Business 320, Fall 1999, Final, Solutions

name ____________________________________________

You may use a calculator and two “cheat sheets”.

You have 3 hours.

I pledge my honor that I have not violated the Honor Code during this examination.
Obviously, you may not discuss the exam with anyone until after exams are over.

Signature___________________________________________

There are 10 questions.
Each part of each question is worth 2 points.

question 1, 8 parts, 16 points ______
question 2, 7 parts, 14 points_____
question 3, 12 parts, 24 points _____
question 4, 6 parts, 12 points ______
question 5, 3 parts, 6 points ______
question 6, 9 parts, 18 points_______
question 7, 4 parts, 8 points________
question 8, 9 parts, 18 points_______
question 9, 5 parts, 10 points_______
question 10, 4 parts, 8 points_______

Total __________________
**Question 1**

Let R1, R2, R3, and R4 represent returns on 4 different assets.

Both of R1 and R2 are independent of both of R3 and R4.
All four assets have mean return .1 and standard deviation .2.
The correlation between R1 and R2 is .7 and the correlation between
R3 and R4 is .5.

Let P1 = .5*R1+.5*R2, P2 = .5*R3+.5*R4, P3 = .25*R1+.25*R2+.25*R3+.25*R4.

(a) What is the covariance between R1 and R2?
\[.028\]

(b) What is E(P1)?
\[.1\]

(c) What is Var(P1)?
\[.034\]

(d) What is E(P2)?
\[.1\]

(e) What is Var(P2)?
\[.03\]

(f) What is the correlation between P1 and P2?
\[0\]

(g) What is E(P3)?
\[.1\]

(h) What is Var(P3)?
\[.016\]
Question 2

In all parts below, refer to this scatter plot of Y vs X.
All parts are multiple choice. Circle your choice.

(a) The average value of x is

(i) .12  (ii) -1.2  (iii) -46.4  (iv) 9.8

(b) The standard deviation of x is

(i) .97  (ii) 1.89  (iii) .021  (iv) -1.6

(c) The correlation between y and x is

(i) .91  (ii) -.91  (iii) -.54  (iv) -.98
(d) The intercept estimate in the regression of Y on X is

(i) **10.0**  (ii) 15.0  (iii) 5.0  (iv) .91

(e) The slope estimate in the regression of Y on X is

(i) **-1.98**  (ii) 1.78  (iii) -.91  (iv) -3.9

(f) The standard error of the mean of the x values is

(i) .01  (ii) 1.01  (iii) 2.7  **(iv) .1**

(g) The estimate of the error standard deviation is

(i) -2.1  (ii) .13  (iii) **1.8**  (iv) 4.8
Question 3

Use the regression output below to answer the questions on the next page.

The regression equation is

\[
\text{norway} = 0.00207 + 0.846 \text{ usa}
\]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>StDev</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.002074</td>
<td>0.006249</td>
<td>(i) 0.741</td>
<td></td>
</tr>
<tr>
<td>usa</td>
<td>0.8459</td>
<td>0.1748</td>
<td>4.84</td>
<td>(ii)</td>
</tr>
</tbody>
</table>

\[ S = (v) \quad R^2 = (iv)\% \quad R^2(\text{adj}) = 17.5\% \]

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>(iii) 0.084011</td>
<td>23.42</td>
<td>(vi)</td>
<td></td>
</tr>
<tr>
<td>Residual Error</td>
<td>105</td>
<td>0.376610</td>
<td>0.003587</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>106</td>
<td>0.460621</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Unusual Observations

<table>
<thead>
<tr>
<th>Obs</th>
<th>usa</th>
<th>norway</th>
<th>Fit</th>
<th>StDev Fit</th>
<th>Residual</th>
<th>St Resid</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-0.030</td>
<td>0.13000</td>
<td>(vii)</td>
<td>0.00955</td>
<td>(viii)</td>
<td>2.59R</td>
</tr>
<tr>
<td>18</td>
<td>0.090</td>
<td>0.04000</td>
<td>0.07820</td>
<td>0.01458</td>
<td>-0.03820</td>
<td>-0.66 X</td>
</tr>
<tr>
<td>24</td>
<td>-0.060</td>
<td>(ix)</td>
<td>-0.04868</td>
<td>0.01408</td>
<td>0.14868</td>
<td>2.55R</td>
</tr>
<tr>
<td>28</td>
<td>0.090</td>
<td>0.10000</td>
<td>0.07820</td>
<td>0.01458</td>
<td>0.02180</td>
<td>0.38 X</td>
</tr>
<tr>
<td>31</td>
<td>-0.090</td>
<td>-0.06000</td>
<td>-0.07405</td>
<td>0.01899</td>
<td>0.01405</td>
<td>0.25 X</td>
</tr>
<tr>
<td>46</td>
<td>-0.040</td>
<td>-0.16000</td>
<td>-0.03176</td>
<td>0.01099</td>
<td>-0.12824</td>
<td>-2.18R</td>
</tr>
<tr>
<td>47</td>
<td>0.110</td>
<td>0.14000</td>
<td>0.09512</td>
<td>0.01784</td>
<td>0.04488</td>
<td>0.79 X</td>
</tr>
</tbody>
</table>
(a) (i) = .33

(b) (ii) = 0

(c) (iii) = .084

(d) (iv) = 18%

(e) (v) = .06

(f) (vi) = 0

(g) (vii) = -.023

(h) (viii) = .1533

(i) (ix) = .1

(j) Give a 95% confidence interval for the true slope.

.8459 +/- .35

(k) Test at level .05 the null hypothesis that the true slope = 0.

reject

(l) For the null hypothesis that the true slope = .5, what, approximately, is the p-value? .05
Question 4

Suppose I am teaching a class of 100 MBA students. For each student, the probability that the student hates Statistics is .75 and the hatred outcome is independent for each student.

(a)

What is the distribution of the number of students who hate Statistics in the class of 100.

\( B(100, .75) \)

(b)

What is the expected value of the number of students who hate Statistics.

75

(c)

What is the variance of the number of students who hate Statistics.

18.75

(d)

Give an approximate 95% interval for the number of students in the class who hate Statistics.

(66, 84)
(e) Suppose I view the 100 students in the class as a sample of all students I might teach and find that 50 hate Statistics (I am no longer willing to assume I know the probability is .75.)
Give a 95% confidence interval for the true probability that a randomly chosen student hates Statistics. What assumptions are you making?

.5 +/- .1
You are assuming that the 100 students are an iid bernoulli sample.

(f) Using the same information as in part (e), test the null hypothesis (at level .05) that the true $p = .75$.

$z = -5.77$, reject
Question 5

Below are dotplots of data c1, c2, and c3. For each data set we tested normality and got the three p-values, .024, .068, and .2.

(a) c1 goes with the p-value__2__ (b) c2 goes with __.068__ (c) c3 goes with ____024____
Question 6

Suppose you are a salesperson. You sell cars (kind of a depressing problem already, it gets worse).

Let $X_1$ denote the number of cars you sell in the first period. and $X_2$ denote the number of cars you sell in the second period.

You believe that $X_1$ will be either 1, 2, or 3 and each possibility is equally likely. You also believe that $X_1$ and $X_2$ are iid.

(a)

What is the expected value of $X_1$?

2

(b)

What is the variance of $X_1$?

$2/3$

(c)

Write out a table (in our usual format) giving the joint distribution of $X_1$ and $X_2$.

1/9 in each place (a three by three table)
(d) Suppose you will get fired if you sell less than 2 cars in the second period. What is the probability that you get fired.

1/3

(e) Suppose you will get fired if you sell less than 2 cars in either period. What is the probability you get fired.

5/9

(f) Suppose you will get fired if the average of $X_1$ and $X_2$ is less than 2. What is the probability that you get fired?

1/3
Given you lose your job there is a .5 probability that your wife will leave you. Otherwise she will not.
Your boss is equally likely to use one of the decision rules in parts (d), (e), and (f) to fire you.

(g)
What is the probability your wife leaves you?

\[(.5) \times \frac{1}{3} \times (\frac{1}{3} + \frac{5}{9} + \frac{3}{9}) = .2037\]

(h)
Suppose we observe that your wife leaves you and we have all the same probabilities in mind. If you had to guess which decision rule (out of the three in (d), (e), and (f)) was used, which one would you guess?

(e)

(i)
Given we observe that your wife has left you, what is the probability that the decision rule you chose in part (h) was used?

\[\frac{5}{11} = \frac{5}{9}/\left(\frac{5}{9} + \frac{1}{3} + \frac{1}{3}\right)\]
Question 7

Suppose your company manufactures Zinglots. You are in charge of the production line. Suppose you have studied the output carefully and believe the process is now such that the lengths of the Zinglots are iid N(50,100).

(a)

You get a phone call at night from a friend who tells you that tomorrow your supervisor is going to take the first Zinglot that comes off the line. If the length is more than 10 away from 50, you will be fired. What is the probability that you are fired?

.32

(b)

Suppose instead your friend tells you that your supervisor is going to take the first 4 Zinglots that come off the line and fire you if the average length of the 4 is more than 10 away from 50. What is the probability that you get fired?

.05
(c)

Suppose your supervisor takes the first 100 Zinglots and uses the average to estimate the true average length of Zinglots.
What is the probability that your supervisor’s estimate is wrong by more than 2?

.05

(d)

Suppose the supervisor will take the first n Zinglots off the line and use their average as an estimate of the true average length.
You get to tell him n.
If you want to have a 95% chance that his estimate is within 1 of the true value what n should you tell him to use?

$2^*\left(\frac{10}{\sqrt{n}}\right) = 1$  so n=400
Question 8

Suppose I want to express the idea that any value between 0 and 1 is equally likely and any value not between 0 and 1 is impossible. Let $X$ be the corresponding random variable.

I will use a probability density function which is $f(x) = c$ if $0 < x < 1$, and 0 for all other values of $x$.

(a)

What is the value of $c$?

1

(b)

What is the expected value of $X$?

.5

(c)

The variance of $X$ is one of the below. Choose one of them (circle your choice).

(i) 1  
(ii) 100  
(iii) 1/100  
(iv) $1/12$

(d)

What is the probability that $X$ is within two standard deviations of its mean?

1
(e)

Suppose we want to have a random variable $Y$ which is equally likely to take on any value between -100 and 100 and must between -100 and 100.

Choose constants $a$ and $b$ such that $Y = a + bX$, where $X$ is the random variable above (uniform on $(0, 1)$).

$a = -100$, $b = 200$

(f)

What is the expected value of $Y$?

0

(g)

What is the variance of $Y$?

$200 * 200 / 12 = 3333.33$
(h) Sketch the pdf of Y.

Height is 1/200 between -100 and 100, 0 everywhere else.

(i) Suppose Z is equally likely to be any value between m and n (m<n), but must be between m and n. What are the mean and variance of Z?

\[ Z = m + (n-m)X, \quad E(Z) = \frac{n+m}{2}, \quad \text{Var}(Z) = \frac{(n-m)^2}{12} \]
Question 9

Suppose \[ Y = X_1 + X_2 + \varepsilon \]

and both X's and \( \varepsilon \) are iid \( N(0,1) \).

(a) What is \( E(Y) \)?

0

(b) What is \( Var(Y) \)?

3

(c) If we regress \( Y \) on \( X_1 \), what is the true error variance?

2

(d) What is the covariance between \( Y \) and \( X_1 \)?

1

(e) What is the covariance between \( Y \) and \( X_1 + X_2 \)?

2
Question 10

Suppose we have the usual regression model:

\[ Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \]

Suppose you are going to observe two Y values corresponding to two known x values \( x_1 \) and \( x_2 \).

Let,

\[ B = \frac{Y_2 - \overline{Y}}{x_2 - x_1} \]

(a) What is \( E(B) \)?
\( \beta_1 \)

(b) What is \( \text{Var}(B) \)?
\[ 2\text{Var}(\varepsilon)/((x_2 - x_1)^2) \]

(c) What is \( \text{Pr}(B > \beta_1) \)?
\[ .5 \]

(d)

If you get to choose the two x values to be any values in the interval (0,1), how would you choose them if you want B close to the slope?

\( x_1 \) close to 0, and \( x_2 \) close to 1