The Pennsylvania State University

Department of Economics

Econ 390, Section 101, Summer 2007

Homework Assignment # 3

Solution

Problem 1

Insurance company has 1500 clients - truck drivers. Each client is insured against a car accident. Probability that an accident happens while a client is insured is 0.10. Find the probability that $100 - 200$ drivers will be involved in accidents.

Hint: Think about approximation to the binomial distribution.

The hint is to use Normal approximation to the binomial distribution. Note that n is quite large, $p = 0.10$, thus $np(1-p) = 135 > 9$. All conditions are satisfied and we can use normall approximation to the binomial distribution.

Let

$$Z = \frac{X - np}{\sqrt{np(1-p)}} = \frac{X - 150}{\sqrt{135}} \approx \frac{X - 150}{11.6190}$$

$$P(100 \leq X \leq 200) = P\left(\frac{100-150}{11.6190} \leq \frac{X-150}{11.6190} \leq \frac{200-150}{11.6190}\right) =$$

$$= P\left(-4.3033 \leq Z \leq 4.3033\right) = F(4.3033) - F(-4.3033) \approx 0.999$$

Problem 2

"Drinkers Inc" purchases an expensive refrigerator. Producer claims that refrigerators of this type work for 5 years on average.

Assume that probability that the unit dies in t years has an exponential distribution. Find the probability that refrigerator will die in $6 - 8$ years.

Hint: Use information on the average to recover parameters of the distribution.

Exponential distribution should be used. $\lambda = 1/5 = 0.2$.

$$P(6 \leq t \leq 8) = F(8) - F(6) = (1 - e^{-0.2*8}) - (1 - e^{-0.2*6}) = e^{-0.2*6} - e^{-0.2*8} = e^{-1.2} - e^{-1.6} \approx 0.0993$$

Problem 3

Kate has collected data on gas prices in two cities. One is located in California and the other is in Pennsylvania. Gas prices depend on a great number of events and could be treated as random. Denote prices in CA as X and prices in PA as Y. The collected data has been recorded in the form of Joint Probability distribution

		Price in PA (Y)				$PDFX$	$E(X)$	σ_X^2
		2.90	2.95	3.00	3.05	$P(X)$	$x * P(x)$	$(x-\mu_X)^2 P(x)$
	2.90	0.05	0.03	0.01	0.00	0.09	0.2610	0.00072
Price in	2.95	0.18	0.09	0.03	0.01	0.31	0.9145	0.00048
CA (X)	3.00	0.03	0.19	0.09	0.01	0.32	0.9600	0.00004
	3.05	0.01	0.05	0.18	0.04	0.28	0.8540	0.00102
$PDF\ of\ Y$	$P(Y)$	0.27	0.36	0.31	0.06	1.00	2.9895	0.0023
$E(Y)$	$y * P(y)$	0.7830	1.0620	0.9300	0.1830	2.9580		
$P(Y\|X=3)$	$= \frac{P(Y=y\ \cap\ X=3)}{P(X=3)}$	$\frac{0.03}{0.32} \approx 0.0938$	$\frac{0.19}{0.32} \approx 0.5938$	0.2813	0.0313	1.00		
$E(Y\|X=3)$	$= y\frac{P(Y=y\ \cap\ X=3)}{P(X=3)}$	0.2720	1.7517	0.8439	0.0955	2.9631		
$\sigma_Y^2 =$	$(y-\mu_Y)^2 P(y)$	0.0009	0.0000	0.00055	0.00051	0.0020		

a) Write down the marginal probability function $P(X)$ of X

The 7-th column of the table above (sum across columns)

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

The 7-th raw of the table above (sum across rows)

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

$$E(X) = \sum_{x \in X} (x * P(x)) = 2.9895$$

$$E(Y) = \sum_{y \in Y} (y * P(y)) = 2.9580$$

d) Find the conditional probability function of Y for $X = 3.00$.

The 9-th raw of the table

$$P(Y = y \mid X = 3) = \frac{P(Y = y\ \cap\ X = 3)}{P(X = 3)}$$

e) Use your answer for d) to calculate conditional expectation of Y, when $X = 3.00$.

The 10-th raw of the table:

(**Hint**: You have to find $E(Y|X = 3.00)$)

$$E(Y|X = 3.00) = \sum_{y \in Y} (y * P(Y = y \mid X = 3)) = \sum_{y \in Y} \left(y * \frac{P(Y = y\ \cap\ X = 3)}{P(X = 3)} \right) \approx 2.9631$$

f) Find correlation between X and Y:

We have to find first

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

$$= 2.90 * 2.90 * 0.05 + 2.95 * 2.90 * 0.03 + 3.00 * 2.90 * 0.01 + 3.05 * 2.90 * 0.00 +$$

$$+ 2.90 * 2.95 * 0.18 + 2.95 * 2.95 * 0.09 + 3.00 * 2.95 * 0.03 + 3.05 * 2.95 * 0.01 +$$

$$+2.90 * 3.00 * 0.03 + 2.95 * 3.00 * 0.19 + 3.00 * 3.00 * 0.09 + 3.05 * 3.00 * 0.01+$$

$$+2.90 * 3.05 * 0.01 + 2.95 * 3.05 * 0.05 + 3.00 * 3.05 * 0.18 + 3.05 * 3.05 * 0.04 =$$

$$= 0.7642 + 2.6786 + 2.8440 + 2.5574 = 8.8442$$

Hence,

$$Cov(X,Y) = E(XY) - \mu_X * \mu_Y = 8.8442 - 2.9895 * 2.9580 = 0.0013$$

The only remaining part is to calculate σ_X and σ_Y, where:

$$\sigma_X = \sqrt{\sum_{x \in X} (x - \mu_X)^2 P(x)} \approx \sqrt{0.0023} \approx 0.0475$$

and

$$\sigma_Y = \sqrt{\sum_{y \in Y} (y - \mu_Y)^2 P(y)} \approx \sqrt{0.0020} \approx 0.0443$$

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{0.0013}{0.0475 * 0.0443} \approx 0.6178$$

Problem 4

Suppose that you toss a pair of dice and write down the values of the faces from each die.

a. What is the population distribution for one die?

Die outcome	*Probability*
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

b. Determine the sampling distribution of the sample means obtained by tossing two dice.

(**Hint:** to answer question b think of all combinations that could lead to 2 points, 3 points and etc.)

Total points	*Sample*	$\overline{x}$	*Prob. of* $\overline{x}$
2	(1 1)	1.0	1/36
3	(1 2), (2 1)	1.5	2/36
4	(1 3), (3 1), (2 2)	2.0	3/36
5	(1 4), (4 1), (2 3), (3 2)	2.5	4/36
6	(1 5), (5 1), (2 4), (4 2), (3 3)	3.0	5/36
7	(1 6), (6 1), (2 5), (5 2), (3 4), (4 3)	3.5	6/36
8	(2 6), (6 2), (3 5), (5 3), (4 4)	4.0	5/36
9	(3 6), (6 3), (4 5), (5 4)	4.5	4/36
10	(4 6), (6 4), (5 5)	5.0	3/36
11	(5 6), (6 5)	5.5	2/36
12	(6 6)	6.0	1/36

Problem 5 (Only if we have time to cover CLT in class)

Given a population with mean $\mu = 100$ and variance $\sigma^2 = 81$, the central limit theorem applies when the sample size $n \geqslant 25$. A random sample of size $n = 25$ is obtained.

a. What are the mean and variance of the sampling distribution for the sample means?

$$E(\bar{X}) = \mu = 100$$

and

$$\sigma_X = \frac{\sigma}{\sqrt{n}} = \frac{9}{5} = 1.8 \;\; or \;\; \sigma_X^2 = 3.24$$

b. What is the probability that $\bar{x} > 102$

$$P(\bar{X} > 102) = P\left(\frac{\bar{X} - 100}{1.8} > \frac{102 - 100}{1.8}\right) = P\left(Z > 1.11\right) = 1 - P(Z \leq 1.11) = 1 - 0.8665 = 0.1335$$

c. What is the probability that $98 \leq x \leq 101$

$$P(98 \leq \bar{X} \leq 101) = P(\frac{98 - 100}{1.8} \leq \frac{\bar{X} - 100}{1.8} \leq \frac{101 - 100}{1.8}) = P(-1.11 \leq Z \leq 0.56) =$$

$$= F(0.56) - F(-1.11) = F(0.56) - (1 - F(1.11)) = 0.7123 - (1 - 0.8665) = 0.5788$$