FORMULAS AND TABLES

Canadian Business Statistics, by Vaidyanathan and Vaidyanathan

Chapter 3 • Describing Data—Measures of Central Tendency

$\bar{x} = \frac{\sum x}{n}$ Mean

$\bar{x} = \frac{\sum fx}{n}$ Mean for Grouped Data

$\bar{x} = \frac{\sum n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$ Combined Mean

$\bar{x}_w = \frac{\sum wx}{\sum w}$ Weighted Mean

The value of the $\left(\frac{n+1}{2}\right)^{th}$ item in ranked data
Median for ungrouped data

$\sqrt[n]{x_1 x_2 \ldots x_n}$ = Geometric Mean

$\sqrt[(n-1)]{\frac{\text{end of period value}}{\text{beg. of period value}}} - 1$ Geometric Mean

Approximate relationship for moderately asymmetrical distributions
Mean – mode ≅ 3 × (mean – median)

Chapter 4 • Measures of Dispersion

$\text{MAD} = \frac{\sum|x - \bar{x}|}{n}$

Sample Variance $s^2 = \frac{\sum (x - \bar{x})^2}{n-1}$

Sample Variance
Computational Method

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$$

Sample variance for grouped data

$$s^2 = \frac{\sum f(x - \bar{x})^2}{n-1}$$

Sample variance by Computational Method for Grouped Data

$$s^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n-1}$$

Standard Deviation for a sample
$s = \sqrt{s^2}$

Coefficient of Variation $CV = \frac{s}{\bar{x}} \times 100$

Percentiles:

If nk is an integer, the location of P_k is $\frac{nk}{100} + 0.5$

If $\frac{nk}{100}$ is not an integer, then the location of P_k is the next higher integer.

Chapter 5 • Index Numbers

Unweighted aggregate index $= \frac{\sum p_k}{\sum p_0} \times 100$

Weighted aggregate index $= \frac{\sum p_k q}{\sum p_0 q} \times 100$

Weighted quantity index $= \frac{\sum p_0 q_k}{\sum p_0 q_0} \times 100$

Laspeyres index $= \frac{\sum p_k q_0}{\sum p_0 q_0} \times 100$

Paasche index $= \frac{\sum p_k q_k}{\sum p_0 q_k} \times 100$

Fisher's ideal index $= \sqrt{\frac{\sum p_k q_0}{\sum p_0 q_0} \times \frac{\sum p_k q_k}{\sum p_0 q_k}} \times 100$

Value index $= \frac{\sum p_k q_k}{\sum p_0 q_0} \times 100$

Real Income $= \frac{\text{money income}}{\text{CPI}} \times 100$

Purchasing power of a dollar $= \frac{\$1}{CPI} \times 100$

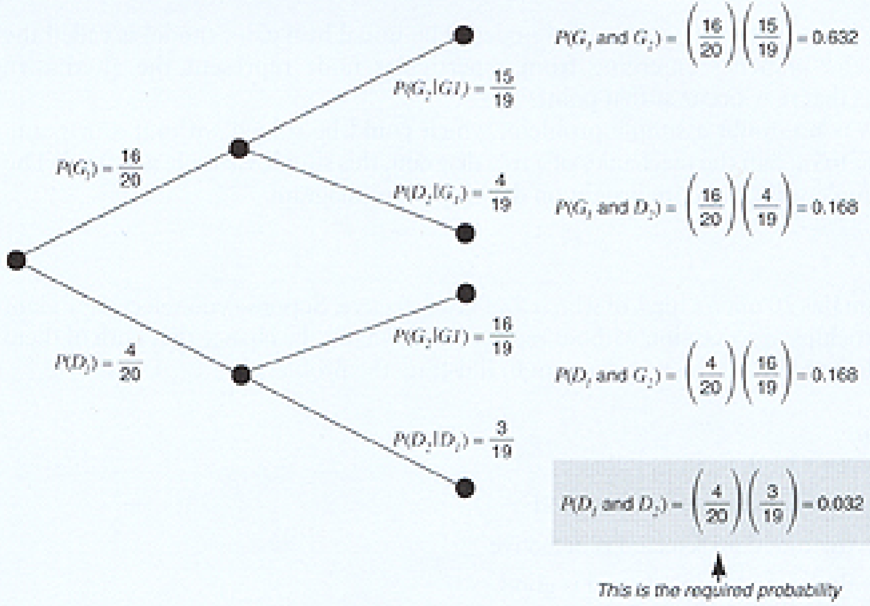

Figure 6-7

Tree Diagram on Selection of Two Microchips

6.21 THE DISTINCTIONS BETWEEN MUTUALLY EXCLUSIVE AND MUTUALLY INDEPENDENT EVENTS

Before concluding this chapter, we wish to clear a common misconception about mutually exclusive and mutually independent events. Students often think these two are the same thing. They are not. The following illustration will clear the misconception.

ILLUSTRATION

Consider the following events:

A. There will be a boom in the economy in the next quarter.

B. There will not be a boom in the economy in the next quarter.

C. The sales of automobiles will increase in the next quarter.

D. The Toronto Blue Jays will win the next World Series.

A and *B* are evidently mutually exclusive. If one occurs, the other cannot occur. Because the occurrence of *A* precludes the occurrence of *B* and vice versa, one event has an influence on the other and thus *A* and *B* are dependent. Events that are mutually exclusive must be dependent.

Events *A* and *C* are also dependent, since if the economy improves, the sales of cars are likely to show an upward trend; they are not mutually exclusive, since both can happen at the same time. Dependent events need not be mutually exclusive.

Events *A* and *C* are not mutually exclusive and dependent. Events *A* and *D* are not mutually exclusive and independent. Events that are not mutually exclusive may be either independent or dependent.

Events *A* and *D* are independent, but they are not (and cannot be) mutually exclusive. Independent events cannot be mutually exclusive. To recap:

7.7 THE HISTOGRAM OF THE BINOMIAL DISTRIBUTION

The shape of the histogram of the binomial distribution depends on the values of p and q and on the number of trials n. The distribution is symmetrical if p and q are equal, as depicted in Figure 7.3.

When p is small, say $p = 0.1$ and therefore $q = 0.9$, the distribution is positively skewed. (See Figure 7.4.)

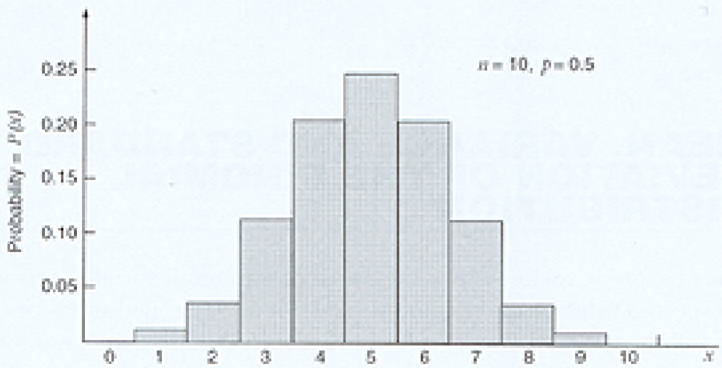

Figure 7-3

Binomial Distribution

When p is large, say $p = 0.9$ and $q = 0.1$, the distribution is negatively skewed. (See Figure 7.5.)

Note that if you keep the value of p constant, say $p = 0.4$, and increase the number of trials, namely n, the non-symmetrical binomial tends to become symmetrical or what is called bell-shaped. That is, for large values of n, even though p and q are not equal, the binomial tends to be symmetrical. This bell-shaped distribution, more commonly known as the normal distribution, will be studied shortly in the next chapter.

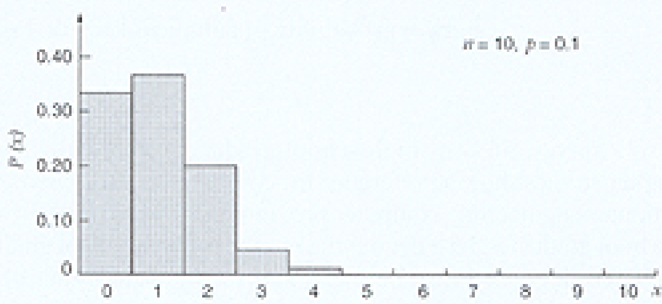

Figure 7-4

Positively Skewed Binomial Distribution

APPENDIX V — TABLE OF BINOMIAL PROBABILITIES

		p									
n	*x*	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
1	0	0.9500	0.9000	0.8500	0.8000	0.7500	0.7000	0.6500	0.6000	0.5500	0.5000
	1	0.0500	0.1000	0.1500	0.2000	0.2500	0.3000	0.3500	0.4000	0.4500	0.5000
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500
	1	0.0950	0.1800	0.2550	0.3200	0.3750	0.4200	0.4550	0.4800	0.4950	0.5000
	2	0.0025	0.0100	0.0225	0.0400	0.0625	0.0900	0.1225	0.1600	0.2025	0.2500
3	0	0.8754	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250
	1	0.1354	0.2430	0.3251	0.3840	0.4219	0.4410	0.4436	0.4320	0.4084	0.3750
	2	0.0071	0.0270	0.0574	0.0960	0.1406	0.1890	0.2389	0.2880	0.3341	0.3750
	3	0.0001	0.0010	0.0034	0.0080	0.0156	0.0270	0.0429	0.0640	0.0911	0.1250
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625
	1	0.1715	0.2916	0.3685	0.4096	0.4219	0.4116	0.3845	0.3456	0.2995	0.2500
	2	0.0135	0.0486	0.0975	0.1536	0.2109	0.2646	0.3105	0.3456	0.3675	0.3750
	3	0.0005	0.0036	0.0115	0.0256	0.0469	0.0756	0.1115	0.1536	0.2005	0.2500
	4	0.0000	0.0001	0.0005	0.0016	0.0039	0.0081	0.0150	0.0256	0.0410	0.0625
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0312
	1	0.2036	0.3281	0.3915	0.4096	0.3955	0.3602	0.3124	0.2592	0.2059	0.1562
	2	0.0214	0.0729	0.1382	0.2048	0.2637	0.3087	0.3364	0.3456	0.3369	0.3125
	3	0.0011	0.0081	0.0244	0.0512	0.0879	0.1323	0.1811	0.2304	0.2757	0.3125
	4	0.0000	0.0004	0.0022	0.0064	0.0146	0.0284	0.0488	0.0768	0.1128	0.1562
	5	0.0000	0.0000	0.0001	0.0003	0.0010	0.0024	0.0053	0.0102	0.0185	0.0312
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156
	1	0.2321	0.3543	0.3993	0.3932	0.3560	0.3025	0.2437	0.1866	0.1359	0.0938
	2	0.0305	0.0984	0.1762	0.2458	0.2966	0.3241	0.3280	0.3110	0.2780	0.2344
	3	0.0021	0.0146	0.0415	0.0819	0.1318	0.1852	0.2355	0.2765	0.3032	0.3125
	4	0.0001	0.0012	0.0055	0.0154	0.0330	0.0595	0.0951	0.1382	0.1861	0.2344
	5	0.0000	0.0001	0.0004	0.0015	0.0044	0.0102	0.0205	0.0369	0.0609	0.0938
	6	0.0000	0.0000	0.0000	0.0001	0.0002	0.0007	0.0018	0.0041	0.0083	0.0156
7	0	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078
	1	0.2573	0.3720	0.3960	0.3670	0.3115	0.2471	0.1848	0.1306	0.0872	0.0547
	2	0.0406	0.1240	0.2097	0.2753	0.3115	0.3177	0.2985	0.2613	0.2140	0.1641
	3	0.0036	0.0230	0.0617	0.1147	0.1730	0.2269	0.2679	0.2903	0.2918	0.2734
	4	0.0002	0.0026	0.0109	0.0287	0.0577	0.0972	0.1442	0.1935	0.2388	0.2734
	5	0.0000	0.0002	0.0012	0.0043	0.0115	0.0250	0.0466	0.0774	0.1172	0.1641
	6	0.0000	0.0000	0.0001	0.0004	0.0013	0.0036	0.0084	0.0172	0.0320	0.0547
	7	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0006	0.0016	0.0037	0.0078
8	0	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0319	0.0168	0.0084	0.0039
	1	0.2793	0.3826	0.3847	0.3355	0.2670	0.1977	0.1373	0.0896	0.0548	0.0312
	2	0.0515	0.1488	0.2376	0.2936	0.3115	0.2965	0.2587	0.2090	0.1569	0.1094
	3	0.0054	0.0331	0.0839	0.1468	0.2076	0.2541	0.2786	0.2787	0.2568	0.2188
	4	0.0004	0.0046	0.0185	0.0459	0.0865	0.1361	0.1875	0.2322	0.2627	0.2734
	5	0.0000	0.0004	0.0026	0.0092	0.0231	0.0467	0.0808	0.1239	0.1719	0.2188
	6	0.0000	0.0000	0.0002	0.0011	0.0038	0.0100	0.0217	0.0413	0.0703	0.1094
	7	0.0000	0.0000	0.0000	0.0001	0.0004	0.0012	0.0033	0.0079	0.0164	0.0312
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0007	0.0017	0.0039

CHAPTER *Objectives*

This chapter enables you to:

- *understand the characteristics of a continuous probability distribution*
- *learn about the normal distribution, its characteristics and its role in statistics as a celebrated distribution with great practical applications*
- *define and calculate z values and use the areas of the standard normal curve in calculating probabilities in various types of problems that occur in nature and business*
- *use the normal distribution to approximate the binomial probability distribution.*

8.1 INTRODUCTION

In the last chapter, we discussed discrete probability distributions and their characteristics. We also highlighted a well-known discrete probability distribution, namely the binomial distribution. In this chapter we will present a famous continuous probability distribution known as the **normal distribution**. Before we go into details of this particular distribution, certain observations about continuous probability distributions are in order.

1. A continuous random variable has infinite number of possible values and can assume any value in the interval between two points a and b ($a < x < b$). This was mentioned in the last chapter. Also recall that discrete random variables typically involve count data whereas continuous random variables involve measurement data, such as length, weight, speed and temperature.
2. The graph of the probability distribution for a continuous random variable x, will be a smooth curve that might appear as shown in Figure 8.1. This curve, a function of x symbolically written as $f(x)$, is variously called a **frequency function** or a **probability density function** or a **probability distribution**.[1]
3. A probability density function must satisfy the following conditions: For all values of x, $f(x) \geq 0$ and the area under $f(x)$ and above the x-axis must be equal to 1.0.
4. The areas under a probability distribution correspond to probabilities for x. For example, the shaded area in Figure 8.1 is the probability that x assumes a value between a and b ($a < x < b$). Note that for a continuous probability distribution, probabilities are always calculated for intervals like x falling between a and b and not for x assuming an individual value. Also note that the events $\{a < x < b\}$ and $\{a \leq x \leq b\}$ are equivalent for continuous random variables, because no probability is assigned to individual points.

[1] Mathematicians and statisticians denote a probability distribution of a discrete random variable by $P(x)$ and that of a continuous random variable by $f(x)$. $P(x)$ and $f(x)$ are comparable in the sense that both of these functions can be thought of as characterizing the probabilistic behaviour of their associated random variable. For a discrete probability distribution, the probability that x equals a specific value can be calculated. For example, in the coin-tossing example given in Table 7.2 in Chapter 7, $P(x = 1)$ is equal to 0.50. In the case of a continuous probability distribution, probabilities are always calculated for a range of values of x like the probability that x falls between a and b and not for x equal to a specific value. The number of points with a continuous distribution is infinite; hence the probability that any one single point occurs is so small that it is assumed to be zero i.e. $P(x$ = a specific value $k)$ is defined to be equal to zero for a continuous distribution.