

**Discrete One-Dimensional Random Variables (rvs)**

**DEFINITION:** A rv,  $X$ , is a function defined on the universe ( $U$ ) of an experiment, where  $X$  assigns a unique real number to every outcome ( $O$ ) in  $U$  (i.e.,  $X(O)$  is a real number for any outcome  $O$  in  $U$ ).

**Example 16.** Let the experiment consist of tossing 3 fair coins. Then  $U = \{HHH, HTH, HHT, THH, HTT, THT, TTH, TTT\}$ . Define the function  $X$  as the number of heads observed in one toss of the 3 coins ( $N$ = Nickel,  $D$ = Dime,  $Q$  =Quarter). Then the range space of  $X$  is the subset of real numbers  $R_x = \{0, 1, 2, 3\}$ .

The assignments that the function  $X$  makes are:  $X(HHH) = 3$ ,  $X(HTH) = 2$ ,  $X(TTH) = 1$ ,  $X(TTT) = 0$ , etc. Note that  $R_x$  is a subset of real numbers.

Further, the function  $X$  maps  $U$  onto  $R_x$  (a subset of real numbers) and for each outcome  $O$  in  $U$ , there exists a unique real value of  $X(O)$ . Note that capital letters must be used to represent random variables.

**Exercise 9.** (a) Two dice are tossed. Let the rv,  $X$ , represent the sum of the 2 up faces. Determine the range space  $R_x$ .

A random variable is said to be discrete if its range space  $R_x$  is finite or countably infinite. If  $R_x$  cannot be put onto a 1 to 1 correspondence with the set of integers, then  $X$  is a continuous rv.

**Example 17.** A certain type of electron tube is placed on life test and its time to failure, TTF in hours, is recorded. Let  $X$  represent the time to failure (TTF) of the tube. Then  $X$  is a continuous rv with

$$R_x = \{x \mid x \text{ is a real number, } 0 \leq x < \infty\} = [0, \infty).$$

Other examples of continuous rvs are provided in Example 1.2 ( $X$  = flexural strength) on page 5 of Devore, Example 1.10 ( $X$  = Adjusted Consumption of BTUs) on page 18 of Devore,  $X$  = Tensile ultimate strength (ksi) in Exercise 13 on page 23 of Devore, etc.

**Definition:** Let the event  $A$  belong to an  $U$  and suppose the rv,  $X$ , maps  $A$  onto

a subset of  $R_x$  denoted by  $R_A$ , i.e.,

$$A = \{\text{All outcomes } O \text{ in } U \mid X(O) = x \in R_A\}.$$

Then  $p(x)$  is said to be a probability mass (or distribution) function if : (1)  $p(x) = P(A)$ ,

(2)  $0 \leq p(x) \leq 1$ , and (3)  $\sum_x p(x) = 1$ .

**Example 18:** Consider the rv,  $X$ , defined in Example 16. Let  $A = \{\text{the 3 coins show exactly 2 heads}\} = \{\text{HHT, HTH, THH}\}$ . Then  $X$  maps  $A$  onto the subset of  $R_x$  denoted by  $R_A = \{2\}$ , and  $P(X = 2) = p_x(2) = p(2) = P(A) = 3/8$ . The entire probability mass function (pmf) is give below:

$$\text{pmf}(x) = p(x) = \begin{cases} 1/8, & x=0, 3 \\ 3/8, & x=1, 2. \end{cases} = 0.125 + 0.375x - 0.125x^2, \quad x = 0, 1, 2, 3.$$

**Exercise 9(b).** Obtain the Pr distribution function (PDF) of the rv defined in Exercise 9(a). (c) In Example 18 above, let the event  $B = \{\text{The 3 coins show at most one head}\}$ . Determine  $R_B$ , express  $P(B)$  in terms of the rv  $X$ , and then compute its Pr.

## The Cumulative Distribution Function (The cdf)

The cdf for a discrete rv,  $X$ , at a specified value  $x$  in  $R_x$  is defined as

$$F_X(x) = P(X \leq x) = \sum_{y=-\infty}^x p(y)$$

Note that, starting from  $-\infty$ , the cdf accumulates all the Prs to the left of and at the value of the specified  $x$ .

For example, for the PDF of Example 18, the cdf is given by

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/8, & x = 0 \\ 4/8, & x = 1 \\ 7/8, & x = 2 \\ 1, & x = 3. \end{cases} = \frac{1}{8} + \frac{7x}{24} + \frac{x^2}{8} - \frac{x^3}{24}, \quad x = 0, 1, 2, 3.$$

For the rv of Exercise 9(b) above, the cdf is given by

$$F(x) = \begin{cases} x(x-1)/72, & 2 \leq x \leq 7, \\ (x-4)(21-x)/72, & 8 \leq x \leq 12. \end{cases}$$

Note that the rv,  $X$ , for this last cdf is the sum of 2 faces of 2 balanced dice.

### Properties of the cdf $F(x)$

(1)  $0 \leq F(x) \leq 1$  for all  $x$  in  $R_X$ , (2)  $F(x)$  must be monotonically non-decreasing, (3)  $F(-\infty) = 0$  and  $F(+\infty) = 1$ , (4) Given that  $R_X = [a, b]$ , where  $a$  and  $b$  are real numbers and  $b > a$ , then for certain  $F(x) = 0$  for all  $x < a$ , and  $F(x) = 1$  for all  $x \geq b$ .

**Exercise 9(c).** Consider the Example 4 on pp. 14-15 of my notes. Define the rv  $X$  = as the number items that have to be drawn at random one at a time from the lot of  $N = 10$  units in order to remove both defective units. For this rv,  $X$ , obtain both the pmf,  $p(x)$ , and the cdf,  $F(x)$ . [In developing the cdf you must use the fact that

the sum of the first  $n$  nonnegative integers is given by  $\sum_{x=0}^n x = n(n+1)/2$ ].

## THE EXPECTED VALUE OF A DISCRETE RANDOM VARIABLE

Suppose  $X$  is a discrete rv over the range space  $R_X$  with PDF  $p(x)$ . Then the weighted average (or population mean) of  $X$  is defined as its expected value and is given below:

$$\mu = E(X) = \sum_{R_X} [x \times p(x)] \quad (2)$$

**Example 18 (Continued).** Find the population mean for the PD of Example 18.

The use of equation (2) yields

$$\mu = E(X) = 0(1/8) + 1(3/8) + 2(3/8) + 3(1/8) = 1.50 \text{ heads.}$$

**Exercise 9(d).** Compute the population mean,  $\mu = E(X)$ , for the rv defined as the sum of 2 balanced dice. Then compute the population mean  $\mu$  for the rv of the Exercise 9(c). For the 2<sup>nd</sup> part of this exercise you must make use of the fact that

$$\sum_{i=1}^n i^2 = n(n+1)(2n+1)/6. \text{ ANS: } E(X) = 7.0, \quad \mu = E(X) = 7.3333\bar{3}. \text{ The formula for}$$

$\sum_{i=1}^n i^2$  can be verified by considering the identity  $\sum_{i=1}^n i^3 - \sum_{i=1}^n (i-1)^3 = n^3$  and expanding  $(i-1)^3$ . Similarly,  $\sum_{i=1}^n i^3 = [n(n+1)]^2 / 4$  and  $\sum_{i=1}^n i^4 = [n(n+1)(2n+1)(3n^2+3n-1)] / 30$ .

## THE VARIANCE OF A DISCRETE RV

A population's Variance, by definition, is the weighted average of deviations from the mean ( $\mu$ ) squared, i.e., the population variance of a rv,  $X$ , is defined as

$$V(X) = \sigma^2 = E[(X - \mu)^2] = \sum_{R_x} (x - \mu)^2 \times p(x) \quad (3)$$

Note that  $\sum$  is used only when  $X$  is discrete. Later in Chapter 4 we will learn that only when  $X$  is a continuous rv, then  $\int$  sign will replace the summation in equation (3), and  $p(x)$  will be replaced by the density function of  $X$ . The positive square root of  $V(x)$  in Eq. (3) is called the population standard deviation and is universally denoted by  $\sigma$ . Students must be cognizant of the fact that whenever the operator  $E$  is applied to any rv, the end result is always invariably a population parameter, i.e.,  $\sigma$  is a population parameter. Further, the variance is simply the 2<sup>nd</sup> central moment.

**Example 18 (Continued).**  $V(X) = E[(X - \mu)^2] = E[(X - 1.5)^2] = (0 - 1.5)^2 \times (1/8) + (1 - 1.5)^2 \times (3/8) + (2 - 1.5)^2 \times (3/8) + (3 - 1.5)^2 \times (1/8) = 0.75$ ,  
 $\rightarrow$  The population standard deviation  $\sigma = +\sqrt{0.75} = 0.86603$ .

**Exercises 9(e).** Obtain the variance and standard deviation [std for Matlab and STDEV for Microsoft Excel] of the rv,  $X$ , that represents the sum of 2 faces of dice. ANS:  $\sigma = 2.41523$ .

## THE PROPERTIES OF THE OPERATOR E

The expected value operator,  $E$ , is linear because

- (1)  $E(CX) = C \times E(X)$  for any constant  $C$  and rv  $X$ ,
- (2)  $E(X + Y) = E(X) + E(Y)$

where  $X$  and  $Y$  are rvs. Clearly  $E(C) = C$  for any constant  $C$ .

We now use the properties of  $E$  to obtain a computing formula for  $V(X)$ .

$$\begin{aligned} V(X) &= E[(X - \mu)^2] = E[X^2 - 2\mu X + \mu^2] = E(X^2) - 2\mu E(X) + E(\mu^2) = \\ &= E(X^2) - \mu^2 = E[(X^2)] - [E(X)]^2, \end{aligned} \tag{4}$$

where  $E(X^2)$  is the weighted average of  $x^2$  values in the range space  $R_x$  given by  $E(X^2) = \sum x^2 \times p(x)$ , and this last sum extends over all  $x \in R_x$ .

**Example 18 Continued.**  $E(X^2) = 0^2 \times 1/8 + 1^2 \times 3/8 + 2^2 \times 3/8 + 3^2 \times 1/8 = 3.0 \rightarrow \sigma^2 = 3.0 - (1.5)^2 = 0.75$ , which is identical to the same value obtained using the definition of  $V(X)$  at the bottom of page 29.

**Exercise 9(e).** Compute the (population) variance and std of the rv of Exercise

9(c). In computing  $E(X^2)$ , you need to make use of the fact that  $\sum_{i=1}^n i^3 = n^2(n+1)^2/4$ .

ANS :  $\sigma_x = 2.2111 \bar{1}$ .

In Chapter 3, we will study several discrete PDFs, all of which are based on the very fundamental distribution called Bernoulli. The range space for a Bernoulli rv is always  $R_x = \{0, 1\}$ , where 0 pertains to the occurrence of a failure and 1 pertains to the occurrence of a “generic” success. The discrete frequency functions that are derived from Bernoulli trials are: Binomial, Geometric, Pascal (or Negative Binomial), and the Poisson.

## DISCRETE FUNCTIONS OF RANDOM VARIABLES

As an example, suppose a contractor is going to bid on a project but the days to completion,  $X$ , is a rv whose PD follows the pmf tabulated below:

<b>x</b>	<b>10 days</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14 days</b>
<b>p(x)</b>	<b>0.10</b>	<b>0.30</b>	<b>0.40</b>	<b>0.10</b>	<b>0.10</b>

The contractor's net profit function per project is  $Y = \$600(13 - X)$ .

(a) What is the PDF of  $Y$ ,  $p(y)$ ? Clearly,  $R_y = \{-600, 0, 600, 1200, 1800\}$ .

$$p_Y(1800) = P(Y = 1800) = P(X = 10) = 0.10$$

$$p_Y(1200) = P(Y = 1200) = P(X = 11) = 0.30, \text{ etc. Hence,}$$

$$\text{pmf}(y) = p(y) = \begin{cases} 0.10, & y = -600, 0, 1800 \\ 0.30, & y = 1200 \\ 0.40, & y = 600. \end{cases}$$

(b) What are the contractor's expected net profit and the variance of  $Y$ ?

$$E(Y) = \sum_{R_y} yp(y) = (-600) \times 0.10 + 0 \times 0.10 + 600 \times 0.40 +$$

$$1200 \times 0.30 + 1800 \times 0.10 = \$ 720.00 / \text{project.}$$

$$V(Y) = E[(Y - 720)^2] = E(Y^2) - 720^2.$$

$$E(Y^2) = \sum y^2 \times p(y) = (-600)^2 \times 0.10 + 0^2 \times 0.10 + 600^2 \times 0.40 + 1200^2 \times 0.30 + 1800^2 \times 0.10 = 936,000.00 \rightarrow \sigma_y^2 = 936,000 - 720^2 = 417,600 \rightarrow \sigma_y = \$ 646.22 / \text{project.}$$

The coefficient of variation of  $Y$  is defined as  $CV_y = \sigma_y / E(Y) = 89.753\%$ .

Note that the reciprocal (or inverse) of  $CV$  is called the Signal-to-Noise Ratio.

We now show that the above  $E(Y)$  and  $V(Y)$ , because  $Y$  is a function of  $X$ , can and should be computed W/O obtaining the PDF of  $Y$ , as shown below.

$$\begin{aligned} E(Y) &= \sum_{R_x} Y(x) \times p(x) = \sum_{x=10}^{14} 600(13-x) \times p(x) = 600 \sum_{x=10}^{14} (13-x) \times p(x) = \\ &= 600 \left[ \sum_{x=10}^{14} 13 \times p(x) - \sum_{x=10}^{14} x \times p(x) \right] = 600 [13 - E(X)] = 600 (13 - 11.8) = \$720.00 \text{ per} \end{aligned}$$

project as before!

Before computing the  $V(Y)$  using the pmf of  $X$ , we need to state the properties of the variance operator  $V$ .

(1)  $V(C) = 0$  for any constant  $C$ , (2)  $V(CX) = C^2V(X)$ , where  $X$  is any rv.

Note that property (2) clearly shows that  $V$  is a nonlinear operator because  $V(CX) \neq C$

$V(X)$  for all real constants  $C$ . (3)  $V(X + C) = V(X)$ , (4) If  $X$  and  $Y$  are any two rvs, then  $V(X \pm Y) = V(X) + V(Y) \pm 2 \text{COV}(X, Y)$ , where the covariance between two random variables  $X$  and  $Y$  is defined as  $\text{COV}(X, Y) = \sigma_{xy} = E[(X - \mu_x)(Y - \mu_y)]$ . If the rvs  $X$  and  $Y$  are independent, then for certain  $\text{COV}(X, Y) = \sigma_{xy} = 0$ . The converse of this last statement is not always true.

**Exercise 10.** Compute  $E(X)$  and  $V(X)$  of the above example and then re-compute  $E(Y)$  and  $V(Y)$  as functions of  $E(X)$  and  $V(X)$ . Note that you must use the properties of the variance operator stated above. (b) Show that

$$\text{COV}(X, Y) = E[(X \times Y)] - E(X) \times E(Y) = E(XY) - \mu_x \mu_y.$$

**Exercise 11.** A well-known inventory problem is the "newsboy problem", described as follows: A newsboy buys papers for 15 cents each and sells them for 35 cents each, and he cannot return unsold papers (i.e., the salvage value of each unsold paper at the day's end is zero). From past experience, the daily demand distribution is given below ( $D$  = daily demand).

$$p(d) = \begin{cases} 0.06, & d=100, 101, 102, 103, 111 \\ 0.08, & d=104, 105 \\ 0.10, & d=106, 107, 108 \\ 0.12, & d=109, 110 \end{cases}$$

(a) Obtain  $E(D)$  and  $V(D)$ . (b) Determine the daily net profit function,  $Y$  in dollars, given that his daily stock level is  $I$  ( $100 \leq I \leq 111$ ). Note that  $Y$  is a function of  $d$  and  $I$ . (c) Determine the optimum value of  $I$  by maximizing  $E(Y)$ . ANS:  $I_0 = 107$ ,  $E[Y(d, 107)] = \$20.763$ .

## THE BERNOULLI DISTRIBUTION

Consider an experiment whose outcomes can be classified as either success or failure, i.e.,  $U = \{\text{Success}, \text{Failure}\}$ .

Define the discrete rv,  $X$ , such that the value of  $X$  is 1 if a success occurs and the

value of  $X$  is 0 if the experiment results in a failure. Thus  $R_X = \{0, 1\}$  and  $X(\text{Failure}) = 0$ ,  $X(\text{Success}) = 1$ . Further, suppose the pr of success in one trial of the experiment is  $p$  and failure pr is  $q$ , i.e.,  $p + q = 1$  so that  $q = 1 - p$ . Hence

$$p(x) = \begin{cases} 1 - p, & x=0 \\ p, & x=1 \end{cases} = p^x q^{1-x}, \quad x = 0, 1.$$

The mean of a Bernoulli rv is

$$\mu = E(X) = 0(1 - p) + 1 \times p = p$$

**Exercise 12.** Show that the variance of a Bernoulli rv is  $\sigma^2 = pq$ .

**Example 19.** A small part of a steel pipe is produced by an automatic machine, but the machine is not perfect and manufactures one NCU in 100 units, i.e., the pr of a randomly selected unit being defective is  $p = 0.01$ . One item is selected at random from a lot of  $N = 5000$  parts. Let  $X$  be the Bernoulli rv with

$$U = \{\text{good, NC}\}, \quad R_x = \{0, 1\}.$$

Then the PDF of  $X$  is given by

$$p(x) = \begin{cases} 0.99, & x=0 \\ 0.01, & x=1 \end{cases} = (0.01)^x (0.99)^{1-x}$$

The proportion  $p = 0.01$  is called the FNC of the process. As a continuation of Exercise 12, Verify that  $E(X) = 0.01$  and  $V(X) = 0.0099$ .

## THE BINOMIAL PDF

Now, consider  $n > 1$  independent Bernoulli trials in successions. The binomial rv,  $X$ , is defined as  $X = \text{“The number of successes observed in } n \text{ trials”}$ , where  $R_x = \{0, 1, 2, 3, \dots, n-1, n\}$ , and  $p = \text{Pr of success at each single trial}$ .

**Example 20.** Two dice are tossed 10 times. Compute the Pr that a total of 8 comes up exactly  $x = 3$  times (out of 10 trials).

Success = {The 2 faces of the dice sum to 8 in one toss}

Success = {(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)}.  $\rightarrow$

$p = \text{Pr of success in a single trial} = 5/36$ ,  $q = \text{Pr of failure} = 1 - p = 31/36$ .

The Binomial rv  $X$  = The number of Eight's observed in 10 tosses. The following table shows one possible way to obtain exactly three successes in 10 trials, where Success = {8} and Failure =  $\{\bar{8} = 8'\}$  = {Non-eight}.

<b>Trials</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
<b>outcomes</b>	$\bar{8}$	$\bar{8}$	8	$\bar{8}$	8	$\bar{8}$	$\bar{8}$	$\bar{8}$	8	$\bar{8}$
<b>probability</b>	31/36	q	5/36	q	p	q	q	q	p	q

Because of the multiplication principle, the Pr for the above specific sequence  $\bar{8} \bar{8} 8 \bar{8} 8 \bar{8} \bar{8} \bar{8} 8 \bar{8}$  is  $(31/36)(31/36)(5/36)(31/36)(5/36)(31/36) \times (31/36)(31/36)(5/36)(31/36) = (5/36)^3 (31/36)^7$ .

However, since we are just interested in exactly three 8's in  $n = 10$  trials and each sequence of 8's and  $\bar{8}$ s are MUEX and have the same exact Pr of  $(5/36)^3(31/36)^7$ , plus the fact that the total number of different ways of placing the three successes in 10 different trials is  ${}_{10}C_3$ , we obtain:

$$b(3; 10, 5/36) = {}_{10}C_3(5/36)^3(31/36)^7 = 0.1128751$$

For the above example, the Pr of observing exactly  $x$  successes in 10 trials is given by  $p(x) = b(x; 10, 5/16) = {}_{10}C_x (5/36)^x(31/36)^{10-x}$ ,  $x = 0, 1, 2, \dots, 10$ . The above PDF is called the Binomial with parameters  $n = 10$  and  $p = 5/36$ , denoted as  $\text{Bin}(n, p) = \text{Bin}(10, 5/36)$ . The general form of the Binomial pmf is given by  $b(x; n, p) = {}_nC_x p^x q^{n-x}$ , (where  $q = 1 - p$ ), and its cdf is given by

$$F(x) = B(x; n, p) = \sum_{i=0}^x b(i; n, p) = \sum_{i=0}^x {}_nC_i \times p^i \times q^{n-i} .$$

It is paramount to observe that the binomial rv,  $X$ , is the sum of  $n$  independent Bernoulli rvs,  $X_i$ , i.e.,

$$X = X_1 + X_2 + \dots + X_n \tag{5}$$

where  $X_i$  represents the Bernoulli rv at the  $i$ th trial whose value is equal to 0 or 1 (0 for failure and 1 for success) so that the  $R_x = 0, 1, 2, \dots, n$ .

**Exercise 13.** Use equation (5) and the fact that the operator  $E$  is linear

to show that  $E(X) = np$  and  $V(X) = npq$  for any binomial rv  $X$ .

**Example 21.** A manufacturing process produces parts which are, on the average, 1% NC to customer specifications. A random sample of  $n = 30$  items is drawn from a conveyor belt.

(a) Compute the Pr that the sample contains exactly  $x = 2$  NCUs.

$$b(2; 30, 0.01) = {}_{30}C_2 (0.01)^2 (0.99)^{28} = 0.03283.$$

(b) Compute the Pr that the sample of 30 units contains at most 2 NCUs.

$$P(X \leq 2) = F(2) = B(2; 30, 0.01) = \sum_{x=0}^2 {}_{30}C_x (0.01)^x (0.99)^{30-x} = 0.996682$$

(c) Compute the Pr that the sample contains at least 2 NCUs.

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - F(1) = 1 - B(1; 30, 0.01) = 0.036148.$$

Note that the statistic  $\hat{p} = X/n$  is called the sample FNC and is used as a point estimate of process FNC,  $p$ .

**Exercise 14.** Prove the 2 properties of the variance operator:  
 $V(CY) = C^2V(Y)$ , and  $V(Y_1 + Y_2) = V(Y_1) + V(Y_2)$ , iff  $Y_1$  and  $Y_2$  are independent.

**Exercise 15.** Work Exercise 21, p. 105, and Exercises 46, 47, 50, 51, 52, 55, 60 on pages 120-122 of Devore.

**Exercise 16.** Prove that  $\sum_{x=0}^n b(x; n, p) = \sum_{x=0}^n {}_n C_x p^x (1-p)^{n-x} = 1$ . Hint: You must use the fact that  $(a + b)^n = \sum_{i=0}^n {}_n C_{n-i} a^{n-i} b^i$ . This last equality is called the binomial expansion of  $(a + b)^n$ .

## The Geometric Distribution

Consider one Bernoulli trial after another and define the rv  $X$  = Number of trials required to achieve the 1st success, and as in the case of  $\text{Bin}(n, p)$ ,  $p$  = Pr of success at each trial. Bear in mind that trials are independent. Then the Pr

that the 1st success occurs at the  $x^{\text{th}}$  trial is given by  $g(x; p) = (1 - p)^{x-1}p = q^{x-1}p$ ,  $x = 1, 2, 3, 4, \dots$ . Note that our author Devore uses the general notation  $p(x)$  and special notation  $nb(x; 1, p)$  for the above pmf but my preference is to use  $g(x; p)$ , where  $g$  denotes Geometric Pr and  $nb$  stands for Negative Binomial.

**Example 21.** Consider launching of rockets in succession where the success pr of a single launch is 0.95. Compute the Pr that the 1st failure occurs at the 8th trial, assuming that trials are independent.

$$g(8; 0.05) = (0.95)^7 (0.05) = 0.034917$$

Compute the Pr that the 1st failure occurs after the 10th trial.

$$P(X > 10) = (0.95)^{10} = 0.598737.$$

Compute the Pr that the 1st failure occurs within the 1st seven trials.

$$P(X \leq 7) = 1 - P(X > 7) = 1 - (0.95)^7 = 0.301663$$

Before computing the mean and variance of a Geometric rv, we will verify that  $g(x; p) = q^{x-1}p = nb(x; 1, p)$  is indeed a PDF. Recall from calculus that the geometric infinite

sum  $\sum_{i=0}^{\infty} ar^i$  converges to  $[a / (1 - r)]$  iff  $|r| < 1$ . Because  $0 < p < 1$ , then it follows that

$$\sum_{x=1}^{\infty} g(x; p) = \sum_{x=1}^{\infty} q^{x-1} p = p \sum_{x=0}^{\infty} q^x = p / (1 - q) = 1 \text{ because } 1 - q = p.$$

The mean of the Geometric distribution can be computed using the definition of  $\mu = E(X)$ .

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} x g(x; p) = \sum_{x=1}^{\infty} x q^{x-1} p = p \frac{d}{dq} \sum_{x=1}^{\infty} q^x = p \frac{d}{dq} \left( \frac{q}{1 - q} \right) = \\ &= p \frac{(1 - q) \times 1 - q(-1)}{(1 - q)^2} = p / (1 - q)^2 = 1 / p. \end{aligned}$$

**Exercise 17.** Compute  $E(X^2)$  and use it to show that the variance of the Geometric rv is given by  $\sigma^2 = q / p^2$ . (b) Show that the finite geometric sum  $S_n =$

$$\sum_{i=0}^n ar^i = \frac{a(1-r^{n+1})}{1-r}. \text{ Then use this last result to prove that if } |r| < 1, \text{ then the Lim } S_n$$

(as  $n \rightarrow \infty$ ) =  $a/(1-r)$ . (c) Prove that the Geometric distribution is memory-less! That is, show that  $P(X > a+b | X > a) = P(X > b)$ . Hint: First obtain the cdf of a Geometric random variable given by  $G(x; p) = 1 - q^x$ .

## THE PASCAL (OR NEGATIVE BINOMIAL) DISTRIBUTION

Again consider Bernoulli trials performed one after another (independently of each other). The Pascal rv occurs when the interest lies in the  $r^{\text{th}}$  success ( $r > 1$ ) occurring at the  $x^{\text{th}}$  trial,  $x = r, r+1, r+2, r+3, \dots$ . Note that the Geometric distribution is a special of the Pascal with  $r = 1$ .

For example, consider the rocket launching problem where the Pr of success was 0.95 at each trial. We wish to compute the Pr that the 3rd failure occurs at the 15th launching.

$$nb(15; 3, 0.05) = [{}_{14}C_2 (0.05)^2 (.95)^{12}] \times (0.05) = 0.0061466.$$

Note that the only way the 3rd failed launching can occur at the 15th trial is that during the first 14 trials exactly 2 failures and 12 successes occur. In other words, we have a Bin(14, 0.05) distribution in the first 14 trials followed by the occurrence of the 3<sup>rd</sup> Geometric event at the  $x = 15$ th trial.

**Exercise 18.** Explain the relation between the Binomial and Pascal distributions, i.e., explain why the Pascal PDF is also called Negative Binomial.

In general, the Pr that the  $r^{\text{th}}$  success will occur at the  $x^{\text{th}}$  Bernoulli trial is first the Pr that exactly  $(r-1)$  successes occur in the 1st  $(x-1)$  trials followed by the  $r^{\text{th}}$  success at the  $x^{\text{th}}$  trial, i.e.,

$$nb(x; r, p) = [{}_{x-1}C_{r-1} p^{r-1} q^{(x-1)-(r-1)}] \times p = {}_{x-1}C_{r-1} p^r q^{x-r}, \quad x = r, r+1, \dots \quad (6a)$$

Note that Devore defines the Negative-Binomial rv as the number of failures,  $Y$ , occurring before the occurrence of the  $r$ th success. Letting  $Y = X-r$  in Eq. (6a) results in

$$nb(y; r, p) = {}_{y+r-1}C_{r-1} p^r q^y, y = 0, 1, 2, \dots \quad (6b)$$

Equation (6b) is identical to the equation atop page 126 of Devore (8e).

**Example 3.38 on page 126 of Devore.** For this example that Devore provides on page 126,  $r = 5$ ,  $p = 0.20$  and  $x = 15$ . Hence the Pr that the doctor has to interview exactly 15 couples in order to recruit 5, from equation (6), is given by

$$nb(15; 5, 0.20) = {}_{14}C_4 (0.20)^5 (0.80)^{10} = 0.0343941.$$

The Pr that the doctor has to interview at most 15 couples to recruit 5 is given by

$$P(X \leq 15) = \sum_{x=5}^{15} {}_{x-1}C_4 (0.20)^5 (0.80)^{x-5} = \text{nbincdf}(10, 5, 0.20) = 0.164234, \text{ where}$$

$\text{nbincdf}(10, 5, 0.20)$  is a built-in Matlab function that gives the cdf of the Pascal at  $x = 15$  with  $p = 0.20$ . Note that MS Excel provides only the point mass Pr as  $\text{negbinomdist}(10, 5, 0.20) = 0.0343941$ .

Since the Geometric distribution is memory-less, a Pascal rv is simply the sum of Geometric rvs as described below.

Let  $X$  be the number trials needed to obtain exactly  $r$  successes. Let  $X_1$  be the number of trials required to obtain the 1st success;  $X_2 =$  number of extra trials (beyond  $X_1$ ) required to obtain the 2nd success;  $X_3$  is the number of extra trials beyond the 2<sup>nd</sup> success to attain the 3<sup>rd</sup> success, and so forth. Finally,  $X_r$  is the number of extra trials required after the  $(r - 1)$ th success to obtain the  $r$ th success. Then the total number trials required to obtain exactly  $r$  successes at the  $X = x$ th trial is given by

$$X = X_1 + X_2 + \dots + X_r, \quad (7)$$

where each  $X_i$  ( $i = 1, 2, \dots, r$ ) is a Geometric rv with parameter  $p$ .

**Exercise 19.** Use the above relationship (7) between the Pascal and the Geometric distributions to show that the mean and variance of a Pascal rv,  $X$ , are given by  $E(X) = r / p$  and  $V(X) = rq / p^2$ . Further, if the the negative-binomial rv is defined as the number of failures,  $Y$ , before the  $r$ th success, then  $E(Y) = rq/p$  and  $V(Y) = rq/p^2$ .

**Exercise 20.** An aircraft has 3 computers, only one of which is needed to

operate the aircraft. The other 2 are on standby redundancy that are activated one at a time in case the one on line fails. From past experience it has been determined that the Pr of failure (for the on-line computer) during any one hour is roughly 0.0004. Let  $X_1$ ,  $X_2$ , and  $X_3$  denote the number of hours of operation before the failure of the 1st, 2nd and 3rd computer, respectively. Then the time to failure (TTF) of the standby system is

$$X = X_1 + X_2 + X_3 \text{ hours.}$$

(a) Assuming that failures can occur only at the end of each hour and hours can be thought of independent Bernoulli trials, compute the Pr of system failure during a 7-hour flight. (b) Compute MTTF (Mean TTF). (c) Compute the variance of TTF. Hint: Again, consider each hour as one Bernoulli trial, where the on-line computer either fails or survives by the end of the hour.

Answers:  $0.22373 \times 10^{-8}$ , MTTF = 7500 hours,  $\sigma^2 = (4329.26091 \text{ hours})^2$ .

## THE HYPERGEOMETRIC DISTRIBUTION

Study pages 122-126 of Devore(8e) and work the following exercise.

**Exercise 21.** Lots of size  $N = 200$  pump shafts are inspected before shipment. The outgoing average lot quality (AOQ) is  $p = 0.02$  over many lots. The inspection plan calls for random samples of  $n = 25$  units (w/o replacement) and accepting an outgoing lot if the number of defectives in the sample is 1 or less. (a) Compute acceptance Pr of a lot, denoted by  $P_a$ . (b) Compute  $P_a$  if FNC increases to 5%. (c) Repeat part (a) if sampling is done with replacement.

ANS: (c)  $P_a = 0.911355$ . For parts (a) and (b) it is sufficient to give answers in terms of  ${}_n C_x$ . The approximate Pr for part (b) is  $P_a = 0.642376$ .

## THE POISSON DISTRIBUTION

Consider numerous many Bernoulli trials (i.e.,  $n \rightarrow \infty$ ) such that the occurrence Pr of success at each trial is small ( $p < 0.15$ ) and average number of successes per time unit  $= \mu = n \times p = \lambda$ , is a constant. Let the rv,  $X$ , denote the number of Poisson events (or generic successes, or the event of interest) that occur during one time interval of unit

length. It can be shown that the PDF (or pmf) of  $X$  is given by

$$p(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x = 0, 1, 2, 3, 4, \dots \quad (8)$$

It will be shown later that a Poisson process is simply the limiting distribution for a  $\text{Bin}(n, p)$  as  $n \rightarrow \infty$  and simultaneously as  $p \rightarrow 0$  such that the product  $n \times p$  stays fixed at a value equal to  $\lambda$ , and is generally required that  $n \times p \leq 20$ .

**Example 22.** The number of no-hit games occurring during a major league season is Poisson distributed with an average rate of  $\lambda = 1.8$  no-hit games per year. (a) Compute the Pr of exactly 2 no-hit games during the next year.  $P(X = 2) = p(2; 1.8) = \frac{1.8^2}{2!} e^{-1.8} = 0.267784$  (b) Compute the Pr of at most 2 no-hit games occurring during

the next season.  $P(X \leq 2) = \text{cdf}(\text{ of } X \text{ at } 2) = F_X(2; \lambda=1.8) = \sum_{x=0}^2 \frac{1.8^x}{x!} e^{-1.8} = 0.73062$ .

Note that Table A.2 in Appendix A-4 of Devore gives the cdf of the Poisson for  $\lambda = 0.10(0.10)1(1)20.0$ . (c) Compute the Pr that at least 2 no-hit games occur during the next baseball season.  $P(X \geq 2) = 1 - P(X \leq 1) = 1 - F_X(1; 1.8) = 1 - p(0; 1.8) - p(1; 1.8) = 1 - 0.46284 = 0.537163$ .

**Exercise 22.** The number of accidents per day in a certain city is Poisson distributed at an average rate of 5 accidents/day. (a) Compute the Pr of no accidents occur in the city during the next day. (b) Compute the Pr of at most 4 accidents occurring during the next day. (c) Compute the Pr of an odd number of accidents occurring during next day.

ANSWERS: (a) 0.006738, (b) 0.4404933, (c) 0.4999773. In order to obtain the answer for part (c), you have to make use of the fact that  $\text{Sinh}(x) = (e^x - e^{-x})/2 = (x/1!) + (x^3/3!) + (x^5/5!) + (x^7/7!) + \dots$ . For your info  $\text{Cosh}(x) = (e^x + e^{-x})/2$ , and hence the  $\text{Tanh}(x) = (e^x - e^{-x})/(e^x + e^{-x})$ .

The Poisson distribution can be used to compute probabilities over intervals of

length  $t$  ( $t \neq 1$  unit of time). Let  $Y$  = number of Poisson events occurring during an interval of length  $t$  ( $t \neq 1$ ) where the average number of Poisson events per unit of time ( $t = 1$ ) is  $\lambda = \mu = np$ . Then the average number of Poisson events per interval of length  $t$  is  $E(Y) = \lambda t$ . As a result the pmf for the rv  $Y$  is given by

$$p(y; \lambda t) = \frac{(\lambda t)^y}{y!} e^{-\lambda t}, \quad y = 0, 1, 2, 3, 4, \dots$$

**Example 23.** Consider the Poisson process of Example 22. We wish to compute the Pr of exactly 6 no-hit games during the next 4 years:

$$\lambda = np = 1.8/\text{year}, t = 4 \text{ years} \rightarrow \lambda \times t = 7.2 \text{ no-hit games/ four years} \rightarrow$$

$$P(Y = 6) = \frac{7.2^6}{6!} e^{-7.2} = 0.1444582, \text{ i.e., } Y \text{ is } p(y; 7.2) \text{ read as Poisson distributed at}$$

an average rate of 7.2. The Pr of at least 6 no-hit games in the next 4 years of major leagues is given by  $P(Y \geq 6) = 1 - P(Y \leq 5) = 1 - F_Y(5; 7.2) = 1 - 0.27590 = 0.72410$ .

**Exercise 22 (Continued).** Compute the Pr that there will be exactly 11 accidents in the city during the next 3 days. Secondly, compute the Pr that there will be at least 11 accidents in the next 3 days. Finally, compute the Pr that there will be at most 18 accidents in the next 4 days. ANSWERS: 0.0662874, 0.881536, 0.381422.

## PROPERTIES OF THE POISSON DISTRIBUTION

(1) We 1st verify that  $p(x; \lambda)$  is indeed a PDF ( or a pmf).

$$\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \left[ 1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots \right] = e^{-\lambda} \times e^{\lambda} = e^0 = 1$$

Recall from calculus that the infinite series in the above bracket is the Maclaurin series (i.e., Taylor's expansion about the origin) for  $e^{\lambda}$ .

(2) Compute the long-term (or long-run) average of  $X$ .

$$\mu = E(X) = \sum_{x=0}^{\infty} x \left( \frac{\lambda^x}{x!} e^{-\lambda} \right) = e^{-\lambda} \sum_{x=1}^{\infty} x \left( \frac{\lambda^x}{x!} \right) = e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} =$$

$$\mu = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} (e^{\lambda}) = \lambda.$$

Therefore, the process average for the Poisson distribution, as expected, is  $\mu = \lambda$ .  
 By now the reader must be cognizant of the fact that the mean of the binomial is larger than its variance ( $CV < 100\%$ ); the mean of the Negative-binomial is less than its variance iff  $q > p$ , while the mean of the Poisson distribution is equal to its variance.

**Exercise 23 (Property 3).** Use the above procedure to show that for a Poisson distribution  $E(X^2) = \lambda^2 + \lambda$  and hence  $V(X) = \lambda = \mu$ . As a result, all Poisson processes have a CV (coefficient of variation) equal to  $(100/\sqrt{\lambda})\%$ .

(4) The Poisson distribution is simply the limiting distribution of the binomial as  $n \rightarrow \infty$  and  $p \rightarrow 0$  such that  $n \times p = \lambda$  stays at a constant value.

**Proof.**

$$\left\{ \begin{array}{l} \text{Lim Bin}(n,p) \\ n \rightarrow \infty \\ p \rightarrow 0, np = \lambda \end{array} \right\} = \left\{ \begin{array}{l} \text{Lim } {}_n C_x p^x q^{n-x} \\ n \rightarrow \infty \\ p \rightarrow 0, np = \lambda \end{array} \right\} =$$

$$= \left\{ \begin{array}{l} \text{Lim } \frac{n!}{x!(n-x)!} (\lambda/n)^x (1-\lambda/n)^{n-x} \\ n \rightarrow \infty \\ p \rightarrow 0, np = \lambda \end{array} \right\} =$$

$$\begin{aligned}
&= \left\{ \begin{array}{l} \text{Lim}_{\substack{n \rightarrow \infty \\ p \rightarrow 0, np = \lambda}} \frac{n(n-1) \dots (n-x+1)}{x!} \times \frac{\lambda^x}{n^x} (1 - \lambda/n)^{n-x} \\ \end{array} \right. = \\
&= \left\{ \begin{array}{l} \frac{\lambda^x}{x!} \text{Lim}_{n \rightarrow \infty} \left[ 1 \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{1-x}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{n-x} \right] \\ \end{array} \right. = \\
&= \left\{ \begin{array}{l} \frac{\lambda^x}{x!} \text{Lim}_{n \rightarrow \infty} \left[ \left(1 - \frac{\lambda}{n}\right)^n \times \left(1 - \frac{\lambda}{n}\right)^{-x} \right] \\ \end{array} \right. = \left\{ \begin{array}{l} \frac{\lambda^x}{x!} \text{Lim}_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \\ \end{array} \right. . \quad (9)
\end{aligned}$$

Recall that by definition  $e = \left\{ \begin{array}{l} \text{Lim}_{m \rightarrow \infty} \left(1 + \frac{1}{m}\right)^m \\ \end{array} \right. \cong 2.71828183$ . Now make the

transformation  $1/m = -\lambda/n$  and substitute into equation (9).

$$\left\{ \begin{array}{l} \text{Lim}_{\substack{n \rightarrow \infty \\ p \rightarrow 0, np = \lambda}} {}_n C_x p^x q^{n-x} \\ \end{array} \right. = \left\{ \begin{array}{l} \frac{\lambda^x}{x!} \text{Lim}_{m \rightarrow \infty} \left(1 + \frac{1}{m}\right)^{-\lambda m} \\ \end{array} \right. =$$

$$\left\{ \begin{array}{l} \frac{\lambda^x}{x!} \left[ \text{Lim}_{m \rightarrow \infty} \left(1 + \frac{1}{m}\right)^m \right]^{-\lambda} \\ \end{array} \right. = \frac{\lambda^x}{x!} e^{-\lambda} . \quad \text{QED!}$$

Finally, consider  $K$  independent sources of Poisson arrivals at a service center with rates  $\lambda_i$  ( $i = 1, 2, 3, \dots, K$ ). Now, consider the total arrival stream, which is formed by merging the inputs from all  $K$  sources. It has been shown in the theory of stochastic

processes that the merged stream  $X = \sum_{i=1}^K X_i$ , where each  $X_i$  ( $i = 1, 2, \dots, K$ ) is Poisson

distributed with  $E(X_i) = \lambda_i$ , is also Poisson distributed with parameter  $\lambda = \sum_{i=1}^K \lambda_i$ . As an

example, the Compass Bank in downtown Auburn has  $K = 2$  doors for customer entrances. During the 10:00- 11:00 am hour, if  $\lambda_1 = 1.3$  customers arrive per minute thru the east door and  $\lambda_2 = 0.90$  customers arrive per minute thru the west door, then the total customer stream into the bank is Poisson distributed at an average rate of  $\lambda = 2.2$  customers per minute during 10:00- 11:00 am.

**Exercise 24.** Verify that the variance of Poisson distribution over  $t$  units of time is given by  $V(Y) = \lambda t$ , where  $\lambda$  is the average number Poisson events occurring per unit of time.

**Exercise 25.** Work Exercises 72, 75, 76, 77, 81, 85 and 87 on pages 127-132 of Devore's 8<sup>th</sup> edition.