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 trials it takes to observe a success.

Example 4

If we are looking for universal donors as in Example 1, then the probability the first universal donor found is the first persion is 0.06.

The probability that the first universal donor is the second person.

$$P\left(1^{\text{st}} \text{ no and } 2^{\text{nd}} \text{ yes}\right) = (1 - 0.06)(0.06) = (0.94)(0.06) = 0.0564$$

The probability that the first universal donor is the third person.

$$P\left(1^{\text{st}} \text{ no and } 2^{\text{nd}} \text{ no and } 3^{\text{rd}} \text{ yes}\right) = (0.94)(0.94)(0.06) = 0.053016$$

The probability that the first universal donor is the n^{th} person.

$$\begin{aligned} P\left(1^{\text{st}} \text{ through } (n-1)^{\text{th}} \text{ no and } n^{\text{th}} \text{ yes}\right) &= (0.94)\cdots(0.94)(0.06) \\ &= (0.94)^{n-1} \cdot 0.06 \end{aligned}$$

Geometric Distribution

If the probability of a success in one trial is p and the probability of failure is $1 - p$, then the probability of finding the first success in the n^{th} trial is given by

$$(1-p)^{n-1} \cdot p$$

The mean, variance and standard deviation of this wait time are

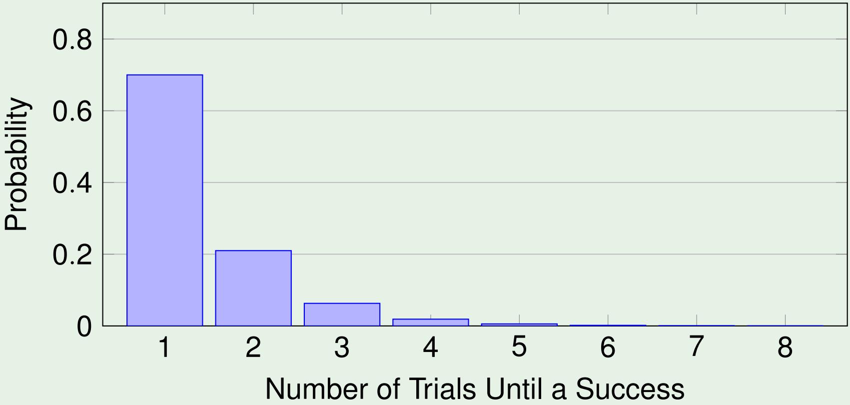
$$\mu = \frac{1}{p} \quad \sigma^2 = \frac{1-p}{p^2} \quad \sigma = \sqrt{\frac{1-p}{p^2}}$$

Note

The trials need to be both independent and identical to use the geometric distribution.

Example 5

The geometric distribution for $p = 0.7$



Note

We call this the geometric distribution because the probabilities decrease exponentially fast as n increases.

Note

The expected value, μ, tells us, on average, how many trials it takes to get a success. The higher p, the fewer it takes.

Example 6

Since the probability of someone being a universal donor is $p = 0.06$, the expected value is:

$$\mu = \frac{1}{p} = \frac{1}{0.06} = 16.666666667$$

It should take, on average, 16.7 people until a universal donor is found.

Example 7

The probability of getting a "heads" on a fair coin is $p = 0.5$, so the expected value is:

$$\mu = \frac{1}{p} = \frac{1}{0.5} = 2$$

On average, it should only take two flips to get a "heads".

Example 8

The probability that a universal donor is in the first three trials is:

$$\begin{aligned} P(\text{first or second or third}) &= P(\text{first}) + P(\text{second}) + P(\text{third}) \\ &= (1-.06)^{1-1}\cdot 0.06 + (1-.06)^{2-1}\cdot 0.06 \\ &\qquad + (1-.06)^{3-1}\cdot 0.06 \\ &= 0.169416 \end{aligned}$$

There is roughly a 16.9% chance a universal donor will be one of the first three people.

Note

We could have also used complements:

$$\begin{aligned} P(\text{first or second or third}) &= 1 - P(\text{none in first three}) \\ &= 1-(1-0.06)^3 \\ &= 0.169416 \end{aligned}$$