Economics 300–01: Quantitative Methods in Economics
Wesleyan University, Spring 2002

Midterm Examination #1
February 21, 2002

Suggested Answers

This document contains the final answers *only* for the first midterm. The purpose of this document is to allow you to compare your answers to the correct ones. You are strongly encouraged to re-work the problems you did not answer correctly on the exam. If you have any questions about how to approach a particular question, please contact me. Note that, as stated in the exam instructions, writing only a final answer was not sufficient to receive credit for a particular question on this exam: all answers must include computations and/or explanations (whichever is most appropriate).

1. On January 1, 2002, the euro became the sole currency in 12 European nations. Britain has not yet decided to join these other nations in using the euro as its currency. A November 2001 survey of 350 British businesses found that, of those who expressed an opinion, 57% favored joining the European Monetary Union and adopting the euro, while 43% did not. *(16 points)*

   (a) Compute the margin of error for this survey at the 98% level of confidence.

   6%

   (Hint: The appropriate value for z is 2.325.)

   (b) Assuming the same proportion of responses and a 98% level of confidence, how many firms would have to be interviewed to yield a margin of error of $\pm 3\%$?

   *Approximately 1473. (Note: must round **up**!)*

2. Consider the following game: For a fee of \$3, you are allowed to toss a fair coin n times. You then get paid $(X^2 - X)$ dollars, where X is the total number of heads tossed. *(20 points)*

   (a) If you are allowed to toss two coins ($n = 2$), what is your expected profit?

   $$\$ -2.50.$$

   (b) How many coins must you toss to break even in expectation? (That is, for your expected profit to be zero?)

   $$n = 4.$$

   (Hint: You can express the expected profit in terms of the mean μ and the variance σ^2 for the appropriate binomial distribution, then solve for the optimal n.)

3. Let A and B be two events defined on some sample space. Suppose $\Pr(B|A) > \Pr(B)$. *(20 points)*

(a) Prove that $\Pr(A) < \Pr(A|B)$.

Key result:

$$\begin{aligned} \Pr(B|A) &= \frac{\Pr(A|B)\Pr(B)}{\Pr(A)} > \Pr(B) \\ &\implies \Pr(A|B)\Pr(B) > \Pr(A)\Pr(B)\,. \end{aligned}$$

(b) Prove that $\Pr(A) > \Pr(A|\bar{B})$. (*Hint: $\bar{B}$ is the complement of B.*)

Solve as above. Key concepts:

1. $\Pr(\bar{B}) = 1 - \Pr(B)$
2. $\Pr(A \cap \bar{B}) = \Pr(A) - \Pr(A \cap B)$
3. $\Pr(B|A) > \Pr(B) \implies 1 - \Pr(B|A) < 1 - \Pr(B)$

4. On a recent (hypothetical) bill before Congress, 400 of the 435 members of the House of Representatives voted as follows: *(22 points)*

X: Party	Y: Vote For (0)	Against (1)
Republican (0)	132	88
Democrat (1)	108	72

(a) Tabulate the joint and marginal probabilities for this vote.

X: Party	Y: Vote 0	1	
0	0.33	0.22	0.55
1	0.27	0.18	0.45
	0.60	0.40	

(b) Compute the correlation between political party and vote.

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}\,. \quad \textit{Since } \sigma_{XY} = 0\,,\ \rho = 0\,.$$

(c) Is support for this bill independent of political affiliation? Justify your answer.

$$p(x) \cdot p(y) = p(x,\, y) \textit{ for all 4 values.}$$

5. The following quote comes from a February 10, 2002, *New York Times* article entitled "The Painful Fact of Medical Uncertainty": *(20 points)*

> For years, women over 40 years old were told that a yearly mammogram could find breast cancer early enough to save them from death. This was medical dogma; it was the truth. But in the early 1990's, doubts grew about whether the test helped women in their 40's, and now some experts say they question whether it saves anyone....
>
> Some of the strongest evidence is from a study begun in the 1960's. It found that after 18 years, 153 out of 30,131 women who had mammograms had died of breast cancer, and 196 out of 30,565 women who did not have the test died of breast cancer. That is a 30 percent difference in breast cancer death rates — but it hinges on the medical histories of just 43 women. Questions about the design and conduct of this study have led some to doubt its conclusion. And similar questions have been raised about other mammography studies....
>
> But many doctors say it is inappropriate for scientists to quibble about the fine points of evidence in the case of a devastating disease like breast cancer. "Unfortunately, the people making these arguments are statisticians," said Dr. Maurie Markman, a specialist in gynecological cancer at the Cleveland Clinic. "I'm not trying to say that statistics are not important," Dr. Markman said. "We can argue about how many angels can dance on the head of a pin. But these are real live patients, and it doesn't help anyone to go through this."

Provide a **brief** reply to Dr. Markman's comments, from a statistical perspective. Keep your comments to *less* than 80 words and *less* than 5 sentences.

Two main issues to address:

(a) Why should doctors be skeptical of the results reported in these early mammogram studies?

(b) Why should they care? (Hint: What if some women undertake chemotherapy or mastectomies unnecessarily as a result?)

6. The noontime air pollution index for downtown Hartford varied as follows for the past 6 months: *(20 points)*

Index Value	Relative Frequency
0 – 10	30%
10 – 20	50%
20 – 30	8%
30 – 40	12%

(a) Calculate the mean value of this index.

$$\bar{X} = 15.2\,.$$

(b) Is the median value of this index above, equal to, or below the mean value computed in part (a)? Justify your answer.

Based on the information given, our best guess of the median is approximately 14.

(c) If you instead had six months of daily observations on the air pollution index, would the variance of these daily observations be greater than or less than a variance calculated from the above table? Explain.

Greater: the above table loses the within-group variation.

7. Let X and Y be independent normally distributed random variables, each with a mean of 2 and a standard deviation of 2. Define the random variable $W = X - Y$. *(32 points)*

(a) Compute the mean of W.

$$\mathrm{E}[W] = 0\,.$$

(b) Compute the variance of W.

$$\mathrm{var}(W) = 8\,.$$

(c) A linear function of normally distributed random variables produces a random variable that itself has a normal distribution. Using this fact, the definition of W, and your answers to parts (a) and (b) above, compute $\Pr(W \geq 2)$.

$$\Pr(W > 2) = 0.239\,.$$