STAT 610 – MIDTERM EXAM (Sample)

Name: ________________________

Instructions:

1. You may use your own calculator.
2. Cell phones, mobile devices, tablets, laptops and other electronic devices with wireless capability must be turned off and put away.
3. Please show your work in the space provided. If you need additional space, use the back of the preceding page, indicating **clearly** that you have done so.
4. This exam has **7** numbered pages (including this cover page). Please make sure you don't miss any questions.
5. **In order to receive full credit for a problem, you should show all of your work and explain your reasoning (concisely and clearly)**. Legible and logical work can receive substantial partial credit even if the final answer is incorrect. You will **not** receive any credit if you do not justify your answers.
6. Please read through all questions before starting to answer any. It is advisable that you determine which questions you are most comfortable with and do those first.
7. Do not dwell too long on any one question. Answer as many questions as you can.

Good Luck!

1. (20 points) A market research firm supplies manufacturers with estimates of the retail sales of their products from samples of retail stores. Marketing managers are prone to look at the estimate and ignore sampling error. Suppose that a simple random sample of 15 stores this month shows mean sales of 53 units of a small appliance, with standard deviation 12 units. During the same month last year, a simple random sample of 14 stores gave mean sales of 50 units, with standard deviation 10 units. An increase from 50 to 53 is a rise of 6%. The marketing manager is happy, because sales are up 6%.

   Perform an independent two-sample T test for $H_0 : \mu_1 = \mu_2$ versus $H_A : \mu_1 > \mu_2$, where μ_1 (μ_2) is the mean number of units sold at all retail stores this month (during the same month last year). Draw conclusions in the context of the study. Consider a 5% significance level.

2. (40 points) Consider the one-way ANOVA model

$$Y_{ij} = \beta_i + \varepsilon_{ij}, \quad i = 1, \ldots, k, j = 1, \ldots, n_i$$

with $\varepsilon_{ij} \sim_{\text{i.i.d.}} N(0, \sigma^2)$.

(a) (10 points) Find the least-square estimate of β_i and σ^2.

(b) (10 points) Now suppose the i-th group mean satisfies

$$\beta_i = \alpha \times i \quad \text{for all} \quad i = 1, \ldots, k.$$

Find the least-square estimate of α and σ^2.

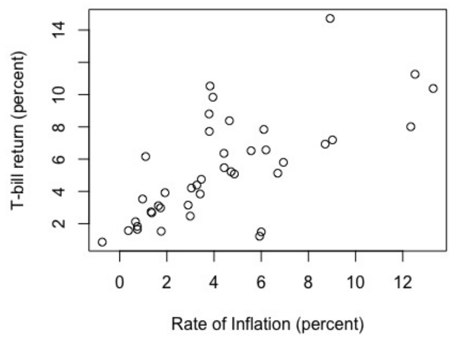
(c) (20 points) Under the model

$$Y_{ij} = \beta_i + \varepsilon_{ij}, \quad i = 1, \ldots, k, j = 1, \ldots, n_i, \quad \varepsilon_{ij} \sim_{\text{i.i.d.}} N(0, \sigma^2),$$

derive the F-test for

$$H_0 : \beta_i = \alpha \times i \quad \text{for all} \quad i = 1, \ldots, k, \quad \text{vs.} \quad H_a : \text{not } H_0.$$

3. (40 points) When inflation is high, lenders require higher interest rates to make up for the loss of purchasing power of their money while it is loaned out. Figure 1 (a) is a scatter plot of the return of one-year Treasury bills (or T-bills) against the rate of inflation as measured by the change in the government's Consumer Price Index in the same year. The data cover 41 years, from 1950 to 1990. Figure 1 (b) displays the first 10 observations in the data set. An inflation rate of 5% means that the same set of goods and service cost 5% more.



(a) Scatter Plot

Year	T-bill percent	Inflation percent
1950	1.22	5.93
1951	1.49	6.00
1952	1.65	0.75
1953	1.83	0.75
1954	0.86	−0.74
1955	1.57	0.37
1956	2.47	2.99
1957	3.15	2.90
1958	1.53	1.76
1959	2.98	1.73

(b) Data

Figure 1: Returns on Treasury bills and rate of inflation

Consider a linear regression of T-bill returns (Y) on inflation rate (X)

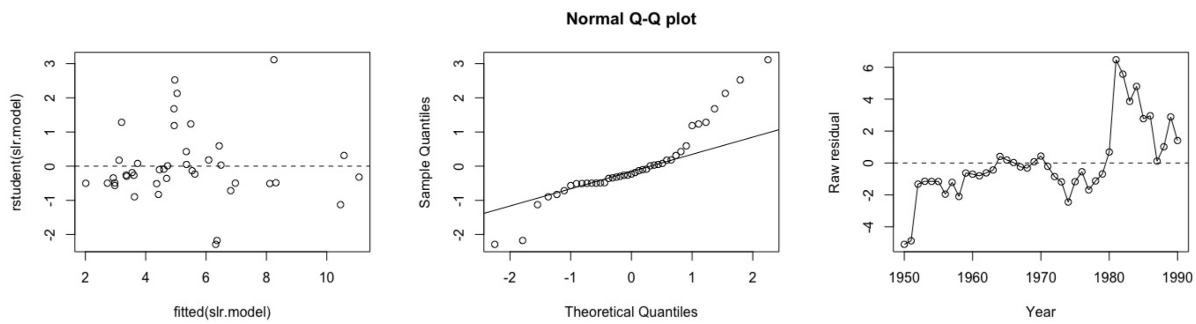
$$Y_j = \beta_0 + \beta_1 X_j + \varepsilon_j, \quad \varepsilon_j \sim \text{ iid } N(0, \sigma^2) \tag{1}$$

for $j = 1, \ldots, 41$. The data were analyzed using **R**. Below are some **R** code and edited **R** output:

```
> mydata = read.csv('Tbill.csv', header=T); attach(mydata)
> head(mydata)
  Year Tbill Inflation
1 1950  1.22      5.93
2 1951  1.49      6.00
3 1952  1.65      0.75
4 1953  1.83      0.75
5 1954  0.86     -0.74
6 1955  1.57      0.37
> mean(Tbill)
[1] 5.315366
> mean(Inflation)
[1] 4.375122
> slr.model = lm(Tbill~Inflation); summary(slr.model)
```

```
Coefficients:
            Estimate   Std. Error    t value    Pr(>|t|)
(Intercept)   2.4913   0.6138        4.058      0.00023 ***
Inflation     0.6455   0.1116        5.784      1.03e-06 ***
---
Residual standard error: 2.382 on 39 degrees of freedom
```

A plot of externally studentized residuals versus fitted values is shown in Figure 2 (a); Figure 2 (b) is the normal QQ plot of externally studentized residuals and Figure 2 (c) depicts the raw residuals against year. The following problems ask you to use this information.



(a) Externally studentized residual vs fitted

(b) Normal Q-Q plot

(c) Raw residual vs year

Figure 2: Returns on Treasury bills and rate of inflation

(a) (10 points) The intercept β_0 in the regression model (1) is meaningful. Explain what β_0 represents. Is there good evidence that β_0 is greater than 1? Draw conclusions in the context of the study.

(b) (10 points) Provide an appropriate ANOVA table and find the coefficient of determination.

(c) (10 points) Give a 90% confidence interval for the mean return on T-bills in all years having 3.7% inflation.

(d) (10 points) Perform model diagnostics.