Discrete Probability Distributions with R

Marianne Huebner, Michigan State University

2015-05-01

### Discrete probability distribution

A discrete random variable has probability distribution

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| x | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| P(X=x) | 0.1 | 0.1 | 0.4 | 0.3 | 0.05 | 0.03 | 0.02 |

Then we can calculate the expectation and cumulative distribution function in R as follows.

x<-c(1,2,3,4,5,6,7) # values of the random variable X  
p<-c(0.1,0.1,0.4,0.3,0.05,0.03, 0.02); # P(X=x)  
mu<-sum(x\*p); mu # expectation

## [1] 3.27

sqrt(sum(p\*(x-mu)^2)) # standard deviation

## [1] 1.231706

cumsum(p) # cumulative distribution

## [1] 0.10 0.20 0.60 0.90 0.95 0.98 1.00

= 0.6 using sum(p[1:3]).

= 0.4 using p[3].

= 0.8 using sum(p[2:4]).

### Binomial distribution

Suppose there are 10 multiple choice questions. Each question has 5 choices, and only one of these is correct. What is the probability getting two or fewer correct answers, if you guess each answer at random.

pbinom(2, size=10, prob=0.2)

## [1] 0.6777995

This is the *cumulative distribution function.*

What is the probability of getting seven or more correct?

1-pbinom(6, size=10, prob=0.2)

## [1] 0.0008643584

What is the probability of getting exactly two answers correct?

dbinom(2, size=10, prob=0.2)

## [1] 0.3019899

This is the *probability density function.*

Calculate the probability of getting two or fewer answers correct with the probability density function.

dbinom(0,10,0.2) +dbinom(1,10,0.2)+dbinom(2,10,0.2)

## [1] 0.6777995

# or  
sum(dbinom(0:2,10,0.2))

## [1] 0.6777995

Calculate for ~ Binomial(30, 0.5)

sum(dbinom(6:9, 30, 0.5))

## [1] 0.02122452

*Inverse cumulative distribution function*

qbinom(0.5,size=15, prob=0.2)

## [1] 3

There is at least a 50% chance of getting 3 or fewer successes in 15 trials with success probability 0.2. Check

pbinom(3, size=15, prob=0.2)

## [1] 0.6481621

pbinom(2, size=15, prob=0.2)

## [1] 0.3980232

### Poisson distribution

A university police department writes, on average, five tickets per day. Suppose the number of tickets written per day follows a Poisson distribution with a mean of 8.8 tickets per day. Find the probability that less than six tickets are written on a randomly selected day from this distribution. Note: All these numbers are made up.

*Cumulative distribution function.*

ppois(5, lambda=8.8)

## [1] 0.1283866

What is the probability of writing seven or more tickets?

1-ppois(6, lambda=8.8)

## [1] 0.7743897

*Probability density function.*

dpois(6, lambda=8.8)

## [1] 0.09722369

Calculate the probability of writing two or fewer tickets with the probability density function.

dpois(0, lambda=8.8) +dpois(1, lambda=8.8)+dpois(2, lambda=8.8)

## [1] 0.007313569

# or  
sum(dpois(0:2,8.8))

## [1] 0.007313569

Calculate P(5< X <10) for X ~ Poisson(lambda=8.8)

sum(dpois(6:9, 8.8))

## [1] 0.485353

*Inverse cumulative distribution function.*

qpois(0.2,lambda=8.8)

## [1] 6

There is at least a 20% chance that the police department issued 6 or fewer tickets. Check

ppois(6, lambda=8.8)

## [1] 0.2256103

ppois(5, lambda=8.8)

## [1] 0.1283866