

Test 2 Stat 461 Fall 1998

Instructions: Be efficient with your time. If you are stuck in one question, skip it and go to the next. Sometimes it is better to get partial credit for all the questions than full credit for a few.

An auto financing company uses a rather complex credit rating system for car loans. The questionnaire requires a substantial time to be completed, using sales staff time and risking alienating the customer. In order to simplify the process, a manager collects a sample of 200 recent applications. The variables recorded for each applicant are (*DATA SET: RATING*):

- Credit rating (from 0 to 100)
- Age (in years)
- Monthly income (in dollars)
- Current debt payments as a fraction of monthly income
- Marital status (S if single, and M if married)

Below are the results of a regression analysis performed on the data.

Response: RATING original

Summary of Fit

RSquare	0.975196
RSquare Adj	0.974687
Root Mean Square Error	2.004888
Mean of Response	62.19
Observations (or Sum Wgts)	200

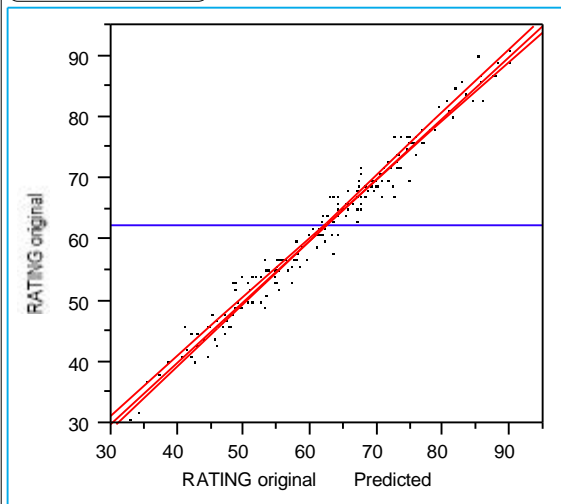
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	55.190631	0.95687	57.68	<.0001
AGE	0.000716	0.01803	0.04	0.9684
INCOME	0.009677	0.000283	34.16	<.0001
DEBTFRAC	-38.28811	1.487937	-25.73	<.0001
STATUS[M-S]	0.6900491	0.217088	3.18	0.0017

Effect Test

Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
AGE	1	1	0.0063	0.0016	0.9684
INCOME	1	1	4689.3825	1166.636	<.0001
DEBTFRAC	1	1	2661.5766	662.1537	<.0001
STATUS	1	1	40.6135	10.1039	0.0017

Whole-Model Test



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob>F
Model	4	30816.963	7704.24	1916.68	
Error	195	783.817	4.02		Prob>F
C Total	199	31600.780			<.0001

a) In terms of the full model, what do these results suggest?

b) Characterize each regressor with respect to its contribution to determining the credit rating?

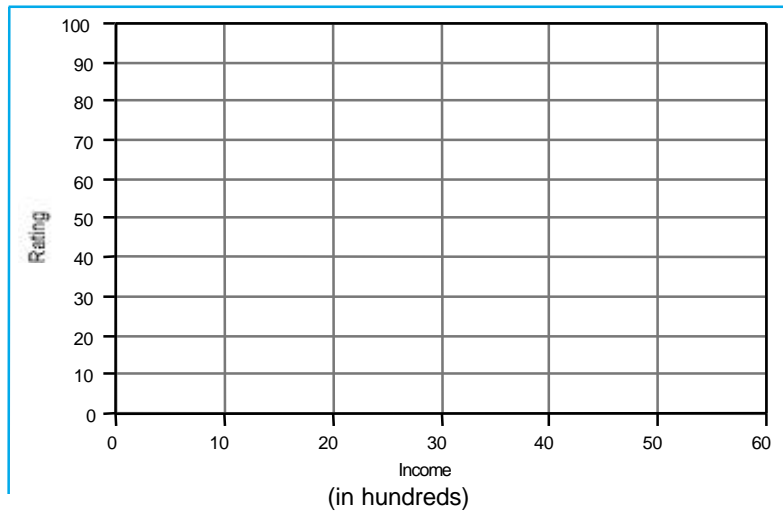
c) Carefully explain the meaning of the coefficient for STATUS?

d) Can these results be used to simplify the credit rating process? Why and how?

e) When the results are presented to the credit manager, he expresses concern about age not being a factor. In his many years in the business, he has found that age is a very important factor in the credit rating (If fact, age is significant when the other variables are not in the model). Explain to the manager why age appears to be not important in the above model.

f) The credit manager also points out that he has observed that low income tends to decrease more the credit rating of a single person than that of a married person.

1. Draw a plot to illustrate this effect (use the graph below and use one line per status)
2. How would you determine if this effect is significant in the model?



g) One observation has a value of DFBETAS for income of 1.5. Explain what this means.

h) The model with only INCOME has an $R^2 = 0.891$. Is it worth keeping the rest of the regressors in the model?