Geometric Distribution

Colby Community College

Definition

When an individual trial only has two possible outcomes, often labeled `success` or `failure`, it is called a **Bernoulli random variable**.

Note

It does not matter which outcome is labeled as `success` or `failure`, just that there are only two outcomes.

Note

Bernoulli random variables are often denoted with

- 1 for `success`
- 0 for `failure`

Note

The events `success` and `failure` are complements.

Example 1

Subjects are randomly selected for the National Health and Nutrition Examination Survey conducted by the National Center for Health Statistics, Centers for Disease Control and Prevention.

A person is a universal donor if they have group O and type Rh blood.

If we think of each subject as a trial then:

- If a person is a universal donor, we label them a `success`.
- If a person is not a universal donor, we label them a `failure`.

If there is a 6% chance that a person is a universal donor, then:

- The probability of a success is $p = 0.06$
- The probability of a failure is $q = 1 - p = 0.94$

Note

`success` and `failure` are not moral descriptions. We could have just as easily labeled the universal donors as `failure`.

Definition

The **sample proportion**, $\hat{p}$, is the sample mean:

$$\hat{p} = \frac{\text{\# of successes}}{\text{\# of trials}}$$

Example 2

Suppose we observe the ten trials of a Bernoulli random variable:

$$1\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 1\ 0$$

The sample proportion of these observations would be:

$$\hat{p} = \frac{1+1+1+0+1+0+0+1+1+0}{10} = 0.6$$

Bernoulli Random Variable

If X is a random variable that takes value 1 with probability p and 0 with probability $q = 1 - p$, then X is a Bernoulli random variable with mean and standard deviation:

$$\mu = p \qquad \sigma = \sqrt{p(1-p)}$$

Example 3

In Example 1, X describes the chances a subject is a universal donor, with probability of success $p = 0.06$.

The mean of X is:

$$\mu = p = 0.06$$

The standard deviation of X is:

$$\sigma = \sqrt{p(1-p)} = \sqrt{0.06(1-0.06)} = \sqrt{0.0564} = 0.237486842$$

Definition

The **geometric distribution** is used to describe how many trails it takes to observe a success.