Discrete Distributions

Now consider a series of *independent* experiments, each of which produces a Bernoulli random variable with probability of success p (p is the same for all of the trials)

Discrete Distributions

Now consider a series of *independent* experiments, each of which produces a Bernoulli random variable with probability of success p (p is the same for all of the trials)

The following discrete probability distributions arise from this model:

If the number of trials n is fixed in advance, the number of successes X has a **binomial** distribution

Discrete Distributions

Now consider a series of *independent* experiments, each of which produces a Bernoulli random variable with probability of success p (p is the same for all of the trials)

The following discrete probability distributions arise from this model:

If the number of trials n is fixed in advance, the number of successes X has a **binomial** distribution

If trials continue indefinitely until the first success is obtained, the number of failures obtained X has a **geometric** distribution.

Discrete Distributions

Now consider a series of *independent* experiments, each of which produces a Bernoulli random variable with probability of success p (p is the same for all of the trials)

The following discrete probability distributions arise from this model:

If the number of trials n is fixed in advance, the number of successes X has a **binomial** distribution

If trials continue indefinitely until the first success is obtained, the number of failures obtained X has a **geometric** distribution.

If trials continue indefinitely until the r^{th} success is obtained, the number of failures obtained X has a **negative binomial** distribution.

The Negative Binomial Distribution

The negative binomial experiment consists of:

- independent Bernoulli trials are performed until r successes are obtained
- The random variable X is the number of **failures** that occur before the r^{th} success.
- The probability of success p is the same for all trials

The Negative Binomial Distribution

The probability mass function is:

$$f(x) = nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x, \quad x = 0, 1, 2, 3, \ldots$$

The Negative Binomial Distribution

The probability mass function is:

$$f(x) = nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x, \quad x = 0, 1, 2, 3, \ldots$$

$$E(X) = \frac{r(1-p)}{p} \quad V(X) = \frac{r(1-p)}{p^2}$$

The Negative Binomial Distribution

The probability mass function is:

$$f(x) = nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x, \quad x = 0, 1, 2, 3, \ldots$$

$$E(X) = \frac{r(1-p)}{p} \quad V(X) = \frac{r(1-p)}{p^2}$$

Computation:

Value	R	Spreadsheet
$P(X = x)$	$dnbinom(x, n, p)$	$= NEGBINOMDIST(x, r, p)$
$P(X \leq x)$	$pnbinom(x, n, p)$	

The Negative Binomial Distribution

Note that the expected value and variance for the negative binomial distribution:

$$E(X) = \frac{r(1-p)}{p} \quad V(X) = \frac{r(1-p)}{p^2}$$

is r times the expected value and variance of the *geometric* distribution:

$$E(X) = \frac{1-p}{p} \quad V(X) = \frac{1-p}{p^2}$$

The Negative Binomial Distribution

Note that the expected value and variance for the negative binomial distribution:

$$E(X) = \frac{r(1-p)}{p} \quad V(X) = \frac{r(1-p)}{p^2}$$

is r times the expected value and variance of the *geometric* distribution:

$$E(X) = \frac{1-p}{p} \quad V(X) = \frac{1-p}{p^2}$$

This reflects the fact that the negative binomial can be viewed as the sum of r independent geometric experiments:

Conduct Bernoulli trials until the first success, and repeat this r times.

The Negative Binomial Distribution

Now we will perform some numerical experiments.

First generate a sample of $1,000,000$ observations for a negative binomial experiment with $r = 3$ and a probability of success $p = 0.4$ at each trial:

x<-rnbinom(1000000,3,0.4)

The Negative Binomial Distribution

Now we will perform some numerical experiments.

First generate a sample of $1,000,000$ observations for a negative binomial experiment with $r = 3$ and a probability of success $p = 0.4$ at each trial:

x<-rnbinom(1000000,3,0.4)

Now plot a histogram of the results:

hist(x)

The Negative Binomial Distribution

Now we will perform some numerical experiments.

First generate a sample of $1,000,000$ observations for a negative binomial experiment with $r = 3$ and a probability of success $p = 0.4$ at each trial:

x<-rnbinom(1000000,3,0.4)

Now plot a histogram of the results:

hist(x)

To get a table of the results enter

table(x)

The Negative Binomial Distribution

Now we will perform some numerical experiments.

First generate a sample of $1,000,000$ observations for a negative binomial experiment with $r = 3$ and a probability of success $p = 0.4$ at each trial:

x<-rnbinom(1000000,3,0.4)

Now plot a histogram of the results:

hist(x)

To get a table of the results enter

table(x)

The results through $X = 6$ should look something like:

0	1	2	3	4	5	
63784	115545	138570	138259	124481	103801	8349

The Negative Binomial Distribution

0	1	2	3	4	5	
63784	115545	138570	138259	124481	103801	834

Now compare the frequencies to the probabilities.

First compute the probability that $X = 0$:

dnbinom(0,3,0.4)

The Negative Binomial Distribution

0	1	2	3	4	5	
63784	115545	138570	138259	124481	103801	8349

Now compare the frequencies to the probabilities.

First compute the probability that $X = 0$:

dnbinom(0,3,0.4)

The result should be something like

[1]   0.064

The Negative Binomial Distribution

0	1	2	3	4	5	
63784	115545	138570	138259	124481	103801	8349

Now compare the frequencies to the probabilities.

First compute the probability that $X = 0$:

dnbinom(0,3,0.4)

The result should be something like

[1] 0.064

To get the probability that $X = 1$ enter

dnbinom(1,3,0.4)

The Negative Binomial Distribution

0	1	2	3	4	5	
63784	115545	138570	138259	124481	103801	8349

Now compare the frequencies to the probabilities.

First compute the probability that $X = 0$:

dnbinom(0,3,0.4)

The result should be something like

[1] 0.064

To get the probability that $X = 1$ enter

dnbinom(1,3,0.4)

This time the results should look something like:

[1] 0.1152

The Negative Binomial Distribution

0	1	2	3	4	5	
63784	115545	138570	138259	124481	103801	8349

Next compute the probability that $X = 2$:

dnbinom(2,3,0.4)

The Negative Binomial Distribution

0	1	2	3	4	5	
63784	115545	138570	138259	124481	103801	8349

Next compute the probability that $X = 2$:

dnbinom(2,3,0.4)

The result should be something like

[1]   0.13824

The Negative Binomial Distribution

0	1	2	3	4	5	
63784	115545	138570	138259	124481	103801	8349

Next compute the probability that $X = 2$:

dnbinom(2,3,0.4)

The result should be something like

[1]   0.13824

To get the probability that $X = 5$ enter

dnbinom(5,3,0.4)

The Negative Binomial Distribution

0	1	2	3	4	5	
63784	115545	138570	138259	124481	103801	8349

Next compute the probability that $X = 2$:

dnbinom(2,3,0.4)

The result should be something like

[1]   0.13824

To get the probability that $X = 5$ enter

dnbinom(5,3,0.4)

This time the results should look something like:

[1]   0.10451

The Negative Binomial Distribution

The expected value $E(X)$ in this case is:

$$E(X) = \frac{r(1-p)}{p} = \frac{3(.6)}{.4} = 4.5$$

The Negative Binomial Distribution

The expected value $E(X)$ in this case is:

$$E(X) = \frac{r(1-p)}{p} = \frac{3(.6)}{.4} = 4.5$$

To compute the sample mean $\overline{x}$, enter

mean(x)

The Negative Binomial Distribution

The expected value $E(X)$ in this case is:

$$E(X) = \frac{r(1-p)}{p} = \frac{3(.6)}{.4} = 4.5$$

To compute the sample mean $\overline{x}$, enter

mean(x) The result should be something like

[1] 4.499121

The Negative Binomial Distribution

The variance $V(X)$ in this case is:

$$V(X) = \frac{r(1-p)}{p^2} = \frac{3(.6)}{.4^2} = 11.25$$

The Negative Binomial Distribution

The variance $V(X)$ in this case is:

$$V(X) = \frac{r(1-p)}{p^2} = \frac{3(.6)}{.4^2} = 11.25$$

To compute the sample variance s^2, enter

var(x)

The Negative Binomial Distribution

The variance $V(X)$ in this case is:

$$V(X) \;=\; \frac{r(1-p)}{p^2} \;=\; \frac{3(.6)}{.4^2} \;=\; 11.25$$

To compute the sample variance s^2, enter

var(x) The result should be something like

[1]   11.2477

The Negative Binomial Distribution

A fair coin is tossed until the second heads comes up.

Find the probability that the second heads comes up on the fifth toss (x=3).

The Negative Binomial Distribution

A fair coin is tossed until the second heads comes up.

Find the probability that the second heads comes up on the fifth toss (x=3).

Solution: 0.125

The Negative Binomial Distribution

A fair coin is tossed until the second heads comes up.

Find the probability that the second heads comes up on the fifth toss (x=3).

Solution: 0.125

$dnbinom(3, 2, 0.5)$

The Geometric Distribution

A fair coin is tossed until the fourth heads comes up.

Find the probability that the fourth heads comes up on the seventh toss or sooner $x \leq 3$.

The Geometric Distribution

A fair coin is tossed until the fourth heads comes up.

Find the probability that the fourth heads comes up on the seventh toss or sooner $x \leq 3$.

Solution: 0.5

The Geometric Distribution

A fair coin is tossed until the fourth heads comes up.

Find the probability that the fourth heads comes up on the seventh toss or sooner $x \leq 3$.

Solution: 0.5

$pnbinom(3, 4, 0.5)$

The Negative Binomial Distribution

A fair coin is tossed until the fifth heads comes up.

Find the probability that this takes more than 8 tosses ($x > 3$)

The Negative Binomial Distribution

A fair coin is tossed until the fifth heads comes up.

Find the probability that this takes more than 8 tosses ($x > 3$)

Solution: 0.63672

The Negative Binomial Distribution

A fair coin is tossed until the fifth heads comes up.

Find the probability that this takes more than 8 tosses ($x > 3$)

Solution: 0.63672

$1 - pnbinom(3, 5, 0.5)$

The Negative Binomial Distribution

A fair coin is tossed until the third heads comes up.

Find the probability that this takes 9 or more tosses ($x > 5$)

The Negative Binomial Distribution

A fair coin is tossed until the third heads comes up.

Find the probability that this takes 9 or more tosses ($x > 5$)

Solution: 0.22656

The Negative Binomial Distribution

A fair coin is tossed until the third heads comes up.

Find the probability that this takes 9 or more tosses ($x > 5$)

Solution: 0.22656

$1 - pnbinom(4, 3, 0.5)$ or $= 1 - GEOMDIST(7, 0.5, TRUE)$

(Spreadsheet function is for GNUMERIC. EXCEL does not have this function)

The Negative Binomial Distribution

A baseball player has a $.300$ batting average.

Find the probability that their second hit in a game occurs on the 5^{th} time at bat. ($x = 3$)

The Negative Binomial Distribution

A baseball player has a $.300$ batting average.

Find the probability that their second hit in a game occurs on the 5^{th} time at bat. ($x = 3$)

Solution: 0.12348

The Negative Binomial Distribution

A baseball player has a $.300$ batting average.

Find the probability that their second hit in a game occurs on the 5^{th} time at bat. ($x = 3$)

Solution: 0.12348

$dnbinom(3, 2, 0.3)$