FINAL EXAM STAT 572
Spring 2000

Instructions : You are to provide your own answers to every problem. Direct references to the notes, book, or other source will not be accepted.

1. Researchers for a meat-processing firm want to study the effect of the amount of time and the temperature in the smokehouse in the flavor of smoked hams. An experiment is designed in which five times (2, 4, 6, 8, 10 hrs) and three temperatures (300, 350, 400 degrees) are tested. The combinations of 2 hrs and 300 degrees and 10 hrs and 400 degrees were not tested because they were thought to produce indesirable results (see table below). Two hams were tested at each of 13 conditions:

		Time				
		2 hrs	4 hrs	6 hrs	8 hrs	10 hrs
Temperature	300	0	2	2	2	2
	350	2	2	2	2	2
	400	2	2	2	2	0

The flavor of each ham was tested by a panel of experts and the average score was used as the response. The higher the score, the better the flavor. A regression model was fitted to the data using the actual levels of the two factors with the following results:

Response: Flavor

Summary of Fit

RSquare	0.825538
RSquare Adj	0.781923
Root Mean Square Error	3.246234
Mean of Response	18.33958
Observations (or Sum Wgts)	26

Lack of Fit

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	VIF
Intercept	-601.8627	81.01573	-7.43	<.0001	0
Time	26.010414	4.079132	6.38	<.0001	277.9
Temperature	3.1544298	0.428143	7.37	<.0001	695.79091
Time*Time	-1.024524	0.122347	-8.37	<.0001	37.48951
Temperat*Time	-0.038854	0.008754	-4.44	0.0003	144.4028
Temperat*Temperat	-0.004158	0.000577	-7.20	<.0001	621.28951

Effect Test

Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
Time	1	1	428.46824	40.6592	<.0001
Temperature	1	1	572.03550	54.2829	<.0001
Time*Time	1	1	738.95318	70.1225	<.0001
Temperature*Time	1	1	207.57290	19.6975	0.0003
Temperature*Temperature	1	1	546.36955	51.8474	<.0001

a) Is the model useful to describe these data? (10 p)

b) The pure error sum of squares is 148.55. Test the presence of lack of fit in this model. (10 p)

c) Explain how the optimal settings of the two factors (at which the response is maximized) could be estimated using the model. (5 p) [Hint: Just delineate the method]

d) Explain how you would construct a confidence interval for the response at the estimated optimal settings and how this interval would be interpreted. (10 p)

e) How would you estimate the variation in the flavor of individual hams at the estimated optimal settings? [Hint: This is related to a prediction interval](5 p)

BONUS

f) Is the level of collinearity in the model surprising? Are you concerned about collinearity in this model? (7 p)

2. The manager of a plant that manufactures high resistance packaging material wants to study how the temperature and the pressure of in a certain phase of the process affect the strength of the material. The data shown in the table below comes from past records and show the levels of temperature and pressure used, along with the number of samples that passed and failed a strength test.

6 Cols

	Temp	Pressure	Passed	Failed
1	120	250	6	3
2	100	205	13	9
3	102	248	15	10
4	82	213	7	2
5	130	190	11	8
6	103	244	18	4
7	140	265	14	0
8	150	243	11	3
9	133	271	20	2
10	121	210	8	5

The data was analyzed in JMP and the results are shown below. In this output the response is

$$\_ID\_ = \begin{cases} F & \text{if failed} \\ P & \text{if passed} \end{cases}$$

Response: _ID_

Iteration History

Converged by Gradient

Whole-Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	6.772332	3	13.54466	0.0036
Full	92.164371			
Reduced	98.936703			

RSquare (U)	0.0685
Observations (or Sum Wgts)	169

Lack of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack of Fit	6	3.958768	7.917536
Pure Error	159	88.205603	Prob>ChiSq
Total Error	165	92.164371	0.2442

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-16.919681	12.701937	1.77	0.1828
Temp	0.16592479	0.1039842	2.55	0.1106
Pressure	0.07523709	0.0553487	1.85	0.1740
Temp*Pressure	-0.0007696	0.0004517	2.90	0.0884

Effect Test

Effect Likelihood-Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
Temp	1	1	2.5966140	0.1071
Pressure	1	1	1.8681080	0.1717
Temp*Pressure	1	1	2.9620151	0.0852

a) Write out the model fitted in terms of p = P[fail] and determine its usefulness. (10 p)

b) Estimate the proportion of failures when the temperature is set at 120 degrees and pressure at 250 psi. (5 p)

c) Speculate as to why the model is significant, but the individual terms are not. (5 p)

d) Another term is added to the model and the -loglikelihood of the full model goes down to 82.6. Test the significance of this term. (10 p)

3. Explain the type of curve that is fitted by the following model:

$$E(y)=ß_0+ß_1x+ß_2x^2+ß_3x^3+ß_4(x-t)_+^2+ß_5(x-t)_+^3$$

where $x_+=\begin{cases}x & \text{if } x>0\\ 0 & \text{otherwise}\end{cases}$

[Hint: Graphs can help] (10 p)