STA 302 Summer 2001 Quiz Four

1. (1 point) Let $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and let $\mathbf{c}$ be a vector of constants. What is the distribution of $\mathbf{X} + \mathbf{c}$? Just write down the answer.

2. (1 point) Let $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and let $\mathbf{A}$ be a matrix of constants. What is the distribution of $\mathbf{AX}$? Just write down the answer.

3. (1 point) For the regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$, what is the distribution of $\mathbf{Y}$? Just write down the answer.

4. (1 point) Give the formula for the least squares estimator $\mathbf{b}$. Just write down the answer.

5. (3 points) What is the distribution of $\mathbf{b}$? Show your work. You may apply your answers to earlier questions without comment.

6. (1 point) For the SMSA data, what proportion of the variation in number of serious crimes is explained by the independent variables in the model?

7. (1 point) For the SMSA data, estimate the change in number of serious crimes if the number of people 65 or older in a district goes up by one percent while all other independent variables in the model remain constant. Give a number, and *make sure to specify whether your estimate is an increase or a decrease.*

8. (1 point) When someone asks you for an explanation of the last result, you reply that it doesn't require explanation. Why? (Zero points for saying, "I just crunch the numbers. I don't have to explain what they mean.")

Jerry's Answers to Quiz 4

(1) $X + c \sim N(\mu + c, \Sigma)$

(2) $AX \sim N(A\mu, A\Sigma A')$

(3) $Y \sim N(X\beta, \sigma^2 I_n)$

(4) $b = (X'X)^{-1}X'Y$

(5) $b = AY$, where $A = (X'X)^{-1}X'$

$E(b) = (X'X)^{-1}X'X\beta = \beta$, and

$$\sigma^2\{b\} = (X'X)^{-1}X'\sigma^2 I_n [(X'X)^{-1}X']'$$

$$= \sigma^2 \underbrace{(X'X)^{-1}X'X}_{I} (X'X)^{-1\prime}$$

$$= \sigma^2 (X'X)^{-1},$$

so $b \sim N(\beta, \sigma^2(X'X)^{-1})$

(6) $R^2 = .97$

(7) Increase of 641.6 ≈ 642 crimes

(8) The estimate is not significantly different from zero.

```
options linesize=79 pagesize=100;
title 'STA 302 Summer 2001: SMSA Data';
title2 'Assignment 4';

data census;
     infile 'smsa.dat';
     input id landarea totpop urban oldfolks doctors hospbeds hsgrads
           labforce income crimes region;

proc reg simple;
     model crimes = landarea -- income;
```

The REG Procedure

Descriptive Statistics

Variable	Sum	Mean	Uncorrected SS	Variance	Standard Deviation
Intercept	141.00000	1.00000	141.00000	0	0
landarea	363721	2579.58156	2024868829	7761577	2785.96078
totpop	132838	942.11348	357271388	1658021	1287.64158
urban	5891.60000	41.78440	289672	310.67575	17.62600
oldfolks	1384.70000	9.82057	14489	6.36293	2.52249
doctors	250719	1778.14894	1715566309	9069647	3011.58547
hospbeds	864913	6134.13475	15583556605	73414741	8568.24024
hsgrads	7689.70000	54.53688	428124	62.51406	7.90658
labforce	60872	431.71773	75116315	348834	590.62138
income	914794	6487.90071	19631090146	97828554	9890.83180
crimes	7925316	56208	1.519761E12	7673537171	87599

The REG Procedure
Model: MODEL1
Dependent Variable: crimes

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	1.048425E12	1.164916E11	589.88	<.0001
Error	131	25870458264	197484414		
Corrected Total	140	1.074295E12			

Root MSE	14053	R-Square	0.9759
Dependent Mean	56208	Adj R-Sq	0.9743
Coeff Var	25.00167		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-9364.05829	11780	-0.79	0.4281
landarea	1	1.13511	0.49204	2.31	0.0226
totpop	1	51.13477	20.10594	2.54	0.0121
urban	1	189.32091	73.95737	2.56	0.0116
oldfolks	1	641.62331	517.64266	1.24	0.2174
doctors	1	18.02869	2.23884	8.05	<.0001
hospbeds	1	-4.40220	0.79999	-5.50	<.0001
hsgrads	1	-95.20798	166.84251	-0.57	0.5692
labforce	1	-36.95021	35.65215	-1.04	0.3019
income	1	2.51972	2.14283	1.18	0.2418