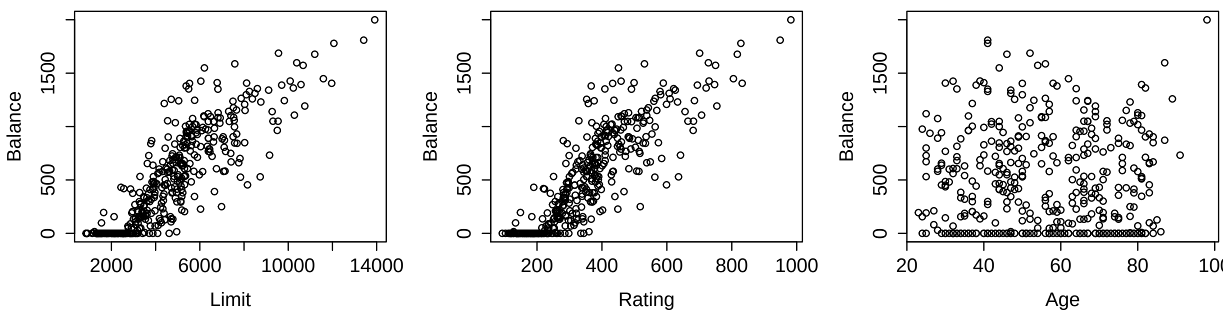
Practice Questions for the Final Exam

STA 371G, Statistics and Modeling, Spring 2016

Problem 1: Credit Card Balance

A dataset contains information about credit card *Balance*, credit *Limit*, credit *Rating* and *Age* for 400 credit card users. Below are the plots of their *Balance* against *Limit*, *Rating* and *Age*.



A multiple regression model is constructed as

$$Balance = \ \beta_0 + \beta_1 Limit + \beta_2 Rating + \beta_3 Age + \epsilon, \ \epsilon \sim \mathcal{N}(0, \sigma^2)$$

and the results are shown below.

Linear Regression						
Regression Statistics						
R	0.8681					
R Square	0.7536					
Adjusted R Square	0.75173					
Standard Error	229.08043					
Total Number Of Cases	400					
Balance =- 259.5175 + 0.0190 * Limit + 2.3105 * Rating - 2.3458 * Age						
ANOVA						
	d.f.	*SS*	*MS*	*F*	*p-level*	
Regression	3.	63,558,685.05779	21,186,228.3526	403.71757	0.E+0	
Residual	396.	20,781,226.85221	52,477.84559			
Total	399.	84,339,911.91				
	Coefficients	*Standard Error*	*LCL*	*UCL*	*t Stat*	*p-level*
Intercept	-259.51752	55.88219	-369.38038	-149.65466	-4.64401	0.
Limit	0.01901	0.06296	-0.10477	0.1428	0.30197	0.76283
Rating	2.31046	0.93953	0.46338	4.15754	2.45918	0.01435
Age	-2.34575	0.66861	-3.66023	-1.03127	-3.50838	0.0005

(a) (**5 points**) Suppose a bank will automatically reject a user's credit card application if they believe the user's credit card Balance is lower than \$500 with a probability larger than 10%. Will a person with $Limit = 8000$, $Rating = 600$ and $Age = 40$ get his credit card application approved by this bank?

Problem 2: Auto Fuel Economy

A dataset of 392 cars is provided to investigate the relationship between the fuel consumption, measured by miles per gallon (MPG), and the following explanatory variables:

- Cylinders: number of cylinders
- Displacement: a number related to the size of the engine
- Horsepower: a measurement of power
- Weight: in pounds
- Acceleration: the number of seconds to go from 0 to 60 miles per hour

(a) A multiple regression model is constructed as

$$\log(MPG) = \beta_0 + \beta_1 Cylinders + \beta_2 \log(Displacement) + \beta_3 \log(Horsepower) + \beta_4 \log(Weight) + \beta_5 \log(Acceleration) + \epsilon, \ \epsilon \sim \mathcal{N}(0, \sigma^2)$$

and the results are shown in the next plot.

Linear Regression

Regression Statistics	
R	0.89795
R Square	0.80631
Adjusted R Square	0.80381
Standard Error	0.15062
Total Number Of Cases	392

log(Mpg) = 9.1809 - 0.0213 * Cylinders - 0.0987 * log(Displacement) - 0.5062 * log(Horsepower) - 0.3011 * log(Weight) - 0.2716 * log(Acceleration)

ANOVA

	d.f.	*SS*	*MS*	*F*	*p-level*
Regression	5.	36.45342	7.29068	321.38326	0.E+0
Residual	386.	8.75654	0.02269		
Total	391.	45.20996			

	Coefficients	*Standard Error*	*LCL*	*UCL*	*t Stat*	*p-level*
Intercept	9.18086	0.46449	8.26762	10.0941	19.76566	0.E+0
Cylinders	-0.02129	0.01356	-0.04795	0.00538	-1.56933	0.11739
log(Displacement)	-0.09869	0.06199	-0.22058	0.0232	-1.59191	0.11222
log(Horsepower)	-0.50618	0.07383	-0.65134	-0.36101	-6.85572	2.82363E-11
log(Weight)	-0.30114	0.10672	-0.51097	-0.09131	-2.82168	0.00502
log(Acceleration)	-0.27164	0.07849	-0.42595	-0.11733	-3.46103	0.0006

(1) For a car with $Cylinders = 4$, $Displacement = 100$, $Horsepower = 100$, $Weight = 3000$, $Acceleration = 15$, what's the probability for the car to have a MPG between 15 and 20?

Problem 3: Freemark Abbey Winery

This problem is based on the Freemark Abbey Winery case. The payoff table below is constructed using the information given in the case.

Table 1: Payoff Table

	0.50*0.40 =0.20 Storm Botrytis	0.50*0.60 =0.30 Storm No Botrytis	0.50*0.40 =0.20 No Storm Sugar 25%	0.50*0.40 =0.20 No Storm Sugre 20%	0.50*0.20 =0.10 No Storm Acidity <0.7%
Harvest Now	2.85*12 =34.2	2.85*12 =34.2	2.85*12 =34.2	2.85*12 =34.2	2.85*12 =34.2
Harvest Later	8*12*0.7 =67.2	2*12/2 =12	3.5*12 =42	3.0*12 =36	2.5*12 =30

(a) Suppose you could buy perfect information regarding the sugar and acid levels of the grapes if the storm does not hit. What is the most you would be willing to pay for this information?

(b) Suppose we could buy a grape expert's opinion on the sugar and acid levels of the grapes if the storm does not hit. Suppose that if the condition of the grapes is such that the grapes will have a sugar level of 25% if the storm does not hit, the grape expert correctly indicates this 80% of the time, and mistakenly indicates a 20% sugar level 10% of the time; if the condition of the grapes is such that the sugar level is 20% if the storm does not hit, the grape expert correctly indicates this 70% of the time, and mistakenly indicates a 25% sugar level 20% of the time; and if the condition of the grapes is such that the acid level is smaller than 0.7% if the storm does not hit, the grape expert correctly indicates this 90% of the time, and mistakenly indicates a 20% sugar level 5% of the time. What is the expected value of the grape expert's opinion?

Problem 4: Simulation

Suppose there are 24 male employees and 24 female employees from a company. They are in similar positions and have similar work experiences. A survey found that among the 24 male employees, 21 of them got promoted at least once during the past two years; but among the 24 female employees, only 14 of them got promoted at least once during the past two years. The difference of the promotion rate is found to as large as $21/24 - 14/24 = 29\%$.

Suppose the truth is that there is no difference at all between female and male employees in terms of promotion, please describe a Simulation procedure to estimate the probability for the difference of promotion rates between two genders to be larger than 29%.