R Tutorial

Discrete Distributions in R

Erik Hintz

erik.hintz@uwaterloo.ca
Department of Statistics and Actuarial Science
University of Waterloo

Agenda

- Accessing the probability function of a random variable
- Accessing the cumulative distribution function of a random variable
- Computing probabilities

Main syntax

- R comes with the pf/cdf (and more!) of many important distributions (`binom`, `nbinom`, `pois`, `geo`, `hyper`, ...)
- The basic syntax for a distribution with name `fun` (e.g., `binom`),
  - `dfun()`: probability function $P(X = x)$; e.g. `dbinom(5, size = 16, prob = 0.5)` gives $P(X = 5)$ for $X \sim Bin(16, 0.5)$;
  - `pfun()`: cumulative distribution function $P(X \leq x)$; e.g., `pbinom(5, size = 16, prob = 0.5)` gives $P(X \leq 5)$ for $X \sim Bin(16, 0.5)$;
  - `rfun()`: returns a random sample $x_1, \ldots, x_n$ from the distribution; e.g., `rbinom(5, size = 16, prob = 0.5)` returns a vector with 5 independent realizations of $X \sim Bin(16, 0.5)$;
  - `qfun()` quantile function of X.

Vectorization

- The probability function `dfun()`, cumulative distribution function `pfun()` and the quantile function `qfun()` are vectorized, i.e., if called with a vector input, they return a result vector of the same length consisting of the componentwise results.
- For instance,
  `dbinom(c(3, 5), size = 16, prob = 0.5)`
  is equivalent to `c(dbinom(3, size = 16, prob = 0.5), dbinom(5, size = 16, prob = 0.5))`
- This can be helpful in computing probabilities, e.g.,
  `sum(dbinom(seq(from = 2, to = 16, by = 2), size = 16, prob = 0.5))`
  gives the probability Michael Scott flips an even number of heads in 16 trials!

Exercise 2

Use R to solve the following questions:

a) Each day, Jim rolls a die to determine how many pranks he plays on Dwight the next day. What's the probability at least on 3 days of a 5-day week, he's playing Dwight at least 5 pranks?

b) Suppose Angela is calling the cat tv show until she gets in. From past experience, she knows that each time she calls, there's a 0.01 probability that she needs to call between 50 and 100 times before she finally gets in.

c) Plot the probability function and the cumulative distribution function of a Poisson distribution with rate $\lambda = 2$.