3.4 The Binomial Probability Distribution

- Consider an experiment consisting of multiple trials, each of which has exactly <u>two</u> outcomes.

<u>Example 1</u>: Suppose 2% of all items produced from an assembly line are defective. We randomly sample 50 items and count how many are defective (and how many are not).

- A <u>binomial experiment</u> has the following characteristics:

1. There are a fixed number, denoted n, of identical trials.
2. Each trial results in one of two outcomes (which we denote "Success" or "Failure").
3. The probability of success is constant across trials and is denoted p $(0 \leq p \leq 1)$. Hence $P[\text{Failure}] =$
4. The trials are independent.

- In such a binomial experiment, the total number of successes (out of n trials) is denoted Y and is called a ____________ r.v.
- Is Example 1 a binomial experiment?

<u>Example 2</u>: Suppose 40% of students at a college are male. We select 10 at random and count how many are male. Is this a binomial experiment?

<u>Binomial Distribution</u>

- If Y (the number of successes) is a binomial r.v., we should derive its probability function, $P(Y=y)$.
- If we have n trials, a typical sample point looks like:

- Of the n slots, suppose we have y successes (where $0 \le y \le n$).

- How many arrangements of those n letters contain y S's and $(n-y)$ F's?

<u>Answer</u>:

- For any sample point containing y S's and $(n-y)$ F's, what is its probability?

- Corresponding to $Y=y$, we have sample points, each with probability
- So for any $y = 0, 1, \dots, n$:

- This is the <u>binomial probability distribution</u>.
- Figure 3.4 shows some probability graphs for various n and p.

- These binomial probabilities are terms in the binomial expansion:

- Theorem: The binomial distribution is a valid probability distribution.

Proof:

Example 1(a): If we sample 10 items from that assembly line, what is the probability that 2 or more are defective?

- For many problems, we use tables or software to find binomial probabilities.

- Table 1, Appendix 3 gives cumulative probabilities, $P(Y \leq a)$:

- Example 2(a): What is the probability that 7 or fewer of the selected students are male?

- What is the probability that at least 4 of the selected students are male?

- See analogous R examples on the course web page.

Theorem (Mean and variance of a binomial r.v.):
Let Y be a binomial r.v. with n trials and success probability p.
Shorthand notation:
Then: (i)

(ii)

Proof (i):

Proof (ii): We know $V(Y) =$

But it is difficult to find $E(Y^2)$ directly.

Recall Example 1(a): Find $E(Y)$, $V(Y)$ and σ.

Recall Example 2(a): Find $E(Y)$, $V(Y)$ and σ.

Note: When we have a single trial $(n=1)$, the binomial probability function is:

- This $Bin(1, p)$ distribution is usually called the ______________ distribution.