3.5 The Geometric Distribution

- Consider an experiment with a series of identical and independent trials, each resulting in either a success or a failure.
- Similar to a binomial experiment, except:
  - There is <u>not</u> a fixed number of trials.
  - Rather, the series concludes after the first success.
- The r.v. of interest, Y, is the number of the trial on which the first success occurs.
- The sample space consists of:

- If $P(\text{Success}) = p$ and $P(\text{Failure}) = q$ on each trial, then:

In general, for any number $y = 1, 2, 3, \ldots$

This is the probability function of the geometric distribution.

Example 1: Suppose the probability of an applicant passing a driving test is 0.75 on any given attempt (and that attempts are independent). What is the probability that his initial pass is on his fourth try?

Lemma (Geometric Series): Let r be a number such that $|r| < 1$. Then

Corollary:

Theorem: If Y is a geometric r.v. [Shorthand: $Y \sim geom(p)$], then:

Proof:

Theorem: If $Y \sim \text{geom}(p)$, then:

Proof: Will be given in a later section.

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

3.6 The Negative Binomial Distribution

- Consider independent and identical trials, each resulting in a success or failure.
- Now we define a r.v. Y that is the number of the trial on which the r-th success occurs.

Note: If $r=1$, then this Y is

a ______________ r.v.

- What is the probability that the r-th success occurs on trial y?

This implies the first $(y-1)$ trials contain ______________, and the y-th trial is a __________.

Probability of this?

So the probability function of this <u>negative</u> <u>binomial</u> distribution is:

<u>Example 2:</u> Suppose 40% of employees at a firm have traces of asbestos in their lungs. The firm is asked to send 3 such employees to a medical center for further testing. Find the probability that exactly 10 employees must be checked to find 3 with asbestos traces.

Theorem: If Y is a negative binomial r.v. [Shorthand: $Y \sim NB(r, p)$], then:

Proof: Will be proved in a later section.

Example 2(a): If each initial test costs \$20, find $E(C)$ and $V(C)$, where C = the total costs of the initial tests for the firm.

Contrast with Binomial Distribution

- Both involve a sequence of "Bernoulli trials".
- For the binomial r.v., we fix the number of trials and count the number of successes obtained.
- For the negative binomial r.v., we fix the number of successes and count the number of trials needed.

Note: Some sources define a "geometric" r.v. or a "negative binomial" r.v. differently than our book:

As "the number of failures before the first success," or "the number of failures before the r-th success," respectively.

- Using those definitions, the geometric and negative binomial probability functions, means, variances, etc. would be somewhat different.

- In this class, we will always use the definitions given in our book.

- Examples of finding geometric and negative binomial probabilities in R are given on the course web page.