Module 4

This is a single, concatenated file, suitable for printing or saving as a PDF for offline viewing. Please note that some animations or images may not work.

Module 4 Study Guide and Deliverables

Topics:	Distributions
Readings:	• Lecture material
Assignments:	• Assignment 4 due **Tuesday, June 6 at 6:00 AM ET**
Assessments:	• Quiz 4 due **Tuesday, June 6 at 6:00 AM ET**
Live Classrooms:	• **Wednesday, May 31 from 7:00-9:00 PM ET** • One-hour open office with facilitator. Day and time will be provided by your facilitator. • Live sessions will be recorded.

Discrete Distributions

Introduction

Numeric variables are either discrete or continuous. The outcomes of a continuous numeric variable come from a measuring process while those of a discrete variable come from a counting process. In this lecture, the probability distributions of discrete variables are explored. A probability distribution for a discrete random variable is defined in terms of a list of all possible numerical outcomes along with the probability of each outcome. The list of all outcomes is known as the support of the random variable. The probability distribution is specified in terms of a function, the probability mass function that maps each outcome to its probability.

The support of a discrete random variable \(X\), \(S_X\), is a finite set or a countably infinite set. The probability ete random variable \(X\) is denoted by:

Loading [Contrib]/a11y/accessibility-menu.js

$$f_X:S_X\to[0,1]$$

and defined as:

$$f_X(x)=P(X=x), x \in S_X$$

The *mean* or the expected value of the probability distribution of the discrete random variable \(X\) is:

$$\mu=E[X]=\sum_{x\in S}x\cdot f_x(x)$$

The variance of the discrete distribution is computed as:

$$\sigma^2=\sum_{x\in S}(x-\mu)^2f_x(x)$$

or, through the alternate formula:

$$\sigma^2=\sum_{x\in S}x^2f_x(x)-\mu^2$$

The standard deviation of the distribution is computed as the square root of the variance.

The cumulative distribution function (CDF) of the random variable \(X\) is defined as:

$$F_X(x)=P(X\le x), -\infty\lt x\lt\infty$$

Random Variable Example—Number of *Heads*

Let the random variable \(X\) be the number of *heads* observed when a fair coin is tossed three times. The support for \(X\) is \(S_X=\{0,1,2,3\}\). For computing the probabilities, examine the sample space for the experiment:

\(S=\{HHH,HHT,HTH,HTT,THH,THT,TTH,TTT\}\)

Out of the eight outcomes, only one outcome (TTT) has the number of *heads* \(0\); three outcomes \((HTT, THT, TTH)\) have the number of *heads* \(1\); three outcomes \((HHT, HTH, THH)\) have the number of *heads* \(2\); and one outcome \((HHH)\) has the number of *heads* \(3\).

The probability mass function for the number of *heads* is shown in the following table.

\(x\in S_X\)	**\(f_X(x)=P[X=x]\)**
0	\(\frac{1}{8}\)
1	\(\frac{3}{8}\)
2	\(\frac{3}{8}\)

3	$\frac{1}{8}$
Total	**1**

The mean value of the above discrete random variable is:

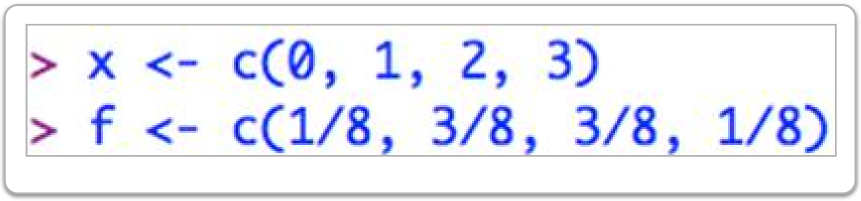
$$\mu=\sum^3_{x=0}x\cdot f_X(x)=0\left(\frac{1}{8}\right)+1\left(\frac{3}{8}\right)+2\left(\frac{3}{8}\right)+3\left(\frac{1}{8}\right)=\frac{12}{8}=1.5$$

Similarly, the variance and the standard deviation of the random variable can be calculated.

The cumulative distribution function (CDF) of the above random variable is shown in the following table.

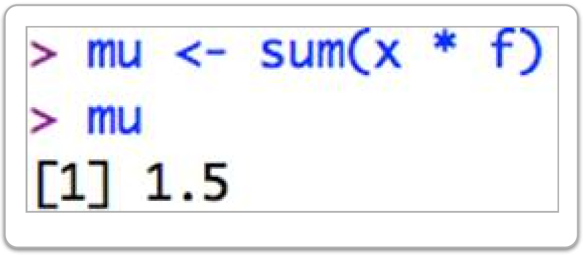
$x\in S_X$	$f_X(x)$	$F_X(x)=P[X\le x]$
0	$\frac{1}{8}$	$\frac{1}{8}$
1	$\frac{3}{8}$	$\frac{1}{8}+\frac{3}{8}=\frac{4}{8}$
2	$\frac{3}{8}$	$\frac{4}{8}+\frac{3}{8}=\frac{7}{8}$
3	$\frac{1}{8}$	$\frac{7}{8}+\frac{1}{8}=1$

The above problem can be modeled in *R* as follows. The support set for the random variable and the probabilities for this distribution are:



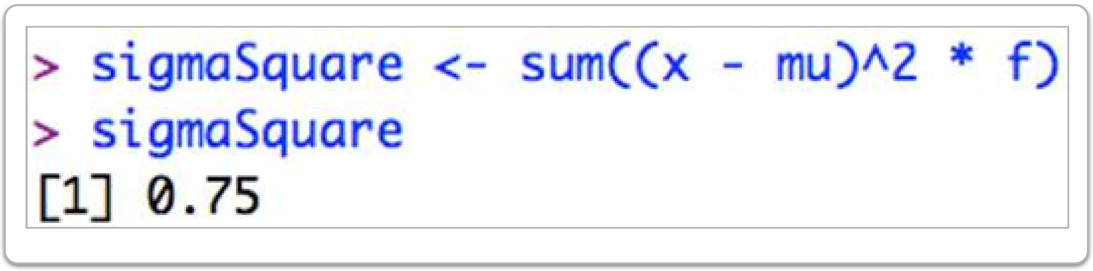
```
> x <- c(0, 1, 2, 3)
> f <- c(1/8, 3/8, 3/8, 1/8)
```

The mean of the distribution is:



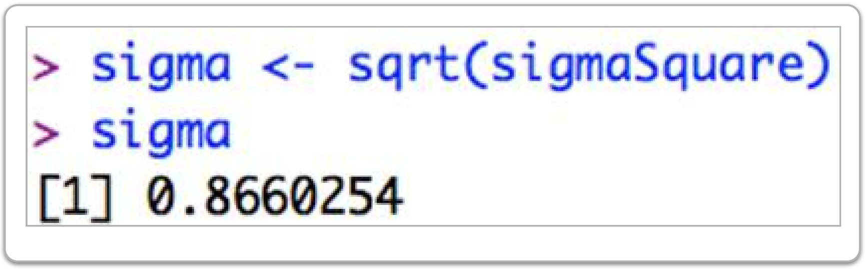
```
> mu <- sum(x * f)
> mu
[1] 1.5
```

Loading [Contrib]/a11y/accessibility-menu.js is:



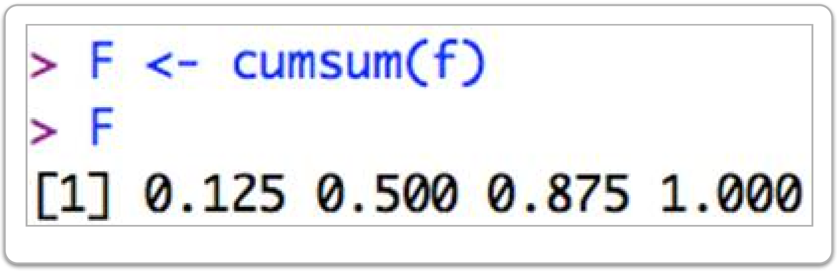
```
> sigmaSquare <- sum((x - mu)^2 * f)
> sigmaSquare
[1] 0.75
```

The standard deviation of the distribution is:



```
> sigma <- sqrt(sigmaSquare)
> sigma
[1] 0.8660254
```

The cumulative distribution for the random variable is:



```
> F <- cumsum(f)
> F
[1] 0.125 0.500 0.875 1.000
```

Random Variable Example—Age of Students

Suppose the ages of ten students enrolled in a class are \(21, 25, 27, 23, 21, 21, 25, 25, 21\), and \(27\). Let the random variable \(X\) be the age of a randomly selected student in the class. The frequency distribution of \(X\) is shown in the following table:

Age \((x)\)	Count
21	4
23	1
25	3

Loading [Contrib]/a11y/accessibility-menu.js

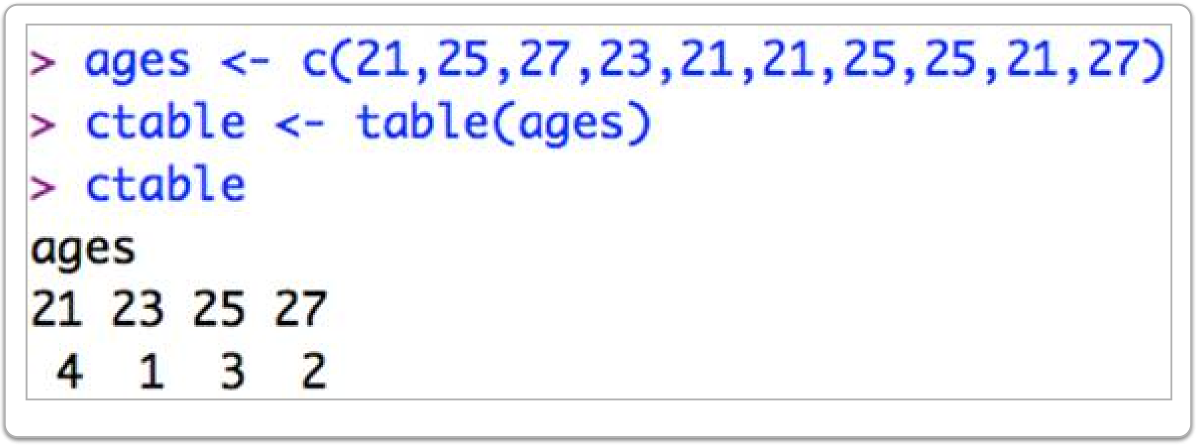
Total	**10**

The support for the random variable X is the set $S_X=\{21,23,25,27\}$. The probability distribution of X is:

Age (x)	**$P(X=x)$**
21	4/10 = 0.4
23	1/10 = 0.1
25	3/10 = 0.3
27	2/10 = 0.2
Total	**1**

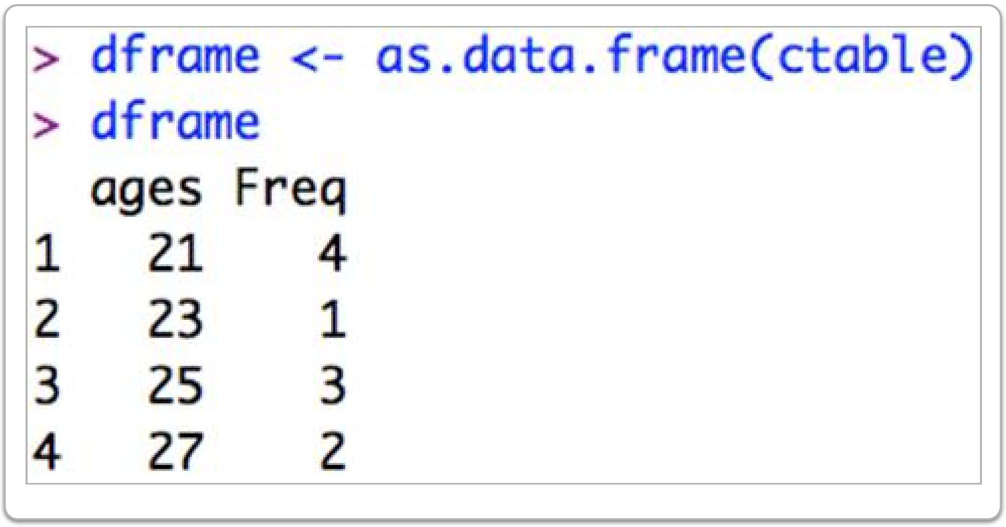
The mean of the above random variable is $21\cdot 0.4+23\cdot 0.1+25\cdot 0.3+27\cdot 0.2=23.6$

The above data can be modeled using *R* as follows: (The `ages` variable is initialized with the given data and passed as an argument to the `table` function.)



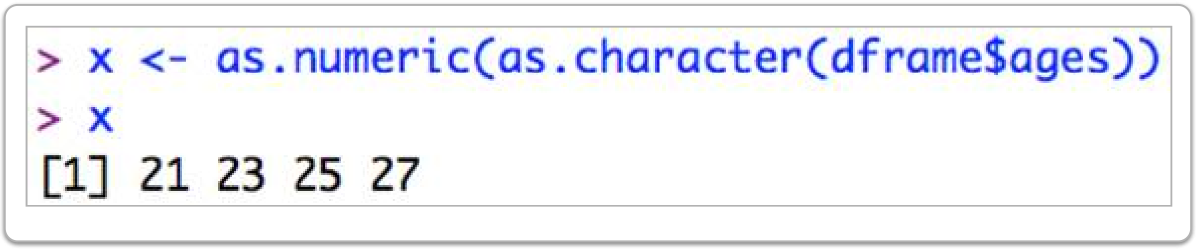
```
> ages <- c(21,25,27,23,21,21,25,25,21,27)
> ctable <- table(ages)
> ctable
ages
21 23 25 27
 4  1  3  2
```

The above contingency table is converted to a data frame as follows:



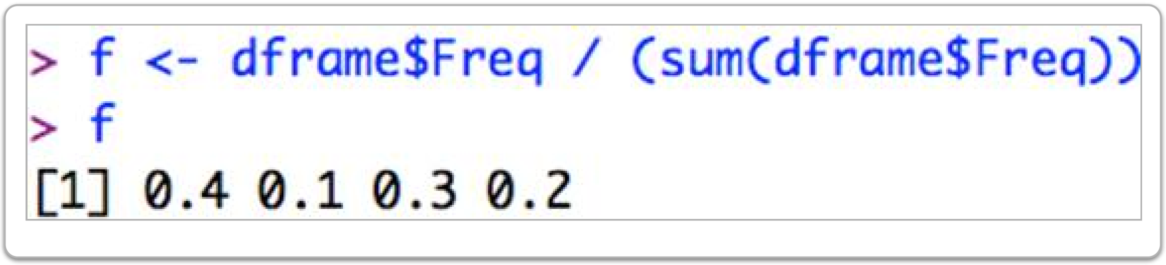
```
> dframe <- as.data.frame(ctable)
> dframe
  ages Freq
1   21    4
2   23    1
3   25    3
4   27    2
```

The inputs for the random variable x are the distinct ages:



```
> x <- as.numeric(as.character(dframe$ages))
> x
[1] 21 23 25 27
```

Similarly, the probability distribution for the random variable is:



```
> f <- dframe$Freq / (sum(dframe$Freq))
> f
[1] 0.4 0.1 0.3 0.2
```

Discrete Uniform Distribution

A discrete uniform distribution is a symmetric probability distribution where a finite number of input values are equally likely. Consider the rolling of a single die. The outcomes $1,2,\ldots ,6$ are equally likely, each occurring with a probability of $1/6$. If a discrete random variable X has m input values $1,2,\ldots ,m$, then X has the discrete uniform distribution when $P(X=x)=\frac{1}{m}$, for all values of x from 1 to m.

The probability mass function (PMF) of a random variable X with discrete uniform distribution is:

$$f_X(x)=\frac{1}{m} x=1,2,\ldots,m$$

Notation

```
X ~ disunif(m)
```

The cumulative distribution function (CDF) of \(X,F_X(x)\), for a given value of \(x\) is the probability \(P(X\le x)\). For uniform distribution,

$$F_X(x)=\frac{x}{m},x=1,2,\ldots,m$$

The *mean* of the discrete uniform distribution is:

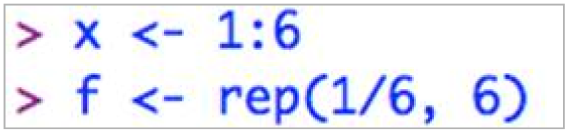
$$\begin{align}\mu&=\sum^m_{x=1}xf_X(x)\\&=\sum^m_{x=1}x\cdot\frac{1}{m}\\&=\frac{1}{m}(1+2+\ldots +m)\\&=\frac{1}{m}\cdot \frac{m(m+1)}{2}\\&=\frac{m+1}{2}\end{align}$$

and the variance of the discrete uniform distribution is:

$$\sigma^2=\sum x^2f_X(x)-\mu^2=\frac{m^2-1}{12}$$

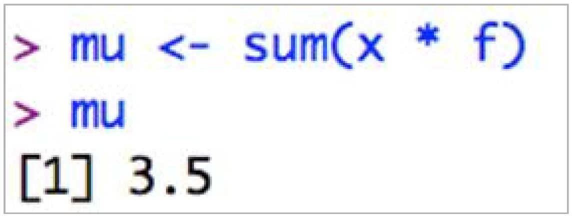
In the case of the single die, \(m=6\), hence the mean is \(3.5\) and the variance is \(2.917\).

The values can be explicitly computed in *R* as follows: (The input sequence is \(1..6\) and each of the \(6\) probabilities are \(1/6\).)



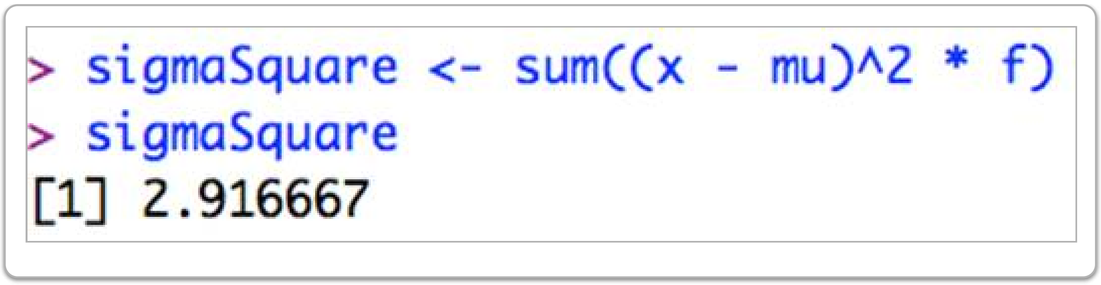
```
> x <- 1:6
> f <- rep(1/6, 6)
```

The mean value is calculated as follows:



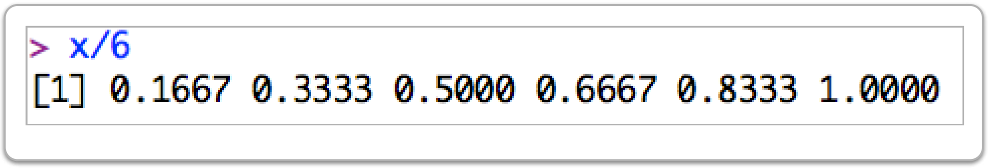
```
> mu <- sum(x * f)
> mu
[1] 3.5
```

The variance is computed using the following formula



```
> sigmaSquare <- sum((x - mu)^2 * f)
> sigmaSquare
[1] 2.916667
```

The cumulative distribution function of the above discrete uniform random variable is calculated as follows:



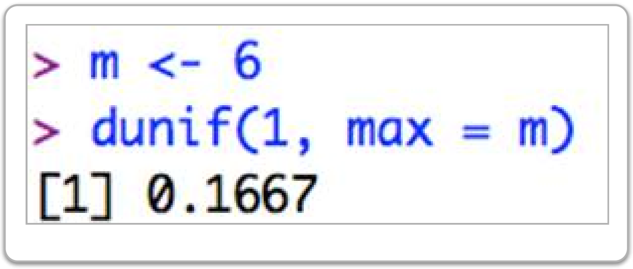
```
> x/6
[1] 0.1667 0.3333 0.5000 0.6667 0.8333 1.0000
```

Sometimes, the input values for the discrete uniform random variable are in the integer range $[a,b]$, inclusive where $a < b$. In this case, the probability mass function is:

$$f_X(x)=\frac{1}{b-a+1},x=a,\ldots,b$$

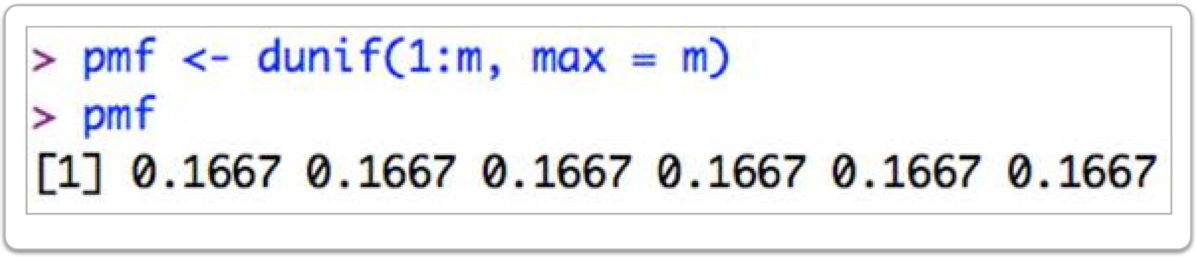
The *R* functions `dunif` and `punif` provide the probability mass function and the cumulative distribution function for a discrete uniform distribution.

The rolling die example can be modeled in *R* as follows. The probability, $P(X=1)$, is:



```
> m <- 6
> dunif(1, max = m)
[1] 0.1667
```

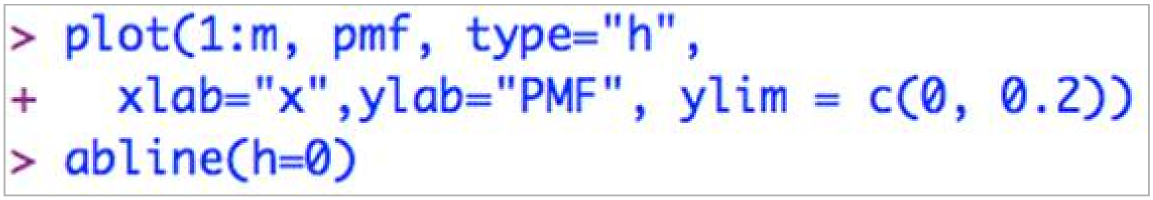
The PMF can be generated as shown below.



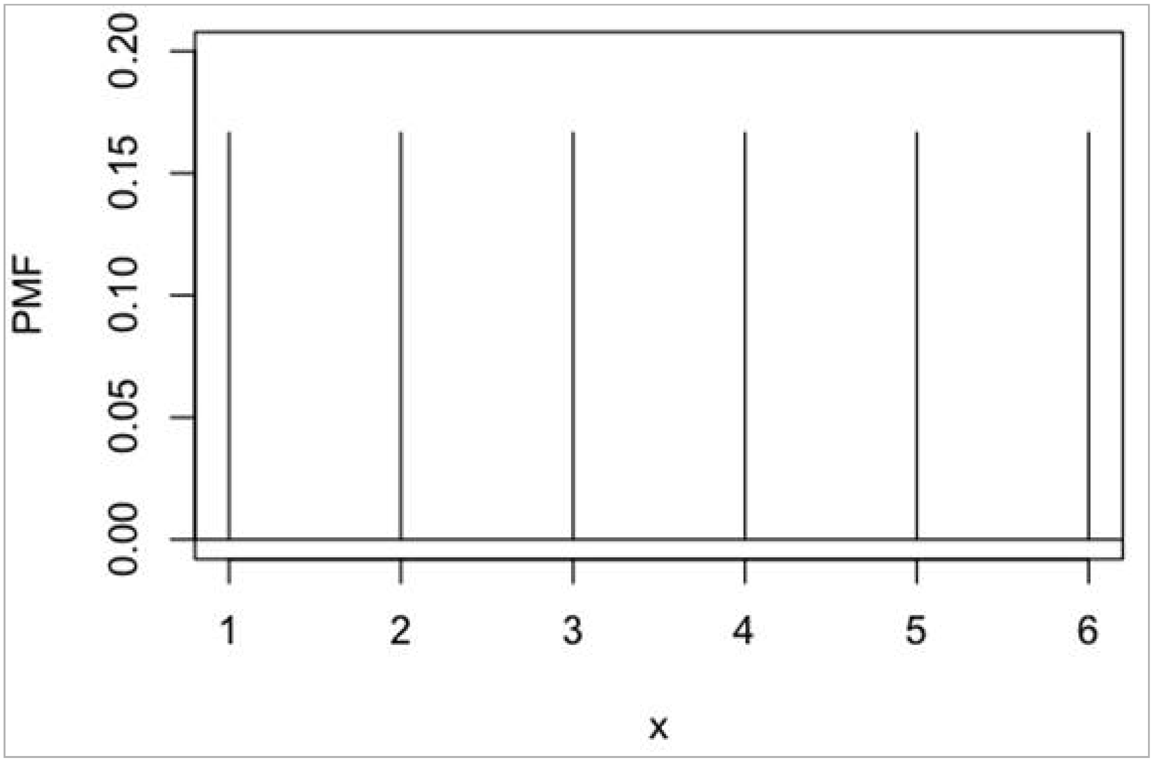
```
> pmf <- dunif(1:m, max = m)
> pmf
[1] 0.1667 0.1667 0.1667 0.1667 0.1667 0.1667
```

A plot of the PMF for the uniform distribution using the following code is shown below.

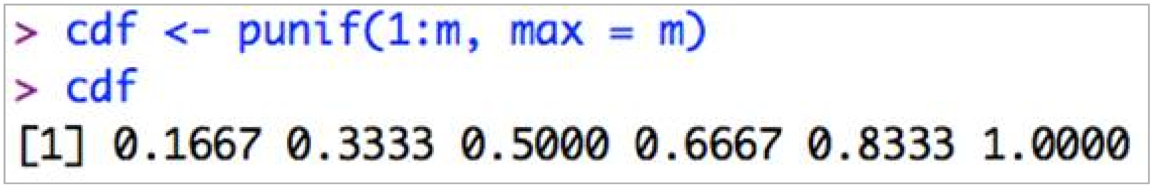
Loading [Contrib]/a11y/accessibility-menu.js



```
> plot(1:m, pmf, type="h",
+   xlab="x",ylab="PMF", ylim = c(0, 0.2))
> abline(h=0)
```

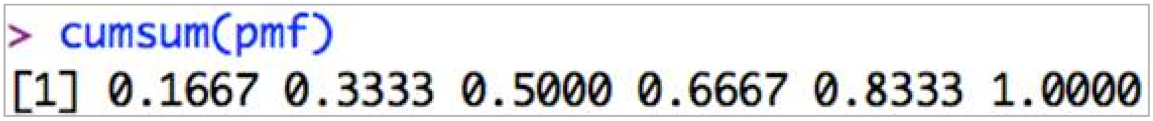


The CDF for the above distribution is:



```
> cdf <- punif(1:m, max = m)
> cdf
[1] 0.1667 0.3333 0.5000 0.6667 0.8333 1.0000
```

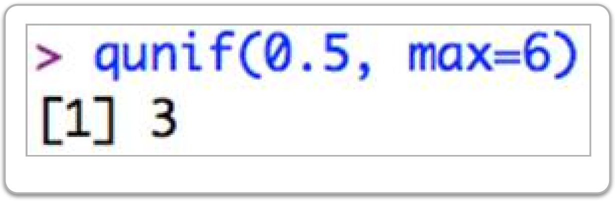
The CDF can also be calculated using the PMF as follows:



```
> cumsum(pmf)
[1] 0.1667 0.3333 0.5000 0.6667 0.8333 1.0000
```

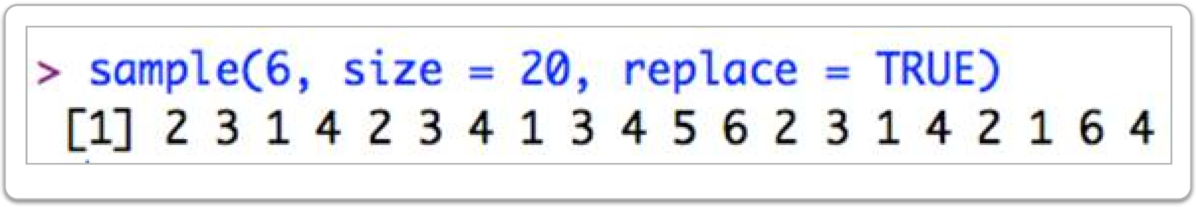
Loading [Contrib]/a11y/accessibility-menu.js

The qunif function is the quantile function that is the inverse of the punif function, i.e., it returns the value for which the CDF is given.



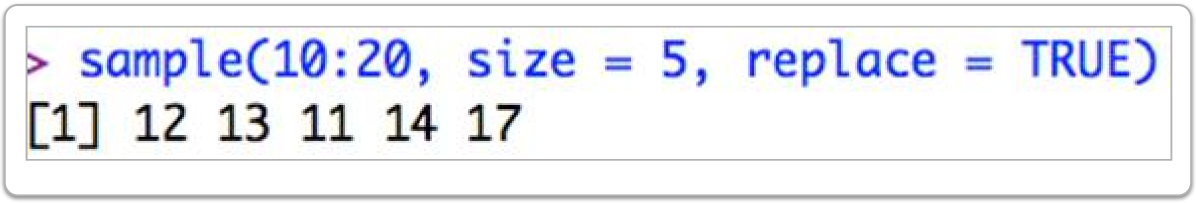
```
> qunif(0.5, max=6)
[1] 3
```

The data for a discrete uniform random variable can be generated in *R* using the sample function with the replace option set to TRUE. The following code generates the data for a fair die rolled twenty times:



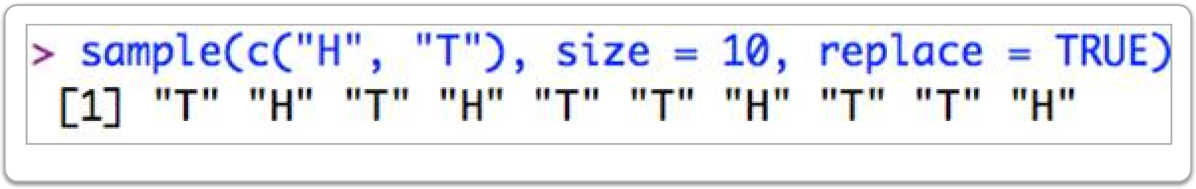
```
> sample(6, size = 20, replace = TRUE)
 [1] 2 3 1 4 2 3 4 1 3 4 5 6 2 3 1 4 2 1 6 4
```

The following code generates five random numbers in the given range from \(10\) to \(20\):



```
> sample(10:20, size = 5, replace = TRUE)
[1] 12 13 11 14 17
```

The following code generates ten outcomes from flipping a fair coin:



```
> sample(c("H", "T"), size = 10, replace = TRUE)
 [1] "T" "H" "T" "H" "T" "T" "H" "T" "T" "H"
```

Binomial Distribution Background

The binomial distribution is based on binomial coefficients and Bernoulli trials. A Bernoulli trial is a random experiment with only two possible outcomes, success and failure.

Loading [Contrib]/a11y/accessibility-menu.js

fficients

The binomial coefficient \(\binom{n}{x}\) is defined as:

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

where \(n\) is a positive integer, and \(x\) is a nonnegative integer less than or equal to \(n\).

The binomial coefficients appear in the binomial expansion of \((a+b)^n\):

$$(a+b)^n=\sum^n_{x=0}\binom{n}{x}a^{n-x}b^x$$

For example:

$$\begin{align}(a+b)^4&=\sum^4_{x=0}\binom{4}{x}a^{n-x}b^x\\&=\binom{4}{0}a^4+\binom{4}{1}a^3b+\binom{4}{2}a^2b^2+\binom{4}{3}ab^3+\binom{4}{4}b^4\\&=a^4+4a^3b+6a^2b^2+4ab^3+b^4\end{align}$$

Bernoulli Trials

A Bernoulli trial is a random experiment in which there are only two possible outcomes—success (\(S\)) and failure (\(F\)). The random variable \(X\) in a Bernoulli trial is defined as follows:

\(X=1\), if the outcome is a success (\(S\)), and \(X=0\), if the outcome is a failure (\(F\)).

Let the probability of success be \(p\). Then the probability of failure is \((1 – p)\).

The probability mass function of \(X\) is:

\(f_X(x)=p^x(1-p)^{1-x}\), for \(x=0\) or \(1\)

The mean is:

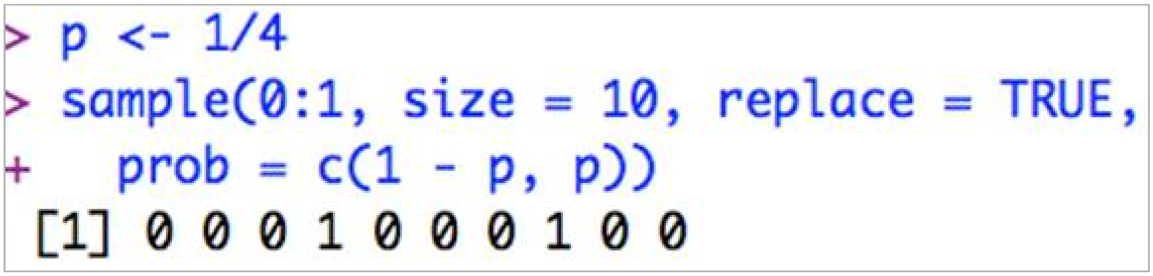
$$\mu=\sum xf_X(x)=p$$

The variance is:

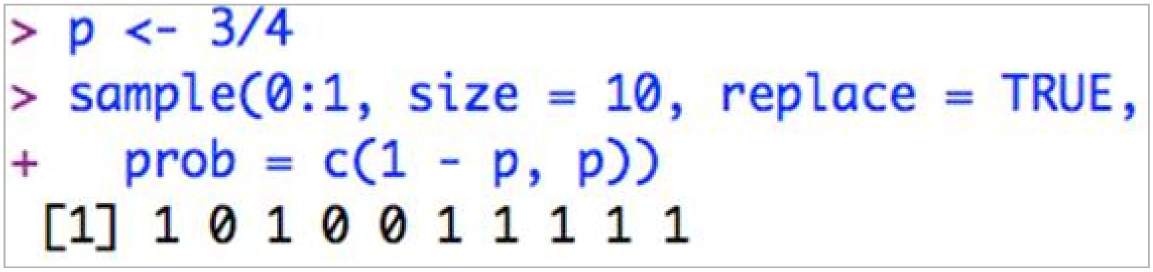
$$\sigma^2=\sum x^2f_X(x)-\mu^2=p-p^2=p(1-p)$$

Repeated trials of an experiment are called *Bernoulli trials*. The trials are independent, with each trial having two possible outcomes (success and failure), and the probability of success remains the same from trial to trial.

Using *R*, the `sample()` function can be used to draw the samples with the specified distribution. The first argument is the list of outcomes \(0,1\). Since the first outcome in this list is the negative outcome, its probability is \(1–p\). The following examples show ten random samples with \(p=\frac{1}{4}\) and \(p=\frac{3}{4}\), respectively:



```
> p <- 1/4
> sample(0:1, size = 10, replace = TRUE,
+   prob = c(1 - p, p))
 [1] 0 0 0 1 0 0 0 1 0 0
```



```
> p <- 3/4
> sample(0:1, size = 10, replace = TRUE,
+   prob = c(1 - p, p))
 [1] 1 0 1 0 0 1 1 1 1 1
```

Binomial Distribution

The *binomial distribution* is the probability distribution for the number of successes in a sequence of Bernoulli trials. The number of outcomes that contain x successes out of n Bernoulli trails is the binomial coefficient $\binom{n}{x}$. A binomial random variable X counts the number of successes in n Bernoulli trials. The distribution of X is described by the two parameters, n (the number of trials) and p (the success probability).

Notation

```
X ~ binom(size = n, prob = p)
```

The probability mass function, $f_X(x)$, of the binomial random variable X is:

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

The sum of all the probability values in the distribution is:

$$\sum^n_{x=0}\binom{n}{x}p^x(1-p)^{n-x}=[(1-p)+p]^n=1^n=1$$

The mean of the binomial distribution is:

$$\mu=\sum xf_X(x)=\sum^n_{x=0}x\binom{n}{x}p^x(1-p)^{n-x}=np$$

Loading [Contrib]/a11y/accessibility-menu.js

The variance of the binomial distribution is:

$$\sigma^2=\sum x^2f_X(x)-\mu^2=np(1-p)$$

Bernoulli Trials—Coin Toss Example

A coin is tossed three times. Let success be the outcome of getting a *head* for each toss. Given that the coin is unfair and has an \(80\%\) chance of landing on *heads,* the probability of success, \(p\), is \(0.8\). The probability of failure is \(0.2\). For each toss, let \(s\) denote the success outcome and \(f\) denote the failure outcome. The possible outcomes for three tosses and their corresponding probabilities are shown in the table below:

Outcome	Probability
\(\text{sss}\)	\(0.8\cdot0.8\cdot0.8 = 0.512\)
\(\text{ssf}\)	\(0.8\cdot0.8\cdot0.2 = 0.128\)
\(\text{sfs}\)	\(0.8\cdot0.2\cdot0.8 = 0.128\)
\(\text{sff}\)	\(0.8\cdot0.2\cdot0.2 = 0.032\)
\(\text{fss}\)	\(0.2\cdot0.8\cdot0.8 = 0.128\)
\(\text{fsf}\)	\(0.2\cdot0.8\cdot0.2 = 0.032\)
\(\text{ffs}\)	\(0.2\cdot0.2\cdot0.8 = 0.032\)
\(\text{fff}\)	\(0.2\cdot0.2\cdot0.2 = 0.008\)

Let the random variable \(X\) be the number of *heads*. \(X\) is a binomial variable with size = \(3\), and prob = \(0.8\). The probability distribution of \(X\) is:

\(P(X=0)=P(\text{fff})=0.008\)
\(P(X=1)=P(\text{sff})+P(\text{fsf})+P(\text{ffs})=0.032+0.032+0.032=0.096\)
\(P(X=2)=P(\text{ssf})+P(\text{sfs})+P(\text{fss})=0.128+0.128+0.128=0.384\)
\(P(X=3)=P(\text{sss})=0.512\)

Using the binomial coefficients formula for the probability mass function, the values correspond to the ones explicitly computed.

$$f_X(0)=\binom{3}{0}(0.8)^0(0.2)^3=1\cdot 1\cdot 0.008=0.008$$

$$f_X(1)=\binom{3}{1}(0.8)^1(0.2)^2=3\cdot 0.8\cdot 0.04=0.096$$

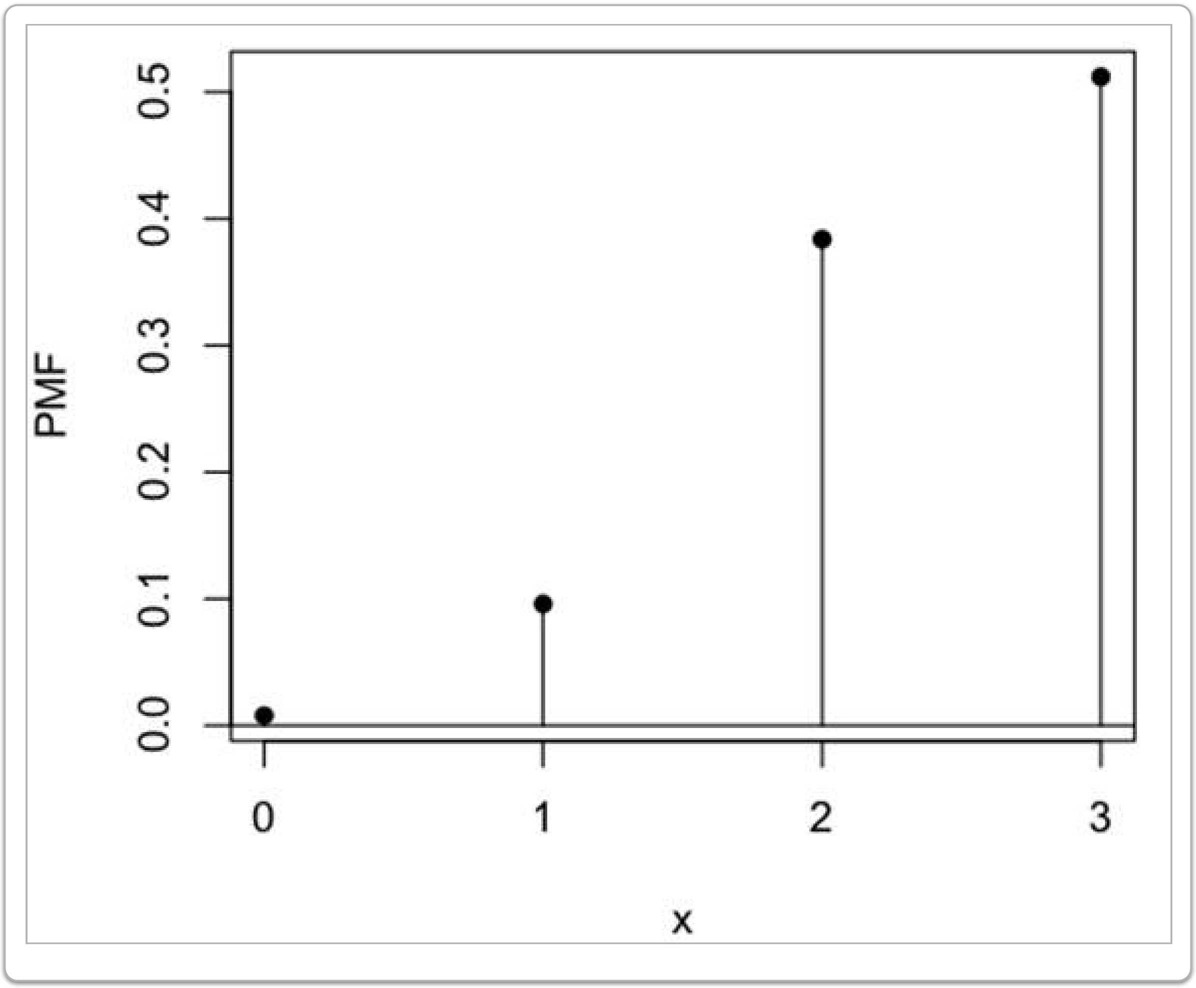
$$f_X(2)=\binom{3}{2}(0.8)^2(0.2)^1=3\cdot 0.64\cdot 0.2=0.384$$

$$f_X(3)=\binom{3}{3}(0.8)^3(0.2)^0=1\cdot 1\cdot 0.008=0.512$$

The above PMF values for the binomial random variable, X, are summarized as shown in the following table:

$x \text{ (\# of \textit{Heads})}$	$f_X(x)=P(X=x)$
0	0.008
1	0.096
2	0.384
3	0.512

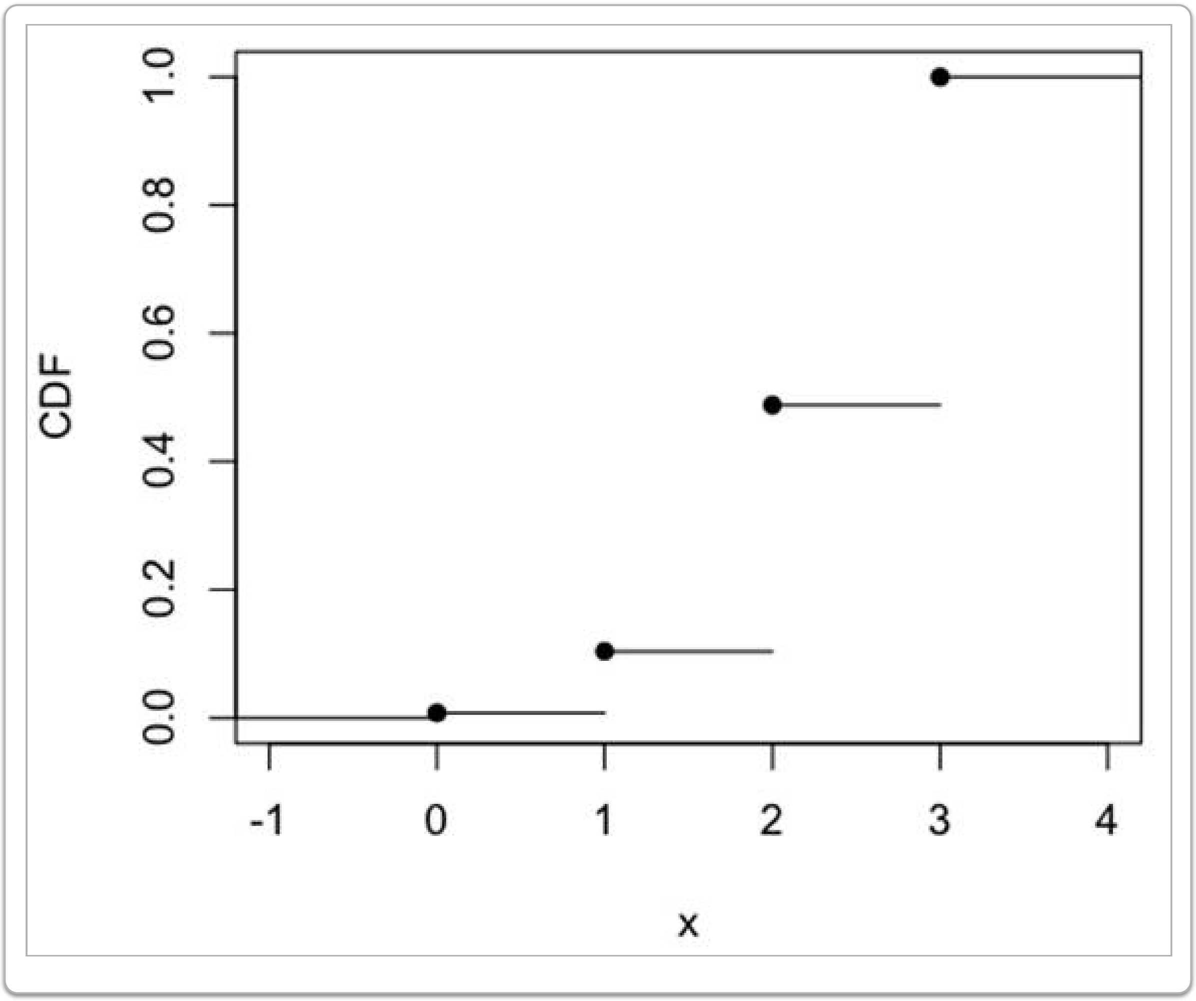
The following figure shows the heights of the probabilities of the above random variable:



The cumulative distribution function of $X,F_X(x)$, for a given value of x is the probability $P(X\le x)$. The CDF for X is defined for the entire real line. Since X cannot be negative, $F_X(0)=0$. For $0\le x\lt 1$, the event \(\{X\le x\}\) is the event \(\{X=0\}\). Similarly, for any \(x\), \(1\le x\lt 2\), the event \(\{X\le x\}\) is the event \(\{X=0 \text{ or } X=1\}\). The following table shows the values for all possible values of \(x\):

\(x\)	**\(F_X(x)=P(X\le x)\)**	**\(F_X(x)=P(X\le x)\)**
\(x\lt 0\)		\(0\)
\(0\le x\lt 1\)	\(P(X=0)\)	\(0.008\)
\(1\le x\lt 2\)	\(P(X=0)+P(X=1)\)	\(0.104\)
\(2\le x\lt 3\)	\(P(X = 0) + P (X = 1) + P(X = 2)\)	\(0.488\)
\(x\ge 3\)	\(P(X=0) + P (X=1) + P(X=2) + P(X=3)\)	\(1\)

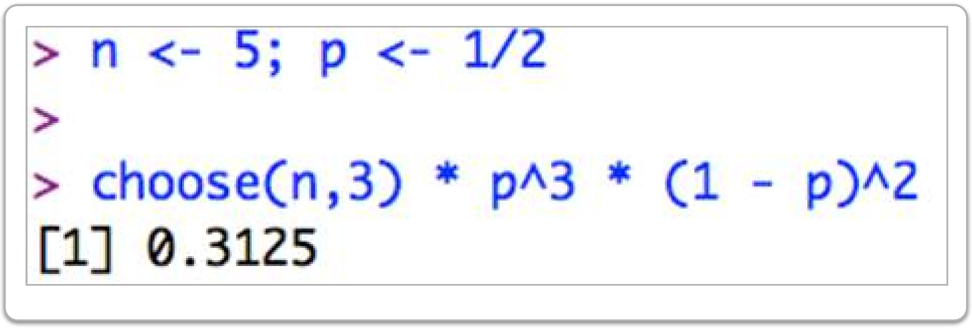
The following figure captures the cumulative values with the solid circle showing the inclusive end:



Example Using *R*—Tossing 5 Coins

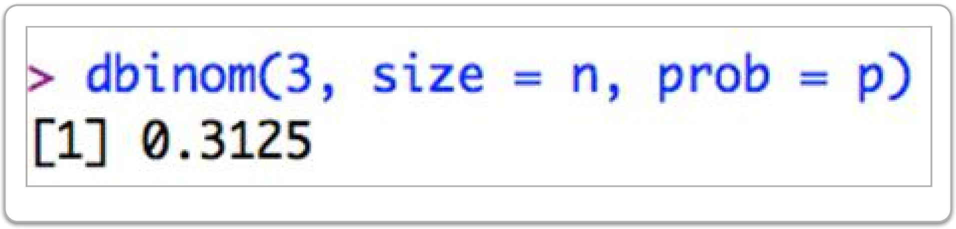
The coin-tossing example can be modeled using *R* as follows. Suppose a fair coin is tossed five times. The random variable \(X\) is the number of *heads*, and success probability is \(\frac{1}{2}\).

The probability that, say, \(X=3\), can directly be calculated using the binomial formula as shown below.



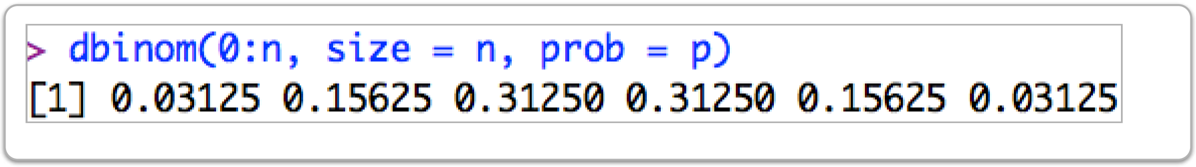
```
> n <- 5; p <- 1/2
>
> choose(n,3) * p^3 * (1 - p)^2
[1] 0.3125
```

The `stats` package provides the `dbinom` function to compute the probability for the given value of \(X\). The probability for \(X=3\) can be calculated as follows.



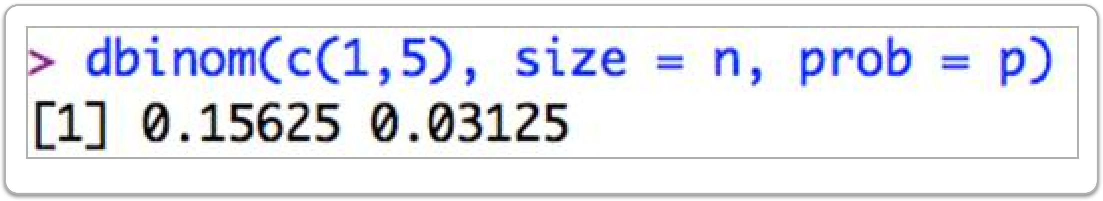
```
> dbinom(3, size = n, prob = p)
[1] 0.3125
```

The probability mass function for a sequence (or list) of values can be directly calculated as shown below. The following arguments for the function calculate all probabilities for \(0..n\).



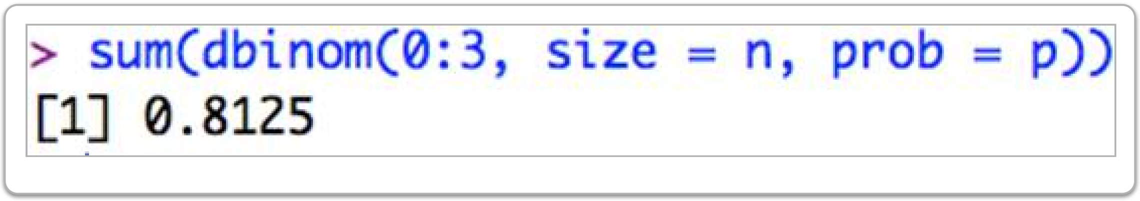
```
> dbinom(0:n, size = n, prob = p)
[1] 0.03125 0.15625 0.31250 0.31250 0.15625 0.03125
```

Similarly, the probabilities for \(X=1\) and \(X=5\) can be calculated by providing the list of values as the first argument.



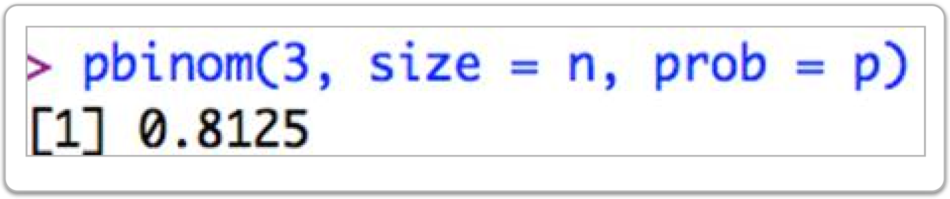
```
> dbinom(c(1,5), size = n, prob = p)
[1] 0.15625 0.03125
```

The probability for three or fewer *heads*, \(P(X\le 3)\), can be calculated as the sum of the probabilities \(P(X=0)+P(X=1)+P(X=2)+P(X=3)\). The cumulative probability is shown below.



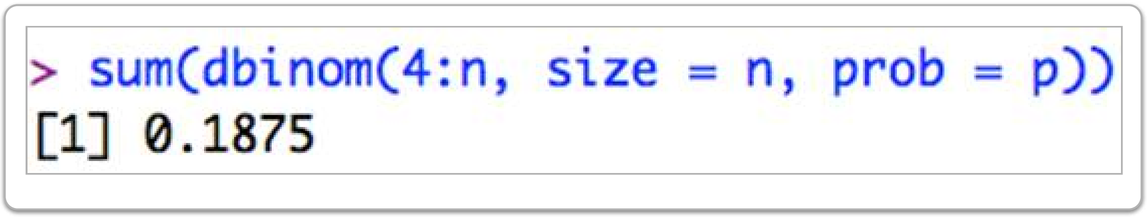
```
> sum(dbinom(0:3, size = n, prob = p))
[1] 0.8125
```

The `pbinom` function provides the alternative to directly compute the cumulative probability. \(P(X\le 3)\) is also calculated as shown below.



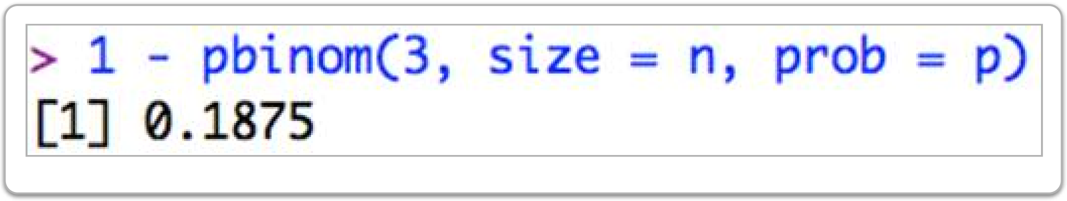
```
> pbinom(3, size = n, prob = p)
[1] 0.8125
```

The probability of \(4\) or more *heads* can be calculated as \(P(X=4)+P(X=5)\) as shown below.



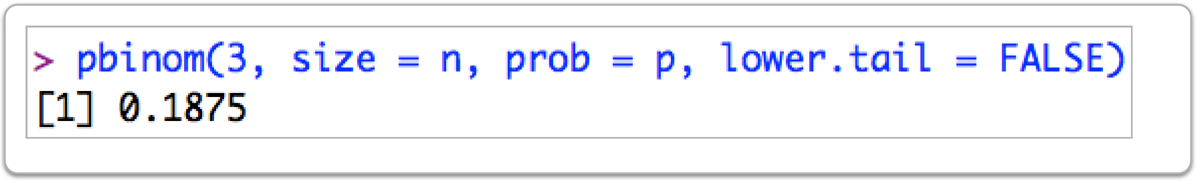
```
> sum(dbinom(4:n, size = n, prob = p))
[1] 0.1875
```

The above probability is the same as \(1-P(X\le 3)\). Using the `pbinom` function, the value is:



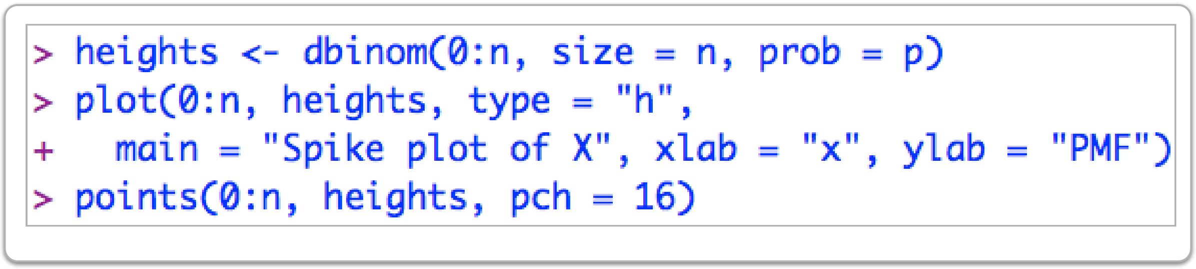
```
> 1 - pbinom(3, size = n, prob = p)
[1] 0.1875
```

Instead of using the complement, the value \(P(X\ge 4)\), or \(P(X\gt 3)\), can be calculated by providing the additional argument `lower.tail = FALSE` in the formula for computing \(P(X\le 3)\) as shown below. This adds the values of the probabilities from the next integer of the first argument to the size of the experiment.



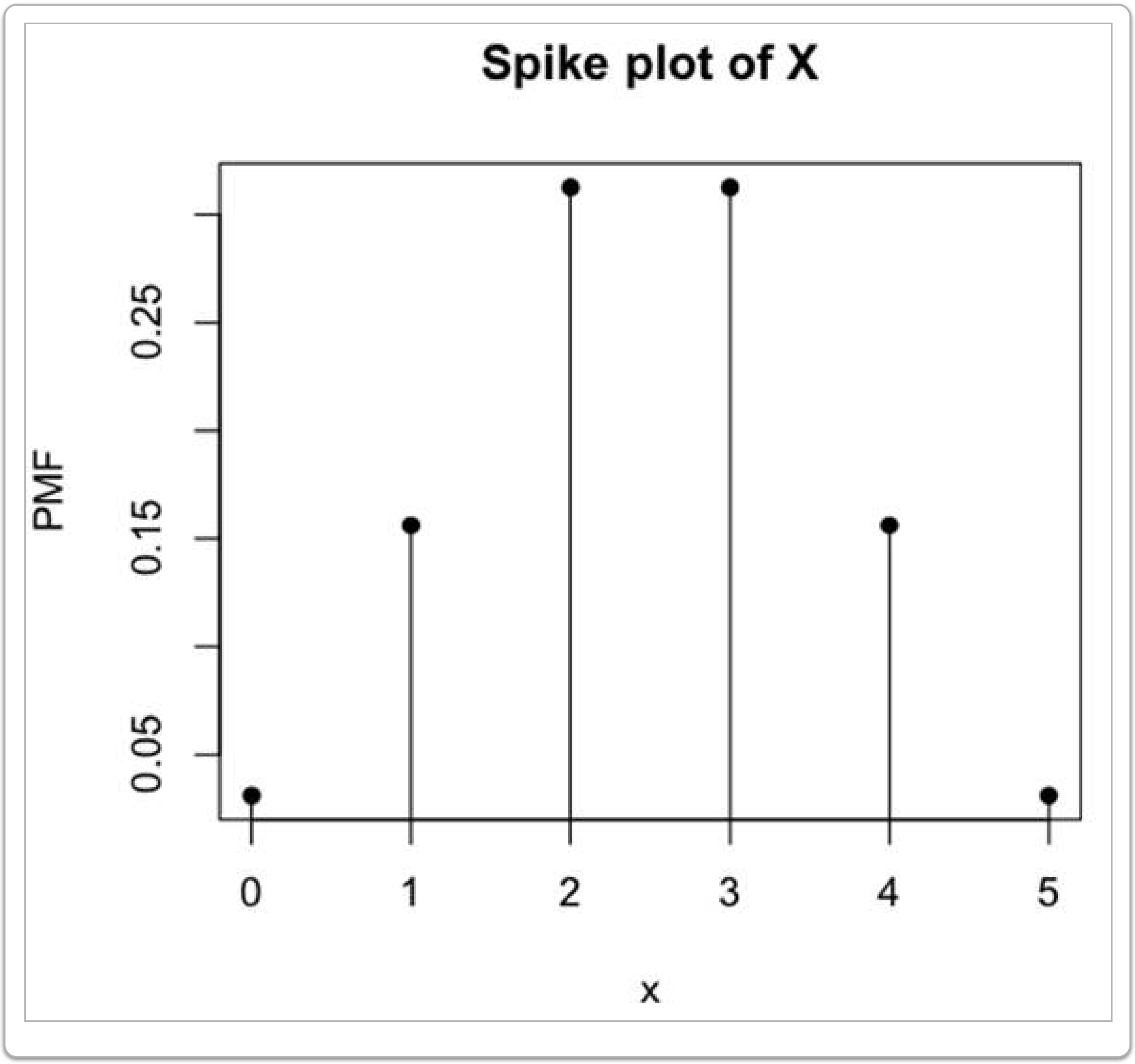
```
> pbinom(3, size = n, prob = p, lower.tail = FALSE)
[1] 0.1875
```

A spike plot of the probability distribution for the binomial random variable can be produced as shown below. The probabilities, computed as `heights`, correspond to the input values \(0..n\). The plot is generated for the sequence using the vertical lines (`type = "h"`). The tops of the lines are shown as points with the plotting character of a filled circle.

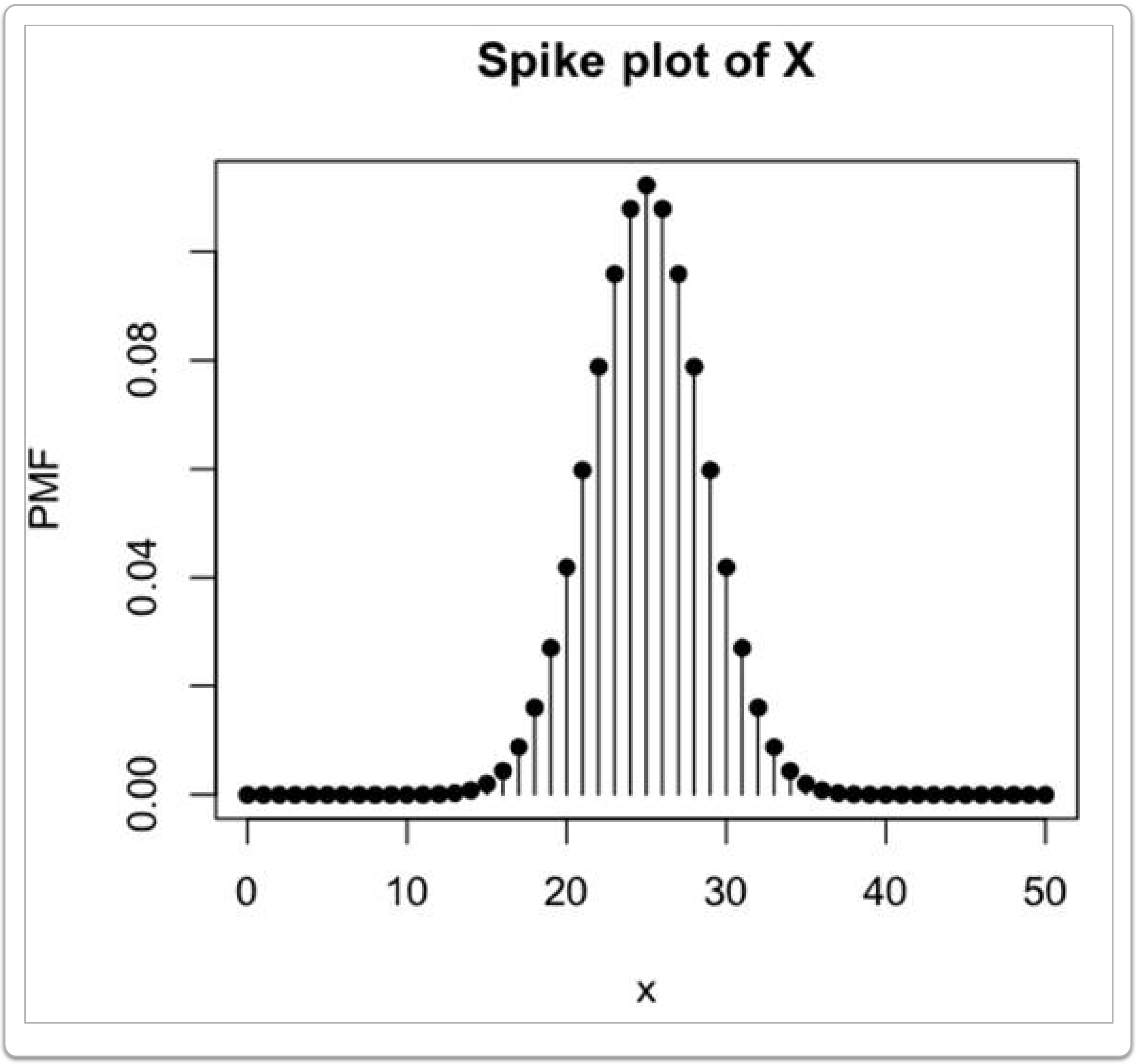


```
> heights <- dbinom(0:n, size = n, prob = p)
> plot(0:n, heights, type = "h",
+   main = "Spike plot of X", xlab = "x", ylab = "PMF")
> points(0:n, heights, pch = 16)
```

The plot of the probability mass function for size = \(5\) and prob = \(\frac{1}{2}\) is shown below.

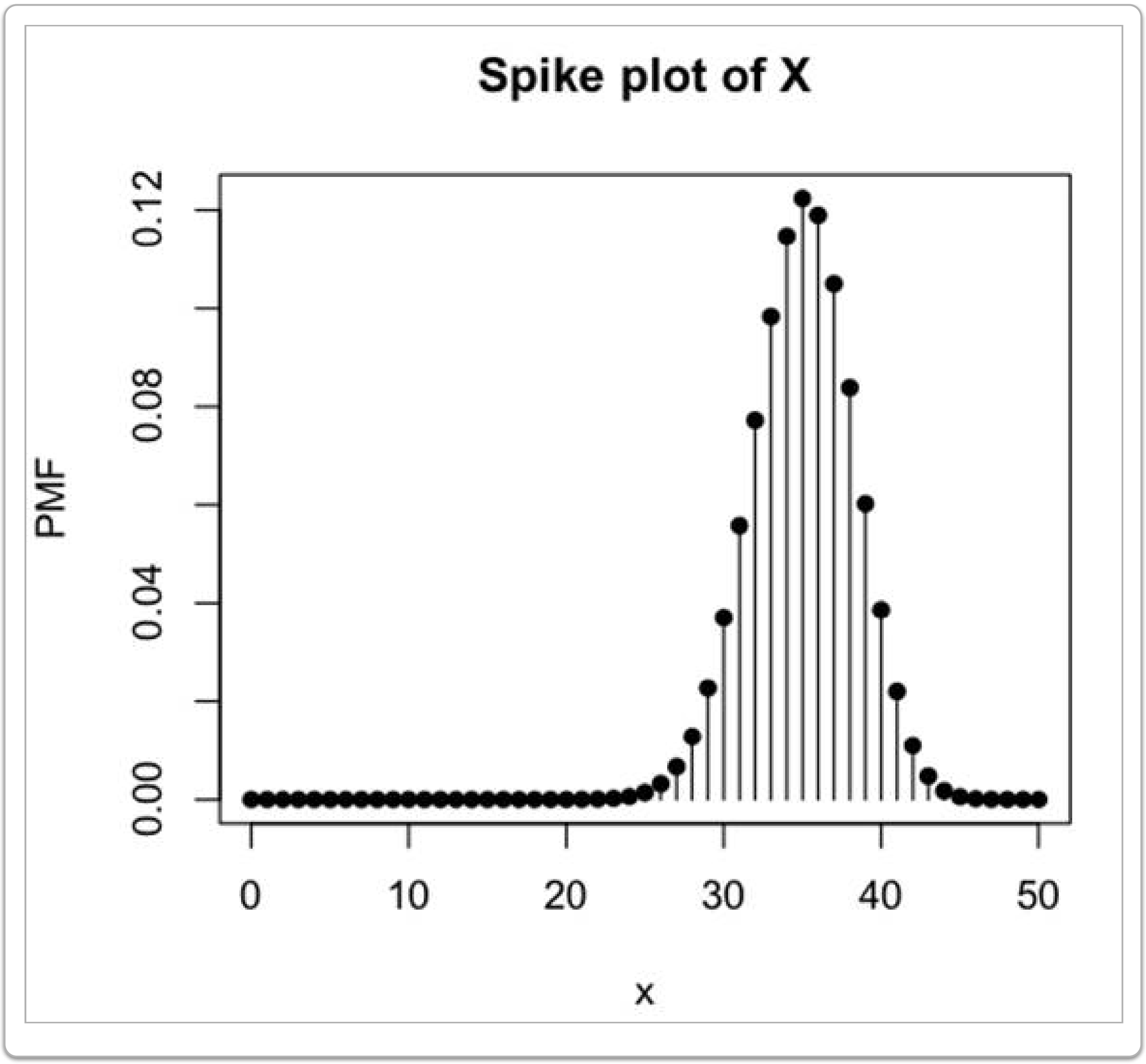


Another plot for size = \(50\) and prob = \(\frac{1}{2}\) is shown in the figure below.



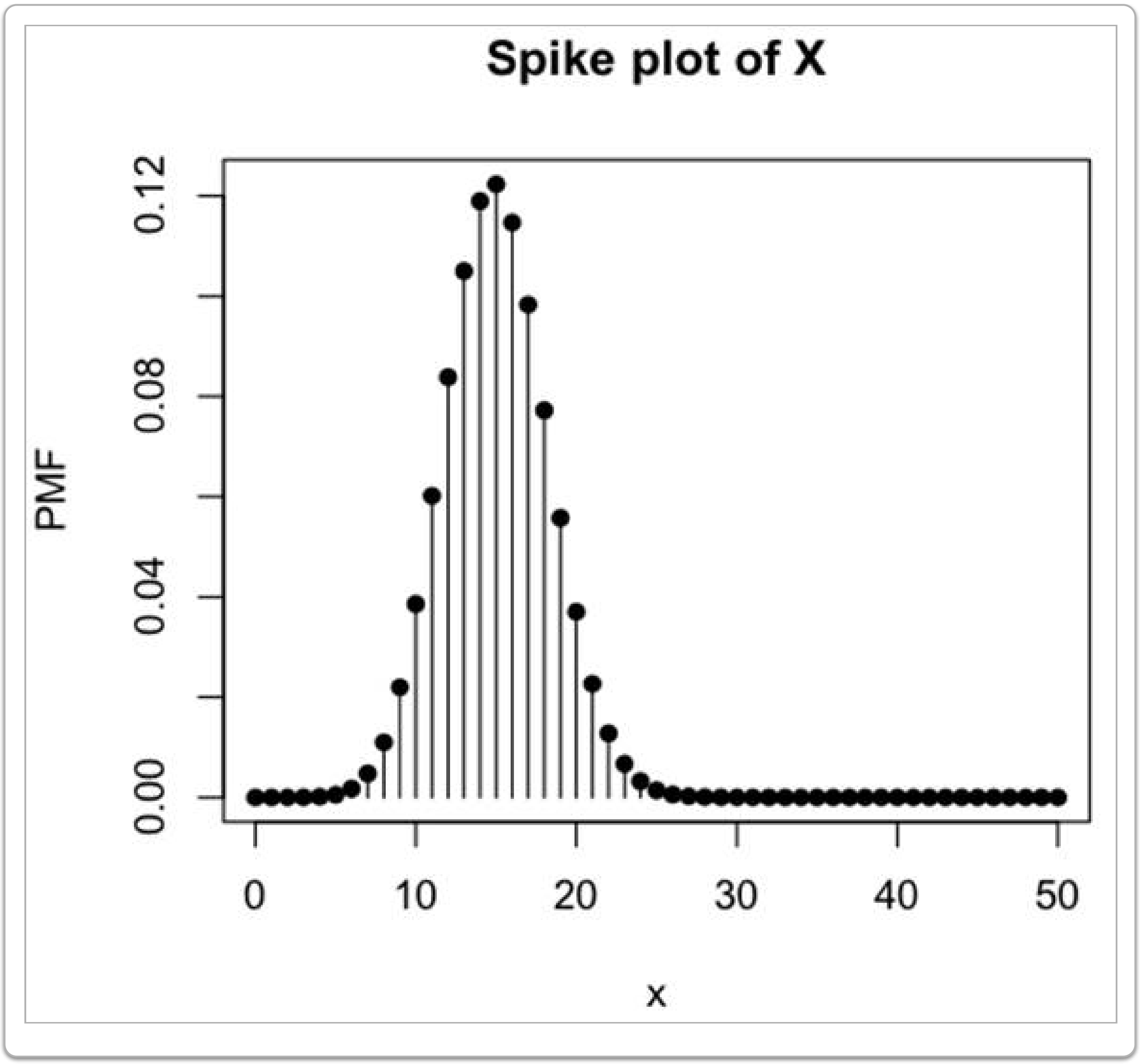
Both the plots in the previous figures had a symmetric distribution, since the probability of success is the same as the probability of failure. However, if the probability of success is more than the probability of failure, the distribution is left-skewed. The following figure shows the distribution for size = \(50\) and prob = \(0.7\).

Loading [Contrib]/a11y/accessibility-menu.js



If the probability of success is less than \(0.5\), the distribution is right-skewed as shown below for size = \(50\) and prob = \(0.3\).

Loading [Contrib]/a11y/accessibility-menu.js

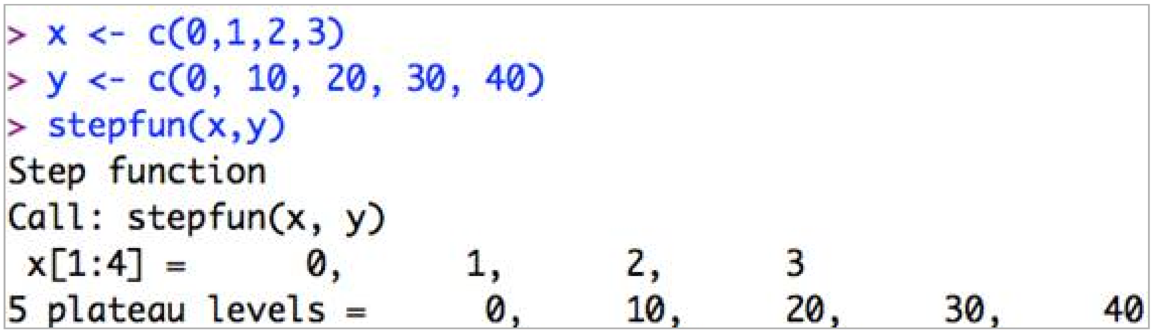


Plotting the CDF

The cumulative distribution function of the binomial random variable can be plotted in *R* by using the `cumsum` and `stepfun` functions. Given the probabilities, the `cumsum` function returns the cumulative probabilities. The first value of \(0\) (corresponding to \(F_X(x), x < 0)\) is inserted into the `cdf` variable.

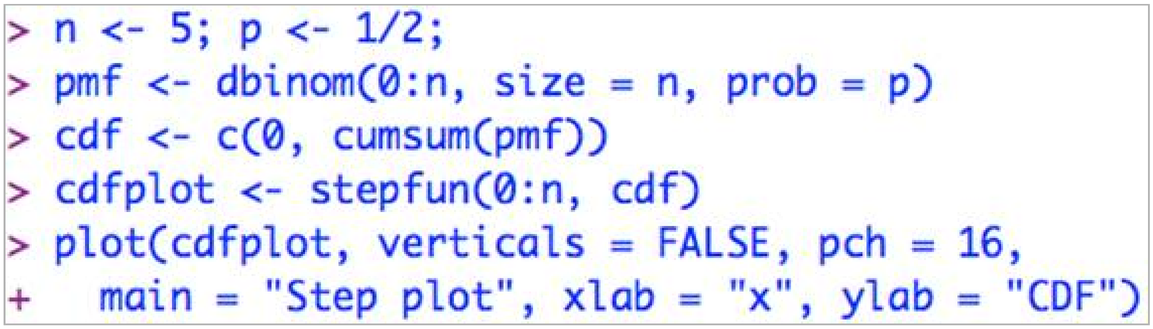
The `stepfun`, for the given values of \((x_1,x_2,\ldots,x_n)\) and \((y_0,y_1,y_2,\ldots,y_n)\), returns piecewise constant functions. The values returned by this function are constant for \([x_0,x_{i+1})\). The following example shows how a stepfun works:

Loading [Contrib]/a11y/accessibility-menu.js



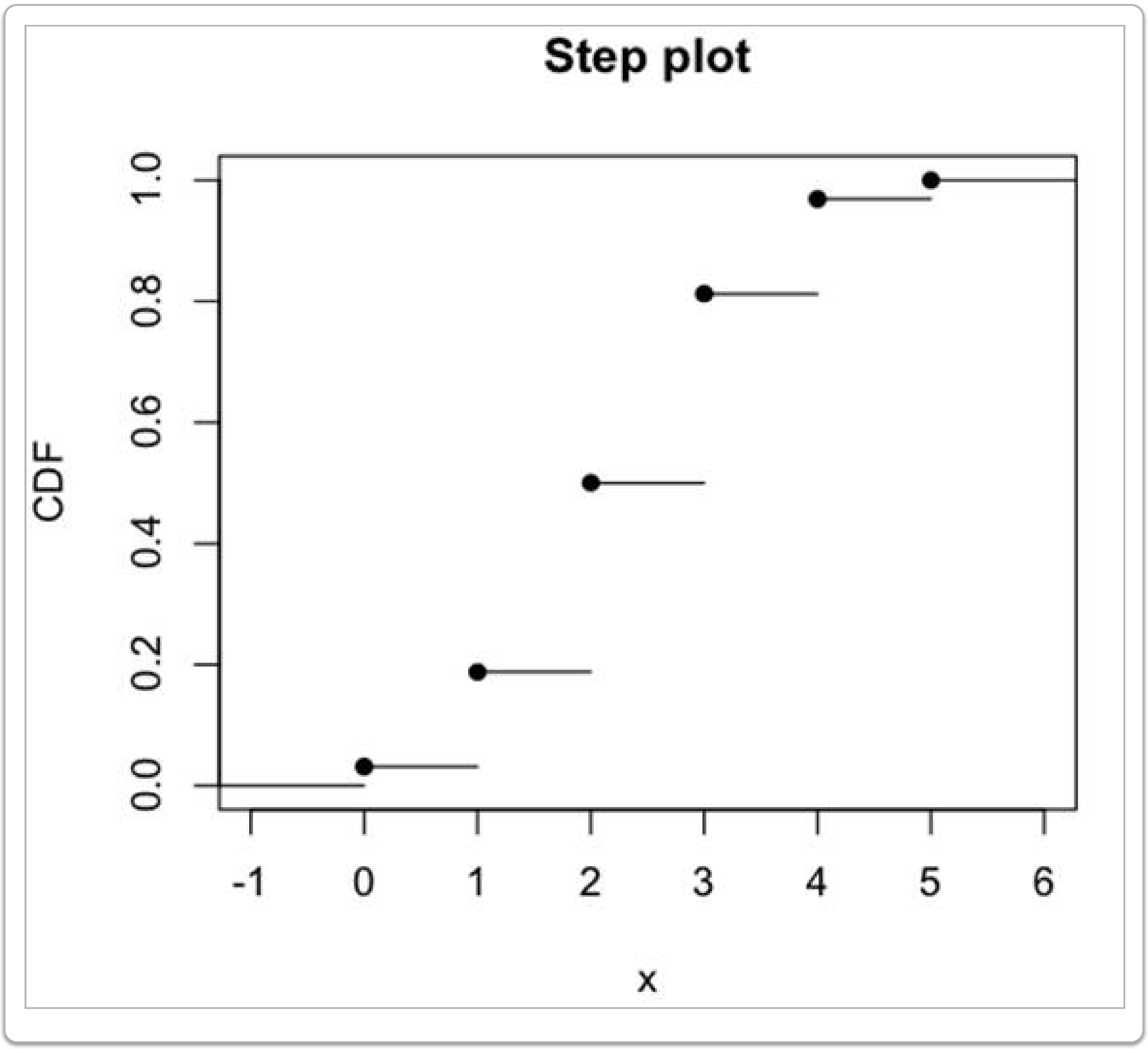
```
> x <- c(0,1,2,3)
> y <- c(0, 10, 20, 30, 40)
> stepfun(x,y)
Step function
Call: stepfun(x, y)
 x[1:4] =      0,      1,      2,      3
5 plateau levels =      0,     10,     20,     30,     40
```

For the coin-tossing examples, the CDF is plotted as shown below. The step function values are plotted using the `plot` function. The `verticals` option is `FALSE` so that vertical lines are not drawn at the steps.



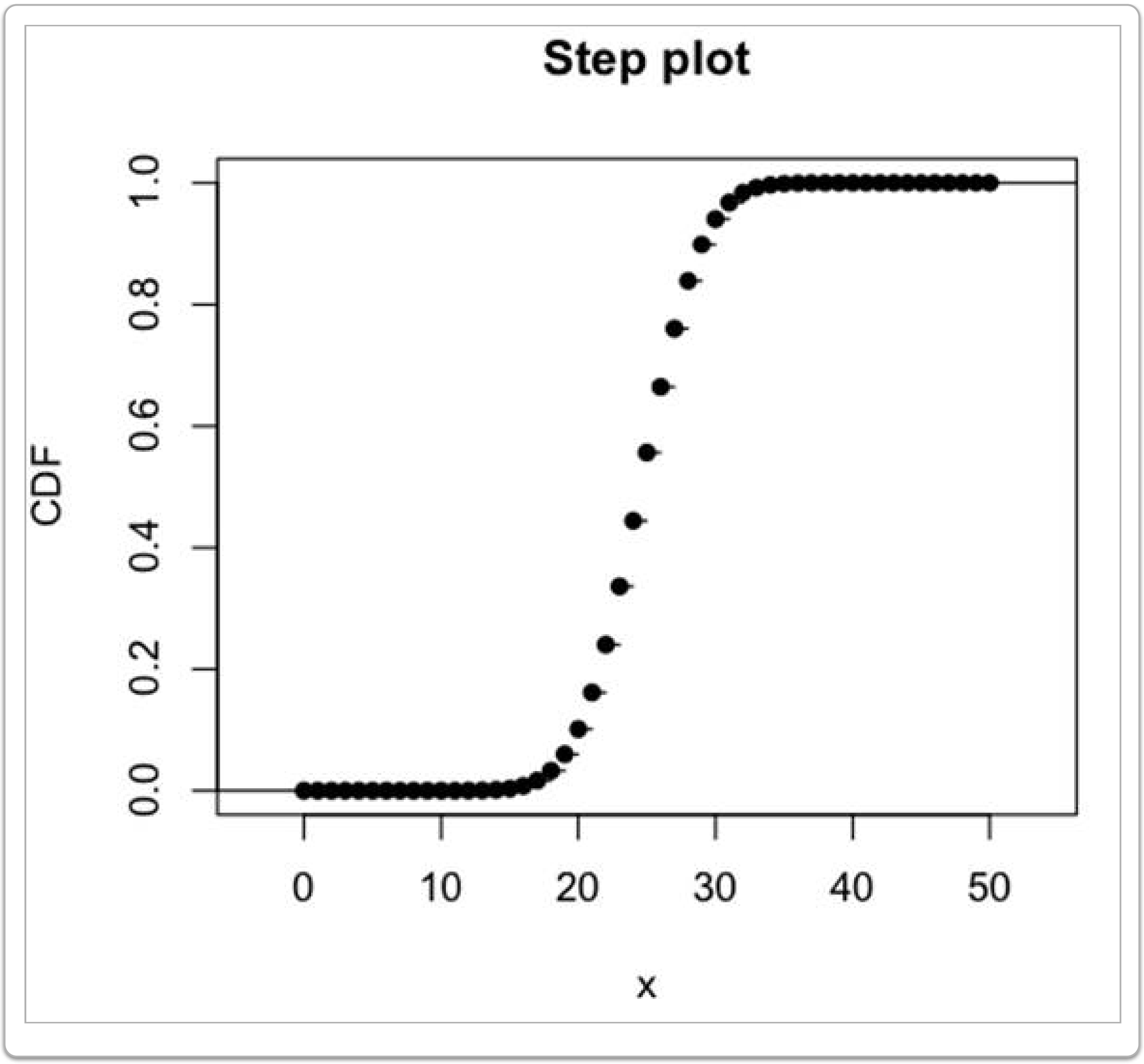
```
> n <- 5; p <- 1/2;
> pmf <- dbinom(0:n, size = n, prob = p)
> cdf <- c(0, cumsum(pmf))
> cdfplot <- stepfun(0:n, cdf)
> plot(cdfplot, verticals = FALSE, pch = 16,
+   main = "Step plot", xlab = "x", ylab = "CDF")
```

The step plot showing the CDF for size = \(5\) and prob = \(\frac{1}{2}\) is shown below.



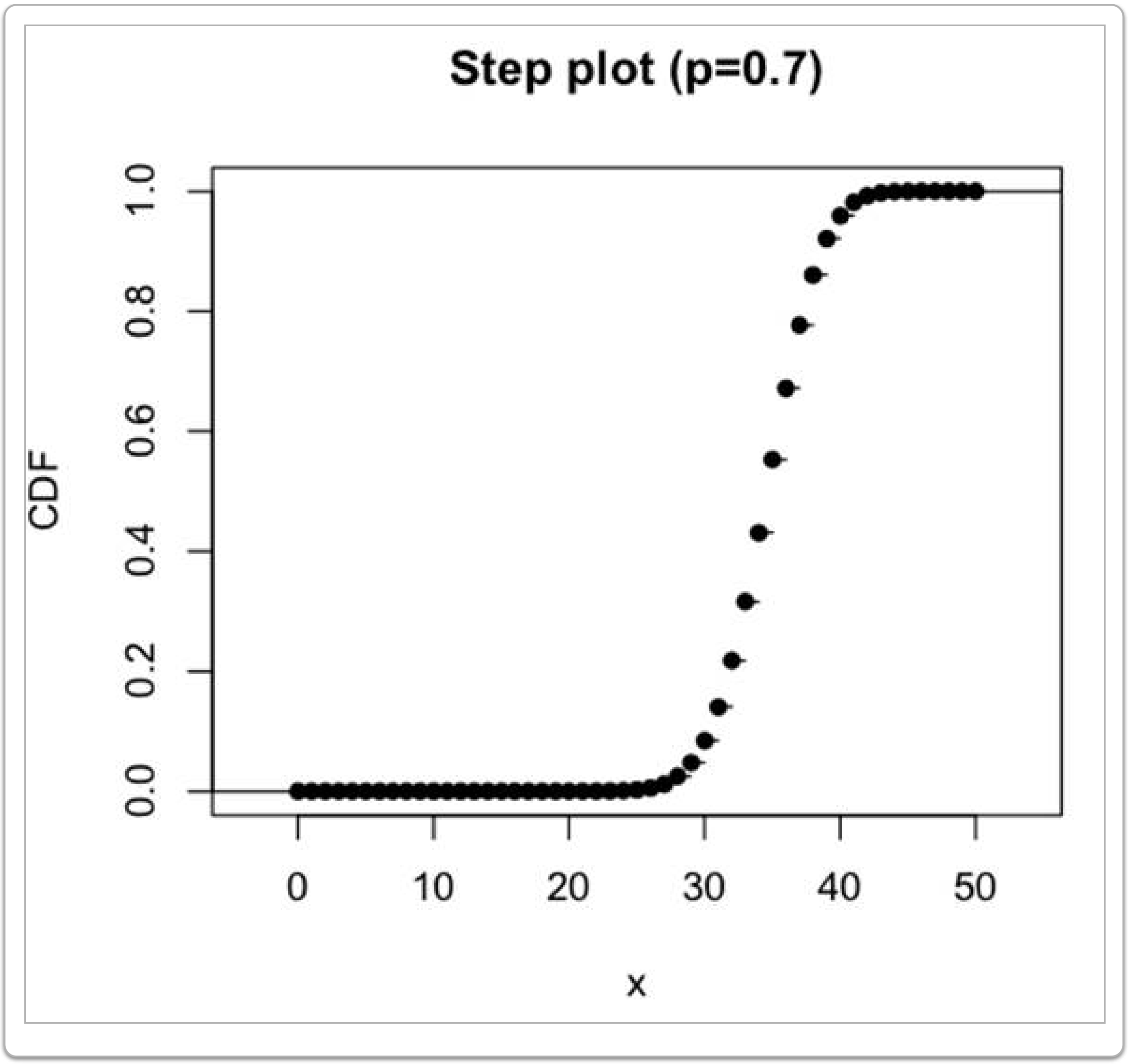
Similarly, the step plot showing the CDF for size = \(50\) and prob = \(\frac{1}{2}\) is as shown below.

Loading [Contrib]/a11y/accessibility-menu.js



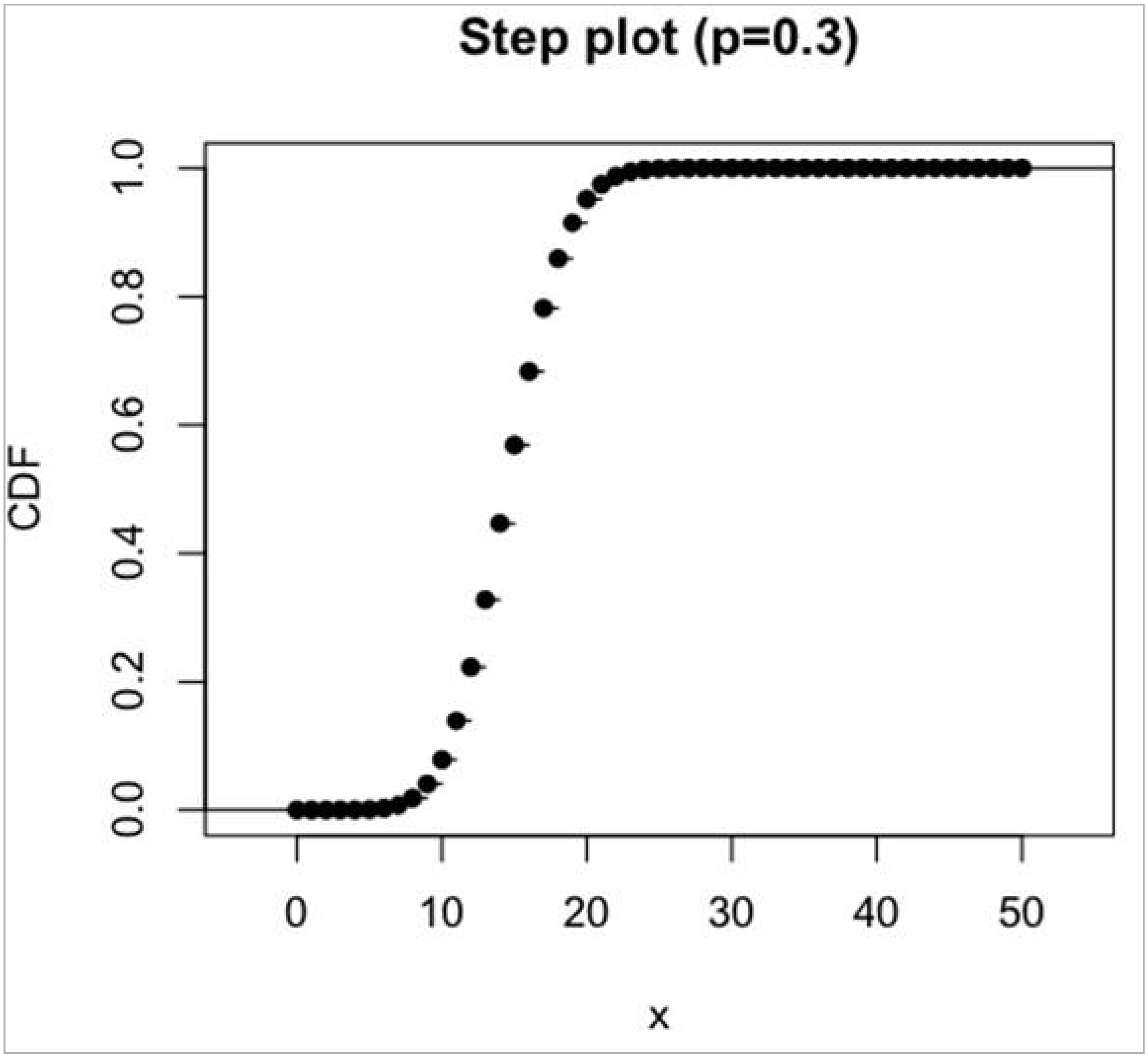
The CDF for the left-skewed distribution with size = \(50\) and prob = \(0.7\) is shown below.

Loading [Contrib]/a11y/accessibility-menu.js

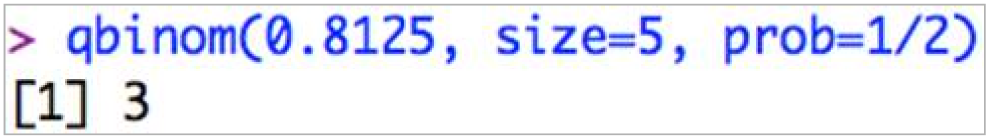


The CDF for the right-skewed distribution with size = \(50\) and prob = \(0.3\) is shown below.

Loading [Contrib]/a11y/accessibility-menu.js

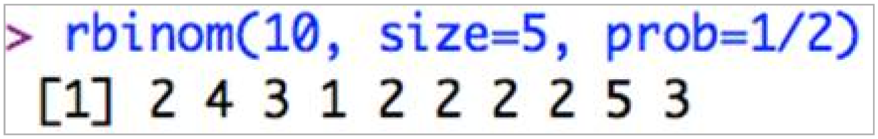


The `qbinom` function is the quantile function that is the inverse of the `pbinom` function, i.e., it returns the value for which the CDF is given.



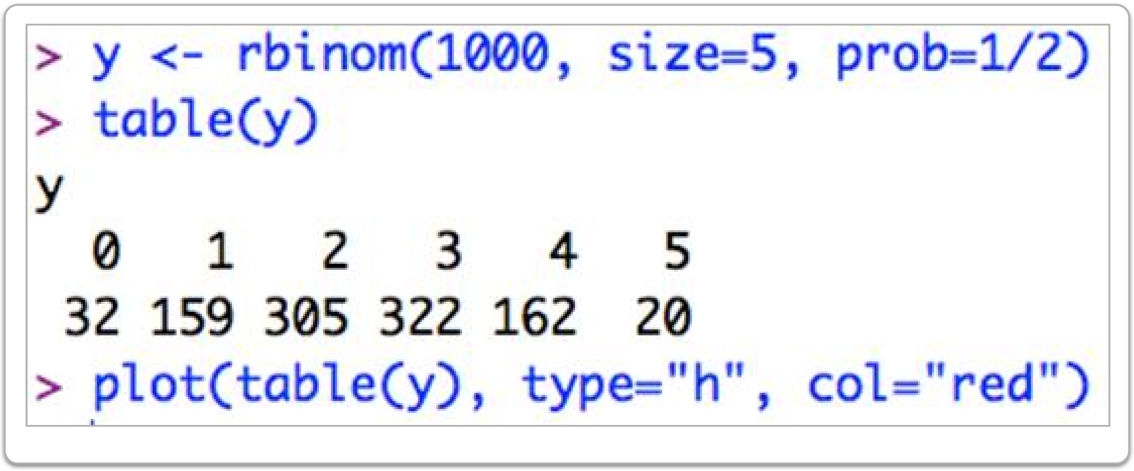
```
> qbinom(0.8125, size=5, prob=1/2)
[1] 3
```

The `rbinom` function can be used to generate random numbers according to the binomial distribution.

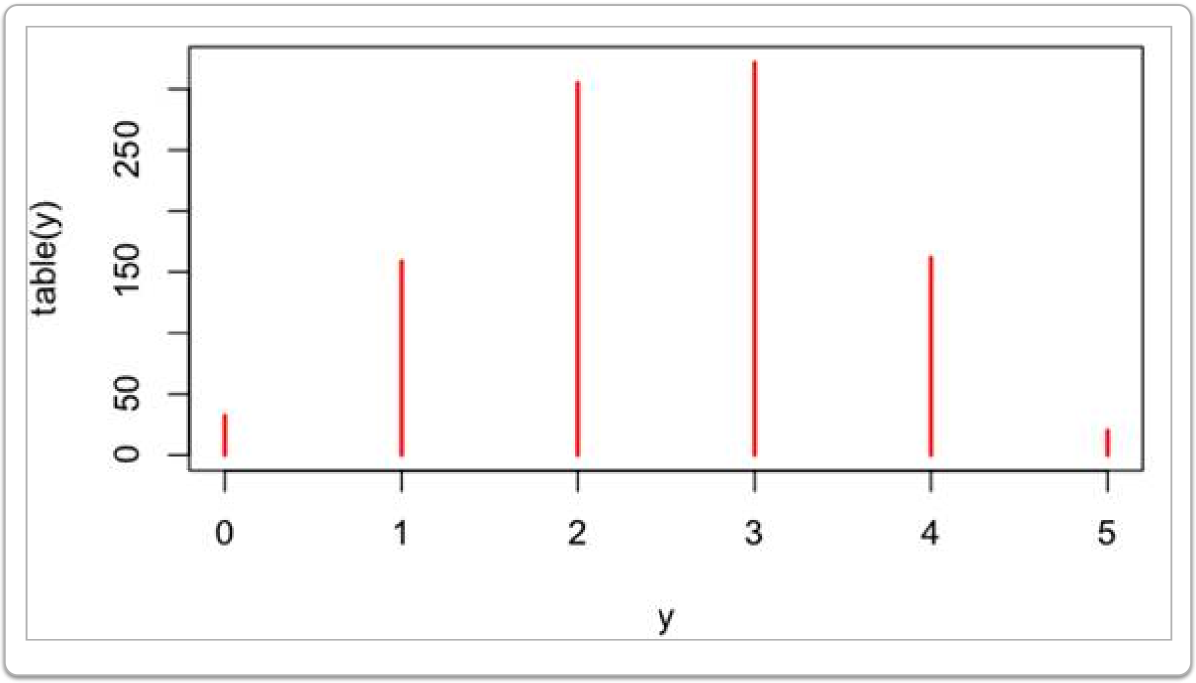


```
> rbinom(10, size=5, prob=1/2)
 [1] 2 4 3 1 2 2 2 2 5 3
```

A plot of the frequency distribution of a sample of \(1000\) numbers following the binomial distribution can be generated as shown below.



```
> y <- rbinom(1000, size=5, prob=1/2)
> table(y)
y
  0   1   2   3   4   5 
 32 159 305 322 162  20 
> plot(table(y), type="h", col="red")
```



Hypergeometric Distribution

In the binomial distribution, the Bernoulli trials are independent of each other and hence the probability of success remains the same for each trial. Also, the sample data is selected with replacement. In the *hypergeometric* distribution, the sample data is selected without replacement. Hence, the outcomes are dependent on the previous observations.

The hypergeometric distribution is illustrated with the following example. Consider an urn with five white balls and three black balls. The random experiment, say, is to select randomly two balls from the urn without replacement.

Loading [Contrib]/a11y/accessibility-menu.js e number of white balls in the selected sample.

The probability of observing a white ball out of the two balls selected would then be:

$P (1 \text{ white ball}, 1 \text{ black ball}) = \frac{\binom{5}{1}\cdot\binom{3}{1}}{\binom{8}{2}}=\frac{5\cdot 3}{28}=0.5357$

The above random variable X has the following distribution (PMF):

$x\text{ (# of white balls)}$	**$P(x\text{ white},2-x\text{ black})$**	**$f_X(x)=P(X=x)$**
0	$P (0 \text{ white}, 2 \text{ black}) = \frac{\binom{5}{0}\cdot\binom{3}{2}}{\binom{8}{2}}=\frac{1\cdot 3}{28}$	0.1071
1	$P (1 \text{ white}, 1 \text{ black}) = \frac{\binom{5}{1}\cdot\binom{3}{1}}{\binom{8}{2}}=\frac{5\cdot 3}{28}$	0.5357
2	$P (2 \text{ white}, 0 \text{ black}) = \frac{\binom{5}{2}\cdot\binom{3}{0}}{\binom{8}{2}}=\frac{10\cdot 1}{28}$	0.3571

In general, the hypergeometric distribution for a random variable X is defined as follows. Given a sample size of $M+N$, where M is the number of events of interest and N is the number of events that are not of interest, and K is the sample size without replacement, the probability of x events of interest is

$$f_X(x)=\frac{\binom{M}{x}\cdot\binom{N}{K-x}}{\binom{M+N}{K}}$$

The probability mass function of X is computed for $x=0,1,\ldots,K$ using the above formula.

Notation

```
X ~ hyper(m = M, n = N, k = K)
```

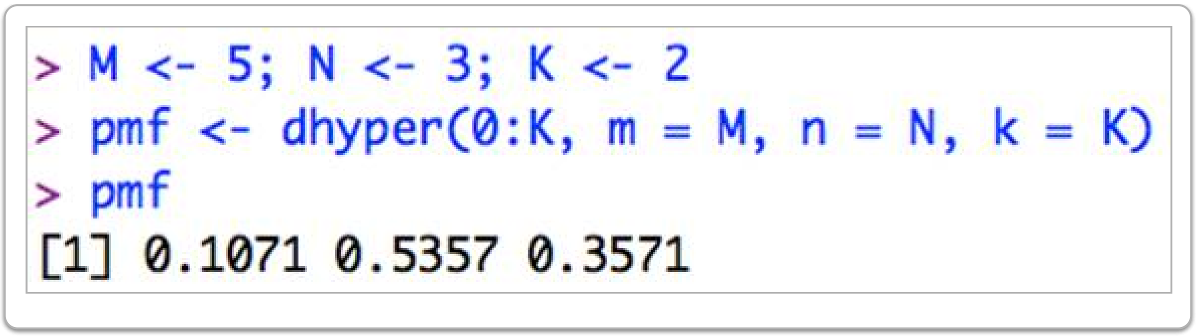
The mean of the hypergeometric distribution is:

$$\mu=\sum xf_X(x)=K\cdot\frac{M}{M+N}$$

The variance of the hypergeometric distribution is:

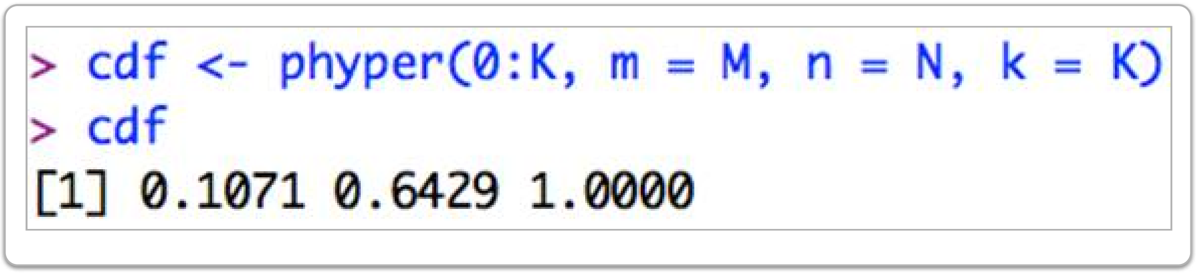
$$\sigma^2=\sum x^2f_X(x)-\mu^2=K\cdot\frac{M\cdot N}{(M+N)^2}\cdot\frac{(M+N-K)}{(M+N-1)}$$

The associated *R* functions for the PMF and CDF of the hypergeometric distribution are `dhyper(x, m, n, k)` and `phyper(x, m, n, k)`, respectively. The above problem with five white balls, three black balls, and a sample size of two is modeled as follows:



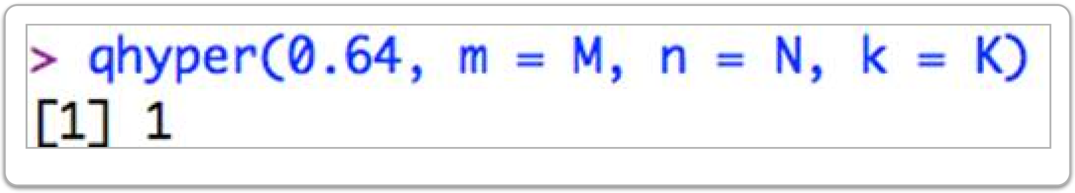
```
> M <- 5; N <- 3; K <- 2
> pmf <- dhyper(0:K, m = M, n = N, k = K)
> pmf
[1] 0.1071 0.5357 0.3571
```

The corresponding cumulative probabilities are shown below.



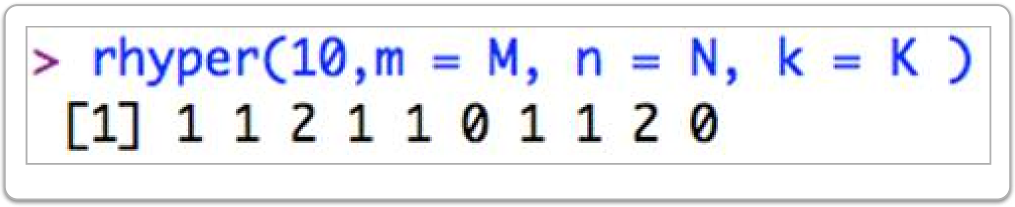
```
> cdf <- phyper(0:K, m = M, n = N, k = K)
> cdf
[1] 0.1071 0.6429 1.0000
```

The quantile function, `qhyper`, is the inverse of the `phyper` function, returning the number for the specified CDF.



```
> qhyper(0.64, m = M, n = N, k = K)
[1] 1
```

The `rhyper` function can be used to generate random numbers based on this distribution.



```
> rhyper(10,m = M, n = N, k = K )
 [1] 1 1 2 1 1 0 1 1 2 0
```

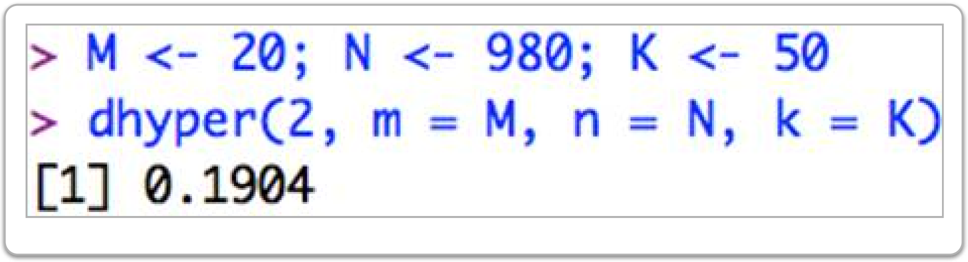
Example Using *R*—Faulty Chips

Suppose there are 20 faulty chips out of manufactured lot of 1000 chips. For quality control, 50 chips are selected at random without replacement. The random variable is the number of faulty chips in the selected sample. In this hypergeometric distribution, $M=20$, $N=980$, and $K=50$.

The probability of exactly 2 faulty chips in the sample is:

$$P(X=2)=f_X(2)=\frac{\binom{20}{2}\cdot\binom{980}{48}}{\binom{1000}{50}}=0.1904$$

Using *R*,

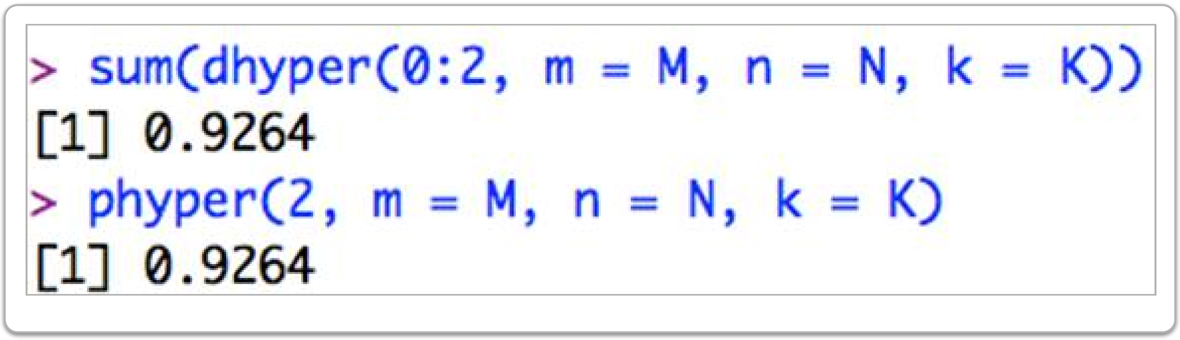


```
> M <- 20; N <- 980; K <- 50
> dhyper(2, m = M, n = N, k = K)
[1] 0.1904
```

Similarly, the probability of at most 2 faulty chips in the sample is:

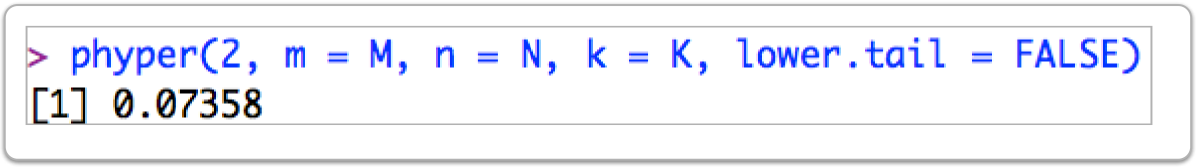
$$P(X\le2)=P(X=0)+P(X=1)+P(X=2)$$

Using *R*, the same value can be calculated in either of the two ways shown below.



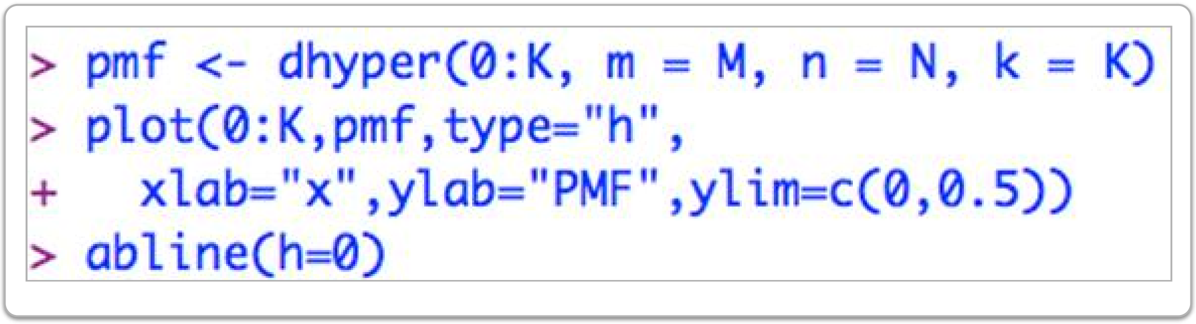
```
> sum(dhyper(0:2, m = M, n = N, k = K))
[1] 0.9264
> phyper(2, m = M, n = N, k = K)
[1] 0.9264
```

To calculate $P(X\ge3)=P(X\gt2)$, the `lower.tail` property of the `phyper` function is used. This is equivalent to $1-P(X\le2)$.



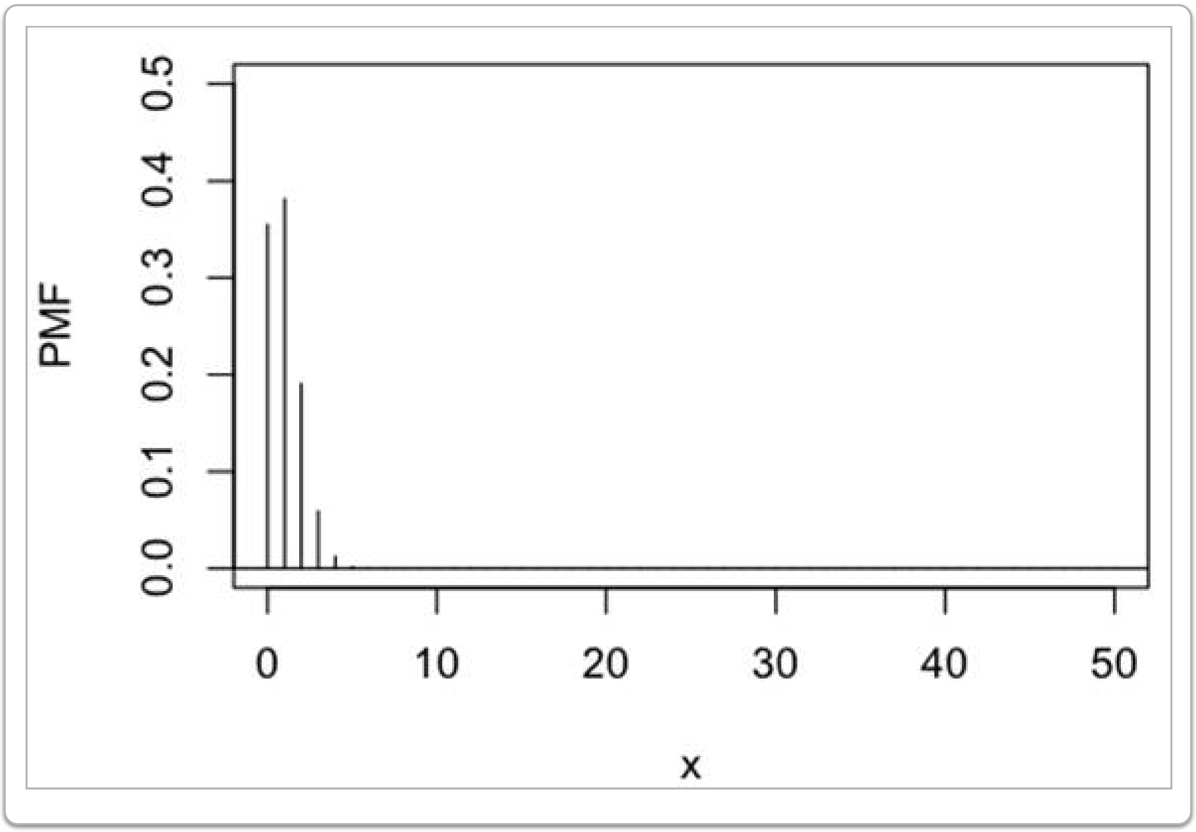
```
> phyper(2, m = M, n = N, k = K, lower.tail = FALSE)
[1] 0.07358
```

The probability distribution for the number of faulty chips can be calculated and plotted as follows:



```
> pmf <- dhyper(0:K, m = M, n = N, k = K)
> plot(0:K,pmf,type="h",
+   xlab="x",ylab="PMF",ylim=c(0,0.5))
> abline(h=0)
```

The plot for the above data is shown below.



Geometric Distribution

The geometric distribution concerns the number of failures before a success occurs in a sequence of Bernoulli trials. In a geometric distribution, the random variable X is the number of failures before a success. If the probability of success is p (hence the probability of failure is $1–p$), the probability a getting a success after 2 failures is:

$$P(X=2)=(1-p)\cdot(1-p)\cdot p$$

The probability mass function of X, $p(X=x)$, is:

$$f_X(x)=p\cdot(1-p)^x,x=0,1,2,\ldots$$

Notation

```
X ~ geom(prob = p)
```

The mean of the geometric distribution is:

$$\mu=\sum xf_X(x)=\frac{1-p}{p}$$

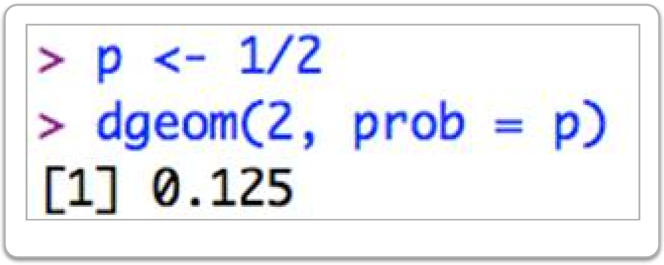
The variance of the geometric distribution is:

$$\sigma^2=\sum x^2f_X(x)-\mu^2=\frac{1-p}{p^2}$$

The associated *R* functions for the PMF and CDF of the geometric distribution are `dgeom(x, prob)` and `pgeom(x, prob)`, respectively. Consider the coin-tossing example with a fair coin where the probability of success (say, getting a *head*) is \(\frac{1}{2}\). The geometric distribution computes the probabilities for the number of *tails* before a *head* occurs.

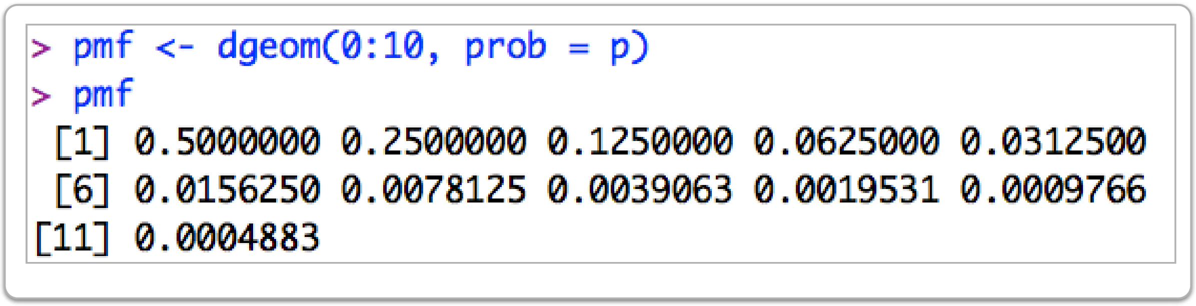
\(x (\text{# of *Tails*})\)	**\(f_X(x)=P(X=x)\)**
\(0\)	\(0.5\)
\(1\)	\(0.5 \cdot 0.5 = 0.25\)
\(2\)	\(0.5 \cdot 0.5 \cdot 0.5 = 0.125\)
\(3\)	\(0.5 \cdot 0.5 \cdot 0.5 \cdot 0.5 = 0.0625\)

The above probabilities are computed in *R* as follows:



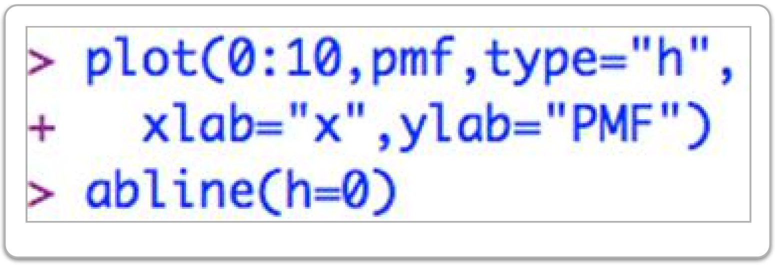
```
> p <- 1/2
> dgeom(2, prob = p)
[1] 0.125
```

Similarly, the PMF can be calculated in *R* as follows:

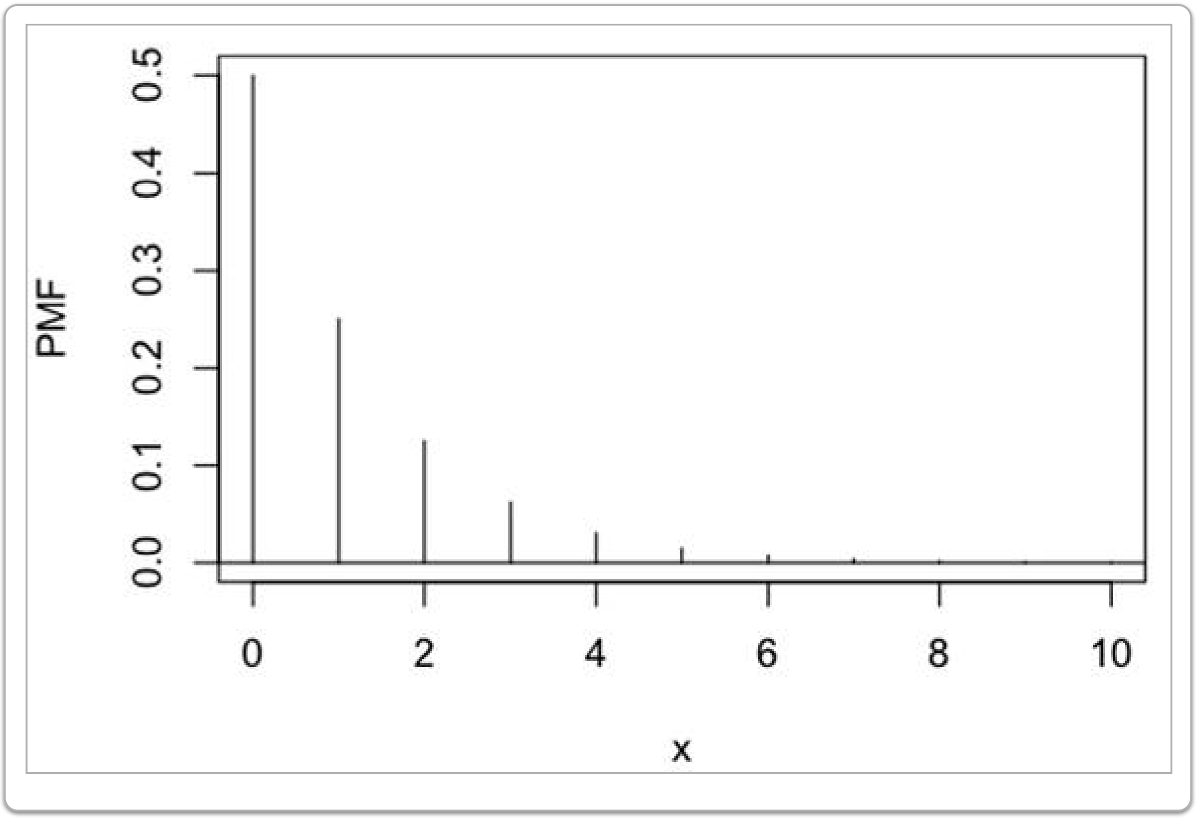


```
> pmf <- dgeom(0:10, prob = p)
> pmf
 [1] 0.5000000 0.2500000 0.1250000 0.0625000 0.0312500
 [6] 0.0156250 0.0078125 0.0039063 0.0019531 0.0009766
[11] 0.0004883
```

The plot of the above distribution is shown below.



```
> plot(0:10,pmf,type="h",
+   xlab="x",ylab="PMF")
> abline(h=0)
```



Negative Binomial Distribution

The negative binomial distribution concerns the number of failures until a total of r successes occur in a sequence of Bernoulli trials. In a negative binomial distribution, the random variable X is the number of failures that precede the rth success. The total number of successes, r, is fixed in the experiment. If the probability of success is p (hence the probability of failure is $1–p$), the probability a getting 3 successes with 5 failures is calculated as follows. In this case, there will be exactly 5 failures before the third success. In the 8 trials, the 8th trial has to a success, and there must be exactly 2 successes among the first 7 trials. There are $\binom{7}{2}$ possibilities.

$$P(X=5)=\left\{\binom{7}{2}(1-p)^5\cdot p^2\right\}\cdot p$$

The probability mass function of X, $P(X=x)$, is:

Loading [Contrib]/a11y/accessibility-menu.js \cdot(1-p)^x,x=0,1,2,\ldots$$

Notation

`X ~ nbinom(size = r, prob = p)`

The mean of the negative binomial distribution is:

$$\mu=\frac{r\cdot(1-p)}{p}$$

The variance of the negative binomial distribution is:

$$\sigma^2=\frac{r\cdot(1-p)}{p^2}$$

In the example illustrated above, if the experiment is tossing a fair coin, and success is getting a *head*, the probability of \(5\) *tails* before getting the third *head* \((r=3)\) is:

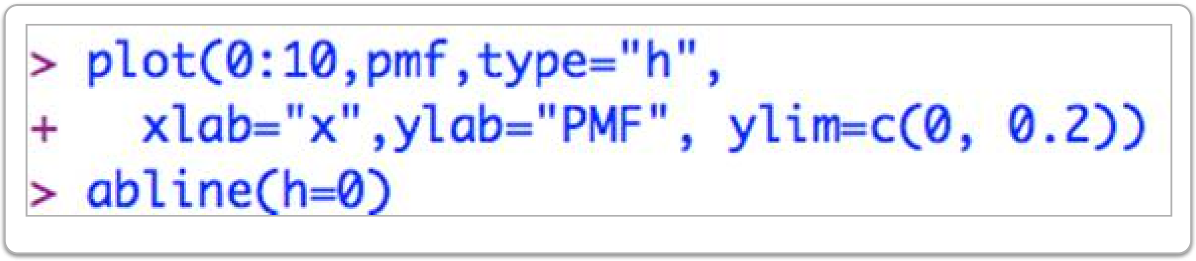
$$P(X=5)=\binom{7}{2}\cdot\left(\frac{1}{2}\right)^3\cdot\cdot\left(\frac{1}{2}\right)^5=0.082$$

The associated *R* functions for the PMF and CDF of the negative binomial distribution are `dnbinom(x, size, prob)` and `pnbinom(x, size, prob)`, respectively.

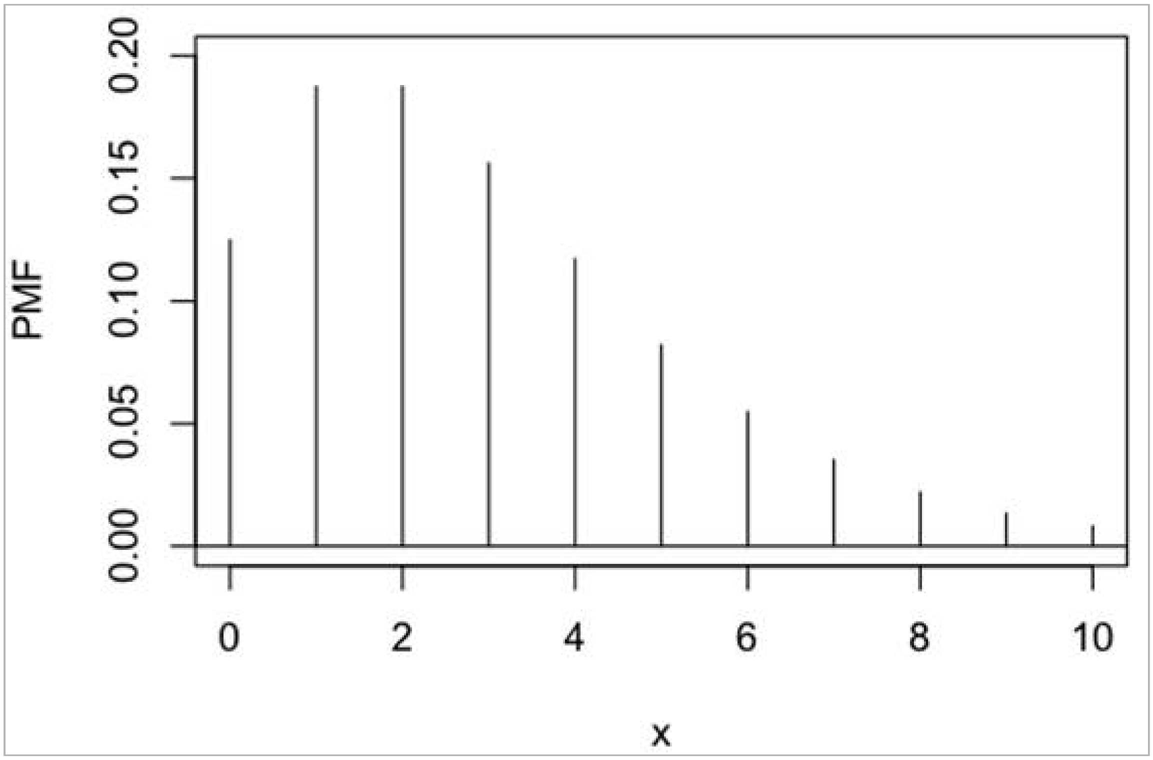
For the above example, the probability is computed in *R* as follows:

The probability mass function of the negative binomial variable is calculated in *R* as follows:

The plot of the above distribution is shown below.



```
> plot(0:10,pmf,type="h",
+   xlab="x",ylab="PMF", ylim=c(0, 0.2))
> abline(h=0)
```



The probability that at most 5 failures are observed is:

$$P(X\le5)=P(X=0)+p(X=1+\ldots+P(X=5))$$

The cumulative probability can be computed in *R* as follows:

The plot for the cumulative distribution function is shown below.

Poisson Distribution

The Poisson distribution is used to model the frequency with which a specified event occurs during a particular

where the Poisson distribution is applicable are:

- The number of patients arriving in an ER between 11 PM and midnight
- The number of customers arriving in a bank between 9 AM and 11 AM
- The number of support calls received per day

The Poisson distribution is identified by a single parameter, \(\lambda\) (*lambda*), which is the mean or the average number of events per time unit \([0,1]\).

In the Poisson distribution, the random variable \(X\) counts the number of events occurring in the unit interval. The probability mass function of the random variable \(X,P(X=x)\), is:

$$f_X(x)=e^{-\lambda}\frac{\lambda^x}{x!},x=0,1,2,\ldots$$

where \(\lambda\) is a positive real number and \(e\approx2.718\).

Notation

`X ~ pois(lambda = \(\lambda\))`

If the random variable \(X\) counts the number of events in the interval \([0, t]\), then \(X\) has the Poisson distribution X ~ *pois*(lambda = \(\lambda\)t):

$$P(X=x)=e^{-\lambda t}\cdot\frac{(\lambda t)^x}{x!},x=0,1,2,\ldots$$

The mean of the Poisson distribution is:

$$\mu=\lambda$$

The variance of the Poisson distribution is:

$$\sigma^2=\lambda$$

The associated *R* functions for the PMF and CDF of the Poisson distribution are `dpois(x, lambda)` and `ppois(x, lambda)`, respectively.

Suppose that, on average, a hospital ER receives \(8\) patients between 11 PM and midnight. With Poisson distribution, \(\lambda=8\). The random variable \(X\) is the number of patients arriving between 11 PM and midnight. The probabilities for \(X\) are:

$$P(X=x)=e^{-8}\frac{8^x}{x!}$$

The probability that exactly \(6\) patients arrive during this interval is:

$$P(X=6)=e^{-8}\frac{8^6}{6!}=0.122$$

The above probability can be computed in *R* using the `dpois` function as follows:

Chances are 12.2% that exactly 6 patients will arrive during this period.

The probability that at most 2 patients arrive during this interval is:

$$\begin{align}P(X\le2)&=P(X=0)+P(X=1)+P(X=2)\\&=e^{-8}\frac{8^0}{0!}+e^{-8}\frac{8^1}{1!}+e^{-8}\frac{8^2}{2!}\\&=e^{-8}(1+8+32)\\&=0.0137\end{align}$$

The above cumulative probability can be computed in *R* using the `ppois` function as follows:

Chances are only 1.37% that 2 or fewer patients will arrive during this interval.

The probability of between 5 and 10 patients (inclusive) arriving is:

$$\begin{align}P(5\le X\le10)&=P(X=5)+P(X=6)+P(X=7)+P(X=8)+P(X=9)+P(X=10)\\&=P(X\le10)-P(X\lt5)\end{align}$$

The above value can be calculated in *R* as follows:

or by using a vector of the two values and taking the difference of the resulting vector.

The probability mass function of the random variable X can be computed as follows:

The above values can be plotted showing the distribution for the first 20 values.

The above figure shows that the Poisson distribution is right-skewed (all Poisson distributions are right-skewed).

Suppose that a fast food drive-thru serves, on average, 3 vehicles per 5-minute interval during lunchtime. In this case, if X is number of vehicles arriving in a 5-minute period, then $X\sim pois(lambda=3)$. If the lunch hour is Noon to 1 PM, the total time period is 60 minutes (*twelve* 5-minute units). If Y is the number of vehicles arriving during the lunch hour, then $Y\sim pois(lambda=3\cdot12=36)$.

Continuous Distributions

Introduction

For continuous random variables, the probability density function (PDF) defines the distribution of values. Common distributions include the normal distribution (also known as Gaussian distribution), uniform distribution (also known as rectangular distribution), and exponential distribution.

The normal distribution is symmetric and has a bell shape. Most of the values tend to cluster around the mean. Since this is symmetrical distribution, the mean equals the median. In the uniform distribution, each value has an equal probability of occurrence. The uniform distribution is also symmetrical, hence the mean equals the median. The exponential distribution is skewed to the right, hence its mean is larger than the median. The density curve is always on or above the x-axis, and the total area under a density curve equals 1.

Continuous Uniform Distribution

For the random variable X with the continuous uniform distribution over the interval $[a,b]$, the probability of occurrence is the same anywhere in the range. The probability density function (PDF) for the random variable X is:

$$f_X(x)=\frac{1}{b-a},a\le x\le b, \text{ and }0\text{ elsewhere}$$

```
X ~ unif(min = a, max = b)
```

The cumulative distribution function for the random variable \(X\) is:

$$F_x(t)=\left\{\begin{array}{cc}0, & t\lt a\\ \frac{t-a}{b-a}, & a\le t\lt b\\ 1, & t\ge b\end{array}\right.$$

The mean of the continuous uniform random variable \(X\) is:

$$\begin{align}\mu=E[X]&=\int^{\infty}_{-\infty}xf_X(x)dx\\&=\int^b_ax\frac{1}{b-a}dx\\&=\frac{1}{b-a}\int^b_axdx\\&=\frac{1}{b-a}\frac{b^2-a^2}{2}\\&=\frac{1}{b-a}\frac{(b+a)(b-a)}{2}\\&=\frac{b+a}{2}\end{align}$$

The variance of the continuous uniform random variable \(X\) is:

$$\sigma^2=\frac{(b-a)^2}{12}$$

Using *R*, the associated functions for the PDF and CDF are `dunif(x, min = a, max = b)` and `punif(x, min = a, max = b)`, respectively.

Example 1—Uniform Distribution

A common use of the uniform distribution is in the selection of random numbers. For random sampling, the uniform distribution \([0,1]\) is used.

The probability of getting a random number between \(0.2\) and \(0.4\), \(P[0.2\lt X\lt0.4]\) can be calculated in *R* as follows:

From the distribution, the above probability is the area of the region between \(x=0.2\) and \(x=0.4\), shown outlined in the figure above.

The mean of the uniform distribution for \(a=0\) and \(b=1\) is:

$$\mu=\frac{b+a}{2}=\frac{1}{2}=0.5$$

and the variance is:

$$\sigma^2=\frac{(b-a)^2}{12}=\frac{1}{12}=0.0833.$$

The standard deviation is $\sigma=\sqrt{0.0833}=0.2887$.

Example 2—Uniform Distribution

Suppose that the download time of songs follows a uniform distribution between 2.5 and 6.5 seconds. The probability that the download time will be more than 5 seconds can be calculated as follows:

$$P[X\gt5]=1-P[X\le5]$$

The value can be calculated in *R* as follows:

Normal Distribution

The normal distribution is the most common continuous distribution in practice. The classic bell-shaped curve represents the normal distribution. The probability is calculated for values that occur within a certain range or interval as the area under the curve. A normal distribution is completely determined by the *mean* (μ) and the *standard deviation* (σ). Two normally distributed random variables having the same mean and standard deviation must have identical distributions.

A normal distribution is symmetric and centered around its mean, while the spread of the curve depends on the standard deviation. As the standard deviation increases, the distribution will be flatter and more spread out. The following figure shows three normal distributions with standard deviations of 0.5, 1, and 2, respectively, all centered around the mean equal to 0.

Similarly, the following figure shows three normal distributions with different means (-3, 0, and 6), but with the same standard deviation of 1.

In a normal distribution, almost all the possible observations of the random variable lie within three standard deviations of either side of the mean. The curve associated with the normal distribution satisfies the following properties:

- The curve is bell shaped
- The curve is centered around the mean (\(\mu\))
- The curve is close to the x-axis for values below \(\mu-3\sigma\) and for values above \(\mu+3\sigma\)

The probability density function of the normal random continuous variable \(X\) is:

$$f_X(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}},-\infty\lt x\lt\infty$$

for the given mean (\(\mu\)) and the standard deviation (\(\sigma\)). (\(e\approx2.718\) *and* \(\pi\approx3.142\))

Notation

`X ~ norm(mean=\(\mu,sd=\sigma\))`

The associated PDF and CDF functions in *R* are `dnorm(x, mean = 0, sd = 1) and pnorm(x, mean = 0, sd = 1)`, respectively.

Example—Normal Distribution

The gestation period for humans is normally distributed with a mean of \(266\) days and a standard deviation of \(16\) days. The normal curve for this distribution can be plotted in *R* as follows:

```
.
```

```
.
```

The cumulative probability of a baby being born within the mean value of \(266\) days is:

```
.
```

The above value shows that \(50\%\) of babies are born within the mean value.

Loading [Contrib]/a11y/accessibility-menu.js

The cumulative probability of a baby being born within less than the 3rd standard deviation, \(\mu-3\sigma\), (218 days) is:

The above value shows that \(0.135\%\) of babies are born with a gestation of \(218\) or fewer days.

The cumulative probability of a baby being born within \(3\) standard deviations from the mean, (\(\mu-3\sigma,\mu+3\sigma\)), (\(218\) days to \(314\) days) is:

The above value shows that \(99.73\%\) of babies are born within a gestation period of \(218\) days to \(314\) days.

The cumulative probability of a baby being born within \(2\) standard deviations from the mean, (\(\mu-2\sigma,\mu+2\sigma\)), (\(234\) days to \(298\) days) is:

The above value shows that \(95.45\%\) of babies are born within a gestation period of \(234\) days to \(298\) days.

Similarly, the cumulative probability of a baby being born within \(1\) standard deviation from the mean, (\(\mu-\sigma,\mu+\sigma\)), (\(250\) days to \(282\) days) is:

The above value shows that \(68.27\%\) of babies are born within a gestation period of \(250\) days to \(282\) days.

The CDF for the gestation period is plotted in *R* as follows:

Standard Normal Distribution

Given a normally distributed random variable X with mean μ and standard deviation σ, the standardized normal random variable Z, is derived as follows:

$$Z=\frac{X-\mu}{\sigma}$$

The above computes a Z value that shows the difference of the X value from the mean in units of the standard deviation. The standardizing converts all normal distributions into the standard normal distribution that always has the mean 0 and standard deviation 1. The standard normal curve (also referred to as the z-curve) is shown below.

The standard normal curve satisfies the following properties:

- The total area under the standard normal curve is 1
- The curve extends indefinitely in both directions along the x-axis
- The curve is symmetric around $x=0$
- Almost all of the area under the standard curve lies between -3 and 3
- 68.27% of the observations lie within one standard deviation of the mean $(-1,1)$
- 95.45% of the observations lie within two standard deviations of the mean $(-2,2)$
- 99.73% of the observations lie within three standard deviations of the mean $(-3,3)$

Notation

```
Z ~ norm(mean = 0, sd = 1)
```

The probability density function of the standard normal random variable Z is:

$$f_Z(z)=\phi(z)=\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}},-\infty\lt z\lt\infty$$

The cumulative distribution function of the standard normal random variable Z is:

$$F_Z(t)=\Phi(t)=\int^t_{-\infty}\phi(z)dz$$

The CDF for the standard normal random variable has the shape shown below.

The proportions of the observations falling within the three ranges around the mean can be calculated in *R* as follows:

Normal Quantiles

Given a number, the cumulative distribution function, `pnorm`, computes the probability that a normal random variable will be less than that number. The quantile function, `qnorm(x, mean=0, sd=1)`, does the reverse. Given the probability, the function returns the number whose cumulative distribution matches the probability. For a standard normal variable with mean \(0\) and standard deviation \(1\), the `qnorm` function takes the probability and returns the associated z-score.

Example

Suppose the scores on a test follow the normal distribution with a mean value of \(80\) and a standard deviation of \(5\). The minimum score for a student to be in the top \(5\%\) can be calculated as follows. \(P(X\lt0.95)\) is the area under the normal curve representing the bottom \(95\%\) of the class. The maximum value of \(X\) that satisfies the condition can be computed as shown below.

If the scores are rounded, a student should have \(88\) or above to be in the top \(5\%\) of the class. Similarly, to be in the top \(1\%\) of the class, the following function gives the minimum score as \(92\).

Generating Random Numbers with Normal Distribution

The `rnorm` function, `rnorm(n, mean = 0, sd = 1)`, generates \(n\) random numbers that follow the normal distribution with the corresponding mean and standard deviation.

The random values for \(20\) students can be generated with a mean value of \(80\) and standard deviation of \(5\) as follows:

The generated values, for say, \(1000\) numbers, can be rounded and plotted as shown below.

Exponential Distribution

The exponential distribution is a continuous distribution and ranges from zero to positive infinity. This distribution is commonly used in queuing theory for waiting time distributions, the length of time between arrivals, patients entering a hospital, etc.

The exponential distribution is defined by a single parameter, \(\lambda\), the mean number of arrivals per unit of time. The probability density function of the random variable \(X\) with exponential distribution is:

$$F_X(x)=\lambda e^{-\lambda x},x\gt0$$

The cumulative distribution function is:

$$F_X(t)=P(X\le t)=1-e^{-\lambda t},t\gt0$$

The mean of the random variable (the mean time between arrivals) is:

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

The variance of the time between arrivals is:

$$\sigma^2=\frac{1}{\lambda^2}$$

Notation

Loading [Contrib]/a11y/accessibility-menu.js

```
        X ~ exp(rate = \(\lambda\))
```

The associated PDF and CDF functions in *R* are `dexp(x, rate = 1)` and `pexp(x, rate = 1)`, respectively.

Example—Exponential Distribution

Suppose customers come to a bank at the rate of \(20\) per hour (\(\lambda=20\)). If a customer has just arrived, the probability that the next customer will arrive within \(1\) minute is calculated as follows. Since the arrival rate is per hour, \(1\) minute corresponds to \(\frac{1}{60}\) of the hour. Hence the required probability is:

$$P\left(t\le\frac{1}{60}\right)=1-e^{-20\cdot\frac{1}{60}}=1-e^{-\frac{1}{3}}=0.2835$$

The above value can be calculated in *R* as follows:

The probability that a customer will arrive within the next minute is \(28.35\%\).

The PDF for this distribution can be plotted in *R* as follows:

The CDF for this distribution is as follows:

Boston University Metropolitan College

Loading [Contrib]/a11y/accessibility-menu.js