

Distribuições discretas na linguagem R

Várias distribuições de probabilidade discretas são parte de um dos pacotes básicos da linguagem R (pacote `stats`). Em geral, se X denota uma variável aleatória, para uma distribuição qualquer identificada por `nome`, existem quatro funções, brevemente descritas abaixo.

(i) `dnome`: se a distribuição for contínua, significa a função densidade de probabilidade $f(x)$; se a distribuição for discreta, significa a função massa de probabilidade $p(x) = P(X = x)$.

(ii) `pnome`: função distribuição acumulada $F(x) = P(X \leq x)$.

(iii) `qnome`: função quantil. De forma simplificada, para $0 < u < 1$, a função retorna o valor x correspondente a $F(x) = u$.

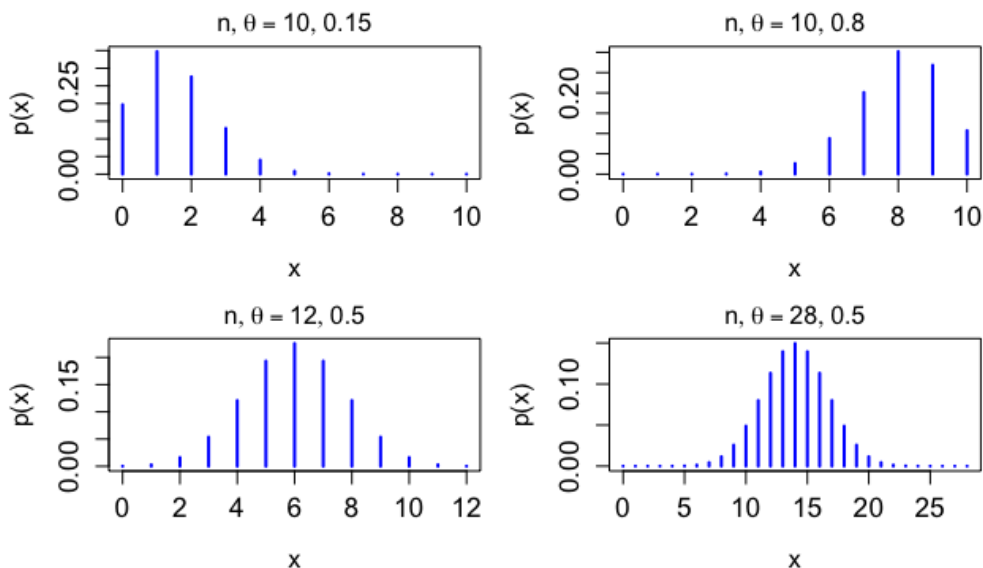
(iv) `rnome`: geração de observações da distribuição.

1. Binomial

Sintaxe básica: `dbinom(x, n, teta)`

```
n <- 10
teta <- 0.15
curve(dbinom(x, n, teta), from = 0, to = n, n = n + 1, type = "h",
      xlab = "x", ylab = "p(x)", col = "blue", lwd = 2,
      cex.main = 1.2, cex.axis = 1.3, cex.lab = 1.2,
      main = substitute(list(n, theta) == list(v1, v2),
                        list(v1 = n, v2 = teta)))
```

Nota 1. Procure entender todos os argumentos do comando acima.



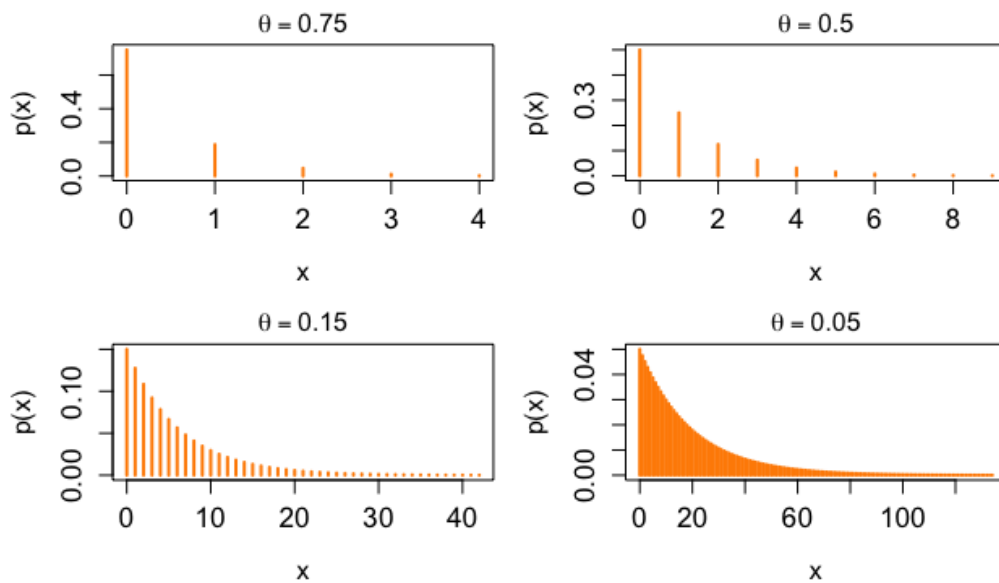
2. Geométrica

Sintaxe básica: `dgeom(x, teta)`

```
teta <- 0.75
xmax <- qgeom(0.999, teta)
```

Na linha acima, `xmax` representa o valor máximo da variável no gráfico e corresponde à probabilidade acumulada de 0,999 (é o quantil de 99,9% da distribuição).

```
curve(dgeom(x, teta), from = 0, to = xmax, n = xmax + 1, type = "h",
      lwd = 2, xlab = "x", ylab = "p(x)", col = "darkorange",
      main = substitute(list(theta) == list(v1), list(v1 = teta)),
      cex.main = 1.2, cex.axis = 1.3, cex.lab = 1.2)
```

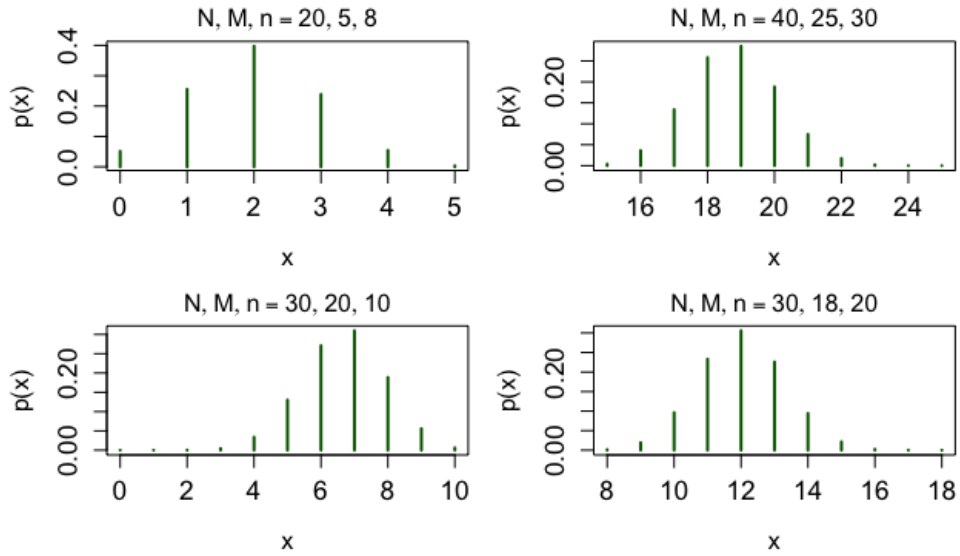


3. Hipergeométrica

Sintaxe básica: `dhyper(x, M, N - M, n)`

```
N <- 20
M <- 5
n <- 8
xmin <- max(0, n - N + M)
xmax <- min(n, M)
```

```
curve(dhyper(x, M, N - M, n), from = xmin, to = xmax,
      n = xmax - xmin + 1, lwd = 2, type = "h", xlab = "x",
      ylab = "p(x)", col = "darkgreen", cex.lab = 1.2, cex.axis = 1.3,
      main = substitute(list(N, M, n) == list(v1, v2, v3),
      list(v1 = N, v2 = M, v3 = n)), cex.main = 1.2)
```

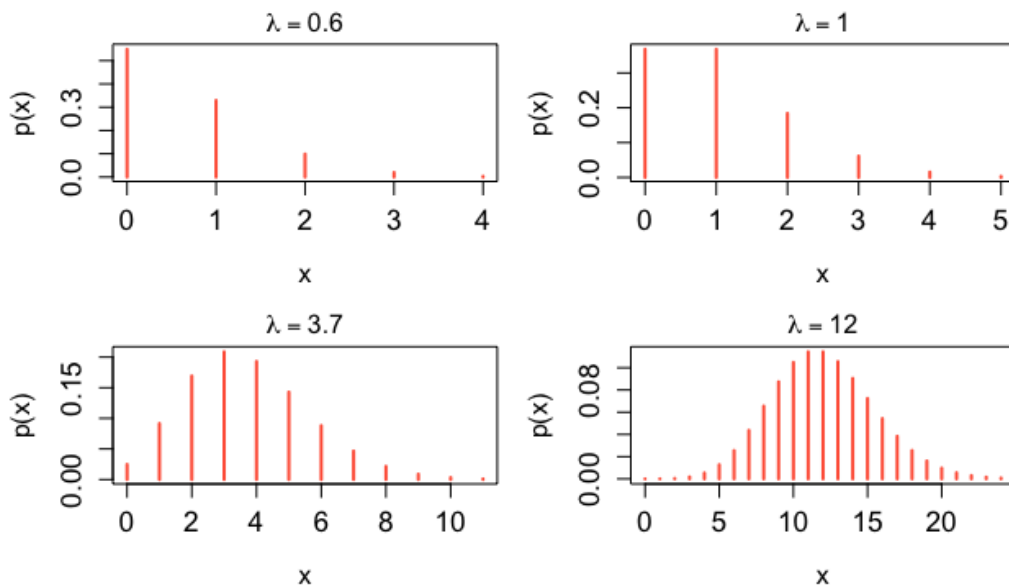


4. Poisson

Sintaxe básica: `dpois(x, lambda)`

```
lambda <- 0.6
xmax <- qpois(0.999, lambda)
```

```
curve(dpois(x, lambda), from = 0, to = xmax, n = xmax + 1, type = "h",
      lwd = 2, xlab = "x", ylab = "p(x)", col = "tomato",
      main = substitute(list(lambda) == list(v1), list(v1 = lambda)),
      cex.main = 1.2, cex.lab = 1.2, cex.axis = 1.3)
```



Nota 2. Utilizando as funções `rbinom`, `rgeom`, `rhyper` e `rpois`, obtenha observações (100, digamos) das distribuições dos itens 1 a 4.

Nota 3. Utilizando as funções `pbinom`, `pgeom`, `phyper` e `ppois`, apresente o gráfico da função distribuição acumulada das distribuições dos itens 1 a 4.

Nota 4. Em linguagem R, escreva a função massa de probabilidade da variável aleatória do exercício 3 da lista 2 (<http://wiki.icmc.usp.br/images/3/33/Lista2SME0800-2018.pdf>).

Resolva o item (b) deste exercício com a função que você escreveu e apresente soluções aproximadas para os itens (a) e (c).

Nota 5. Diversas distribuições de probabilidade, tanto discretas quanto contínuas, estão disponíveis na linguagem R e estão descritas na página *CRAN Task View: Probability Distributions* (<https://cran.r-project.org/web/views/Distributions.html>).