R doc: Discrete probability models

Josep L. Carrasco
Bioestadística. Departament de Fonaments Clínics
Universitat de Barcelona

Binomial

Example: In a population the 25% of people are smokers. A sample of 30 subjects were randomly selected from this population. The variable "count of smokers in the sample" is a random variable whose values are distributed as a Binomial model with parameters $n = 30$ and $p = 0.25$.

Probability function

The probability function of a binomial model can be computed using the *dbinom(x,size,prob)* function. The arguments are

- x: count of successes;
- size: number of trials:
- p: probability of a success

For example, let's compute the probability of exactly 10 smokers.

```
dbinom(10,30,0.25)
```

```
[1] 0.09086524
```

Distribution function

The distribution function is computed by applying the *pbinom(q,size,prob)* function. Where q stands for the value whose accumulated probability is computed.

The probability of less than 10 smokers is computed as

$P(X < 10) = P(X \leq 9)$

```
pbinom(9,30,0.25)
```

```
[1] 0.8034066
```

The probability of having less than 10 smokers but more than 3 is

$P\left(X > 3 \bigcap X < 10\right) = P\left(X \geq 4 \bigcap X \leq 9\right) = P\left(X \leq 9\right) - P\left(X \leq 3\right)$

```
pbinom(9,30,0.25)-pbinom(3,30,0.25)
```

```
[1] 0.7659573
```

It is also possible to compute quantiles using the *qbinom(p,size,prob)* function. Here p stands for cumulative probability.

For example, what are the quartiles?

```
qbinom(p=c(0.25,0.5,0.75),30,0.25)
```

```
[1] 6 7 9
```

Geometric

Example. Number of times a die is rolled until a "5" appears. This variable follows a geometric distribution with parameter $p = \frac{1}{6}$

Probability function

The probability function of a geometric model can be computed using the *dgeom(x,prob)* function. Where:

- x: count of failures;
- p: probability of sucess.
- Probability of 10 rolls. This means 9 failures plus one success $P\left(X = 9\right)$

```
dgeom(9,1/6)
```

```
[1] 0.03230112
```

Distribution function

The distribution function is computed by applying the *pgeom(q,prob)* function. Where q stands for the value whose accumulated probability is computed.

- Probability that "5" appears before 10 rolls. It involves less than 9 failures.

$P\left(X < 9\right) = P\left(X \leq 8\right)$

```
pgeom(8,1/6)
```

```
[1] 0.8061933
```

To compute the quantiles we can apply the *qgeom(p,prob)* function. For example, let's compute the median.

```
qgeom(0.5,1/6)
```

```
[1] 3
```

This result means the half of times it will be needed 3 failures as much before obtaining one success.

Negative Binomial

Example. A particular hardware is obsolete and must be replaced if it fails 3 times when using it. The probability of a failure every time is used is 0.1. If the failures are independent, the variable X="number of times that the hardware can be used before replacing it" follows a negative binomial distribution.

Note: the count variable in a negative binomial is the number of failures before r successes. However, in this example failure means "the hardware works" whilst success would be "the hardware fails".

Probability function

The probability function of a negative binomial model can be computed using the *dnbinom(x,size,prob)* function.

- Probability of 8 uses. That means the 3rd failure occurs at 11th try (8 uses working plus 3 fails) $P(X = 8)$

```
dnbinom(8,3,0.1)
```

```
[1] 0.01937102
```

Distribution function

The distribution function is computed by applying the *pnbinom(q,size,prob)* function.

- Probability of replacement before 10 uses. $P(X < 10) = P(X \leq 9)$

This means 9 uses plus 3 fails, so it involves 12 tries as much.

```
pnbinom(9,3,0.1)
```

```
[1] 0.11087
```

To compute the quantiles we can apply the *qnbinom(p,size,prob)* function. Let's compute the quantiles 20 and 80.

```
qnbinom(c(0.2,0.8),3,0.1)
```

```
[1] 13 39
```

Poisson

Example. Let X be the number of flaws per mm in a copper wire. Let's suppose that X follows a Poisson distribution with mean of 2.3.

Probability function

The probability function of a Poisson model is obtained using the *dpois(x,lambda)* function, where *lambda* stands for the mean of the variable.

- Probability of finding 2 flaws in 1mm of wire, $P(X = 2)$

```
dpois(2,2.3)
```

```
[1] 0.2651846
```

Distribution function

The distribution function is computed by applying the *ppois(x,lambda)* function.

- Probability of finding one flaw as much in 1mm of wire

```
ppois(1,2.3)
```

```
[1] 0.3308542
```

We can use the function *qpois(p,lambda)* to compute the quantiles. For example, with a probability of 95%, how many flaws will be found in each mm of wire as much?

```
qpois(0.95,2.3)
```

```
[1] 5
```