Business 320, Final Exam, Spring 99

Name__________________________________________

You are reminded that it is assumed that all work will be your own.

You have 3 hours.
You may use two "cheat" sheets and a calculator.
Each part of each question is worth two points.

q1, 6 parts, 12 points________
q2, 5 parts, 10 points _______
q3, 8 parts, 16 points________
q4, 4 parts, 8 points________
q5, 6 parts, 12 points________
q6, 5 parts, 10 points _______
q7, 6 parts, 12 points________
q8, 5 parts, 10 points_______
q9, 8 parts, 16 points________
q10, 8 parts, 16 points_______
We have 50 observations on two variables, stuff and junk.

Below are dot plots for each (labeled) and the scatterplot. Note that in the scatterplot the axes are not labeled.
(a) The average value of stuff is
   (i) -5.1   (ii) 5    (iii) .56  (iv) 25.7

(b) The standard deviation of stuff is
   (i) 2.2    (ii) 16.3 (iii) 0    (iv) 1.01

(c) The average value of junk is
   (i) 9.67   (ii) .967 (iii) 21.4 (iv) -8.67

(d) The standard deviation of junk is
   (i) 10.32  (ii) -4.56 (iii) 5.21 (iv) 1.67

(e) In the scatter plot is the variable stuff on the vertical or horizontal axis?

   vertical

(f) The correlation between stuff and junk is?

   (i) .21    (ii) 7.3  (iii) .92  (iv) .73
Question 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>StDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>france</td>
<td>107</td>
<td>0.01383</td>
<td>0.01000</td>
<td>0.05494</td>
</tr>
</tbody>
</table>

The output above gives the mean and standard deviation of 107 monthly returns on the French portfolio. Assume they are roughly normal.

(a) Give a 95% confidence interval for the true mean return ($\mu$).

$$se(\text{mean}) = \frac{.053}{\sqrt{107}} = .0053$$

$$ci = .0138 \pm 2 \cdot (.0053) = (.0032, .0244)$$

(b) Test the hypothesis that the true mean is 0 at level .05.

$$t = \frac{.0138}{.0053} = 2.6, \ \text{reject}$$

(c) For the test of part (b) the p-value is

(i) .0000001  (ii) .97  (iii) .011  (iv) .06

Now suppose (this has nothing to do with the previous parts of the question) that in a random sample of 100 workers, 58 are judged to be following correct procedures.

(d) Give a 95% confidence interval for the true fraction ($p$) of workers following correct procedure.

$$\hat{p} = .58, \ se(\hat{p}) = .05, \ ci = (.48,.68)$$

(e) Test the hypothesis that half the worker follow correct procedure at level .05.

$$\frac{.58-.5}{.05} = 8/5 = 1.6, \ \text{fail to reject}$$
Question 3

Let R1, R2, R3, and R4 represent return 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 .1.
The correlation between R1 and R2 is .9 and the correlation between R3 and R4 is .5.


(a) What is the covariance between R1 and R2?
\[ \text{Cov}(R1, R2) = .1 \times .1 \times .9 = .009 \]

(b) What is E(P1)?
\[ E(P1) = .2 \times .1 + .8 \times .1 = .1 \]

(c) What is Var(P1)?
\[ \text{Var}(P1) = .2 \times .2 \times .01 + .8 \times .8 \times .01 + 2 \times .2 \times .8 \times .009 = 0.00968 \]

(d) What is E(P2)?
\[ E(P2) = .5 \times .1 + .5 \times .1 = .1 \]

(e) What is Var(P2)?
\[ \text{Var}(P2) = .5 \times .5 \times .01 + .5 \times .5 \times .01 + 2 \times .5 \times .5 \times .005 = 0.0075 \]

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

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

(h) What is Var(P3)?
\[ \text{Var}(P3) = .25 \times .25 \times .01 + 2 \times .25 \times .25 \times .009 + .25 \times .25 \times .01 + 2 \times .25 \times .25 \times .005 = 0.00425 \]
Question 4

Below are plots of the $x$ vs $p_X(x)$ for the binomial distribution. Each plot corresponds to a different pair of values of $n$ and $p$. The four pairs of values are $(n=10, p=.9)$, $(n=10, p=.5)$, $(n=10, p=.3)$ and $(n=7, p=.5)$. Match the pair with the plot:

(a) $p1$ matches $n=10$ $p=.5$
(b) $p2$ matches $n=10$ $p=.3$
(c) $p3$ matches $n=10$ $p=.9$
(d) $p4$ matches $n=7$ $p=.5$
**Question 5**

Each of the histograms below depicts iid draws from a distribution. Match the histogram with the distribution:
(a) 100 iid draws from Binomial(10,.5) are in d_6____
(b) 20 iid draws from the N(0,1) are in d_4____
(c) 100 iid draws from the t distribution with 3 degrees of freedom are in d_5____
(d) 200 iid draws from the N(0,1) are in d_1____
(e) 100 iid draws from the N(2,.01) are in d_2____
(f) 100 iid draws from the Bernoulli p=.9 are in d_3____
Question 6

Below is the regression output from the regression of monthly returns of a portfolio of French assets on monthly returns of a portfolio of German assets

The regression equation is
france = 0.00488 + 0.694 germany

<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>0.004879</td>
<td>0.003862</td>
<td>(i)</td>
<td>0.209</td>
</tr>
<tr>
<td>germany</td>
<td>(ii) 0.06737</td>
<td>10.30 (iii)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

S = 0.03892     R-Sq = (v)%     R-Sq(adj) = 49.8%

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>0.16085</td>
<td>0.16085</td>
<td>106.16</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>105</td>
<td>(iv) 0.00152</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>106</td>
<td>0.31993</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Five items have been deleted and replace with (i), (ii),(iii), (iv),(v). What are the deleted values?

(i) =1.26
(ii) =.694
(iii) =0
(iv) =.159
(v) =50.3 (or .503)
Question 7

Suppose a student is writing a final exam. The exam has 50 multiple choice questions. Each question has 4 possible answers, one of which must be selected. For each wrong answer a student gets 1 point and for each correct answer a student get 5 points.

Suppose a student guesses by randomly picking one of the 4 possible answers for each of the 50 questions.

(a) What is the distribution of the number of correct answers ?
Binomial(50,.25)

(b) What is the distribution of the number of incorrect answers ?
Binomial(50,.75)

(c) What is the expected value of the number of correct answers ?
50*.25=12.5

(d) What is the variance of the number of correct answers ?
9.375 = 50*.25*.75

(e) What is the expected value of the total points ?
\[ TP = 5N + 1(50-N) = 50+4N \]
\[ E(TP) = 50 + 4*12.5 = 100 \]

(f) What is the variance of the total points ?
\[ 16*9.375 = 150 \]
Question 8

The Annoying Corp introduced a new product to its customers. All customers were mailed promotional material. A random selection of 20% of customers were also called on the telephone.

45% of customers purchased the new product. Marketing determined that out of those customers who purchased the new product one out of nine were called.

(a) What is the probability that a customer chosen randomly out of those who purchased the product was not called?

\[ 1 - \left( \frac{1}{9} \right) = \frac{8}{9} \]

(b) What is the probability that a randomly chosen customer purchased the new product and was called?

\[ p(\text{purchase}) \times p(\text{called} \mid \text{purchased}) = 0.45 \times \frac{1}{9} = 0.05 \]

(c) Marketing says that since one out of nine customers who made a purchase were called, there is evidence that the calls were helpful. (I think any fraction bigger than 0 would lead them to the same conclusion). Their reasoning seems faulty.

What probabilities should be calculated to assess the effect of the calls on sales?

\[ p(\text{purchase} \mid \text{called}), \ p(\text{purchase} \mid \text{not called}) \]

(d) Calculate the probabilities you chose in part (c).

<table>
<thead>
<tr>
<th>call</th>
<th>no</th>
<th>yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>purchase</td>
<td>.4</td>
<td>.15</td>
</tr>
<tr>
<td>yes</td>
<td>.4</td>
<td>.05</td>
</tr>
<tr>
<td>.8</td>
<td>.2</td>
<td></td>
</tr>
</tbody>
</table>

\[ p(\text{purchase} \mid \text{called}) = 0.05 / 0.2 = 0.25, \ p(\text{purchase} \mid \text{not called}) = 0.5 \]

You are more likely to purchase, if not called!!!!

(e) For a randomly chosen customer, is whether they were called independent of whether or not they purchased the new product?

no
Question 9

We are interested in the relationship between the height of a student (shgt below) and the heights of the student’s mother and father (mhgt and fhgt). Below are the outputs from three regressions: shgt on mhgt and fhgt, shgt on fhgt, and shgt on mhgt. There are 83 observations.

The regression equation is
shgt = 17.3 + 0.255 fhgt + 0.528 mhgt

<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>17.260</td>
<td>8.994</td>
<td>1.92</td>
<td>0.059</td>
</tr>
<tr>
<td>fhgt</td>
<td>0.25485</td>
<td>0.09890</td>
<td>2.58</td>
<td>0.012</td>
</tr>
<tr>
<td>mhgt</td>
<td>0.5283</td>
<td>0.1362</td>
<td>3.88</td>
<td>0.000</td>
</tr>
</tbody>
</table>

S = 3.188     R-Sq = 29.7%     R-Sq(adj) = 28.0%

The regression equation is
shgt = 41.6 + 0.398 fhgt

<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>41.574</td>
<td>6.988</td>
<td>5.95</td>
<td>0.000</td>
</tr>
<tr>
<td>fhgt</td>
<td>0.39782</td>
<td>0.09942</td>
<td>4.00</td>
<td>0.000</td>
</tr>
</tbody>
</table>

S = 3.454     R-Sq = 16.5%     R-Sq(adj) = 15.5%

The regression equation is
shgt = 26.6 + 0.659 mhgt

<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>26.645</td>
<td>8.505</td>
<td>3.13</td>
<td>0.002</td>
</tr>
<tr>
<td>mhgt</td>
<td>0.6591</td>
<td>0.1307</td>
<td>5.04</td>
<td>0.000</td>
</tr>
</tbody>
</table>

S = 3.297     R-Sq = 23.9%     R-Sq(adj) = 23.0%
(a) Based on the regression of shgt on mhgt and fhgt, give a 95% confidence interval for the coefficient of mhgt.

\[ 0.5283 \pm 2(0.1362) \]

(b) Based on the regression of shgt on mhgt, give a 95% confidence interval for the coefficient of mhgt.

\[ 0.659 \pm 2(0.1307) \]

(c) In both regressions, test whether the coefficient for mhgt=0 at level .05.

The t's are 3.88 and 5.04, clear rejects.

(d) For the two tests in part (c), what are the p-values?

0

(e) Do you think that the correlation between mhgt and fhgt is positive or negative.

Taller people tend to marry taller people, positive

(f) Intuitively, why is the coefficient for mhgt in the regression on mhgt and fhgt smaller than in the regression on mhgt alone?

When you just include the mother's height it has to proxy for the overall height of the couple so the coefficient is bigger.
(g) Suppose a student has a mother whose height is 65.
   Give an approximate 95% predictive interval for the height of the student.

26.6 + .659 * 65 = 69.435  
69.435 +/- 2(3.3) = (62.8350, 76.0350)

(h) Suppose a student has a mother whose height is 65 and a father whose height is 70.
   Give an approximate 95% predictive interval for the height of the student.

17.3 + .255 * 70 + .528 * 65 = 69.47  
69.47 +/- 2(3.2) = (63.07, 75.87)
**Question 10**

Suppose a sequence of prices follow a normal random walk:

\[ P_{t+1} = P_t + 5 + \varepsilon_{t+1} \quad t = 1,2,\ldots \]

Let the initial price be 100.

Let \( \varepsilon_t \sim N(0,25) \text{ iid} \)

Note that this implies that,

\[ P_t = 100 + 5t + \sum_{i=1}^{t} \varepsilon_i \]

The current price is the initial price + the sum of all the changes.

(a) What is the expected value of the second price (= \( P_2 \))?

\[ P_2 = 110 + \varepsilon_1 + \varepsilon_2 \text{, expected value} = 110. \]

(b) What is the variance of \( P_2 \)?

\[ \text{Var}(P_2) = \text{Var}(\varepsilon_1 + \varepsilon_2) = 2*25 = 50. \]

(c) Given the first price turns out to be 102, what is the expected value of the second price?

\[ P_2 = 102 + 5 + \varepsilon_2 = 107 + \varepsilon_2 \text{, so the expected value} = 107. \]
(d) Given the first price turns out to be 102, what is the variance of the second price?

\[ P_2 = 107 + \varepsilon_2, \quad \text{Var}(P_2) = \text{Var}(\varepsilon_2) = 25 \]

(e) What is the covariance between the first and second price?

\[
\begin{align*}
E((P_1 - E(P_1))(P_2 - E(P_2))) &= E(\varepsilon_1(\varepsilon_1 + \varepsilon_2)) \\
&= E(\varepsilon_1^2 + \varepsilon_1\varepsilon_2) = 25
\end{align*}
\]

(f) What is the expected value of \( P_t \)?

\[ E(P_t) = 100 + 5t \]

(g) What is the variance of \( P_t \)?

\[ \text{Var}(P_t) = \text{Var}\left(\sum_{i=1}^{t} \varepsilon_i\right) = t \times 25 \]

(h) What is the covariance between \( P_t \) and \( P_{t+1} \)?

\[ t \times 25 \]