Second Midterm Exam  
(MATH1070 Spring 2012)

Instructions: This is a one hour exam. You can use a notecard. Calculators are allowed, but other electronics are prohibited.

1. [60pts] Multiple Choice Problems

   1) A researcher obtained the average SAT scores of all students in each of the 50 states, and the average teacher salaries in each of the 50 states. She found a negative correlation between these variables. The researcher concluded that a lurking variable must be present. By lurking variable she means B
   (A) any variable that produces a large residual.
   (B) a variable that is not among the variables studied but which affects the response variable.
   (C) the true variable, which is explained by the explanatory variable.
   (D) the true cause of a response.

   2) The television program Nightline once asked viewers whether the United Nations should continue to have its headquarters in the United States. Two phone numbers were given. Viewers were asked to call one telephone number to respond “Yes” and the other to respond “No”. More than 186,000 callers responded and 67% said “No”. This is an example of A
   (A) voluntary response sampling.
   (B) a survey with little bias because a large SRS was used.
   (C) a survey with little bias since someone who called would know his or her opinion.
   (D) all of the above.
A researcher studying the effect of price cuts on consumers’ expectations makes up two different histories of the store price of a hypothetical brand of laundry detergent for the past year. Eight students in a class view one or the other price history on a computer. Some students see a steady price, while others see regular sales that temporarily cut the price. Students are asked the price they would expect to pay.


Using the list of random digits: 41842 81868 71035 09001 43367 49497, start at the beginning of this list and use single-digit labels to assign the first four subjects selected to have the steady price group, and the remaining four to the fluctuating price group. Based on this example answer questions 3)—6).

3) The subjects assigned to the FLUCTUATING price group are C
   (A) Rust, Walsh, Gofberg, and Williams.
   (B) Franklin, James, Wright, and Edwards.
   (C) Wright, Rust, Walsh, and Gofberg.
   (D) Edwards, Franklin, Williams, and James.

4) The response is A
   (A) the price they would expect to pay.
   (B) the laundry detergent.
   (C) the business students.
   (D) the sale prices.

5) The experimental units are B
   (A) the business students who were in the fluctuating price group.
   (B) the eight business students who participated.
   (C) the price they would expect to pay.
   (D) all business students at the college.

6) This is an example of A
   (A) a randomized comparative experiment.
   (B) a matched pairs experiment.
   (C) a randomized observational study.
   (D) a two-factor design.
A congressman wants to know how voters in his district feel about a government bailout of the auto industry. He mails a questionnaire on this subject to an SRS of 1200 voters in his district. Of the 1200 questionnaires mailed, 221 were returned and of these 182 were opposed to the bailout. Based on this example answer questions 7)—8).

7) The population is B
   (A) the 182 voters opposed to the bailout.
   (B) the voters in his district.
   (C) the 221 questionnaires returned.
   (D) the 1200 voters receiving the questionnaire.

8) This is an example of C
   (A) a survey with little bias since people understand whether they support the bailout.
   (B) a survey with little bias because 1200 voters represent an important part of his district.
   (C) a survey containing nonresponse.
   (D) all of the above.

9) A campus organization has 80 student members and 10 faculty members. They can send 5 people to a national convention and they would like to send 3 students and 2 faculty members. Of the 80 students, 3 are selected at random and of the 10 faculty members, 2 are selected at random. This is an example of A
   (A) stratified random sampling.
   (B) a census.
   (C) simple random sampling.
   (D) nonvoluntary response sampling.

10) I toss a penny and observe whether it lands heads up or tails up. Suppose the penny is fair, i.e., the probability of heads is 1/2 and the probability of tails is 1/2. This means C
    (A) regardless of the number of flips, half will be heads and half tails.
    (B) every occurrence of a head must be balanced by a tail in one of the next two or three tosses.
    (C) if I flip the coin many, many times, the proportion of heads will be approximately 1/2, and this proportion will tend to get closer and closer to 1/2 as the number of tosses increases.
    (D) all of the above.
11) A North American roulette wheel has 38 slots, of which 18 are red, 18 are black, and 2 are green. If you bet on red, the probability of winning is $18/38 = .4737$. The probability .4737 represents (B) the proportion of times this event will occur in a very long series of individual bets on red.

12) Suppose you decide to bet on red on each of 10 consecutive spins of the roulette wheel described in problem 11), above. Suppose you lose all 5 of the first wagers, which means that on the first 5 spins of the wheel, red has not come up once. Which of the following is true? (C) What happened on the first 5 spins tells us nothing about what will happen on the next 5 spins.
2. [15pts] You roll a pair of dice and compute the number of spots on the two sides facing up. Denote this total by $X$. The probability distribution of $X$ is

<table>
<thead>
<tr>
<th>$X$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>$\frac{1}{36}$</td>
<td>$\frac{2}{36}$</td>
<td>$\frac{3}{36}$</td>
<td>$\frac{4}{36}$</td>
<td>$\frac{5}{36}$</td>
<td>$\frac{6}{36}$</td>
<td>$\frac{5}{36}$</td>
<td>$\frac{4}{36}$</td>
<td>$\frac{3}{36}$</td>
<td>$\frac{2}{36}$</td>
<td>$\frac{1}{36}$</td>
</tr>
</tbody>
</table>

1) What is the probability that $X = 12$?
2) What is the probability that $X$ is a 2, 11, or 12?
3) What is the probability that $X$ is at least 7?

1) Since all the probabilities in the above table should sum to 1, we obtain that

$$P(X = 12) = 1 - \left( \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} \right)$$

$$= \frac{1}{36}.$$

2)

$$P(X = 2, \text{ or } X = 11, \text{ or } X = 12)$$

$$= P(X = 2) + P(X = 11) + P(X = 12)$$

$$= \frac{1}{36} + \frac{2}{36} + \frac{1}{36}$$

$$= \frac{4}{36} = \frac{1}{9}.$$

3)

$$P(X \geq 7) = P(X = 7) + P(X = 8) + P(X = 9)$$

$$+ P(X = 10) + P(X = 11) + P(X = 12)$$

$$= \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36}$$

$$= \frac{21}{36} = \frac{7}{12}.$$
3. [15pts] A student wonders if people of similar heights tend to date each other. She measures herself, her dormitory roommate, and the women in the adjoining rooms; then she measures the next man each woman dates. Here are the data (heights in inches).

<table>
<thead>
<tr>
<th>Women (x)</th>
<th>66</th>
<th>64</th>
<th>66</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men (y)</td>
<td>72</td>
<td>68</td>
<td>70</td>
</tr>
</tbody>
</table>

We already found that the correlation $r=0.866$, the standard deviation of women heights is 1.1547, and the standard deviation for men heights is 2.

1). Find the regression line.
2). If a woman is 67 inches tall, what is your best guess of the height of her next date?
3). How good is your prediction for 2) and why?

1) slope:

$$b = r \cdot \frac{s_y}{s_x} = 0.866 \cdot \frac{2}{1.1547} = 1.5000$$

y-intercept:

$$a = \bar{y} - b \cdot \bar{x} = 70 - 1.5(65.333) = -28$$

Therefore the regression line is

$$\hat{y} = 1.5x - 28$$

2) The prediction for a woman 67 inches tall is

$$\hat{y} = 1.5(67) - 28 = 72.5$$

3) We can use $r^2$ to measure the effectiveness of our predication. Since $r = 0.866$, $r^2 = 0.75$. This means 75% of the changes in $y$ can be explained by changes in $x$. Therefore this is a good predication.
4. [10pts] The distribution of actual weights of 8 oz. wedges of cheddar cheese produced at a dairy is Normal with mean 8.1 ounces and standard deviation 0.1 ounces. Consider a random sample of NINE of these cheese wedges.

1). What is the probability distribution of the average weight of a random sample of NINE of these cheese wedges?

2). What is the probability that their average weight is less than 8 oz?

1) Since the population is normally distributed with mean 8.1 and standard deviation 0.1, the sample mean of a size 9 sample is a random variable with normal distribution with mean 8.1 and standard deviation $0.1/\sqrt{9} = 0.0333$, where 9 is the sample size. If we let $\bar{X}$ to denote the sample mean, then $\bar{X} \sim N(8.1, 0.0333)$.

2) Since $\bar{X} \sim N(8.1, 0.0333)$, the $z$-score for 8 is

$$z = \frac{8 - 8.1}{0.0333} = -3.$$

Therefore, according to the normal table, the probability that their average weight is less than 8 oz is

$$P(\bar{X} < 8) = 0.0013.$$