

## § 4.4 The Binomial Random Variable

Here we deal with repeated trials of the same experiment, where each trial has only two possible outcomes

### Examples

- (i) flip a fair coin <sup>10 times</sup> and count the number of times, out of 10, that the coin lands with tails face up.
- (ii) Ask  $n$  people if they approve of a policy.
- (iii) test  $n$  people with a new drug and look for the occurrence of a particular side effect.

Characteristics of a Binomial Random Variable:

1. The experiment consists of  $n$  identical trials.
2. There are only two possible outcomes on each trial (Success or Failure).
3. The probability of  $S$  remains the same for each trial. Usually denote  $p(S) = p$  and so  $p(F) = 1 - p$ .
4. The trials are independent.
5. The binomial variable  $x$  is the number of  $S$ 's in  $n$  trials.

Example We flip a fair coin four times. Let  $S$  = heads and  $F$  = tails (not heads). What values does  $x$  take? What is the PDF for this experiment?

The sample space is  $x$  take values in  $\{0, 1, 2, 3, 4\}$

The ways  $x$  equals each of these values are given by

$x=0$  : TTTT

$x=1$  : HTTT, THTT, TTHT, or TTTTH

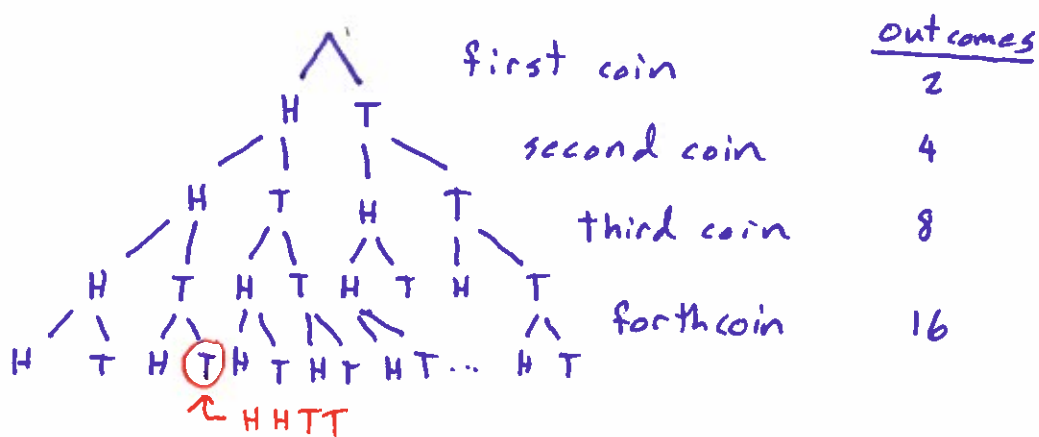
$x=2$  : A lot. See below \*

$x=3$  : THHH, HTHH, HHTH, or HHHT

$x=4$  : HHHH

Notice the symmetry between  $x=0/x=4$  and  $x=1/x=3$ .

How many total ways (outcomes) do we have



\* Since there are a total of 16 possible outcomes and  $x \in \{0, 1, 3, 4\}$  account for 10 of them, that means the number of ways of getting  $x=2$  is 6.

The PDF or Relative Frequency is

| $x$    | 0              | 1              | 2              | 3              | 4              |
|--------|----------------|----------------|----------------|----------------|----------------|
| $P(x)$ | $\frac{1}{16}$ | $\frac{4}{16}$ | $\frac{6}{16}$ | $\frac{4}{16}$ | $\frac{1}{16}$ |

## Binomial Coefficients

Pascal's triangle:

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & 1 & & 1 & & \\ & 1 & & 2 & & 1 & \\ 1 & & 3 & & 3 & & 1 \\ & 1 & & 4 & & 6 & & 4 & & 1 \\ 1 & & 5 & & 10 & & 10 & & 5 & & 1 \end{array}$$

Binomial expansion:

$$(x+y)^0 = 1$$

$$(x+y)^1 = x+y$$

$$(x+y)^2 = x^2 + 2xy + y^2$$

$$(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$

$$(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$$

$\binom{n}{k}$  or  ${}_nC_k = \frac{n!}{k!(n-k)!}$  are the binomial coefficients

which is the number of ways of getting  $k$  successes out of  $n$  trials.

Note:  $0! = 1$ ,  $1! = 1$ ,  $2! = (2)(1) = 2$ ,  $3! = (3)(2)(1) = 6$ ,  
 $4! = (4)(3)(2)(1) = 24, \dots$

In the previous example the number of ways of getting  $x=2$  would be  $\binom{4}{2} = {}_4C_2 = \frac{4!}{2!(4-2)!} = \frac{24}{(2)(2)} = 6$

The Binomial Probability Distribution:

$$p(x) = \binom{n}{x} p^x q^{n-x} \quad \text{for } x = 0, 1, 2, \dots, n$$

where  $p$  is the probability of a success in each trial.

$q = 1-p$  is the probability of a failure.

$n$  = Number of trials

$x$  = Number of desired successes

$n-x$  = Number of failures

$$\binom{n}{x} = {}_nC_x = \frac{n!}{x!(n-x)!}$$

Example Ten, six-sided dice are rolled and the number of six's are noted. Find the probability ~~of~~ that exactly seven of the dice are six's. What is the probability that there are fewer than 4 six's? What is the probability that there are four or more six's?

Here  $n=10$ ,  $p=\frac{1}{6}$ ,  $q=1-\frac{1}{6}=\frac{5}{6}$ . Thus,

$$P(7) = P(x=7) = \binom{10}{7} p^7 q^{10-7} = 120 \left(\frac{1}{6}\right)^7 \left(\frac{5}{6}\right)^3 \approx 2.48 \times 10^{-4} \\ = .0248\%$$

$$P(x < 4) = P(x=0) + P(x=1) + P(x=2) + P(x=3) \\ = \binom{10}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{10} + \binom{10}{1} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^9 + \binom{10}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^8 + \binom{10}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^7 \\ \approx 0.93027 = 93.03\%$$

$$P(x \geq 4) = 1 - P(x < 4) = 1 - 0.93027 \approx 0.06972 \\ = 6.97\%$$

Notice that calculating  $P(x < 4)$  involved calculating  $P(x=0)$ ,  $P(x=1)$ ,  $P(x=2)$  and  $P(x=3)$ . Imagine the difficulty if we had 100 trials and we wanted to calculate  $P(x < 35)$ . For this reason we will develop an approximation to the Binomial Distribution that will allow us to estimate probabilities of this type. To do that we need:

Mean, Variance and Standard Deviation for Binomial Random Variable

$$\text{Mean: } \mu = np$$

$$\text{Variance: } \sigma^2 = npq$$

$$\text{Standard Deviation: } \sigma = \sqrt{npq}$$

Example In the above example:

$$\mu = np = (10)\left(\frac{1}{6}\right) \approx 1.667$$

$$\sigma^2 = npq = (10)\left(\frac{1}{6}\right)\left(\frac{5}{6}\right) \approx 1.389$$

$$\sigma = \sqrt{npq} = \sqrt{\frac{50}{36}} \approx 1.179$$

Note: On pages 812-815 are tables for Binomial Distributions with  $n \in \{5, 6, 7, 8, 9, 10, 15, 20, 25\}$  and  $p \in \{.01, .05, .1, .2, .3, .4, .5, .6, .7, .8, .9, .95, .99\}$

Example Using the table to approximate the above example, we get:

We use the table for  $n=10$ , but  $p=\frac{1}{6}=0.167$  is between  $p=.1$  and  $p=.2$  in the table. Also note that the tables give  $\sum_{x=0}^k p(x)$ . So

$$P(x=7) = \sum_{x=0}^7 p(x) - \sum_{x=0}^6 p(x) \stackrel{\text{Table}}{=} \begin{cases} 0 & \text{for } p=.1 \\ .001 & \text{for } p=.2 \end{cases}$$

So we average these two values to get

$$P(x=7) \approx .0005 \quad \text{recall we got .000248 above}$$

$$\text{For } P(x < 4) = \sum_{x=0}^3 p(x) \stackrel{\text{Table}}{=} \begin{cases} .987 & \text{for } p=.1 \\ .879 & \text{for } p=.2 \end{cases}$$

Again if we take a straight average of these values we get:

$$P(x < 4) \approx \frac{.879 + .987}{2} = 0.933 \quad \text{recall we got .9303}$$