Midterm I

Econ 253 Spring 2002

You have 75 minutes to answer the following 5 questions. The maximum number of points on this exam is 75. You may use a calculator and a single-sided sheet of paper with notes and formulas. Any collaboration is considered a violation of the honor code. Good luck!

1. A recent survey found that 70% of the business community read *The Wall Street Journal*, 40% read *The Financial Times* and 20% read both. Answer the following questions:

   (a) (5 points) What is the probability that a randomly chosen member of the business community does not read any of the two newspapers?

   Answer: Let WSJ denote the event that a person reads the Wall Street Journal and FT denote the event that a person reads Financial Times. The probability that a person doesn't read either of the newspapers is equal to one minus the probability that a person reads WSJ or FT. $P(notWSJ \cdot notFT) = 1 - P(WSJ + FT) = 1 - [P(WSJ) + P(FT) - P(WSJ \cdot FT)] = 1 - 0.9 = 0.1$. This can also be seen from the Venn diagram.

   (b) (5 points) What is the probability that a person who does not read *The Financial Times* reads *The Wall Street Journal*?

   Answer: $P(WSJ|notFT) = \frac{P(WSJ \cdot notFT)}{P(notFT)}$. The probability that a person reads WSJ but not FT is 0.5. Of the 70% who read WSJ, only 20% also read FT. The probability that a person does not read FT is 0.6 since 40% people read FT. Hence, $P(WSJ|notFT) = \frac{P(WSJ \cdot notFT)}{P(notFT)} = \frac{0.5}{0.6} = 0.83$.

2. The table below shows the joint probability distribution of years of education(Y) and annual income (X) in thousands of dollars. High school graduates have 12 years of education, while college graduates have 16.

	income		
$Y \setminus X$	30	50	70
12	0.2	0.2	0
16	0	0.3	0.3

   Answer the following questions:

(a) (5 points) Suppose you find a college graduate in this population. What is the probability that he or she has an annual income of $70 thousand?

Answer: Write down the marginal probabilities of X and Y.

	income			
$Y \setminus X$	30	50	70	
12	0.2	0.2	0	0.4
16	0	0.3	0.3	0.6
	0.2	0.5	0.3	

$P(X = 70|Y = 16) = \frac{0.3}{0.6} = \frac{1}{2}$.

(b) (5 points) What is the expected income in this population?

Answer: Using marginal probabilities of X, $E(X) = 0.2 \cdot 30 + 0.5 \cdot 50 + 0.3 \cdot 70 = 6 + 25 + 21 = 52$.

(c) (5 points) What is the expected income of a college graduate?

Answer: The conditional PDF:

x_i	$P(X = x_i\|Y = 16)$
30	0
50	0.3/0.6
70	0.3/0.6

Hence, $E(X|Y = 16) = 25 + 35 = 60$.

(d) (5 points) What is the variance of income in this population?

Answer: To simplify computations use the fact that $var(X) = E(X^2) - \mu_X^2$:

x_i	$P(X = x_i)$	x_i^2	$P(X = x_i) \cdot x_i^2$
30	0.2	900	180
50	0.5	2500	1250
70	0.3	4900	1470

We see that $E(X)^2 = 180+1250+1470 = 2900$. Hence, $var(X) = E(X)^2 - \mu_X^2 = 2900 - 52^2 = 196$.

(e) (5 points) If income was measured in terms of dollars instead of thousands of dollars would it change the variance of income? Would it change the standard deviation?

Answer: Measuring income in dollars instead of thousands of dollars is equivalent to multiplying income X by a thousand. Therefore, the variance would increase a million times, since $var(1,000 \cdot X) = 1,000,000 \cdot var(X)$. The standard deviation would increase a thousand times, since $\sqrt{var(1000 \cdot X)} = 1,000 \cdot \sqrt{var(X)}$.

(f) (5 points) What is the covariance between years of education and annual income?(Hint: to simplify computations remember that $cov(X, Y-a) = cov(X, Y)$).

Answer: Note that $cov(X, Y) = cov(X, Y - 12) = E(X \cdot (Y - 12)) - \mu_X \cdot \mu_{Y-12}$.

x_i	$y_i - 12$	$x_i \cdot y_i - 12$	$P(X = x_i, Y = y_i)$
$\vdots$	0	0	$\vdots$
30	4	120	0
50	4	200	0.3
70	4	280	0.3

$E(X \cdot (Y - 12)) = 60 + 84 = 144$

$\mu_{Y-12} = 4 \cdot 0.6 = 2.4$. Hence, $cov(X, Y) = 144 - 52 \cdot 2.4 = 19.2$.

(g) (5 points) If income were measured in dollars instead of thousands of dollars, would the covariance change? Would the correlation coefficient change?

Answer: The covariance would be 1000 times greater since $cov(1000 \cdot X, Y) = 1000 \cdot cov(X, Y)$. The correlation coefficient would not change since $\rho_{1000 \cdot X,Y} = \frac{cov(1000 \cdot X,Y)}{\sqrt{var(1000 \cdot X)var(Y}} = \frac{1000 \cdot cov(X,Y)}{1000 \cdot \sqrt{var(X)var(Y)}} = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}} = \rho_{X,Y}$.

3. (10 points) A disease occurs in 1 out 1000 people. Suppose that 10% of the population tested positively for the disease. The test is 90% accurate. This means that if a person has the disease he or she will test positively with the probability of 0.9. Suppose that you are one of the people who tested positively. What it the probability that you have the disease?

Answer: Let D be the event that a person has the disease, then P(D)=0.001. Let P be the event that a person tests positively for the disease, then P(P)=0.1. The probability of getting a positive result given that a person has the disease is $P(P|D) = 0.9$. In order to find the probability that a person who tested positively has the disease we need $P(D|P) = \frac{P(P \cdot D)}{P(P)}$. Since $P(P|D) = \frac{P(P \cdot D)}{P(D)} = 0.9 \Rightarrow P(P \cdot D) = 0.0009$. Hence, $P(D|P) = \frac{P(P \cdot D)}{P(P)} = \frac{0.0009}{0.1} = 0.009$.

4. Suppose that the life of a tire on an automobile is a normally distributed random variable with a mean of 36,000 miles and a variance $2000^2 = 4,000,000$. Answer the following questions:

(a) (5 points) What is the probability that a tire will last less than 32,000 miles?

Answer: If X is the life of a tire then we need $P(X < 32,000)$. Since $X \sim N(3,000, 2000^2)$ we need:

$$\begin{aligned} P(X < 32,000) &= ? \\ &= P(\frac{X - 36,000}{2000} < \frac{32,000 - 36,000}{2000}) \\ &= P(Z < -2) \\ &= 0.5 - P(0 < Z < 2) = 0.5 - 0.4772 = 0.0228 \end{aligned}$$

(b) (10 points) What is the number of miles that the car maker should guarantee so that the probability that a tire will burst before the guarantee expires, is 10%.

Answer: We need to find number N so that $P(X <$N$) = 0.1$ or equivalently $P(\frac{X-36,000}{2000} < \frac{N-36,000}{2000}) = 0.1$ which can be written as $P(Z < \frac{N-36,000}{2000}) = 0.1$ where Z is N(0,1). In the table we find that $P(Z < 1.28) = 0.9$, hence $P(Z < -1.28) = 0.1$. We can see that $\frac{N-36,000}{2000}$ must equal -1.28. Solving for N yields 33,440. If the guarantee on the tire is 33,440 miles, the tire will burst before the guarantee with the probability of 10%.

5. (5 points) The amount of soup poured into a can by an automatic machine is distributed normally with a mean 15.9 ounces and a standard deviation of 1.5 ounces. What is the probability of finding a sample of 400 cans with an average greater than 16 ounces?

Answers: We know that $\overline{x} \sim N(15.9, \frac{1.5^2}{400})$. Hence, $P(\overline{x} > 16) = P(\frac{\overline{x}-15.9}{1.5/20} > \frac{16-15.9}{1.5/20}) = P(Z > 1.33) = 0.0918$.