STA 110B
Spring 2000

Name ____________________
Section ____________________

Midterm, Form A

07MAR2000

1. This test is CLOSED BOOK; however, you may use one 8.5” × 11” sheet (front and back) of notes.
2. You may use a calculator.
3. You have 75 minutes to complete the exam; you must finish by the end of the period (12:10PM).

Page	1	2	3	4	5	Total
Possible Points	20	24	17	21	18	100
Points Earned						

1. Twenty-five heat lamps are connected in a greenhouse so that when one lamp fails, another takes over immediately. (Only one lamp is turned on at any time.) The lamps operate independently; the lifetime of each lamp is normally distributed with a mean life of 50 hours and a standard deviation of 4 hours.

a. (6 pts) What is the probability of any one such heat lamp burning for more than 56 hours but less than 58 hours?

$$
\begin{aligned}
P(56 < X < 58) &= P(1.5 < Z < 2.0) \\
&= P(Z > 1.5) - P(Z > 2.0) \\
&= 0.067 - 0.023 \\
&= 0.044
\end{aligned}
$$

b. (9 pts) If the greenhouse is not checked for 1300 hours after the lamp system is turned on, what's the probability that a lamp will be burning at the end of the 1300 hour period?

$$
\begin{aligned}
P(\sum_{i=1}^{25} X_i > 1300) &= 1300) \\
&= P(\bar{X} > \frac{1300}{25}) \\
&= P(Z > \frac{52 - 50}{\frac{4}{\sqrt{25}}}) \\
&= P(Z > 2.5) \\
&= 0.006
\end{aligned}
$$

2. Let's say that I have a tendency to lose the dry erase marker that I use in class. After each lecture, I lose the marker while on the way back to my office with probability 0.2. After losing a marker I get a new one for the next class. Whether I lose a marker on any given day is independent of whether I lose one on any other day.

a. (5 pts) What is the probability that after 4 lectures I have lost exactly 2 markers?

$$P(X = 2) = \binom{4}{2}(0.2)^2(0.8)^2 = 0.1536$$

b. (7 pts) What is the probability that after 4 lectures, I will have lost at least 2 markers?

$$
\begin{aligned}
P(X \geq 2) &= P(X = 2) + P(X = 3) + P(X = 4) \\
&= \binom{4}{2}(0.2)^2(0.8)^2 + \binom{4}{3}(0.2)^3(0.8) + \binom{4}{4}(0.2)^4
\end{aligned}
$$

$$
\begin{aligned}
&= 0.1536 + 0.0256 + 0.0016 \\
&= 0.1808
\end{aligned}
$$

c. (9 pts) After 2 semesters, I will have given 100 lectures. What is the approximate probability that I will have lost 16 or fewer markers?

You can approximate (notice the it says "approximate probability") the exact answer, which is given by the binomial formula, using the normal approximation to the binomial formula.

$$
\begin{aligned}
P(X \leq 16) &= P(Z < \frac{\frac{16.5}{100} - \frac{20}{100}}{\sqrt{\frac{(0.2)(0.8)}{100}}}) \\
&= P(Z < -0.875) \\
&\approx P(Z < -0.88) \\
&\approx 0.189
\end{aligned}
$$

3. The apartment units in a certain complex have varying numbers of bedrooms and bathrooms. In fact, the percentages of units satisfying various combinations of bedrooms and bathrooms are given by the following joint probability distribution $p(x, y)$, where X is the number of bedrooms and Y is the number of bathrooms.

	Y	
X	1	2
1	0.20	0
2	0.35	0.30
3	0	0.15

a. (8 pts) Find $Cov(X, Y)$.

$$
\begin{aligned}
Cov(X, Y) &= E(XY) - E(X)E(Y) \\
&= 3 - (1.95)(1.45) \\
&= 0.1725
\end{aligned}
$$

$$E(XY) = (1)(1)(0.20) + (1)(2)(0.35) + (2)(2)(0.30) + (3)(2)(0.15) = 3$$

$$E(X) = (1)(0.20) + (2)(0.35 + 0.30) + (3)(0.15) = 1.95$$

$$E(Y) = (1)(0.55) + (2)(0.45) = 1.45$$

b. (4 pts) Are X and Y independent? Why or why not? (State in mathematical expression and/or in a brief phrase.)

X and Y are NOT independent because $Cov(X, Y) \neq 0$.

c. (8 pts) Let the monthly rent for an apartment unit be denoted by R, where $R = 500 + 75X + 25Y$. Find the average cost of an apartment unit in this complex.

$$\begin{aligned} E(R) &= 500 + 75E(X) + 25E(Y) \\ &= 500 + (75)(1.95) + (25)(1.45) \\ &= 682.50 \end{aligned}$$

4. Of the travelers arriving at a small airport, 60% fly on major airlines, 30% fly on privately owned planes, and the remainder fly on commercially owned planes not belonging to a major airline. Of those traveling on major airlines, 50% are traveling for business reasons, whereas 60% of those arriving on private planes and 90% of those arriving on other commercially owned planes are traveling for business reasons. Suppose that we randomly select one person arriving at this airport. What is the probability that the person

a. (5 pts) is not traveling on business?

B: event traveling on business
M: event fly major airline
C: event fly other commercial airline (not major)
R: event fly privately owned plane

$$\begin{aligned} P(B^c) &= 1 - P(B) \\ &= 1 - [P(B|M)P(M) + P(B|R)P(R) + P(B|C)P(C)] \\ &= 1 - [(0.50)(0.60) + (0.60)(0.30) + (0.90)(0.10)] \\ &= 1 - 0.57 \\ &= 0.43 \end{aligned}$$

b. (4 pts) is traveling for business on a privately owned plane?

$$\begin{aligned} P(B \text{ and } R) &= P(B|R)P(R) \\ &= (0.60)(0.30) \\ &= 0.18 \end{aligned}$$

c. (9 pts) arrived on a privately owned plane given that the person is traveling on business?

$$\begin{aligned} P(R|B) &= \frac{P(B|R)P(R)}{P(B)} \\ &= \frac{0.18}{0.57} \\ &\approx 0.3158 \end{aligned}$$

5. A researcher randomly selected 18 households, each of which had at least one child and one parent living in the home. He was interested in determining whether the number of hours of TV a child watches affects his/her grade point average (GPA) in school. In each of the households he administered a questionnaire to the parent(s) and asked whether or not the child watched more than 20 hours of TV per week. (Assume the parent gave an accurate report. Also, if there was more than one child, he used a random process to pick one of the children to ask questions about.) The researcher also got permission to get the child's GPA from school. The two datasets that he obtained are:

GPAs for kids who watched more than 20 hours of TV: 1.9, 2.0, 2.0, 2.1, 2.2, 2.3, 3.5, 3.5, 3.9
GPAs for the kids who watched less TV: 1.2, 3.3, 3.3, 3.5, 3.6, 3.8, 3.8, 3.9, 4.0

a. (2 pts) Was this an observational study or an experimental one?

observational study

b. (3 pts) Find the median GPA for students who watched more than 20 hours of TV per week.

2.2

c. (3 pts) Find the range of the GPAs for students who watched less than 20 hours of TV per week.

4.0 - 1.2 = 2.8

d. (4 pts) The researcher concluded that watching a lot of TV causes a child to perform poorly in school. Do you agree? Explain your answer in 2 sentences or less.

Since this is an observational study, we cannot draw conclusions about cause and effect. We don't know if A causes B, if B causes A, or if some other confounding factor C is responsible for A and B. For instance, low levels of parental involvement could be responsible for both increased TV watching and lowered grades.

6. Suppose that two economists estimate μ (the average expenditure of American families on food), with two unbiased (and statistically independent) estimates U and V. The second economist is less careful than the first - the standard deviation of V is 3 times as large as the standard deviation of U. When asked how to combine U and V to get a publishable overall estimate, three proposals are made:

(i) $W_1 = \frac{3}{4}U + \frac{1}{4}V$ (weighted average)
(ii) $W_2 = 1U + 0V$ (drop the less accurate esimate)

a. (6 pts) Which of these two, if any, are unbiased?

We know that U and V re unbiased, so the following must be true:

$$E(U) = E(V) = \mu$$

For W_1 and W_2 to be unbiased, their expected values must also be equal to μ.

$$\begin{aligned} E(W_1) &= E(\frac{3}{4}U) + E(\frac{1}{4}V) \\ &= \frac{3}{4}E(U) + \frac{1}{4}E(V) \end{aligned}$$

$$\begin{aligned} &= \frac{3}{4}\mu + \frac{1}{4}\mu \\ &= \mu \end{aligned}$$

$$\begin{aligned} E(W_2) &= E(1U) + E(0V) \\ &= E(U) + OE(V) \\ &= \mu \end{aligned}$$

So, both W_1 and W_2 are unbiased.

b. (8 pts) Which is the best estimator?

Since both W_1 and W_2 are unbiased, to choose the best estimator we just need to know which has lower variance. (This is equivalent to determining which has the lower mean squared error.) We will need to be able to compare the variances of U and V. We know that the standard deviation of V is 3 times that of U. So, if the standard deviation of U is σ, the standard deviation of V is 3σ.

$$\begin{aligned} Var(U) &= \sigma^2 \\ Var(V) &= 9\sigma^2 \\ Cov(U,V) &= 0 \end{aligned}$$

$$\begin{aligned} MSE(W_1) &= Var(W_1) + (\text{bias of } W_1)^2 \\ &= Var(\frac{3}{4}U + \frac{1}{4}V) + 0 \\ &= \frac{9}{16}Var(U) + \frac{1}{16}Var(V) + (2)(\frac{3}{4})(\frac{1}{4})Cov(U,V) \\ &= \frac{9}{16}\sigma^2 + \frac{1}{16}(9\sigma^2) + 0 \\ &= \frac{18}{16}\sigma^2 \\ &= \frac{9}{8}\sigma^2 \end{aligned}$$

$$\begin{aligned} MSE(W_2) &= Var(W_2) + (\text{bias of } W_2)^2 \\ &= Var(1U + 0V) + 0 \\ &= Var(U) + 0Var(V) + (2)(1)(0)Cov(U,V) \\ &= \sigma^2 + 0 \\ &= \sigma^2 \end{aligned}$$

Since the mean squared error of W_2 is less than that of W_1, W_2 is the better estimator.