Statistics 496, Applied Statistics for Industry II
Final Exam, Spring 2001

Name:____________________________
Site: ____________________________

INSTRUCTIONS: You will have 2 hours to complete the exam. There are 6 questions worth a total of 135 points. Not all questions have the same point value so gauge your time appropriately. Read the questions carefully and completely. Answer each question and show work in the space provided on the exam. Turn in the **entire** exam when you are done or when time is up. For essay questions, think before you write.

1. [**10 pts**] Statistical thinking consists of the recognition that variability always is, and always will be, present and that learning is an iterative procedure (one must go through several cycles to truly learn about something). How do statistically designed experiments fit into the framework of statistical thinking? Be sure to mention both aspects, variability and the iterative nature of learning.

2. [**10 pts**] The three fundamental principles of experimental design are the control of outside variables, replication within the experiment and randomization. All of these are important, but which do you think is the most important? Briefly support your choice and indicate why the other two principles are not the most important. **You have to make a choice and support that one choice.**

3. **[40 pts]** A fractional factorial experiment is performed to assess the effect of six factors on the cutting force produced during the machining of an aluminum alloy. The factors, and their levels, are given below.

Factor	Low Level (−)	High Level (+)
A: Feedrate (in/rev)	0.005	0.010
B: Depth of cut (in)	0.050	0.100
C: Cutting speed (sf/min)	2000	2500
D: Tool Geometry	Positive	Negative
E: Machine	American	Monarch
F: Cutting Fluid	Absent	Present

The experiment uses the two design generators: E=ABC and F=BCD to get the plus/minus levels for the last two factors. The data are given below.

A	B	C	D	E=ABC	F=BCD	Force
−	−	−	−	−	−	21
+	−	−	−	+	−	59
−	+	−	−	+	+	57
+	+	−	−	−	+	122
−	−	+	−	+	+	29
+	−	+	−	−	+	52
−	+	+	−	−	−	45
+	+	+	−	+	−	110
−	−	−	+	−	+	31
+	−	−	+	+	+	70
−	+	−	+	+	−	65
+	+	−	+	−	−	119
−	−	+	+	+	−	35
+	−	+	+	−	−	61
−	+	+	+	−	+	57
+	+	+	+	+	+	121

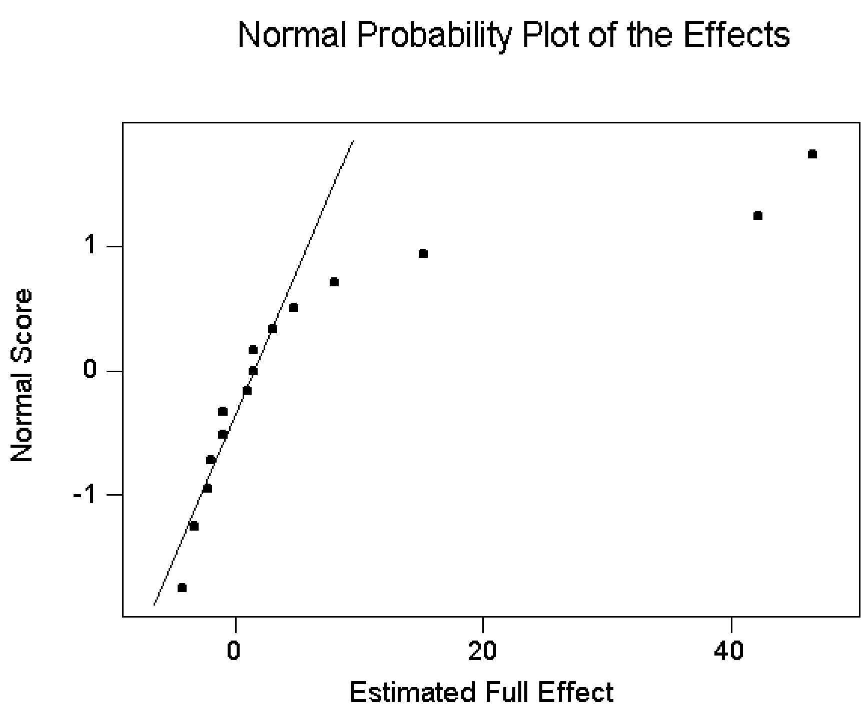
(a) **[4]** Give the full defining relation for this experiment.

(b) **[2]** What is the resolution of this experiment?

(c) **[9]** A partial list of the estimated full effects (identified by name but not aliases) is given below. Complete the list by calculating the remaining estimated full effects.

A	46.75	AB	15.25	BD	−1.00
B	42.25	AC	______	BF	1.50
C	−4.25	AD	−1.00	ABD	−2.00
D	8.00	AE	−3.25	ABF	1.50
E	4.75	AF	1.00		
F	______				

(d) **[5]** On the normal plot of estimated full effects, given below, clearly label (with names and aliases) only those points that correspond to effects that appear to be significant.



(e) **[5]** Below is output from Minitab for a reduced model in factors A, B and D. Give a prediction equation using only those variables that are clearly statistically significant, use t=3.

Term	Effect	Coef	Std Coef	t-value	P
Constant		65.875	1.487	44.30	0.000
A: FeedRate	46.750	23.375	1.487	15.72	0.000
B: DepthCut	42.250	21.125	1.487	14.21	0.000
D: Tool	8.000	4.000	1.487	2.69	0.027
AB	15.250	7.625	1.487	5.13	0.000
AD	–1.000	–0.500	1.487	–0.34	0.745
BD	–1.000	–0.500	1.487	–0.34	0.745
ABD	–2.000	–1.000	1.487	–0.67	0.520

(f) **[5]** Suppose that for a 0.10 inch cut depth we wish to keep the cutting force at 100. What feed rate should be used?

(g) **[10]** What levels of the 6 factors would you recommend to produce the lowest cutting force? Give the predicted cutting force and prediction interval for your recommended levels. The value of MS_{Error} is 46.92.

4. **[30 pts]** A study made to investigate the roles of three printing machine variables, speed (coded as X_1), pressure (coded as X_2), and distance (coded as X_3), upon the application of coloring inks onto package labels is run as a 3^3 factorial design. The response Y is a measure of quality with higher values indicating better quality. Refer to the Minitab output given for each part.

```
Response Surface Regression: Full Quadratic

Estimated Regression Coefficients for Y

Term            Coef       StDev          T       P
Constant      340.48       37.87      8.992   0.000
X1            173.39       17.53      9.892   0.000
X2            125.67       17.53      7.169   0.000
X3            146.33       17.53      8.348   0.000
X1*X1          26.39       30.36      0.869   0.397
X2*X2         -34.44       30.36     -1.134   0.272
X3*X3         -24.44       30.36     -0.805   0.432
X1*X2          48.00       21.47      2.236   0.039
X1*X3          83.58       21.47      3.893   0.001
X2*X3          51.75       21.47      2.411   0.028

Source            DF      Seq SS      Adj SS      Adj MS       F       P
Regression         9     1369348     1369348      152150   27.51   0.000
  Linear           3     1210847     1210847      403616   72.98   0.000
  Square           3       14882       14882        4961    0.90   0.463
  Interaction      3      143619      143619       47873    8.66   0.001
Residual Error    17       94023       94023        5531
Total             26     1463370
```

(a) **[5]** The response surface regression that fits a "full" quadratic (X1, X2, X3, X1*X1, X2*X2, X3*X3, X1*X2, X1*X3, X2*X3) has seventeen degrees of freedom for residual error. Where is that residual error coming from? Is it pure or replication error, lack of fit, or a combination of both. Explain your reasoning briefly.

(b) **[7]** According to the t-ratios, which of the nine terms are clearly statistically significant? Possibly significant? Not significant?

Response Surface Regression: Linear + Interaction

Estimated Regression Coefficients for Y

Term	Coef	StDev	T	P
Constant	318.81	14.20	22.450	0.000
X1	173.39	17.39	9.969	0.000
X2	125.67	17.39	7.225	0.000
X3	146.33	17.39	8.413	0.000
X1*X2	48.00	21.30	2.253	0.036
X1*X3	83.58	21.30	3.924	0.001
X2*X3	51.75	21.30	2.429	0.025

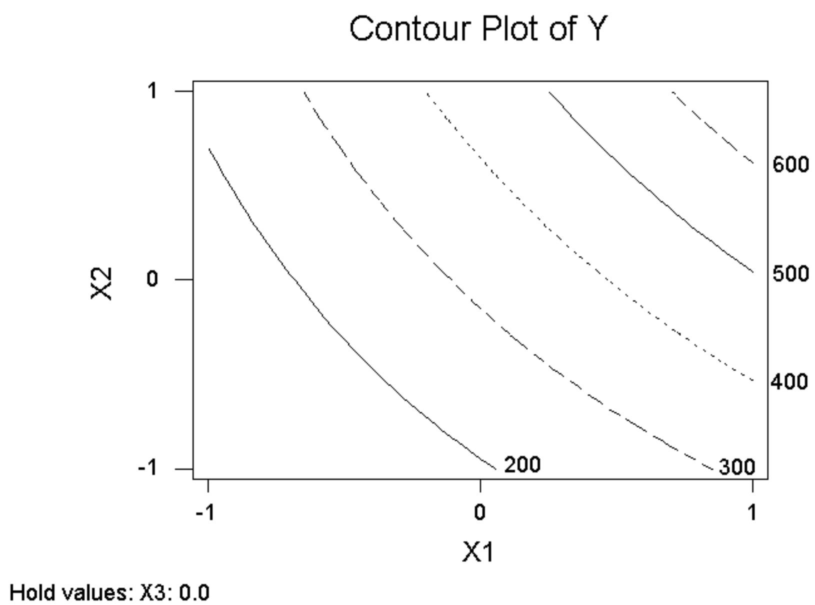
Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	6	1354466	1354466	225744	41.46	0.000
Linear	3	1210847	1210847	403616	74.12	0.000
Interaction	3	143619	143619	47873	8.79	0.001
Residual Error	20	108905	108905	5445		
Total	26	1463370				

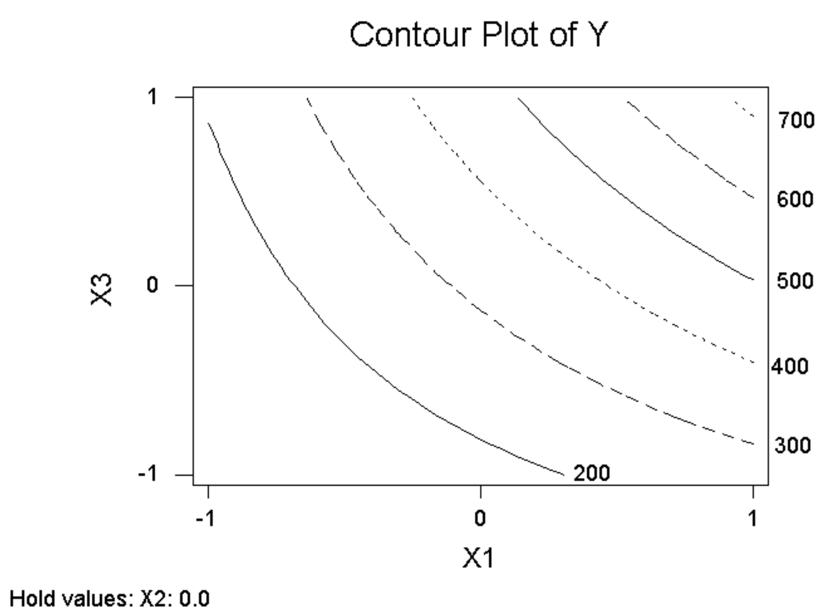
(c) **[3]** In the second response surface regression that does not include squared terms the residual error now has 20 degrees of freedom. Where did the additional 3 degrees of freedom come from?

(d) **[5]** What is the value of R^2 for the model that includes Linear + Interaction terms? What does this value indicate?

(e) **[5]** Using the model that includes Linear + Interaction terms, what settings would you choose for each of the three factors so as to maximize Y? What is the maximum value of Y for these settings?

(f) [**5**] Looking at the contour plots below and the response surface regression equation with Linear + Interaction terms, what recommendation would you make for setting X2 and X3 if X1=0.5 and you wanted a response of 500?



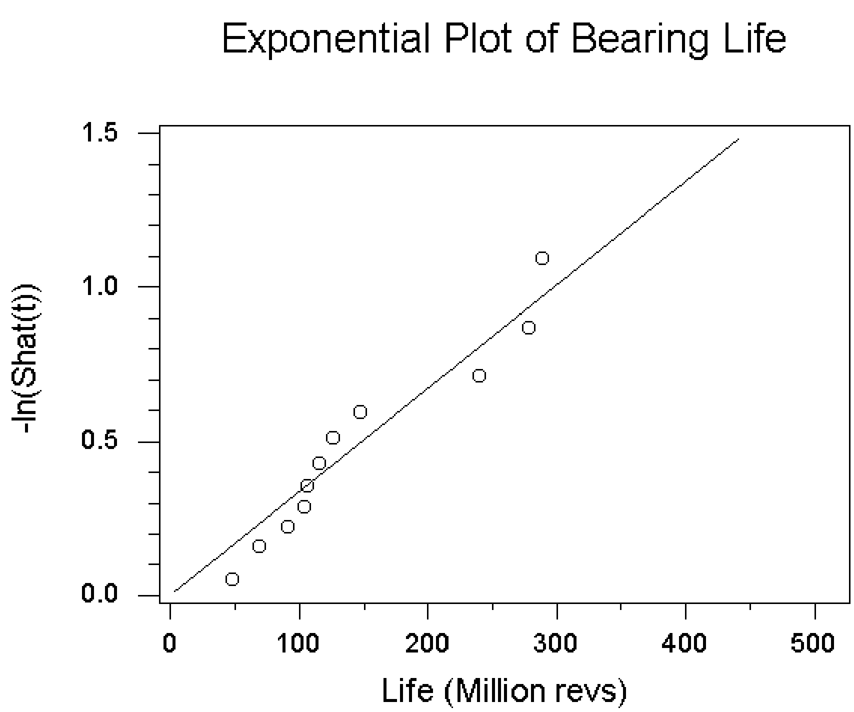


5. **[35 pts]** In an article in the *Journal of the American Society of Lubrication Engineers* are the results of a low speed endurance test on hybrid angular contact ball bearings having silicon nitride balls and rings made of vacuum induction-melted, vacuum arc-remelted (VIMVAR) M50 steel. Below are the data for the twenty bearings tested at a 6.45 kN load. If the test was suspended before failure, the resulting censored observation is denoted by *life*.

Life, t (million revs)	$\hat{S}_t(t)$	Life, t (million revs)	$\hat{S}_t(t)$
47.1	0.95	*238.0*	
68.1	——	240.0	
68.1	0.85	*274.0*	
90.8	0.80	278.0	
103.6	0.75	*285.0*	
106.0	0.70	289.0	
115.0	0.65	*367.0*	
126.0		*385.9*	
146.6		*392.0*	
229.0		*505.0*	

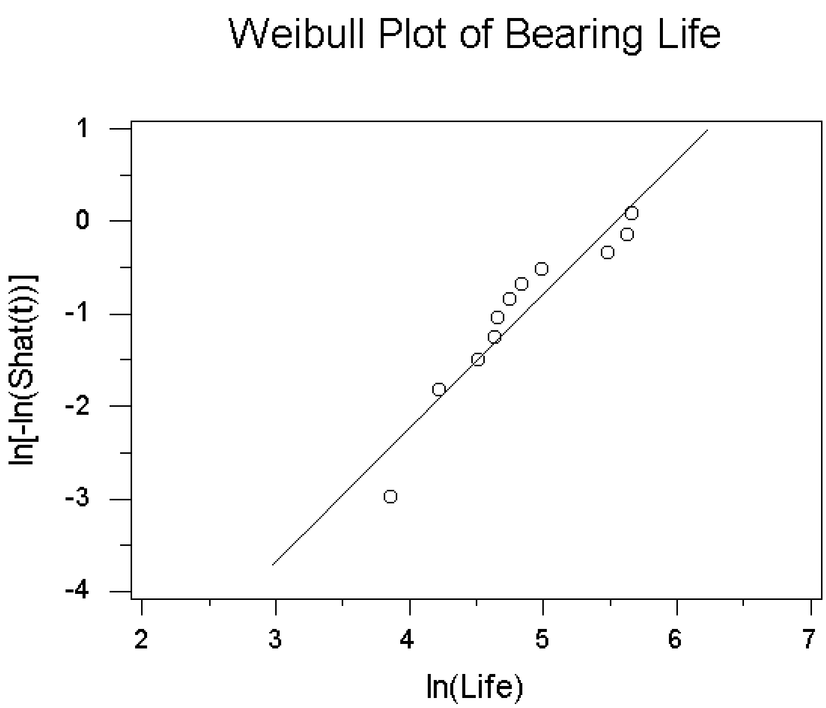
(a) **[5]** Complete the table above by calculating the estimate of the survivor function $\hat{S}_T(t)$ for each of the observed failures.

(b) **[5]** From the exponential plot, given below, estimate the parameter λ.



(c) [**5**] Use the value of λ from part (b) and the exponential model to compute the probability that a bearing will survive longer than 450 million revolutions.

(d) [**5**] From the Weibull plot, given below, estimate the parameter λ.



(e) [**6**] Use the value of λ from part (d), the fact that the slope of the line on the Weibull plot is approximately 1.43 and the Weibull model to compute the probability that a bearing will survive longer than 450 million revolutions.

(f) [**6**] According to the Weibull model, what is the median number of revolutions to failure?

(g) **[3]** Which provides a better fit to the data, the Exponential model or the Weibull model? Support your answer by referring to the probability plots.

6. **[10]** Give an example of a situation related to your work or your graduate study where performing a designed experiment would be helpful. Do not give away trade secrets but explain how you might use the ideas of experimental design in your work or graduate study.

Notice for students at remote sites.
Because most of you will take the final exam and turn in your project after grades are due at the Registrar's Office, you may receive notification from the Registrar's Office that either no grade has been submitted or an Incomplete grade has been submitted. Please ignore this if it is sent to you. I will return graded final exams, critiques of the project and course grades by May 25, 2001.