



# Chapter 4

Dr Wayne Stewart



# Chapter 4

---

## Discrete Random Variables

### OBJECTIVE

*To explain what is meant by a discrete random variable, its probability distribution, and corresponding numerical descriptive measures; to present some useful discrete probability distributions and show how they can be used to solve practical problems*

### CONTENTS

- 4.1 Discrete Random Variables
- 4.2 The Probability Distribution for a Discrete Random Variable
- 4.3 Expected Values for Random Variables
- 4.4 Some Useful Expectation Theorems
- 4.5 Bernoulli Trials
- 4.6 The Binomial Probability Distribution
- 4.7 The Multinomial Probability Distribution
- 4.8 The Negative Binomial and the Geometric Probability Distributions
- 4.9 The Hypergeometric Probability Distribution
- 4.10 The Poisson Probability Distribution
- 4.11 Moments and Moment Generating Functions (*Optional*)

# Skills

- You need to be able to categorize a problem into one of the 7 distributions.
- Once done you will need to be able to use R to calculate the probability that the random variable takes a value or values.
  - **This requires that you can reparametrize the problem**
  - You must master the 4 functions
    - d- stem
    - p – stem
    - q – stem
    - r – stem

**Example:**  
dbinom()  
pbinom()  
qbinom()  
rbinom()

**Definition 4.1**

A random variable  $Y$  is a numerical-valued function defined over a sample space. Each simple event in the sample space is assigned a value of  $Y$ .

**Definition 4.2**

A discrete random variable  $Y$  is one that can assume only a countable number of values.

# Assigning probabilities to random variables

## Example 4.1

Probability Distribution for Coin Tossing Experiment

Solution

A balanced coin is tossed twice, and the number  $Y$  of heads is observed. Find the probability distribution for  $Y$ .

Let  $H_i$  and  $T_i$  denote the observation of a head and a tail, respectively, on the  $i$ th toss, for  $i = 1, 2$ . The four simple events and the associated values of  $Y$  are shown in Table 4.1. You can see that  $Y$  can take on the values 0, 1, or 2.

**TABLE 4.1 Outcomes of Coin-Tossing Experiment**

Simple Event	Description	$P(E_i)$	Number of Heads
			$Y = y$
$E_1$	$H_1H_2$	$\frac{1}{4}$	2
$E_2$	$H_1T_2$	$\frac{1}{4}$	1
$E_3$	$T_1H_2$	$\frac{1}{4}$	1
$E_4$	$T_1T_2$	$\frac{1}{4}$	0

**Definition 4.3**

The probability distribution for a discrete random variable  $Y$  is a table, graph, or formula that gives the probability  $p(y)$  associated with each possible value of  $Y = y$ .

## Requirements for a Discrete Probability Distribution

1.  $0 \leq p(y) \leq 1$

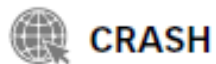
2.  $\sum_{\text{all } y} p(y) = 1$



# Go through all examples

## Example 4.2

### Probability Distribution for Driver-Side Crash Ratings



The National Highway Traffic Safety Administration (NHTSA) has developed a driver-side “star” scoring system for crash-testing new cars. Each crash-tested car is given a rating ranging from one star (\*) to five stars (\*\*\*\*\*); the more stars in the rating, the better the level of crash protection in a head-on collision. Recent data for 98 new cars are saved in the **CRASH** file. A summary of the driver-side star ratings for these cars is reproduced in the MINITAB printout, Figure 4.3. Assume that one of the 98 cars is selected at random and let  $Y$  equal the number of stars in the car’s driver-side star rating. Use the information in the printout to find the probability distribution for  $Y$ . Then find  $P(Y \leq 3)$ .

**FIGURE 4.3**

MINITAB summary of driver-side  
star ratings

### Tally for Discrete Variables: DRIVSTAR

DRIVSTAR	Count	Percent
2	4	4.08
3	17	17.35
4	59	60.20
5	18	18.37
N=	98	

### THEOREM 4.1

Let  $Y$  be a discrete random variable with probability distribution  $p(y)$  and let  $c$  be a constant. Then the expected value (or mean) of  $c$  is

$$E(c) = c$$

### THEOREM 4.2

Let  $Y$  be a discrete random variable with probability distribution  $p(y)$  and let  $c$  be a constant. Then the expected value (or mean) of  $cY$  is

$$E(cY) = cE(Y)$$

### THEOREM 4.3

Let  $Y$  be a discrete random variable with probability distribution  $p(y)$ , and let  $g_1(Y), g_2(Y), \dots, g_k(Y)$  be functions of  $Y$ . Then

$$E[g_1(Y) + g_2(Y) + \cdots + g_k(Y)] = E[g_1(Y)] + E[g_2(Y)] + \cdots + E[g_k(Y)]$$

#### THEOREM 4.4

Lets  $Y$  be a discrete random variable with probability distribution  $p(y)$  and mean  $\mu$ . Then the variance of  $Y$  is:

$$\sigma^2 = E(Y^2) - \mu^2$$

**Proof of Theorem 4.4** From Definition 4.6, we have the following expression for  $\sigma^2$ :

$$\sigma^2 = E[(Y - \mu)^2] = E(Y^2 - 2\mu Y + \mu^2)$$

Applying Theorem 4.3 yields

$$\sigma^2 = E(Y^2) + E(-2\mu Y) + E(\mu^2)$$

We now apply Theorems 4.1 and 4.2 to obtain

$$\begin{aligned}\sigma^2 &= E(Y^2) - 2\mu E(Y) + \mu^2 = E(Y^2) - 2\mu(\mu) + \mu^2 \\ &= E(Y^2) - 2\mu^2 + \mu^2 \\ &= E(Y^2) - \mu^2\end{aligned}$$

## The Bernoulli Probability Distribution

Consider a Bernoulli trial where

$$Y = \begin{cases} 1 & \text{if a success (S) occurs} \\ 0 & \text{if a failure (F) occurs} \end{cases}$$

The probability distribution for the Bernoulli random variable  $Y$  is given by

$$p(y) = p^y q^{1-y} \quad (y = 0, 1)$$

where

$p$  = Probability of a success for a Bernoulli trial

$$q = 1 - p$$

The mean and variance of the Bernoulli random variable are, respectively,

$$\mu = p \quad \text{and} \quad \sigma^2 = pq$$



In R, if  $Y \sim \text{Bern}(p)$   
then

---

$$P(Y = y) = \text{dbinom}(y, \text{size} = 1, \text{prob} = p)$$

### Characteristics That Define a Binomial Random Variable

1. The experiment consists of  $n$  identical Bernoulli trials.
2. There are only two possible outcomes on each trial:  $S$  (for Success) and  $F$  (for Failure).
3.  $P(S) = p$  and  $P(F) = q$  remain the same from trial to trial. (Note that  $p + q = 1$ .)
4. The trials are independent.
5. The binomial random variable  $Y$  is the number of  $S$ 's in  $n$  trials.

## The Binomial Probability Distribution

The probability distribution for a binomial random variable  $Y$  is given by

$$p(y) = \binom{n}{y} p^y q^{n-y} \quad (y = 0, 1, 2, \dots, n)$$

where

$p$  = Probability of a success on a single trial

$q = 1 - p$

$n$  = Number of trials

$y$  = Number of successes in  $n$  trials

$$\binom{n}{y} = \frac{n!}{y!(n-y)!}$$

The mean and variance of the binomial random variable are, respectively,

$$\mu = np \quad \text{and} \quad \sigma^2 = npq$$

In R, if  $Y \sim \text{Bin}(n, p)$  then:

$$P(Y = y) = \text{dbinom}(y, \text{size} = n, \text{prob} = p)$$



**Example 4.11** (optional)

Derivation of Binomial  
Expected Value

Solution

Derive the formula for the expected value for the binomial random variable,  $Y$ .

By Definition 4.4,

$$\mu = E(Y) = \sum_{\text{all } y} yp(y) = \sum_{y=0}^n y \frac{n!}{y!(n-y)!} p^y q^{n-y}$$

The easiest way to sum these terms is to convert them into binomial probabilities and then use the fact that  $\sum_{y=0}^n p(y) = 1$ . Noting that the first term of the summation is equal to 0 (since  $Y = 0$ ), we have

$$\begin{aligned} \mu &= \sum_{y=1}^n y \frac{n!}{[y(y-1) \cdots 3 \cdot 2 \cdot 1](n-y)!} p^y q^{n-y} \\ &= \sum_{y=1}^n \frac{n!}{(y-1)!(n-y)!} p^y q^{n-y} \end{aligned}$$

Because  $n$  and  $p$  are constants, we can use Theorem 4.2 to factor  $np$  out of the sum:

$$\mu = np \sum_{y=1}^n \frac{(n-1)!}{(y-1)!(n-y)!} p^{y-1} q^{n-y}$$

Let  $Z = (Y - 1)$ . Then when  $Y = 1$ ,  $Z = 0$  and when  $Y = n$ ,  $Z = (n - 1)$ ; thus,

$$\begin{aligned} \mu &= np \sum_{y=1}^n \frac{(n-1)!}{(y-1)!(n-y)!} p^{y-1} q^{n-y} \\ &= np \sum_{z=0}^{n-1} \frac{(n-1)!}{z![(n-1)-z]!} p^z q^{(n-1)-z} \end{aligned}$$

The quantity inside the summation sign is  $p(z)$ , where  $Z$  is a binomial random variable based on  $(n - 1)$  Bernoulli trials. Therefore,

$$\sum_{z=0}^{n-1} p(z) = 1$$

and

$$\mu = np \sum_{z=0}^{n-1} p(z) = np(1) = np$$

# The quantile function for a discrete random variable.

- Suppose  $X \sim \text{Bin}(n = 5, p = 0.5)$
- Then how do we explain the following R code?

```
> cp <- pbinom(0:5, 5, 0.5) # cumulative prob
> cp
[1] 0.03125 0.18750 0.50000 0.81250 0.96875 1.00000
> qbinom(cp, 5, 0.5)
[1] 0 1 2 3 4 5
> cpr <- round(cp, 2)
> cpr
[1] 0.03 0.19 0.50 0.81 0.97 1.00
> qbinom(cpr, 5, 0.5)
[1] 0 2 2 3 5 5
> |
```

# Summary of q-stem for discrete distributions

- To find  $x$  when lower tail probability is given as  $lt$  – do the following
- Choose smallest  $x$  so that  $F(x) \geq lt$  where  $F$  is the cumulative probability distribution function

- Predict what `qbinom(0.1, 5, 0.5)` will equal

```
> cp <- pbinom(0:5, 5, 0.5) # cumulative prob
```

```
> cp
```

```
[1] 0.03125 0.18750 0.50000 0.81250 0.96875 1.00000
```

Answer!

```
> qbinom(0.1,5,0.5)
```

```
[1] 1
```



Did you get it right??



# Application of binomial to airline seat overbooking

- On any given flight there could be passenger “no shows” – airline companies will want to have “full flights”.
- One way to help fill a flight is to sell more tickets than the number of seats on the plane.
- Unfortunately, depending on the nature of the flight, timing and number of seats sold it is possible that on occasion more people with tickets will show than there are seats on the flight. We don't want this to happen too many times.
- If  $N = 200$  is the number of seats on the flight and  $n$  is the number of seats sold and  $\gamma = 0.02$  is the probability of overbooking and  $p = 0.95$  is the probability a passenger will show. Find an expression for “ $n$ ”!

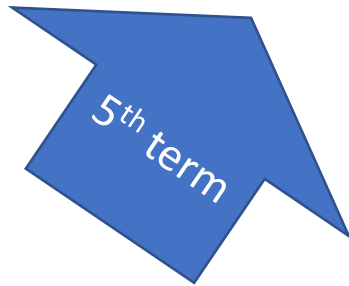
# The equation

$$N = qbinom(1 - \gamma, n, p)$$

$$\mathbf{200} = qbinom(1 - 0.02, n, 0.95), n?$$

```
> qbinom(1-0.02,200:210, 0.95)
```

```
[1] 196 197 198 199 200 201 202 203 204 204 205
```



**n = 204**



### Properties of the Multinomial Experiment

1. The experiment consists of  $n$  identical trials.
2. There are  $k$  possible outcomes to each trial.
3. The probabilities of the  $k$  outcomes, denoted by  $p_1, p_2, \dots, p_k$ , remain the same from trial to trial, where  $p_1 + p_2 + \dots + p_k = 1$ .
4. The trials are independent.
5. The random variables of interest are the counts  $Y_1, Y_2, \dots, Y_k$  in each of the  $k$  classification categories.

The multinomial distribution, its mean, and its variance are shown in the following box.

### The Multinomial Probability Distribution

$$p(y_1, y_2, \dots, y_k) = \frac{n!}{y_1! y_2! \dots y_k!} (p_1)^{y_1} (p_2)^{y_2} \dots (p_k)^{y_k}$$

where

$p_i$  = Probability of outcome  $i$  on a single trial

$$p_1 + p_2 + \dots + p_k = 1$$

$n = y_1 + y_2 + \dots + y_k$  = Number of trials

$y_i$  = Number of occurrences of outcome  $i$  in  $n$  trials

The mean and variance of the multinomial random variable  $y_i$  are, respectively,

$$\mu_i = np_i \quad \text{and} \quad \sigma_i^2 = np_i(1 - p_i)$$

In R, if  $\mathbf{Y} \sim \text{Multinom}(n, \mathbf{p})$  then  $P(\mathbf{Y} = \mathbf{y}) = \text{dmultinom}(\mathbf{y}, \mathbf{p})$  (bold means vector)

- **Example**
- $P(y_1 = 2, y_2 = 4, y_3 = 4) = \text{dmultinom}(x = c(2, 4, 4), prob = c(0.3, 0.4, 0.3))$
- `> dmultinom(x = c(2,4,4), prob = c(0.3,0.4,0.3))`
- `[1] 0.05878656`



### The Negative Binomial Probability Distribution

The probability distribution for a negative binomial random variable  $Y$  is given by

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r} \quad (y = r, r+1, r+2, \dots)$$

where

$p$  = Probability of success on a single Bernoulli trial

$q = 1 - p$

$y$  = Number of trials until the  $r$ th success is observed

The mean and variance of a negative binomial random variable are, respectively,

$$\mu = \frac{r}{p} \quad \text{and} \quad \sigma^2 = \frac{rq}{p^2}$$

From the box, you can see that the negative binomial probability distribution is a function of two parameters,  $p$  and  $r$ . For the special case  $r = 1$ , the probability distribution of  $Y$  is known as a **geometric probability distribution**.

### The Geometric Probability Distribution

$$p(y) = pq^{y-1} \quad (y = 1, 2, \dots)$$

where

$Y$  = Number of trials until the first success is observed

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

$$\sigma^2 = \frac{q}{p^2}$$

In R:

The geometric distribution with prob =  $p$  has density

$$p(x) = p(1-p)^x$$

for  $x = 0, 1, 2, \dots$ ,  $0 < p \leq 1$ .

So  $y-1 = x$

$$P(Y=3) = P(X=2)$$

# Parameterization for R

## The Negative Binomial Probability Distribution

The probability distribution for a negative binomial random variable  $Y$  is given by

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r} \quad (y = r, r+1, r+2, \dots)$$

**The density in R is**

$$\Gamma(x+n)/(\Gamma(n) x!) p^n (1-p)^x$$

Parameterization for R

$$y-r=x,$$

$$r = n$$

### Example 4.13

#### Negative Binomial Application—Motor Assembly

To attach the housing on a motor, a production line assembler must use an electrical hand tool to set and tighten four bolts. Suppose that the probability of setting and tightening a bolt in any 1-second time interval is  $p = .8$ . If the assembler fails in the first second, the probability of success during the second 1-second interval is .8, and so on.

- Find the probability distribution of  $Y$ , the number of 1-second time intervals until a complete housing is attached.
- Find  $p(6)$ .
- Find the mean and variance of  $Y$ .

Solution

- Since the housing contains  $r = 4$  bolts, we will use the formula for the negative binomial probability distribution. Substituting  $p = .8$  and  $r = 4$  into the formula for  $p(y)$ , we obtain

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r} = \binom{y-1}{3} (.8)^4 (.2)^{y-4}$$

- To find the probability that the complete assembly operation will require  $Y = 6$  seconds, we substitute  $y = 6$  into the formula obtained in part a and find

$$p(y) = \binom{5}{3} (.8)^4 (.2)^2 = (10)(.4096)(.04) = .16384$$

- For this negative binomial distribution,

$$\mu = \frac{r}{p} = \frac{4}{.8} = 5 \text{ seconds}$$

and

$$\sigma^2 = \frac{rq}{p^2} = \frac{4(.2)}{(.8)^2} = 1.25$$

# Parameterization for R

## The Negative Binomial Probability Distribution

The probability distribution for a negative binomial random variable  $Y$  is given by

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r} \quad (y = r, r+1, r+2, \dots)$$

The density in R is

$$\frac{\Gamma(x+n)}{\Gamma(n) x!} p^n (1-p)^x$$

In R  $x$  is the number of fails  
Size the number of successes  
 $Y$  = number of trials =  $x + \text{size}$

## Parameterization for R

$$6-4=2=x,$$

$$4 = n$$

$$P(Y = 6) = \text{dnbinom}(x=2, \text{size} = 4, \text{prob} = 0.8)$$

$$> \text{dnbinom}(x=2, \text{size} = 4, \text{prob} = 0.8)$$

$$[1] 0.16384$$

### Characteristics That Define a Hypergeometric Random Variable

1. The experiment consists of randomly drawing  $n$  elements without replacement from a set of  $N$  elements,  $r$  of which are  $S$ 's (for Success) and  $(N - r)$  of which are  $F$ 's (for Failure).
2. The sample size  $n$  is large relative to the number  $N$  of elements in the population, i.e.,  $n/N > .05$ .
3. The hypergeometric random variable  $Y$  is the number of  $S$ 's in the draw of  $n$  elements.

### The Hypergeometric Probability Distribution

The hypergeometric probability distribution is given by

$$p(y) = \frac{\binom{r}{y} \binom{N-r}{n-y}}{\binom{N}{n}}, \quad y = \begin{matrix} \text{Maximum } [0, n - (N - r)], \dots, \\ \text{Minimum } (r, n) \end{matrix}$$

where

$N$  = Total number of elements

$r$  = Number of  $S$ 's in the  $N$  elements

$n$  = Number of elements drawn

$y$  = Number of  $S$ 's drawn in the  $n$  elements

The mean and variance of a hypergeometric random variable are, respectively,

$$\mu = \frac{nr}{N} \quad \sigma^2 = \frac{r(N-r)n(N-n)}{N^2(N-1)}$$



In R

- In R the probability function is:  
 $p(x) = \frac{\text{choose}(m, x) \text{choose}(n, k-x)}{\text{choose}(m+n, k)}$
- $m=r$
- $x=y$
- $n=N-r$
- $k-x=n-y$
- $m+n=N$
- $k=n$

#### Characteristics That Define a Hypergeometric Random Variable

1. The experiment consists of randomly drawing  $n$  elements without replacement from a set of  $N$  elements,  $r$  of which are  $S$ 's (for Success) and  $(N - r)$  of which are  $F$ 's (for Failure).
2. The sample size  $n$  is large relative to the number  $N$  of elements in the population, i.e.,  $n/N > .05$ .
3. The hypergeometric random variable  $Y$  is the number of  $S$ 's in the draw of  $n$  elements.

#### The Hypergeometric Probability Distribution

The hypergeometric probability distribution is given by

$$p(y) = \frac{\binom{r}{y} \binom{N-r}{n-y}}{\binom{N}{n}}, \quad y = \begin{matrix} \text{Maximum } [0, n - (N - r)], \dots, \\ \text{Minimum } (r, n) \end{matrix}$$

where

$N$  = Total number of elements

$r$  = Number of  $S$ 's in the  $N$  elements

$n$  = Number of elements drawn

$y$  = Number of  $S$ 's drawn in the  $n$  elements

The mean and variance of a hypergeometric random variable are, respectively,

$$\mu = \frac{nr}{N} \quad \sigma^2 = \frac{r(N-r)n(N-n)}{N^2(N-1)}$$

### Example 4.15

#### Hypergeometric Application— EDA Catalyst Selection

An experiment is conducted to select a suitable catalyst for the commercial production of ethylenediamine (EDA), a product used in soaps. Suppose a chemical engineer randomly selects 3 catalysts for testing from among a group of 10 catalysts, 6 of which have low acidity and 4 of which have high acidity.

- a. Find the probability that no highly acidic catalyst is selected.
- b. Find the probability that exactly one highly acidic catalyst is selected.



Let  $Y \sim \text{hyp}(N = 10, r = 4, n = 3)$

Find

$a) P(Y = 0) = \text{dhyper}(x=0, m=4, n=6, k=3)$

`> dhyper(x=0,4,6,3)`

`[1] 0.1666667 (1/6)`

$b) P(Y = 1) = \text{dhyper}(x = 1, m=4, n=6, k=3)$

`> dhyper(x = 1, m=4, n=6, k=3)`

`[1] 0.5 (1/2)`

### Characteristics of a Poisson Random Variable

1. The experiment consists of counting the number of times  $Y$  a particular (rare) event occurs during a given unit of time or in a given area or volume (or weight, distance, or any other unit of measurement).
2. The probability that an event occurs in a given unit of time, area, or volume is the same for all the units. Also, units are mutually exclusive.
3. The number of events that occur in one unit of time, area, or volume is independent of the number that occur in other units.

The formulas for the probability distribution, the mean, and the variance of a Poisson random variable are shown in the next box. You will note that the formula involves the quantity  $e = 2.71828 \dots$ , the base of natural logarithms. Values of  $e^{-y}$ , needed to compute values of  $p(y)$ , are given in Table 3 of Appendix B.

### The Poisson Probability Distribution

The probability distribution\* for a Poisson random variable  $Y$  is given by

$$p(y) = \frac{\lambda^y e^{-\lambda}}{y!} \quad (y = 0, 1, 2, \dots)$$

where

$\lambda$  = Mean number of events during a given unit of time, area, or volume

$e = 2.71828 \dots$

The mean and variance of a Poisson random variable are, respectively,

$$\mu = \lambda \quad \text{and} \quad \sigma^2 = \lambda$$

### Example 4.18

#### Poisson Application—Cracks in Concrete

Suppose the number  $Y$  of cracks per concrete specimen for a particular type of cement mix has approximately a Poisson probability distribution. Furthermore, assume that the average number of cracks per specimen is 2.5.

- Find the mean and standard deviation of  $Y$ , the number of cracks per concrete specimen.
- Find the probability that a randomly selected concrete specimen has exactly five cracks.
- Find the probability that a randomly selected concrete specimen has two or more cracks.
- Find  $P(\mu - 2\sigma < Y < \mu + 2\sigma)$ . Does the result agree with the Empirical Rule?

We will do b) and c)  $\lambda = 2.5 \text{ cracks/}$   
*specimen*

- b)  $P(Y = 5) = \text{dpois}(5, 2.5)$

```
> dpois(5, 2.5)
```

```
[1] 0.06680094
```

- c)  $P(Y \geq 2) = 1 - P(Y \leq 1)$

```
> 1 - ppois(1, 2.5)
```

```
[1] 0.7127025
```

```
> ppois(1, 2.5, lower.tail = FALSE) ← Don't do it this way!  $P(Y > 1)$ 
```

```
[1] 0.7127025
```

# Moment generating functions

**Definition 4.7**

The  $k$ th moment of a random variable  $Y$ , taken about the origin, is denoted by the symbol  $\mu'_k$  and defined to be

$$\mu'_k = E(Y^k) \quad (k = 1, 2, \dots)$$

**Definition 4.8**

The  $k$ th moment of a random variable  $Y$ , taken about its mean, is denoted by the symbol  $\mu_k$  and defined to be

$$\mu_k = E[(Y - \mu)^k]$$

**Definition 4.9**

The **moment generating function**,  $m(t)$ , of a discrete random variable  $Y$  is defined to be

$$m(t) = E(e^{tY})$$

### THEOREM 4.5

If  $m(t)$  exists, then the  $k$ th moment about the origin is equal to

$$\mu'_k = \left. \frac{d^k m(t)}{dt^k} \right]_{t=0}$$



### Example 4.21

MGF for a Binomial Random Variable

Solution

Derive the moment generating function for a binomial random variable  $Y$ .

The moment generating function is given by

$$m(t) = E(e^{tY}) = \sum_{y=0}^n e^{ty} p(y) = \sum_{y=0}^n e^{ty} \binom{n}{y} p^y q^{n-y} = \sum_{y=0}^n \binom{n}{y} (pe^t)^y q^{n-y}$$

We now recall the binomial theorem (see Exercise 4.36, p. 154).

$$(a + b)^n = \sum_{y=0}^n \binom{n}{y} a^y b^{n-y}$$

Letting  $a = pe^t$  and  $b = q$  yields the desired result:

$$m(t) = (pe^t + q)^n$$

### Example 4.22

First Two Moments for a  
Binomial Random Variable

Solution

Use Theorem 4.5 to derive  $\mu'_1 = \mu$  and  $\mu'_2$  for the binomial random variable.

From Theorem 4.5,

$$\begin{aligned}\mu'_1 = \mu &= \left. \frac{dm(t)}{dt} \right]_{t=0} = n(pe^t + q)^{n-1}(pe^t) \Big]_{t=0} \\ &= n(pe^0 + q)^{n-1}(pe^0)\end{aligned}$$

But  $e^0 = 1$ . Therefore,

$$\mu'_1 = \mu = n(p + q)^{n-1}p = n(1)^{n-1}p = np$$

Similarly,

$$\begin{aligned}\mu'_2 &= \left. \frac{d^2m(t)}{dt^2} \right]_{t=0} = np \frac{d}{dt} [e^t(pe^t + q)^{n-1}] \Big]_{t=0} \\ &= np[e^t(n-1)(pe^t + q)^{n-2}pe^t + (pe^t + q)^{n-1}e^t] \Big]_{t=0} \\ &= np[(1)(n-1)(1)p + (1)(1)] = np[(n-1)p + 1] \\ &= np(np - p + 1) = np(np + q) = n^2p^2 + npq\end{aligned}$$

### Example 4.23

Using Moments to Derive  
the Variance of a Binomial  
Random Variable

Solution

Use the results of Example 4.22, in conjunction with Theorem 4.4, to derive the variance of a binomial random variable.

By Theorem 4.4,

$$\sigma^2 = E(Y^2) - \mu^2 = \mu'_2 - (\mu'_1)^2$$

Substituting the values of  $\mu'_2$  and  $\mu'_1 = \mu$  from Example 4.22 yields

$$\sigma^2 = n^2 p^2 + npq - (np)^2 = npq$$

# Quick Review

## Key Terms

*Note: Starred (\*) terms are from the optional section in this chapter.*

Bernoulli distribution 174

Bernoulli random variable 146

Bernoulli trials 146

Binomial distribution 159

Binomial experiment 148

Binomial random variable 174

Discrete random variable 135

Expected value 140

Geometric distribution 163

Hypergeometric distribution 167

Hypergeometric random variable 164

Mean 140

\*Moments 175

\*Moment generating function 175

Multinomial distribution 155

Multinomial experiment 155

Negative binomial distribution 159

Poisson distribution 159

Poisson random variable 169

Probability distribution 135

Random variable 134

Sampling with replacement 164

Standard deviation 141

Variance 141

## Key Formulas

*Note: Starred (\*) formulas are from the optional section in this chapter.*

Random Variable	$p(y)$	$\mu$	$\sigma^2$	* $m(t)$
Discrete (general)	$p(y)$	$E(Y) = \sum yp(y)$	$E(Y^2) - \mu^2$	
Bernoulli	$p(y) = p^y q^{1-y}$ where $q = 1 - p$ , $y = 0, 1$	$p$	$pq$	$pe^t + q$
Binomial	$p(y) = \binom{n}{y} p^y q^{n-y}$ where $q = 1 - p$ , $y = 0, 1, \dots, n$	$np$	$npq$	$(pe^t + q)^n$

Random Variable	$p(y)$	$\mu$	$\sigma^2$	$*m(t)$
Hypergeometric	$p(y) = \frac{\binom{r}{y} \binom{N-r}{n-y}}{\binom{N}{n}}$	$\frac{nr}{N}$	$\frac{r(N-r)n(N-n)}{N^2(N-1)}$	Not given
Poisson	$p(y) = \frac{\lambda^y e^{-\lambda}}{y!} \quad y = 1, 2, \dots$	$\lambda$	$\lambda$	$e^{\lambda(e^t-1)}$
Geometric	$p(y) = p(1-p)^{y-1} \quad y = 1, 2, \dots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Negative binomial	$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r} \quad y = r, r+1, \dots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left( \frac{pe^t}{1-(1-p)e^t} \right)^r$
Multinomial	$p(y_1, y_2, \dots, y_k) = \frac{n!}{y_1! y_2! \dots y_k!} (p_1)^{y_1} (p_2)^{y_2} \dots (p_k)^{y_k}$	$np_i$	$np_i(1-p_i)$	Not given

## LANGUAGE LAB

Symbol	Pronunciation	Description
$p(y)$		Probability distribution of the random variable $Y$
$E(Y)$	Expected value of $Y$	Mean of the probability distribution for $Y$
$S$		The outcome of a Bernoulli trial denoted a “success”
$F$		The outcome of a Bernoulli trial denoted a “failure”
$p$		The probability of success ( $S$ ) in a Bernoulli trial
$q$		The probability of failure ( $F$ ) in a Bernoulli trial, where $q = 1 - p$
$\lambda$	lambda	The mean (or expected) number of events for a Poisson random variable
$e$		A constant used in the Poisson probability distribution, where $e = 2.71828 \dots$
$m(t)$	“m” of “t”	Moment generating function

## Chapter Summary Notes

- A **discrete random variable** can assume only a countable number of values.
- Requirements for a discrete probability distribution:  $p(y) \geq 0$  and  $\sum p(y) = 1$
- Probability models for discrete random variables: **Bernoulli**, **binomial**, **multinomial**, **negative binomial**, **geometric**, **hypergeometric**, and **Poisson**
- Characteristics of a **Bernoulli random variable**: (1) two mutually exclusive outcomes,  $S$  and  $F$ , in a trial, (2) outcomes are exhaustive, (3)  $P(S) = p$  and  $P(F) = q$ , where  $p + q = 1$
- Characteristics of a **binomial random variable**: (1)  $n$  identical trials, (2) two possible outcomes,  $S$  and  $F$ , per trial, (3)  $P(S) = p$  and  $P(F) = q$  remain the same from trial to trial, (4) trials are independent, (5)  $Y$  = number of  $S$ 's in  $n$  trials
- Characteristics of a **multinomial random variable**: (1)  $n$  identical trials, (2)  $k$  possible outcomes per trial, (3) probabilities of  $k$  outcomes remain the same from trial to trial, (4) trials are independent, (5)  $Y_1, Y_2, \dots, Y_k$  are counts of outcomes in  $k$  categories



# Example of skills

4.53 *Distribution of slugs.* The distributional pattern of pulmonate slugs inhabiting Libya was studied in the *AIUB Journal of Science and Engineering* (Aug. 2003). The number of slugs of a certain species found in the survey area was modeled using the negative binomial distribution. Assume that the probability of observing a slug of a certain species (say, *Milax rusticus*) in the survey area is .2. Let  $Y$  represent the number of slugs that must be collected in order to obtain a sample of 10 *Milax rusticus* slugs.

- Give the probability distribution for  $Y$  as a formula.
- What is the expected value of  $Y$ ? Interpret this value.
- Find  $P(Y = 25)$ .

# SKILL

## Categorize the problem!!

### The Negative Binomial Probability Distribution

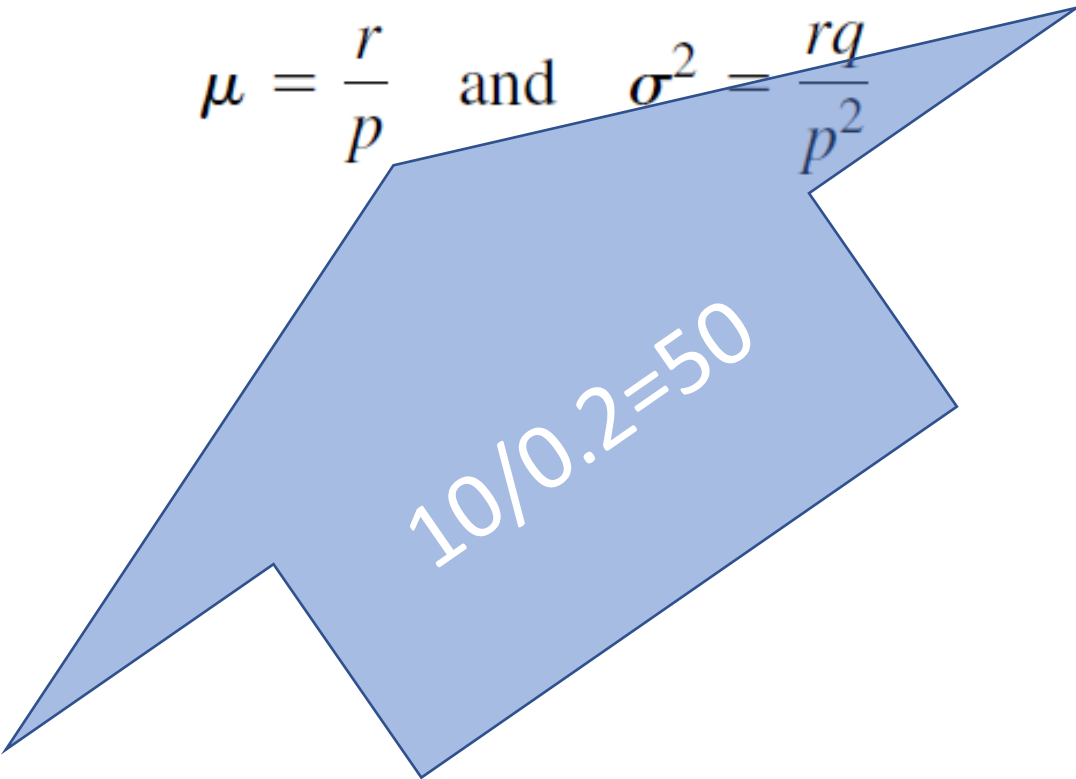
The probability distribution for a negative binomial random variable  $Y$  is given by

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r} \quad (y = r, r+1, r+2, \dots)$$

The expected value is  $\mu$

The mean and variance of a negative binomial random variable are, respectively,

$$\mu = \frac{r}{p} \quad \text{and} \quad \sigma^2 = \frac{rq}{p^2}$$


$$10/0.2=50$$

The expected number of slugs to be collected in order to observe 10 *Milax Rusticus* slugs is 50

# SKILL

$$P(Y = 25)$$

Since we need a single probability we should use a dstem() function namely dnbinom()

The next problem to overcome is to **reparameterize**.

In R we read:

The negative binomial distribution with size = n and prob = p has density

$$\frac{\Gamma(x+n)}{\Gamma(n) x!} p^n (1-p)^x$$

$$\mathbf{n=r, x=y-r}$$

for  $x = 0, 1, 2, \dots$ ,  $n > 0$  and  $0 < p \leq 1$ .


This represents the number of failures which occur in a sequence of Bernoulli trials before a target number of successes is reached.

# The calculation in R


```
> dnbinom(x=25-10,size = 10,prob=0.2)  
[1] 0.00471078
```

## Using the book formula


$$P(Y = 25) = \binom{25-1}{10-1} (0.2)^{10} (0.8)^{25-10} = \frac{24!}{9!15!} (0.2)^{10} (0.8)^{15} = 0.0047$$




Account



Dashboard




Courses




Calendar


392




Inbox




History



Commons



Help



Fall 2022

[Home](#)

[Announcements](#)

[Assignments](#)

[Discussions](#)

[Grades](#)

[People](#)

[Pages](#)

[Files](#)

[Syllabus](#)

[Quizzes](#)

[Modules](#)

[Gradescope](#)

[Multi-Tool](#)

[Rubrics](#)

$$P(3 \leq Y \leq 7)$$

0.82	<a href="#">20 respondents</a>	80 %	<div></div>	80% answered correctly
Something Else	<a href="#">5 respondents</a>	20 %	<div></div>	

Attempts: 24 out of 25

$$P(4 \leq Y < 8)$$

0.78	<a href="#">17 respondents</a>	68 %	<div></div>	68% answered correctly
Something Else	<a href="#">7 respondents</a>	28 %	<div></div>	
No Answer	<a href="#">1 respondent</a>	4 %	<div></div>	