STA 5172 Midterm II

March 28, 2013

Name:

FSUID:

Please sign the following pledge and read all instructions carefully before starting the exam.

Pledge: I have neither given nor received any unauthorized aid in completing this exam, and I have conducted myself within the guidelines of the University Honor Code.

Signature: ______________________________

INSTRUCTIONS:

- This is an open-book, open-notes exam. You can refer to your notes, the text, or any other books. You may use a calculator. qf and qt denote the upper f and t quantiles.
- Total time is 75 minutes (9:30 A.M to 10:45 A.M.)
- **Show all work**, clearly and in order, if you want to receive full credit. When you use your calculator, explain all relevant mathematics. I reserve the right to take off points if I cannot see how you arrived at your answer (even if your final answer is correct).
- Circle or otherwise indicate your final answers.
- **Answer all the questions in the space provided. You may attach additional sheets if necessary.**
- This test has 4 problems and is worth 80 points. It is your responsibility to make sure that you have all of the problems.
- **Good luck!**

Prob. No.	Max Points	Earned Pts.
1	20	
2	20	
3	20	
4	20	

TOTAL: __________

Question 1. (20 pts.) Data was collected on 100 houses recently sold in a city. It consisted of the sales price (in \$), house size (Size) (in square feet) and the lot size (Lot) (in square feet). A multiple linear regression for estimating the price in terms of house size and lot size is fit. (price= $\alpha + \beta_1$ Size + β_2 Lot). Answer the following questions.

```
Call:
lm(formula = Price ~ Size + Lot, data = Housing)
Residuals:
   Min      1Q     Median   3Q     Max
-81681 -19926   2530  17972  84978
Coefficients:
              Estimate     Std. Error   t value   Pr(>|t|)
(Intercept) -1.054e+04  9.436e+03  -1.117    0.267
Size         5.378e+01  6.529e+00   8.237    8.39e-13 ***
Lot          2.840e+00  4.267e-01   6.656    1.68e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 30590 on 97 degrees of freedom
Multiple R-squared: 0.7114,     Adjusted R-squared: 0.7054
F-statistic: 119.5 on 2 and 97 DF,  p-value: < 2.2e-16
```

1. What statistical conclusion should you make about the effect of Lot size on house sale price?
   A. Since 11.30869 $>$ table value, reject the null hypothesis.
   B. Since 12.54858 $>$ table value, reject the null hypothesis.
   C. Since 9.874 $<$ table value, reject the null hypothesis.
   D. Since 6.656 $>$ table value, reject the null hypothesis.

2. Which one of the following assumptions is incorrectly stated?
   A. The house price is normally distributed for any lot size and house size.
   B. The house price has different variability for different lot sizes.
   C. The mean house price for any dividend rate is a linear function of lot size and house size.

3. The estimate of the standard deviation of $\hat{\beta}_1$ is:
   A. 3.36284
   B. 3.14983
   C. 6.529
   D. 0.4267

Question 2. (20 pts.) A study was conducted among a group of people who underwent coronary angiography. A group of 1493 people with coronary-artery disease were identified and were compared with a group of 707 people without the disease (controls). Risk factor information was collected on each group. Among cases, the mean serum cholesterol was 234.8 mg/dL with standard deviation = 48.3 mg/dL. Among controls, the mean serum cholesterol was 215.5 mg/dL with standard deviation = 47.3 mg/dL. The standard deviation of the case and the control group taken together is 48 mg/dL. Determine whether the true mean serum cholesterol is different between the two groups at a 5% level of significance. (qf(0.95,1492,706) = 1.113479, qt(0.975,2198) =1.961044)

Question 3. (20 pts.) Suppose you conducted a drug trial on a group of animals and you hypothesized that the animals receiving the drug would show increased heart rates compared to those that did not receive the drug. You conduct the study and collect the following data on the number of animals in Table 1. Conduct the following test at 5% level of significance given that qchisq(0.95,1) = 3.841459.
H_0: The proportion of animals whose heart rate increased is independent of drug treatment.
H_1: The proportion of animals whose heart rate increased is associated with drug treatment.

Table 1: Drug trial results.

	Heart rate increased	heart rate decreased	Total
Treated	36	14	50
Non-treated	30	25	55
Total	66	39	105

Question 4. (20 pts.) It is thought that the glycaemic index (GI) of food is an indicator of how sustaining or satisfying it is and may influence appetite. A sample of 33 children were provided with a breakfast of low GI foods on one day and high GI foods on another. The two breakfasts contained the same quantities of carbohydrate, fat and protein. On each day a buffet lunch was provided and the number of calories eaten at lunchtime were recorded. On the first day the children ate a low GI breakfast and on the second day a high GI breakfast. The mean calorie intakes after low GI and high GI breakfasts are 607.96 and 771.6061 respectively and the standard deviation of the difference of calorie intake after low GI and high GI breakfasts is 100. Determine whether the kind of breakfast eaten has an effect on mean calorie intake at 5% of significance. (Given qt(0.975,32) = 2.03)