3-16. Some parts of California are particularly earthquake prone. Suppose that in one such area, 20% of all homeowners are insured against earthquake damage. Four homeowners are to be selected at random; let X denote the number among the four who have earthquake insurance.

a. Find the probability distribution of X. [Hint: Let S denotes a homeowner who has insurance and F one who does not. Then 1 possible outcome is SFSS, with probability (.2)(.8)(.2)(.2) and associated X value 3. There are 15 other outcomes.]

b. Draw the corresponding probability histogram.

c. What is the most likely value for?

d. What is the probability that at least two of the four selected have earthquake insurance?

e. Calculate the expected value E (x) or μ and the standard deviation.

f. Graph the Cumulative Distribution Function and the Probability Mass Function on separate graphs.

TERMS

"Histogram"- A histogram is a pictorial representation of a frequency distribution. Examples of histograms can be found on page 13 of Devore.

"Probability Histogram"- A probability histogram is pictorial representation of the Probability Mass Function. Page 96 of the text describes how to construct this graph.

"E(x) or μ"- E(x) or μ stand for the expected value or mean value of X. The text warns that the expected value will not always be value we expect to see. More information on E(x) or μ can be found on page 104 of the text.

"Standard Deviation"- The standard deviation for this problem is not the same as the standard deviation we calculated in Chapter 1. We are calculating the standard deviation for a population and not for a set of data. The definition for "Standard Deviation" and examples can be found on page 109.

"Cumulative Distribution Function"- The Cumulative Distribution Function gives the probability that X falls in a specified interval. The CDF for discrete random variables can be graphically illustrated in a step function graph. More information about CDF can be found starting on page 98 in the text.

"Probability Mass Function"- The Probability Mass Function or Probability Distribution says that for every possible value x of the random variable, the probability of observing that value during the experiment will be given by the pmf. Certain conditions are needed for the pmf and can be found on page 94 in the text.

EQUATIONS

pmf: p(x) = P(X=x) = P(all s$\in$ S: x(s)=x)

"P(X=x) is read the probability that rv X assumes the value x. For example, P(X=2) denotes the probability that the resulting X value is 2."

CDF: P(X$\leq$x) = $\sum_{y:y\leq x} p(y)$

The definition and example can be found on page 98 of the text.

Standard Deviation: $\sigma_x^2 = \sqrt{\sigma_x^2}$

The definition can be found on page 104 of the text.

E(x) or μ : E(x) = μ = $\sum_{x\in D} x \cdot p(x)$

The definition can be found on page 104 of the text.

Now lets solve the problem.

a. Find the probability distribution of X. [Hint: Let S denote a homeowner who has insurance and F one who does not. Then 1 possible outcome is SFSS, with probability (.2)(.8)(.2)(.2) and associated X value 3. There are 15 other outcomes.]

Part a. asks us to calculate the pmf of x. In the problem statement X is defined as number among the four homeowners selected that have insurance. We know in the area where the experiment is conducted that 20% of the homeowners have insurance. When a homeowner is selected there are two possible outcomes. We will let S denote a homeowner who is insured and F be a homeowner who does not have insurance. At this point we know.

P(S) = .2
P(F) = .8
X = number among four selected that have insurance
S = a homeowner with insurance
F = a uninsured homeowner
16 possible outcomes (given in problem statement)

One way to present the 16 possible outcomes is as follows:

p(0) = P(X=0) = P(FFFF)
The probability shows no homeowners with insurance.

$p(1) = P(X=1) = P(FSFF \cup FFSF \cup FFFS \cup SFFF)$
The probability shows only one homeowner with insurance.

Note: There are four ways only one homeowner can have insurance. Each of these four events all are mutually exclusive (they can not occur at the same time).
$p(2) = P(X=2) = P(FFSS \cup FSSF \cup SSFF \cup FSFS \cup FSFS \cup SFSF)$
The probability shows only the two homeowners with insurance.

$p(3) = P(X=3) = P(FSSS \cup SFSS \cup SSFS \cup SSSF)$
The probability shows only three homeowners with insurance.

$p(4) = P(X=4) = P(SSSS)$
The probability shows all four homeowners with insurance.

Calculations
$p(0) = (.8)^4 = 0.4096$

p(1) = # combinations x [probability of whether only 1 owner is insured]

$p(1) = 4[(.2)(.8)^3] = 0.4096$

$p(2) = 6[(.2)^2 [(.8)^2 = 0.1536$

$p(3) = 4[(.2)^3(.8)] = 0.0256$

$p(4) = (.2)^4 = 0.0016$

The sum of the probabilities should be 1 and it is.

Another way to display this in the form of a chart See Figure 1.

X	0	1	2	3	4
p(x)	0.4096	0.4096	0.1536	0.0256	0.0016

Figure 1. pmf display

b. Draw the corresponding probability histogram. Figure 3 displays a pictorial representation of the pmf.

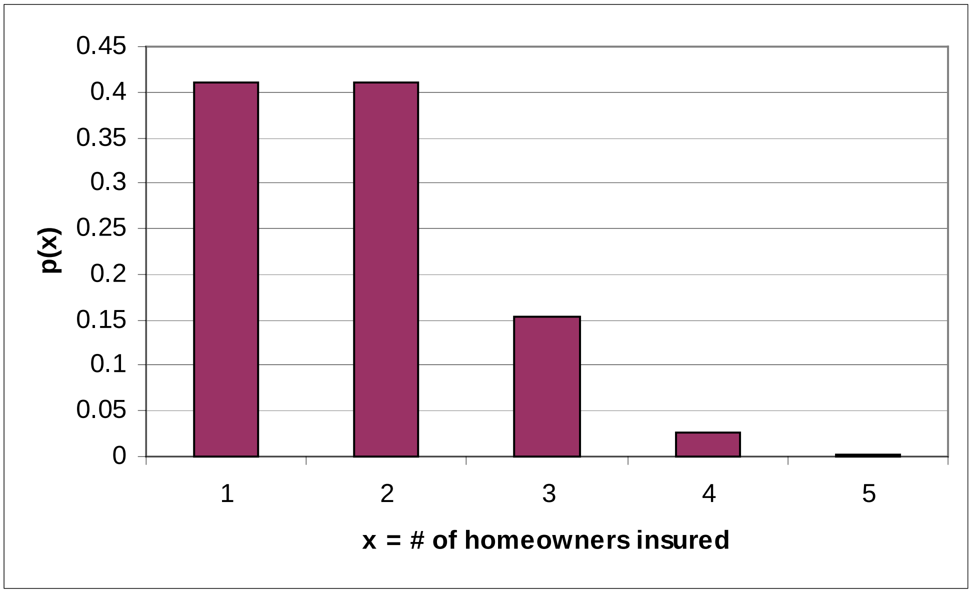

Figure 3. Probability Histogram

Note: X scale should read 0,1,2,3,4 instead of 1,2,,3,4,5.

c. What is the most likely value for? This question just asks us to interpret are probability histogram. From the histogram it is easy to determine that 0 and 1 are the most likely values for X.

d. What is the probability that at least two of the four selected have earthquake insurance? The key word here is *least.* That means we will sum the probabilities of p(2), p(3), p(4).

$p\ (x \geq 2) = p(2) + p(3) + p(4)$

$p(x \geq 2) = 0.1536 + 0.0256 + 0.0016$

$p(x \geq 2) = 0.1808$

This can also be written as the complement.

$1-(p < 2) = 1-[p(0) + p(1)] = 0.1808$

e. Calculate the expected value E(x) or μ and the standard deviation. The expected value can be calculated by using:

$$E(X) = \mu = \sum_{x \in D} x \cdot p(x)$$

$E(X) = 0(0.4096) + 1(0.4096) + 2(0.1536) + 3(0.0256) + 4(0.0016)$

$E(X) = 0.8$

The standard deviation can be calculated by using:

$$\sigma_x^2 = \sum_{D} (x-\mu)^2 \cdot p(x) \text{ and } \sigma_x^2 = \sqrt{\sigma_x^2}$$

$$\sigma_x^2 = [(0-.8)^2 \text{ x } 0.4096] + [(1-.8)^2 \text{ x } 0.4096] + [(2-.8)^2 \text{ x } 0.1536] + [(3-.8)^2 \text{ x } 0.0256] + [(4-.8)^2 \text{ x } 0.0016]$$

$$\sigma_x^2 = \sqrt{1.0211}$$

$$\sigma_x = 1.01$$

f. Graph the Cumulative Distribution Function and the Probability Mass Function on separate graphs.

Both of these graphs are line graphs. They were created in Excel by selecting a XY scatter plot and then plotting the points only. After that excel draw tools were used to add lines on the graphs.

Figure 4 shows the Cumulative Distribution Function for this problem. The cdf could be used to determine probability of seeing values of x greater than or less than certain given values of x.

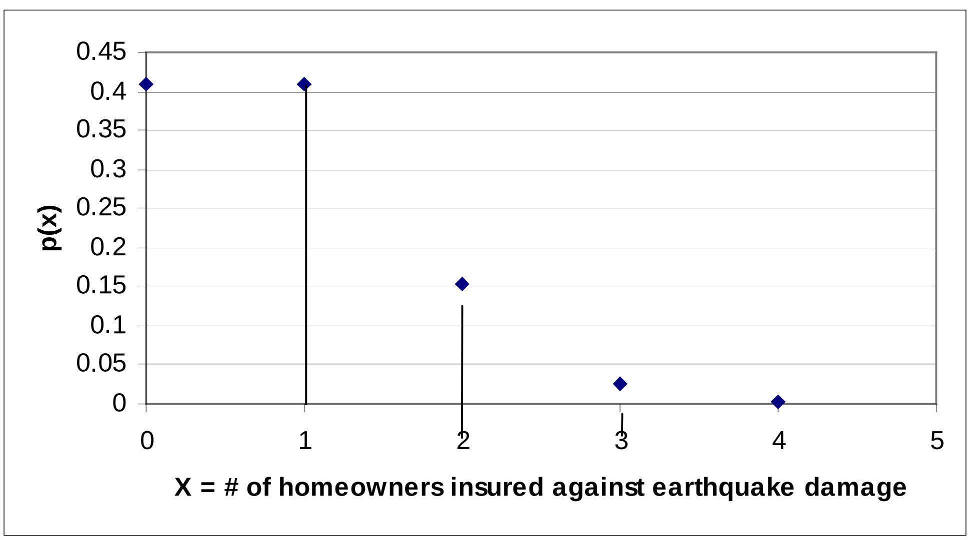

Figure 4. pmf line graph

Figure 5 is a line graph, which gives a pictorial representation of the pmf.

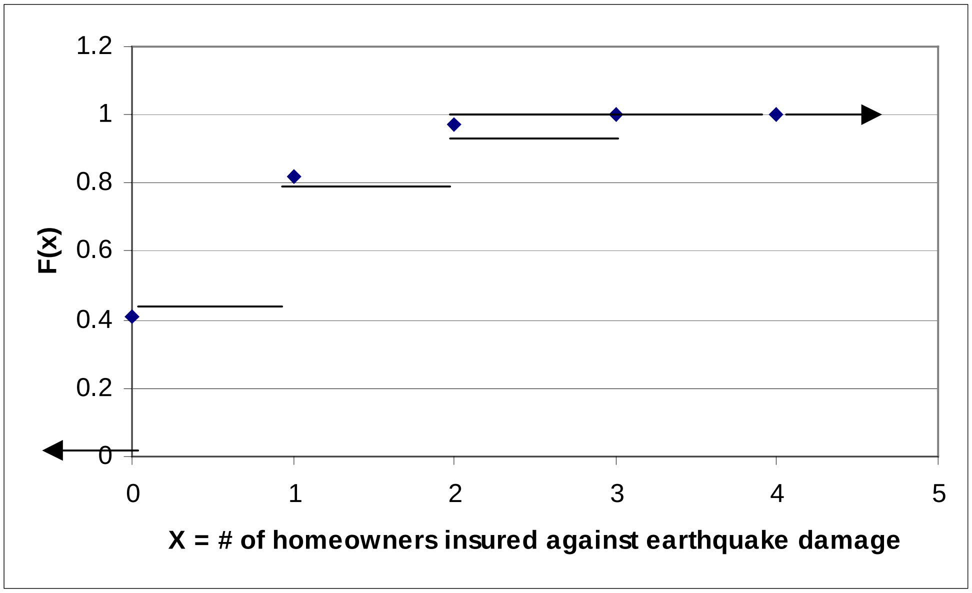

Figure 5. CDF for problem 3-16