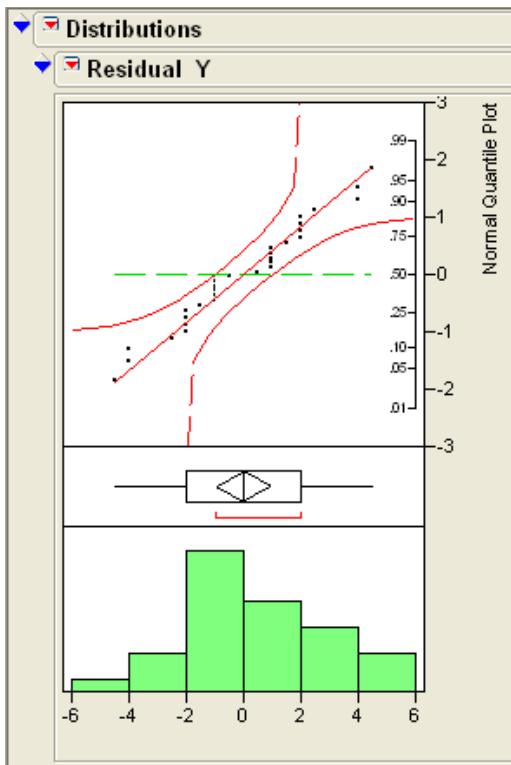
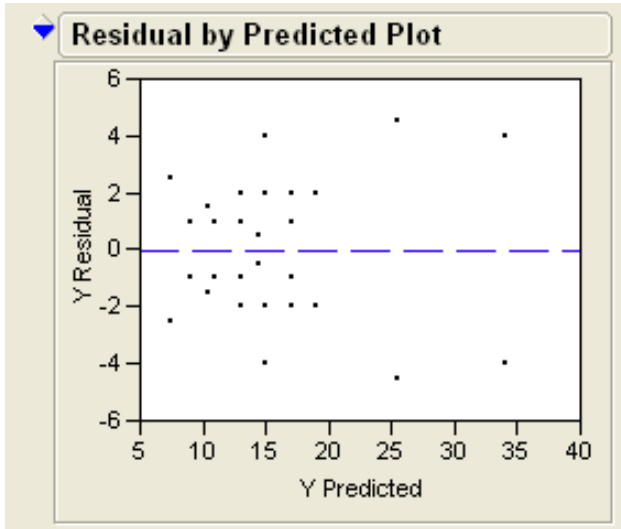
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	14	1261.2000	90.0857	8.2395	
Error	15	164.0000	10.9333		Prob > F
C. Total	29	1425.2000			0.0001

Parameter Estimates

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Bar	2	2	278.60000	12.7409	0.0006
Time	4	4	385.53333	8.8155	0.0007
Bar*Time	8	8	597.06667	6.8262	0.0008



13. Hospitals are graded based on their success rates in treating different categories of patients. We want to compare two hospitals – A, which is a university affiliated research hospital, and B, which is a general community hospital – with respect to success rates for a certain complicated surgery. The data classified by low risk patients and high risk patients are shown in the following table.

	Low Risk				High Risk		
	Success	Failure	Total		Success	Failure	Total
Hospital A	400	100	500	Hospital A	160	640	800
Hospital B	300	200	500	Hospital B	20	180	200
Total	700	300	1000	Total	180	820	1000

- a. Calculate the success rates for each category of patients for both hospitals. Which hospital is better?

	Success Rates	
Hospital	Low	High
A		
B		

- b. Aggregate the data over the two categories of patients and calculate the overall success rates for both hospitals. Now which hospital is better?

Hospital	Success Rate
A	
B	

- c. Explain the discrepancy between the results obtained in (a) and (b).

- d. What are the adjusted (standardized) success rates for each hospital?

	Success Rates		
Risk	Hospital A	Hospital B	Total
Low	80%	60%	1000
High	20%	10%	1000
			2000

14. A sample of employed men aged 18 and 67 were asked if they had carried out work on their home in the preceding year for which they would have previously employed a craftsman. The following table gives the summary of responses of 906 homeowners.

Work	Home Repair	Age			Row Total
		<30	31-45	46+	
Skilled	Yes	56	56	35	147
	No	12	21	8	41
Unskilled	Yes	23	52	49	124
	No	9	31	51	91
Office	Yes	54	191	102	347
	No	19	76	61	156
Column Totals		173	427	306	906

How would you analyze these data if you are interested in determining if home repair and age are associated?