Chapter 5: Some Discrete Probability Distributions

Juhyung Lee

Department of Statistics
University of Florida

Binomial Distribution

Bernoulli Process

An experiment is called a **Bernoulli process** if

1. The experiment consists of repeated trials.
2. Each trial (called a **Bernoulli trial**) results in an outcome that may be classified as a success or a failure.
3. The probability of success, denoted by p, remains constant from trial to trial.
4. The repeated trials are independent.

(e.g., tossing a fair coin n times independently, where a head is a success)

Binomial Distribution

Definition (Binomial Random Variable)

The number of successes in n Bernoulli trials is called a **binomial random variable**.

Definition (Binomial Distribution)

A Bernoulli trial can result in a success with probability p and a failure with probability $1 - p$. Then the probability distribution of the binomial random variable X, the number of successes in n independent trials, is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \; x = 0, 1, 2, \ldots, n.$$

We write $X \sim Bin(n, p)$.

Example (Example 5.2 of WMMY)

The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that (a) at least 10 survive, (b) from 3 to 8 survive, and (c) exactly 5 survive?

Binomial Distribution

Example (Example 5.2 of WMMY)

```
> # a
> 1 - pbinom(9, size = 15, prob = 0.4)
[1] 0.0338333
> # equivalently,
> pbinom(9, size = 15, prob = 0.4, lower.tail = FALSE)
[1] 0.0338333
>
> # b
> sum(dbinom(3:8, 15, 0.4))
[1] 0.8778386
> # equivalently,
> pbinom(8, 15, 0.4) - pbinom(2, 15, 0.4)
[1] 0.8778386
>
> # c
> dbinom(5, 15, 0.4)
[1] 0.1859378
```

Binomial Distribution

Theorem (Theorem 5.1 of WMMY)

The mean and variance of the binomial distribution $Bin(n, p)$ are

$$\mu = np \text{ and } \sigma^2 = np(1 - p).$$

Example (Example 5.5 of WMMY)

Find the mean and variance of the binomial random variable of Example 5.2.

Negative Binomial and Geometric Distributions

Negative Binomial Experiment

Consider an experiment where the trials are repeated until a fixed number of successes occur. We are interested in the probability that the kth success occurs on the xth trial. Experiments of this kind are called **negative binomial experiments**.

Definition (Negative Binomial Random Variable)

The number X of trials required to produce k successes in a negative binomial experiment is called a **negative binomial random variable**.

Negative Binomial and Geometric Distributions

Definition (Negative Binomial Distribution)

If repeated independent trials can result in a success with probability p and a failure with probability $1-p$, then the probability distribution of the random variable X, the number of the trial on which the kth success occurs, is

$$f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \ x = k, k+1, k+2, \ldots$$

We write $X \sim NB(k, p)$.

Negative Binomial and Geometric Distributions

Example (Example 5.14 of WMMY)

In an NBA championship series, the team that wins four games out of seven is the winner. Suppose that teams A and B face each other in the championship games and that team A has probability 0.55 of winning a game over team B.

a. What is the probability that team A will win the series in 6 games?
b. What is the probability that team A will win the series?
c. If teams A and B were facing each other in a regional playoff series, which is decided by winning three out of five games, what is the probability that team A would win the series?

Negative Binomial and Geometric Distributions

Negative Binomial Random Variable in R

R defines the negative binomial random variable

$$Y = \textit{number of failures before the } k\textit{th success}$$

and

$$P(Y = y) = \binom{y+k-1}{k-1} p^k (1-p)^y, \; y = 0, 1, 2, \ldots$$

This is statistically equivalent to the previous definition

$$X = \textit{number of the trial on which the } k\textit{th success occurs}$$

since $Y = X - k$.

Negative Binomial and Geometric Distributions

Example (Example 5.14 of WMMY)

```
> # a
> dnbinom(2, size = 4, prob = 0.55)
[1] 0.1853002
>
> # b
> pnbinom(3, 4, 0.55)
[1] 0.6082878
>
> # c
> pnbinom(2, 3, 0.55)
[1] 0.5931269
```

Negative Binomial and Geometric Distributions

When $k = 1$, the negative binomial random variable becomes the geometric random variable.

Definition (Geometric Distribution)

If repeated independent trials can result in a success with probability p and a failure with probability $1 - p$, then the probability distribution of the random variable X, the number of the trial on which the first success occurs, is

$$f(x) = p(1-p)^{x-1}, \ x = 1, 2, 3, \ldots$$

We write $X \sim Geom(p)$.

Example (Example 5.15 of WMMY)

For a certain manufacturing process, it is known that, on the average, 1 in every 100 items is defective. What is the probability that the fifth item inspected is the first defective item found?

Negative Binomial and Geometric Distributions

Geometric Random Variable in R

R defines the geometric random variable

$$Y = \textit{number of failures before the first success}$$

and

$$P(Y = y) = p(1-p)^y, \; y = 0, 1, 2, \ldots$$

This is statistically equivalent to the previous definition

$$X = \textit{number of the trial on which the first success occurs}$$

since $Y = X - 1$.

Example (Example 5.15 of WMMY)

```
> dgeom(4, prob = 0.01)
[1] 0.00960596
```

Negative Binomial and Geometric Distributions

Theorem (Theorem 5.3 of WMMY)

The mean and variance of a random variable following the geometric distribution are

$$\mu = \frac{1}{p} \text{ and } \sigma^2 = \frac{1-p}{p^2}.$$

Example

In Example 5.15 of WMMY, what is the expected number of items to be inspected to find a defective item?