1. Consider the regression model

$$y_i = 2.0 + 5.0\,x_i + \varepsilon_i \qquad \varepsilon_i \text{ iid } \mathrm{N}(0,\sigma^2).$$

Suppose we have $n = 100$ data points. If $x_{10} = 2.0$, what is the probability that the value of y that is observed (i.e. y_{10}) is more than 1.5σ above the true regression line?

2. Consider the regression model

$$Sales_t = \alpha + \beta Adv_t + \varepsilon_t \qquad \varepsilon_t \text{ iid } N(0, \sigma_\varepsilon^2 = 9)$$

Note that $\sigma_\varepsilon^2 = 9$ is known. Suppose we decide to set *Adv* to 1, 1.5, 3 and 4 in the four quarters for which we will collect data.

(a) What is the probability that the *Sales* figure in the first quarter is more than two units above the true regression line?

(b) What is the probability that the average discrepancy between the four *Sales* figures we obtain and the true regression line is bigger than one?

3. Consider the true regression model

$$y_i = \alpha + \varepsilon_i \qquad \varepsilon_i \text{ iid } N(0, 1).$$

Note that $\beta = 0$ and $\sigma_\varepsilon^2 = \text{Var}(\varepsilon) = 1$ are known.

Suppose we collect a sample of $n = 100$ observations on y to estimate α. The *Descriptive Statistics* for this sample are given below.

Descriptive Statistics

Y	
Mean	0.0399547
Standard Error	0.102021002
Median	0.08434
Mode	#N/A
Standard Deviation	1.020210019
Sample Variance	1.040828482
Kurtosis	0.121153568
Skewness	-0.130734752
Range	5.39217
Minimum	-3.01008
Maximum	2.38209
Sum	3.99547
Count	100

(a) Using the information in the *Descriptive Statistics*, what is the estimate of α in the regression model?

Now suppose we decide to collect a second sample of $n = 100$ observations.

(b) What is the probability that the estimate of α we obtain from the second sample of $n = 100$ observations will be within 0.15 of the true α?

4)

(21 points) Suppose a company has three divisions and that annual sales for each division are measured in millions of dollars. Further, suppose the annual sales for each division are independent and that they have the same $N(\mu, \sigma^2)$ distribution. Also, annual sales for a specific division are independent for each year.

Unfortunately, the company does not provide annual sales broken down by division. However, they do provide total annual sales for the company, or equivalently, average annual sales per division (because average sales per division can be obtained by dividing total company sales by three) for the past ten years.

Let $\overline{Sales_1}$ represent average annual sales per division for the first year, $\overline{Sales_2}$ represent average annual sales per division for the second year, etc., and let $\overline{Sales_{10}}$ represent average annual sales per division for the tenth year.

(a) (7 points) What is the probability that average annual sales per division for the company next year (i.e. $\overline{Sales_{11}}$) will be greater than $\mu + \sigma$?

The average sales per division (in millions of dollars) for each year are given in the following table for the ten year period for which we have data.

Year	Average Division Sales
1	52.911
2	51.411
3	48.785
4	49.250
5	50.783
6	50.741
7	51.796
8	51.119
9	51.105
10	50.648

The *Descriptive Statistics* for these data are given below.

Average Division Sales	
Mean	50.85487937
Standard Error	0.372234802
Median	50.94375486
Mode	#N/A
Standard Deviation	1.177109799
Sample Variance	1.385587478
Kurtosis	0.630363633
Skewness	-0.278976336
Range	4.125142637
Minimum	48.78549935
Maximum	52.91064198
Sum	508.5487937
Count	10

Using the data and their *Descriptive Statistics*, answer parts (b) and (c).

(b) (7 points) What is the estimate of μ?

(c) (7 points) What is the estimate of σ^2? Hint: You are given the *Descriptive Statistics* for the average annual sales per division for the company (i.e. you are given the *Descriptive Statistics* for $\overline{Sales_1}$, $\overline{Sales_2}$, …, $\overline{Sales_{10}}$). However, you are asked to provide an estimate of σ^2 where σ^2 represents the uncertainty associated with the annual sales of a specific division.

5. (15 points) Consider the following multiple regression model

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \qquad \varepsilon_i \text{ iid } N(0, \sigma_\varepsilon^2).$$

The $n = 30$ data points given below (with several observations removed to conserve space) are used to estimate the regression model. Using the Excel output on the next page, answer parts (a) and (b).

(a) (7 points) What is the estimate of the expected change in y if the x_1-value increases from 1 to 2 and the x_2-value is held constant at $x_2 = 0$.

For part (b), suppose we estimate the model

$$y_i = \alpha + \beta_1 x_{1i} + \varepsilon_i \qquad \varepsilon_i \text{ iid } N(0, \sigma_\varepsilon^2)$$

i.e. we inadvertently omit x_2 from the regression model.

(b) (8 points) Is there enough information available to determine if the R^2 value obtained from the regression model with a single explanatory variable will be larger, smaller or the same as the R^2 value from the multiple regression model?

If so, will the R^2 value obtained from the regression model with a single explanatory variable be larger, smaller or the same as the R^2 value from the multiple regression model? You must provide a brief justification of your conclusion to receive credit.

If not, explain what additional information you would need to make this determination.

Data

Obs. #	y	x1	x2
1	3.787	0.771	-0.730
2	7.273	1.194	-2.334
3	-3.941	0.460	3.086
4	1.555	1.315	0.629
5	1.106	1.016	1.122
6	-9.692	-0.744	4.492
7	12.015	3.596	-1.899
⋮	⋮	⋮	⋮
24	9.080	1.249	-2.807
25	2.551	3.562	3.285
26	1.086	4.451	4.345
27	-4.804	-0.117	2.404
28	-10.506	0.795	6.193
29	16.267	3.911	-3.634
30	1.276	-1.615	-2.101

Excel output for question #5

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.988840295
R Square	0.977805128
Adjusted R Square	0.976161064
Standard Error	1.023796853
Observations	30

ANOVA

	df	*SS*	*MS*
Regression	2	1246.783406	623.3917028
Residual	27	28.3003199	1.048159996
Total	29	1275.083725	

	Coefficients	*Standard Error*	*t Stat*
Intercept	0.783489137	0.239665537	3.269093866
x1	2.142654548	0.094033639	22.78604312
x2	-2.031349388	0.070717874	-28.72469547