Practice Exam 1
Statistics for Business (22S:008, Bogner) **September 14, 2006**

Name: Key

TA: ____________________

Section: ____________________

Make sure you understand the instructions!

- **Instructions:**
  - Use only a number 2 pencil.
  - Fill in your name, TA, and discussion section on this test form.
  - Fill in your name, student ID number, discussion code (see below), and test form (see below) on the answer sheet.
  - Read *all* of the answers, choose the *best* answer, grid your choice on the answer sheet, and circle your answer on this test form.
  - Turn in both the answer sheet and test form when you are finished.
  - Keep your crib sheet; you will be able to use it on the final exam.
- **Discussion Code:** Fill in your discussion code in columns J and K on the answer sheet. Example: Students in Joe's 8:30–9:20 section ($A03$) should fill in 0 in column J and 3 in column K. Students in Sorah's 3:30–4:20 section (SCA) should fill in 2 in column J and 1 in column K.
  - $A03$ = Joe (8:30–9:20) .... $J = 0, K = 3$
  - $A04$ = Greg (8:30–9:20) .... $J = 0, K = 4$
  - $A06$ = Zhanchong (8:30–9:20) .... $J = 0, K = 6$
  - $A07$ = Katherine (9:30–10:20) .... $J = 0, K = 7$
  - $A08$ = Greg (9:30–10:20) .... $J = 0, K = 8$
  - $A09$ = Joe (10:30–11:20) .... $J = 0, K = 9$
  - $A10$ = Wade (10:30–11:20) .... $J = 1, K = 0$
  - $A11$ = Barry (11:30–12:20) .... $J = 1, K = 1$
  - $A12$ = Wade (11:30–12:20) .... $J = 1, K = 2$
  - $A13$ = Katherine (12:30–1:20) .... $J = 1, K = 3$
  - $A14$ = Dan (12:30–1:20) .... $J = 1, K = 4$
  - $A15$ = Yiwen (1:30–2:20) .... $J = 1, K = 5$
  - $A16$ = Dan (1:30–2:20) .... $J = 1, K = 6$
  - $A17$ = Yiwen (2:30–3:20) .... $J = 1, K = 7$
  - $A18$ = Jie (2:30–3:20) .... $J = 1, K = 8$
  - SCA = Sorah (3:30–4:20) .... $J = 2, K = 1$
  - SCB = Jie (3:30–4:20) .... $J = 2, K = 2$
  - SCC = Jay (4:30–5:20) .... $J = 2, K = 3$
- **Exam Form:** An exam form is printed at the bottom of this page. Fill in the type of form (i.e. A, B, C, or D) on your answer sheet.
- **Scoring:** Each of the 25 questions on the exam is worth 4 points (100 points total). Incorrectly following these instructions will result in a 3 point deduction.

Exam Form: A

1. Suppose the median is equal to 100 in a graph which is skewed to the right. Which of the following best describes the mean?

   (a) The mean will be less than 100
   (b) The mean will be greater than 100
   (c) The mean will be equal to 100
   (d) There is not enough information to determine the location of the mean relative to the median
   (e) The median does not exist

2. Which of the following is/are true?

   (a) In general, surveys that use convenience sampling yield results that accurately reflect the views of the population
   (b) A study is biased if it favors a certain result
   (c) Stratified sampling is analogous to drawing names from a hat
   (d) Both (b) and (c)
   (e) All of the above

3. Consider the following stem and leaf plot. What is the interquartile range and median?

Stem	Leaves
7	6
8	
9	1
10	5
11	33

$\underbrace{76, 91}\quad \underset{Q_2}{105} \quad \underbrace{113, 113}$

$Q_1 = \frac{76+91}{2} = 83.5$

$Q_3 = 113$

$IQR = Q_3 - Q_1 = 29.5$

   (a) IQR = 22, median = 91
   (b) IQR = 29.5, median = 105
   (c) IQR = 22, median = 105
   (d) IQR = 29.5, median = 99.6
   (e) IQR = 37, median = 99.6

4. The following data describe the heights (in inches) of a mens basketball team: 78, 73, 76, 77. What is the standard deviation of these data?

   (a) 0
   (b) 1.87
   (c) 2.16
   (d) 3.74
   (e) 4.67

$n = 4$

$\bar{x} = 76$

$s = \sqrt{\frac{(x_1-\bar{x})^2 + \cdots + (x_n-\bar{x})^2}{n-1}} = \sqrt{\frac{(78-76)^2+(73-76)^2+(76-76)^2+(77-76)^2}{4-1}} = \sqrt{\frac{4+9+0+1}{3}} = \sqrt{4.667}$

$= 2.160$

5. As motivated in class, the ability to read mathematical equations is very important. Consider the following equation: $SS = \sqrt{x_1^2/2} + \sqrt{x_2^2/2} + \cdots + \sqrt{x_n^2/2}$. Compute SS for the following dataset: 2, 4, 5.

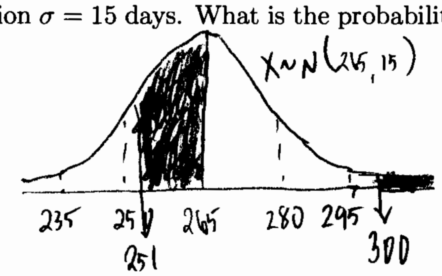
   (a) $SS = 5.50$
   (b) $SS = 22.50$
   (c) $SS = 45.00$
   (d) $SS = 4.74$
   (e) $SS = 7.78$

$SS = \sqrt{x_1^2/2} + \cdots + \sqrt{\frac{x_n^2}{2}} = \sqrt{2^2/2} + \sqrt{4^2/2} + \sqrt{5^2/2} = \sqrt{2} + \sqrt{8} + \sqrt{12.5}$

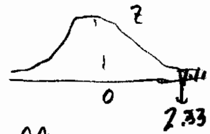
$= 7.78$

6. The length of pregnancies in the U.S. are normally distributed with a mean $\mu = 265$ days and standard deviation $\sigma = 15$ days. What is the probability of a pregnancy lasting 300 days or longer?

   (a) 0.9901
   (b) 0.0099
   (c) 0.9750
   (d) 0.0025
   (e) 0.1587



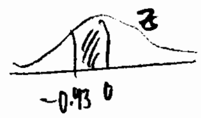
$P(X>300) = P\left(\frac{X-265}{15} > \frac{300-265}{15}\right)$
$= P(Z>2.33)$
$= 1-P(Z<2.33)$
$= 1-0.9901 = 0.0099$



7. In reference to question (6), what percentage of pregnancies last between 251 and 265 days?

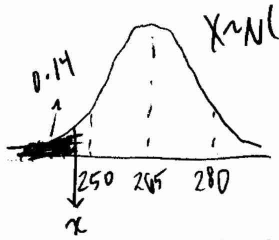
   (a) 17.62%
   (b) 15.87%
   (c) 82.38%
   (d) 32.38%
   (e) 16.00%

$P(251<X<265) = P\left(\frac{251-265}{15} < \frac{X-265}{15} < \frac{265-265}{15}\right) = P(-0.93<Z<0)$
$= 0.5 - 0.1762$
$= 0.3238$



8. In reference to question (6), suppose 14% of babies are born premature. How many days must a pregnancy last before the baby is no longer considered premature?

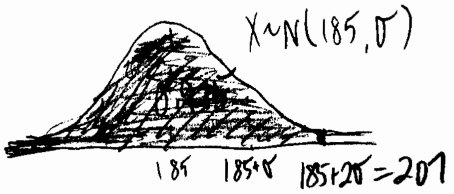
   (a) 281.2 days
   (b) 280.0 days
   (c) 259.5 days
   (d) 250.0 days
   (e) 248.8 days



Find x. For the 14th %ile, $z = -1.08$
$\Rightarrow z = \frac{x-\mu}{\sigma} \Rightarrow x = \mu + z\sigma$
$= 265 + (-1.08)(15)$
$= 248.8$

9. Cholesterol levels in men are normally distributed with mean $\mu = 185$. Out of 2000 randomly chosen men, 1950 had a cholesterol level less than 207. Use the empirical rule to find the standard deviation σ.

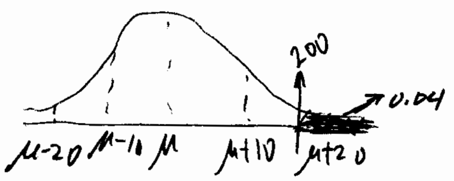
   (a) $\sigma = 22$
   (b) $\sigma = 11$
   (c) $\sigma = 33$
   (d) $\sigma = 8$
   (e) Impossible to determine with the given information



$\frac{1950}{2000} = 0.975$
$\Rightarrow 185 + 2\sigma = 207$
$\Rightarrow \sigma = \frac{207-185}{2} = 11$

10. Cholesterol levels in women are normally distributed with standard deviation $\sigma = 10$. It is known that 4% of women have a cholesterol level more than 200. Find the mean cholesterol level μ.

   (a) $\mu = 210.0$
   (b) $\mu = 217.5$
   (c) $\mu = 182.5$
   (d) $\mu = 185.0$
   (e) Impossible to determine with the given information



~~The~~ For the 96th %ile, $z = 1.75$

$z = \frac{x-\mu}{\sigma} \Rightarrow \mu = x - z\sigma$
$= 200 - 1.75(10)$
$= 182.5$

11. Suppose a die is rolled one time. Consider the following events:

$$A = \text{an even is rolled}$$
$$B = \text{a 3 or 6 is rolled}$$

What is $P(B|A)$?

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{P(\text{roll } 6)}{P(A)} = \frac{1/6}{3/6} = 0.333$$

(a) 0.333
(b) 0.500
(c) 0.250
(d) 0.400
(e) None of the above

12. In reference to question (11), are A and B independent? Why?

(a) Not independent, since $P(B|A) \neq P(A)$
(b) Not independent, since $P(A|B) = P(A)$
(c) Not independent, since $P(A \text{ and } B) \neq P(A)$
(d) Independent, since $P(A \text{ and } B) \neq P(A)$
(e) Independent, since $P(B|A) = P(B)$

Know that $P(B|A) = 0.333$.
And $P(B) = 2/6 = 0.333$.
Since $P(B|A) = P(B)$ then indep

13. In reference to question (11), are A and B mutually exclusive? Why?

(a) Mutually exclusive, since $P(A \text{ or } B) > 0$
(b) Mutually exclusive, since $P(A \text{ and } B) > 0$
(c) Not mutually exclusive, since $P(A \text{ or } B) > 0$
(d) Not mutually exclusive, since $P(A \text{ and } B) > 0$
(e) None of the above

14. A box contains 18 black, 12 red, and 10 green balls. Two balls are chosen from the box *without* replacement. What is the correct equation needed to find the probability that the first ball is green and the second ball is green (i.e. $P(G_1 \text{ and } G_2)$)?

(a) $P(G_1 \text{ and } G_2) = P(G_1)P(G_2)$
(b) $P(G_1 \text{ and } G_2) = P(G_2)P(G_2|G_1)$
(c) $P(G_1 \text{ and } G_2) = P(G_2|G_1)$
(d) $P(G_1 \text{ and } G_2) = P(G_1)/P(G_2|G_1)$
(e) None of the above

Not indep!.
Use gen. multi. rule.
Be Careful! $P(G_1 \text{ and } G_2) = P(G_1)P(G_2|G_1)$

15. It is known that 14% of all adults have diabetes. It is also known that 4% of all adults have diabetes *and* high blood pressure. Given that a randomly selected person has diabetes, find the probability that he/she has high blood pressure.

(a) 0.006
(b) 0.040
(c) 0.286
(d) 0.140
(e) 0.860

$P(D) = 0.14$
$P(D \text{ and } H) = 0.04$
$$P(H|D) = \frac{P(H \text{ and } D)}{P(D)} = \frac{0.04}{0.14} = 0.286$$

16. Bowl A contains 2 red and 1 black chips. Bowl B contains 3 red and 2 black chips. Suppose a single chip is chosen from each bowl. What is the probability that both chips are red?

(a) 0.40
(b) 0.44
(c) 0.50
(d) 0.63
(e) 0.67

A: R R B B: R R R B B

$P(R_A \text{ and } R_B) \overset{\text{indep}}{=} P(R_A)P(R_B) = \frac{2}{3}\cdot\frac{3}{5} = \frac{6}{15} = 0.40$

17. Referring to question (23), what is the probability that the first chip is red, the second chip is red, or both chips are red?

(a) 1.27
(b) 0.40
(c) 0.63
(d) 0.54
(e) 0.87

Don't be a knucklehead and choose this.

$P(R_A \text{ or } R_B) \overset{\text{not me.}}{=} P(R_A)+P(R_B)-P(R_A \text{ and } R_B)$

$= \frac{2}{3} + \frac{3}{5} - 0.40$

$= 0.87$

18. Referring to question (23), what is the probability that exactly one red chip is drawn?

(a) 0.47
(b) 0.27
(c) 0.20
(d) 0.44
(e) 1.06

$P(\text{exactly } 1\ R) = P(R_A \text{ and } B_B) + P(B_A \text{ and } R_B)$

$\overset{\text{indep}}{=} P(R_A)P(B_B) + P(B_A)P(R_B)$

$= \frac{2}{3}\cdot\frac{2}{5} + \frac{1}{3}\cdot\frac{3}{5} = \frac{7}{15} = 0.47$

19. In the U.S., 70% of adults have health insurance. If 5 adults are randomly selected, find the probability *none* have health insurance.

(a) 0.1681
(b) 3.5
(c) 0.1400
(d) 0.0024
(e) 0.2288

Don't be a knucklehead.

$P(NH_1 \text{ and } NH_2 \cdots \text{ and } NH_5)$

$\overset{\text{indep}}{=} P(NH_1)\cdots P(NH_5)$

$= 0.30^5$

$= 0.0024$

20. Suppose a bowl has 2 chips; one of the chips is labeled "1", and the other chip is labeled "3". Suppose **three** chips are selected at random *with* replacement. Suppose the random variable X equals the *sum* of the three draws. What is the probability that the **sum** of the three draws is equal to 7 (i.e. what is $P(X = 7)$)? *Hint: write down all outcomes of this experiment, and use the technique given in class to find the probability.*

(a) 0.500
(b) 0.375
(c) 0.222
(d) 0.625
(e) 0.333

$S = \{111, 113, 131, 133, 311, 313, 331, 333\}$

All outcomes are equally likely! Makes things nice!

X = sum of 3 draws.

$P(X=7) = 3/8 = 0.375$

21. Which of the following is/are true?

(a) Compared to a case-control study (such as smoking/lung cancer studies), it is difficult to prove cause and effect relationships in randomized experiments

(b) A relative frequency probability is based upon ones experience or intuition

(c) An event is a collection of 1 or more outcomes (i.e. an event is a subset of S)

(d) Two factors are said to be "conglomerated" if their effects can not be distinguished from one another

(e) Both (c) and (d)

22. Suppose a UI professor wants to inquire about the percentage of UI undergraduate students that voted in the 2004 election. Of the 1,000 undergraduate students randomly selected (using simple random sampling) from the UI registrar list, 380 voted. Which of the following is/are true?

(a) The population is the 1,000 questioned students.

(b) The sample is (almost certainly) representative because it was obtained by scientific means.

(c) It is known that some students claimed that they voted when in fact they didn't. This will cause this study to be biased.

(d) The sample consists of the 380 students that voted.

(e) Both (b) and (c).

23. Suppose a bowl has 3 chips; one of the chips is green, one of the chips is red, and one of the chips is black. Suppose two chips are selected at random *with* replacement. What is the probability that a green chip is obtained on both draws? *Hint: Write down all of the outcomes in this experiment, and use the theorem discussed in class.*

$S = \{GG, GR, GB, BB, BG, BR, RR, RG, RB\}$

(a) 2/3

(b) 1/3

(c) 1/2

(d) 1/9

(e) 8/9

$P(G_1 \text{ and } G_2) = \frac{1}{9}$

24. In reference to question (23), what is the probability that no red chips are obtained in the two draws? *Hint: Write down all of the outcomes in this experiment, and use the theorem discussed in class.*

$P(\text{no R in 2 draws}) = \frac{4}{9}$

(a) 2/3

(b) 1/3

(c) 1/2

(d) 1/9

(e) [redacted] 4/9

25. As motivated in class, the ability to read mathematical equations is very important. Consider the following equation: $SS = x_1^2 + x_2^2 + \cdots + x_n^2$. Compute SS for the following dataset: 2, 4, 5.

(a) $SS = 45$

(b) $SS = 11$

(c) $SS = 4$

(d) $SS = 18$

(e) $SS = 35$

$SS = x_1^2 + x_2^2 + x_3^2 = 2^2 + 4^2 + 5^2$
$= 4 + 16 + 25$
$= 45$