The Pennsylvania State University

Department of Economics

Econ 390, Section 101, Summer 2007

Midterm Examination # 2

Answer Key

You have 1 hour and 15 minutes to complete the exam. You are not allowed to use any textbooks, notes and etc., except the list of the formulas provided. You may use calculators, but not PDAs or laptops or cell phones. No communication with other students is allowed. You can earn up to 115 point on this exam. However, 15 points is a bonus and is not required to get an A for the course. Each problem has its value in points. If you spend approximately 1 minute per 1 point you will be done just in time.

Please show your work step by step. Use the space provided to answer questions.

Problem 1 (35 points)

For several weeks Kevin and David were recording prices of a premium ice cream in California and in Florida, they found out that all their observations could be summarized by the joint probability distribution function:

		Price in CA (Y)					
		3.00	3.50	4.00	$P(x)$	$x * P(x)$	$(x-\mu_X)^2 P(x)$
Price in	3.00	0.09	0.18	0.07	0.34	1.020	0.0704
Florida	3.50	0.05	0.14	0.22	0.41	1.435	0.0008
(X)	4.00	0.01	0.07	0.17	0.25	1.000	0.0743
$P(y)$		0.15	0.39	0.46	1	$\mu_X = 3.455$	$\sigma_X^2 = 0.1455$
$y * P(y)$		0.450	1.365	1.840	$\mu_Y = 3.655$		$\sigma_X = 0.3814$
$(y-\mu_Y)^2 P(y)$		0.0644	0.0094	0.0548	$\sigma_Y^2 = 0.1285$	$\sigma_Y = 0.3585$	

(For instance, probability that $X = 3.50$ and $Y = 3.00$ is 0.05 or $P(X = 3.50 \cap Y = 3.00) = 0.05$)

a) Write down the marginal probability function $-$ $P(X)$ for X.

(**Hint:** You can use free cells to answer this and the following questions)

See the column 6 of the table - $P(x)$

b) Write down the marginal probability function $-$ $P(Y)$ for Y

See the row 6 of the table - $P(y)$

c) Find the mean of X and the mean of Y (Hint: Use your marginal probability functions)

Refer to column 7 and to row 7:

$$E(X) = \sum_{x \in X} (x * P(x)) = 3.455$$

$$E(Y) = \sum_{y \in Y} (y * P(y)) = 3.655$$

d) Find the conditional probability function of Y for $X = 4.00$.

$$P(Y = y \mid X = 4) = \frac{P(Y = y \ \cap \ X = 4)}{P(X = 4)}$$

$$P(Y = 3.00 \mid X = 4) = \frac{P(Y = 3.00 \ \cap \ X = 4)}{P(X = 4)} = \frac{0.01}{0.25} \approx 0.04$$

$$P(Y = 3.50 \mid X = 4) = \frac{P(Y = 3.50 \ \cap \ X = 4)}{P(X = 4)} = \frac{0.07}{0.25} \approx 0.28$$

$$P(Y = 4.00 \mid X = 4) = \frac{P(Y = 4.00 \ \cap \ X = 4)}{P(X = 4)} = \frac{0.17}{0.25} \approx 0.68$$

e) Use your answer for pard d to calculate conditional expectation of Y, when $X = 4.00$.

(**Hint**: You have to find $E(Y|X = 4.00)$)

$$E(Y|X = x) = \sum_{y \in Y} (y * P(Y = y \mid X = x)) = \sum_{y \in Y} \left(y * \frac{P(Y = y \ \cap \ X = x)}{P(X = x)} \right)$$

$$E(Y|X = 4.00) = 3 * \frac{P(Y = 3.00 \ \cap \ X = 4)}{P(X = 4)} + 3.50 * \frac{P(Y = 3.50 \ \cap \ X = 4)}{P(X = 4)} + 4 * \frac{P(Y = 4.00 \ \cap \ X = 4)}{P(X = 4)} =$$

$$= 3 * 0.04 + 3.50 * 0.28 + 4 * 0.68 = 3.82$$

f) Find the covariance between X and Y

$$Cov(X, Y) = E(XY) - \mu_X * \mu_Y = \sum_{x \in X} \sum_{y \in Y} x * y * P(x, y) - \mu_X * \mu_Y$$

$$E(XY) = \sum_{x \in X} \sum_{y \in Y} x * y * P(x, y) =$$

$$= 3 * 3 * 0.09 + 3 * 3.5 * 0.18 + 3 * 4 * 0.07 +$$

$$+3.5 * 3 * 0.05 + 3.5 * 3.5 * 0.14 + 3.5 * 4 * 0.22 +$$

$$+4 * 3 * 0.01 + 4 * 3.5 * 0.07 + 4 * 4 * 0.17 = 12.68$$

$$Cov(X, Y) = E(XY) - \mu_X * \mu_Y = 12.68 - 3.455 * 3.655 = 0.052$$

g) Find the correlation between X and Y

$$\sigma_X = \sqrt{\sum_{x \in X} (x - \mu_X)^2 P(x)} \approx \sqrt{0.1455} \approx 0.3814$$

$$\sigma_Y = \sqrt{\sum_{y \in Y} (y - \mu_Y)^2 P(y)} \approx \sqrt{0.1285} \approx 0.3585$$

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{0.052}{0.3814 * 0.3585} \approx 0.38$$

Problem 2 (20 points)

Bob owns two shares. The stock market determines price of each share at the end of each business day. We assume that price today does not depend on the price yesterday and does not depend on the preceding history at all. Price of the stock A at a given day is just a draw from a normal distribution with a mean of 50 and a standard deviation of 4. Price of share B is normal with mean 65 and standard deviation equal to 5. Correlation between prices of share A and share B is (-0.70). Thus, $P_A \sim N(50, 4^2)$, $P_B \sim N(65, 5^2))$ and $Corr(P_A, P_B) = -0.70$.

One day Bob discovers that he has to sell both of his shares to cover some emergency expenses. Find the probability that if he sells both of his shares he would have more that 110 dollars.

(**Hint:** Recall that if A and B are normally distributed then $A \pm B$ are also normally distributed. Also think about why the information on correlation you have is crucial).

$$E(P_A + P_B) = E(P_A) + E(P_B) = 50 + 65 = 115$$

$$Var(P_A + P_B) = Var(P_A) + Var(P_B) - 2 * Corr * \sqrt{Var(P_A)}\sqrt{Var(P_B)} =$$

$$= \sigma_X^2 + \sigma_Y^2 - 2 * \rho * \sigma_X * \sigma_Y = 16 + 25 - 2 * 0.70 * 4 * 5 = 13$$

$$P((P_A + P_B) > 100) = P\left(\frac{(P_A + P_B) - E(P_A + P_B)}{\sqrt{Var(P_A + P_B)}} > \frac{110 - E(P_A + P_B)}{\sqrt{Var(P_A + P_B)}}\right) =$$

$$= P\left(\frac{(P_A + P_B) - 115}{\sqrt{13}} > \frac{110 - 115}{\sqrt{13}}\right) = P\left(Z > -1.3868\right) = 1 - P(Z < -1.3868) =$$

$$= 1 - F(-1.3868) = 1 - (1 - F(1.3868)) = F(1.3868) \approx 0.92$$

Problem 3 (15 points)

The local department of transportation counted the number of drivers arriving to renew their driver's licenses and to do other paperwork. The department counted the number of drivers for the last 10 business days:

# of clients:	75	90	105	110	115	80	100	92	97	104

Find 95% confidence interval for the average number of clients that will attend the office during the next year.
(**Hint:** To simplify calculation I have calculated S^2 for you: $S^2 = 162.4$)

Since population variance is unknown you have to use t-distribution (student distribution) to find the interval:

$$\bar{X} = \frac{75 + 90 + 105 + 110 + 115 + 80 + 100 + 92 + 97 + 104}{10} \approx 96.8$$

$$S^2 = \frac{1}{n-1} \sum_{x \in X} (x - \bar{X})^2$$

Observation	75	90	105	110	115	80	100	92	97	104	$\frac{1}{9}\sum(x-\bar{X})^2$
$(x-\bar{X})^2$	475.24	46.24	67.24	174.24	331.24	282.24	10.24	23.04	0.04	51.84	162.4

$$S = \sqrt{162.4} \approx 12.74$$

Confidence interval for the mean in the case of unknown mean could be calculated as follows:

$$\bar{X} \pm t_{9,0.025} \frac{S}{\sqrt{10}}$$

The necessary assumption that allows to use t-distribution is that population is normally distributed.

$$t_{9,0.025} = 2.26$$

$$\text{95\% confidence interval is } \; 96.8 \pm 2.26 \frac{12.74}{\sqrt{10}} \quad \text{or } [87.7, 105.9]$$

Problem 4 (10 points)

Dennis is a professional photographer. In the last three months he took 625 pictures. Only 15% of them were published. Under population we will consider all of the photos he takes in his life. Find a 70% confidence interval for the population share of successful photos.

$\hat{p} = 0.15$, thus

$$\sigma_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.15 * 0.85}{625}} \approx 0.0143$$

Since $\hat{p}(1-\hat{p})n \approx 80 > 9$ we can use Central Limit Theorem to find the interval:

$$100(1-\alpha)\% \text{ confidence interval for p is } \hat{p} \pm Z_{\alpha/2} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where $Z_{\alpha/2} = Z_{0.15} \approx 1.04$

$$\hat{p} \pm Z_{\alpha/2} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.15 \pm 1.04 * 0.0143 \text{ or } [0.1351,\ 0.1649]$$

Problem 5 (10 points)

"Best LCD TV Ever" company produces LCD TVs. In January the company produced $20,000$ units. Probability that a given TV is defective is 0.01 or 1%. Find the probability that from 230 to 250 TVs produced in January are defective.

(**Hint**: Think about approximation to binomial distribution)

Since $p(1-p)n = 198 > 9$ we can use Central Limit Theorem to find the probability:

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \sim N(0,1)$$

$$P(230 \le X \le 250) = P\left(\frac{230 - np}{\sqrt{np(1-p)}} \le \frac{X - np}{\sqrt{np(1-p)}} \le \frac{250 - np}{\sqrt{np(1-p)}}\right) =$$

$$= P\left(\frac{230 - 0.01 * 20000}{\sqrt{20000 * 0.01 * (1 - 0.01)}} \le Z \le \frac{250 - 0.01 * 20000}{\sqrt{2000 * 0.01 * (1 - 0.01)}}\right) =$$

$$= P\left(\frac{30}{\sqrt{198}} \le Z \le \frac{50}{\sqrt{198}}\right) = P(2.13 \le Z \le 3.55) = F(3.55) - F(2.13) \approx$$

$$\approx 1 - 0.9834 = 0.0166$$

Problem 6 (10 points)

Suppose that you toss three fair coins. Each head gives you one point and each tail gives zero. Write down the sampling distribution of the sample means.

(**Hint**: Note that you can have from zero to three points at each trial).

Number of points	Outcomes that lead to this number of points:	Sampling distribution of the sample means: Mean	P(of the mean)
0	(T, T, T)	0.0	1/8
1	$(H, T, T), (T, H, T), (T, T, H)$	0.33	3/8
2	$(H, H, T), (H, T, H), (T, H, H)$	0.67	3/8
3	(H, H, H)	1.0	1/8

Problem 7 (15 points - bonus question)

"Ball-bearing international" produces steel balls. Producer claims that mean diameter (D) of the ball is 20 mm. and standard deviation is 0.5 mm. Distribution is normal. Independent consulting firm purchases 101 balls to check the producer's claim.

a) Find 95% acceptance interval for the sample mean.

b) Find probability that the sample standard deviation is more than 0.46

a) $d \sim N(20, 0.5^2)$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \tilde{} N(0,1)$$

Since Z is normal with a mean of zero and a standard deviation of 1, $100(1-\alpha)\%$ acceptance interval is:

$$\mu \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$Z_{0.025} \approx 1.96$

$$95\%\ confidence\ interval\ is\ \ 20 \pm 1.96 \frac{0.5}{\sqrt{101}} \quad \text{or } [19.902,\ 20.098]$$

b) Use the fact that:

$$\frac{(n-1)\,S^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$P(S^2 > 0.46^2) = P\left(\frac{(n-1)\,S^2}{\sigma^2} > 0.46^2 \frac{(n-1)}{\sigma^2}\right) = P\left(\chi^2_{100} > 0.46^2 \frac{100}{0.5^2}\right) =$$

$$= P\left(\chi^2_{100} > 83.8\right) \approx 0.9$$