# 3.4 The Binomial Probability Distribution

# The Binomial Probability Distribution

There are many experiments that conform either exactly or approximately to the following list of requirements:

**1.** The experiment consists of a sequence of $n$ smaller experiments called *trials,* where $n$ is fixed in advance of the experiment.

**2.** Each trial can result in one of the same two possible outcomes (dichotomous trials), which we generically denote by success ($S$) and failure ($F$).

**3.** The trials are independent, so that the outcome on any particular trial does not influence the outcome on any other trial.

# The Binomial Probability Distribution

**4.** The probability of success $P(S)$ is constant from trial to trial; we denote this probability by $p$.

**Definition**

An experiment for which Conditions 1–4 are satisfied is called a **binomial experiment.**

# Example 27

The same coin is tossed successively and independently *n* times.

We arbitrarily use *S* to denote the outcome *H* (heads) and *F* to denote the outcome *T* (tails). Then this experiment satisfies Conditions 1–4.

Tossing a thumbtack *n* times, with *S* = point up and *F* = point down, also results in a binomial experiment.

# The Binomial Random Variable and Distribution

In most binomial experiments, it is the total number of *S*'s, rather than knowledge of exactly which trials yielded *S*'s, that is of interest.

**Definition**

The **binomial random variable *X*** associated with a binomial experiment consisting of *n* trials is defined as

$$X = \text{the number of } S\text{'s among the } n \text{ trials}$$

# The Binomial Random Variable and Distribution

Suppose, for example, that $n = 3$.

Then there are eight possible outcomes for the experiment:

*SSS SSF SFS SFF FSS FSF FFS FFF*

From the definition of $X$, $X(SSF) = 2$, $X(SFF) = 1$, and so on. Possible values for $X$ in an $n$-trial experiment are $x = 0, 1, 2, \ldots, n$.

We will often write $X \sim \text{Bin}(n, p)$ to indicate that $X$ is a binomial rv based on $n$ trials with success probability $p$.

# The Binomial Random Variable and Distribution

**Notation**

Because the pmf of a binomial random variable *X* depends on the two parameters *n* and *p,* we denote the pmf by

$b(x; n, p)$.

$$b(x;n,p) = \binom{n}{x} p^x (1-p)^{1-x}, \ x = 0,1,\cdots,n.$$

# Example 31

Each of six randomly selected cola drinkers is given a glass containing cola *S* and one containing cola *F*. The glasses are identical in appearance except for a code on the bottom to identify the cola.

Suppose there is actually no tendency among cola drinkers to prefer one cola to the other.

Then $p = P$(a selected individual prefers $S$) = .5, so with
$X$ = the number among the six who prefer $S$,
$X \sim \text{Bin}(6,.5)$.

Thus

$$P(X = 3) = b(3;\, 6,\, .5) = \binom{6}{3}(.5)^3(.5)^3 = 20(.5)^6 = .313$$

# Example 31

cont'd

The probability that at least three prefer *S* is

$$P(3 \le X) = \sum_{x=3}^{6} b(x; 6, .5)$$

$$= \sum_{x=3}^{6} \binom{6}{x} (.5)^x (.5)^{6-x}$$

$$= .656$$

and the probability that at most one prefers *S* is

$$P(X \le 1) = \sum_{x=0}^{1} b(x; 6, .5)$$

$$= .109$$

# Using Binomial Tables

# Using Binomial Tables

Even for a relatively small value of *n,* the computation of binomial probabilities can be tedious.

Appendix Table A.1 tabulates the cdf $F(x) = P(X \leq x)$ for $n = 5, 10, 15, 20, 25$ in combination with selected values of *p.*

Various other probabilities can then be calculated using the proposition on cdf's.

A table entry of 0 signifies only that the probability is 0 to three significant digits since all table entries are actually positive.

# Using Binomial Tables

**Notation**

For $X \sim \text{Bin}(n, p)$, the cdf will be denoted by

$$B(x; n, p) = P(X \leq x) = \sum_{y=0}^{x} b(y; n, p) \qquad x = 0, 1, \ldots, n$$

# The Mean and Variance of *X*

# The Mean and Variance of $X$

For $n = 1$, the binomial distribution becomes the Bernoulli distribution.

The mean value of a Bernoulli variable is $\mu = p$, so the expected number of $S$'s on any single trial is $p$.

Since a binomial experiment consists of $n$ trials, intuition suggests that for $X \sim \text{Bin}(n, p)$, $E(X) = np$, the product of the number of trials and the probability of success on a single trial.

$V(X)$ is not so intuitive.

# The Mean and Variance of $X$

**Proposition**

If $X \sim \text{Bin}(n, p)$, then $E(X) = np$, $V(X) = np(1 - p) = npq$, and $\sigma_X = \sqrt{npq}$ (where $q = 1 - p$).

# Example 34

If 75% of all purchases at a certain store are made with a credit card and $X$ is the number among ten randomly selected purchases made with a credit card, then $X \sim \text{Bin}(10, .75)$.

Thus $E(X) = np = (10)(.75) = 7.5$,

$V(X) = npq = 10(.75)(.25)$

$= 1.875$,

and $\sigma = \sqrt{1.875}$

$= 1.37$.

# Example 34

cont'd

Again, even though $X$ can take on only integer values, $E(X)$ need not be an integer.

If we perform a large number of independent binomial experiments, each with $n$ = 10 trials and $p$ = .75, then the average number of $S$'s per experiment will be close to 7.5.

The probability that $X$ is within 1 standard deviation of its mean value is

$$P(7.5 - 1.37 \leq X \leq 7.5 + 1.37) = P(6.13 \leq X \leq 8.87)$$

$$= P(X = 7 \text{ or } 8)$$

$$= .532.$$