Estatística Computacional

Patrícia de Siqueira Ramos

PPGEAB
UNIFAL-MG

2 de Maio de 2018

Métodos para geração de variáveis aleatórias não uniformes

- Transformação inversa (baseado no teorema da probabilidade integral)
- Aceitação-Rejeição
- Composição
- Convolução
- Caracterização

Métodos para geração de variáveis aleatórias não uniformes

- Realizações de v.a.s de outras distribuições de probabilidade são obtidas pela transformação de sequências de números aleatórios uniformes entre 0 e 1, ou seja, U(0,1)
- Se tivermos números aleatórios uniformes i.i.d. podemos gerar números de muitas outras distribuições

Geração de realizações de v.a.s usando o R

- O R consegue gerar números de várias distribuições de probabilidade
- Basta acrescentarmos os prefixos:
  "d" para *density function* (densidade)
  "p" para *c.d.f.* (função distribuição)
  "q" para quantis
  "r" para *random generation* (realizações de v.a.s)

Algumas das distribuições disponíveis no R

Distribuição	Nome	Parâmetros
beta	beta	shape1, shape2, ncp
binomial	binom	size, prob
Cauchy	cauchy	location, scale
χ^2	chisq	df, ncp
exponencial	exp	rate
F	f	df1, df1, ncp
gama	gamma	shape, scale
geométrica	geom	prob
hipergeométrica	hyper	m, n, k
lognormal	lnorm	meanlog, sdlog
logística	logis	location, scale
binomial neg.	nbinom	size, prob
normal	norm	mean, sd
Poisson	pois	lambda
t de Student	t	df, ncp
uniform	unif	min, max
Weibull	weibull	shape, scale

Método da transformação inversa

Método da transformação inversa

O método da transformação inversa é baseado no resultado conhecido a seguir:

Teorema fundamental da transformação de probabilidades: Sejam U uma variável aleatória uniforme $U(0,1)$ e X uma variável aleatória com densidade $f_X(x)$ e função de distribuição $F_X(x)$ contínua e invertível, então $X = F_X^{-1}(U)$ possui densidade $f_X(x)$, sendo F^{-1} a função inversa de F.

Método da transformação inversa

Uma forma mais simples do teorema é:

Teorema fundamental da transformação de probabilidades: Se X é uma variável aleatória contínua com função de distribuição $F_X(x)$, então $U = F_X(x) \sim U(0, 1)$.

Método da transformação inversa

Uma prova possível do teorema é:

Assuma F_X contínua e crescente. Defina $U = F_X(x)$ e, sendo $U \sim U(0,1)$, U assume valores no intervalo $[0,1]$. Então:

$$F_U(x) = P(F_X(x) \leq x) = P(X \leq F_X^{-1}(x)) = F_X(F_X^{-1}(x)) = x.$$

Note que $F(F^{-1}(x)) = F^{-1}(F(x)) = x$.

Método da transformação inversa

- $F_X^{-1}(U)$ possui a mesma distribuição de X
- Então, para gerar uma realização de uma v.a. X, deve-se primeiro gerar um valor u a partir de $U(0,1)$ e obter $F_X^{-1}(u)$
- O método é simples de ser aplicado se F_X^{-1} for fácil de ser obtida
- Pode ser utilizado para variáveis discretas e contínuas (no caso discreto deverá ser feita uma adaptação)
- Algumas distribuições para as quais é possível aplicar o método: exponencial, Weibull, logística, Cauchy
- Para outras, é mais eficiente utilizar outros métodos

Método da transformação inversa - caso contínuo

Algoritmo:

1. Obter a função inversa $F_X^{-1}(u)$
2. Escrever uma função/comando para obter $F_X^{-1}(u)$
3. Para cada valor a ser gerado:
   - Gerar uma observação u de $U(0,1)$
   - Aplicar $x = F_X^{-1}(u)$

Método da transformação inversa - caso contínuo

Exemplo 1: Usar o método da transformação inversa para simular uma amostra aleatória da distribuição $f_X(x) = 3x^2$, $0 < x < 1$.

Resp.: $F_X(x) = x^3$, $0 < x < 1$ e $F_X^{-1}(u) = u^{1/3}$.

Então é preciso gerar uma sequência de números uniformes, o vetor:

$$\mathbf{u} = u_1, u_2, \ldots, u_n,$$

aplicar $u^{1/3}$ e essa será a amostra:

$$x_1, x_2, \ldots, x_n$$

da distribuição desejada $f_X(x)$.

Método da transformação inversa - caso contínuo

Exemplo 1 no R: O histograma e o gráfico esperado para $f_X(x)$ concordam?

```
n = 1000
u = runif(n)  # geração do vetor u
x = u ** (1/3)  # transformação
# histograma da amostra
hist(x, prob=TRUE, main=expression(f(x)==lambda*exp(-lambda*x)))
lines(density(x))   # estimação kernel

# para gerar o gráfico da função teórica
y = seq(0, 1, .01)
# curva teórica
lines(y, 3*y**2, col='red')
```

Método da transformação inversa - caso contínuo

Exemplo 2: Usar o método da transformação inversa para simular uma amostra aleatória da distribuição exponencial. Se $X \sim exponencial(\lambda)$:

$$f_X(x; \lambda) = \lambda e^{-\lambda x}, \qquad x > 0, \lambda > 0.$$

Método da transformação inversa - caso contínuo

Exemplo 2: Usar o método da transformação inversa para simular uma amostra aleatória da distribuição exponencial. Se $X \sim exponencial(\lambda)$:

$$f_X(x; \lambda) = \lambda e^{-\lambda x}, \qquad x > 0, \lambda > 0.$$

Resp.: $F_X(x) = 1 - e^{-\lambda x}$ e $F_X^{-1}(u) = -\frac{1}{\lambda} \ln(1 - u)$.

Dada a simetria da distribuição uniforme, podemos trocar $1 - u$ por u.

Dica: página que pode ajudar na obtenção de integrais (e outras funções): www.wolframalpha.com

Método da transformação inversa - caso contínuo

Exemplo 2: Para a exponencial,

a) Faça o mesmo que foi feito para o exemplo 1 apenas trocando a expressão para obter a amostra $\mathbf{x}$. Usar $\lambda = 2$, por exemplo.

b) Compare o resultado obtido (histograma e estimação kernel) com o teórico para a distribuição exponencial, como no exemplo 1.

c) Crie uma função no `R` chamada `rexpon(n, lamb)` para gerar uma amostra aleatória da distribuição exponencial usando o método da transformação inversa e retorne os valores gerados (e histogramas) usando $\lambda = 2$ e $\lambda = 5$.

d) Sabendo que a média da exponencial é $1/\lambda$, compare os valores teóricos das médias para $\lambda = 2$ e $\lambda = 5$ com os valores obtidos com as amostras da letra c.

Método da transformação inversa - caso discreto

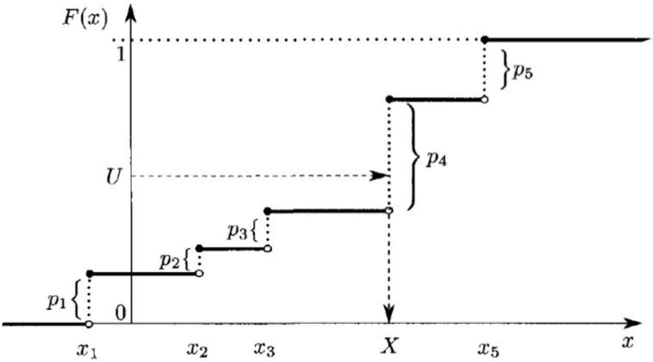
- Pode-se utilizar o método da transformação inversa também para distribuições discretas
- Se X é uma v.a. discreta e

$$\cdots < x_{i-1} < x_i < x_{i+1} < \ldots$$

sejam os pontos de descontinuidade de $F_X(x)$, então a transformação inversa é $F_X^{-1}(u) = x_i$, em que $F_X(x_{i-1}) < u \leq F_X(x_i)$

Método da transformação inversa - caso discreto

1. Gerar uma observação u de $U(0,1)$
2. Encontrar o menor inteiro tal que $F(x_{i-1}) < u \leq F(x_i)$ e retornar $x = x_i$



Método da transformação inversa - caso discreto

Ex. 3: Considere o caso da distribuição Bernoulli:

$$f_X(x;p) = p^x(1-p)^{1-x}, \qquad x \in \{0,1\}, \qquad 0 \leq p \leq 1$$

$$f_X(x;p) = \begin{cases} 1-p, & x = 0 \\ p, & x = 1 \end{cases}$$

$$\mu = p$$
$$\sigma^2 = p(1-p)$$

Sendo $p = 0,40$, por exemplo:

x	0	1
$f(x)$	0,60	0,40

Método da transformação inversa - caso discreto

Ex. 3: Distribuição Bernoulli:

$$F_X(x;p) = \begin{cases} 0, & x < 0 \\ 1-p, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

- Nesse exemplo, $F_X(0) = f_X(0) = 1 - p$ e $F_X(1) = 1$
- Então, $F_X^{-1}(u) = 1$ se $u > 0,6$ e
- $F_X^{-1}(u) = 0$ se $u \leq 0,6$

Método da transformação inversa - caso discreto

Ex. 3: Distribuição Bernoulli:

$$F_X(x;p) = \begin{cases} 0, & x < 0 \\ 1-p, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

- Nesse exemplo, $F_X(0) = f_X(0) = 1 - p$ e $F_X(1) = 1$
- Então, $F_X^{-1}(u) = 1$ se $u > 0,6$ e
- $F_X^{-1}(u) = 0$ se $u \le 0,6$

- Assim, gerar valores da $U(0,1)$ e:
  - se $u \le 1 - p = 0,6$ será considerado 0
  - se $u > 1 - p = 0,6$, será considerado 1 (fazer o gráfico da F da Bernoulli).

Método da transformação inversa - caso discreto

Ex. 3: Bernoulli

```
n = 10   # amostra pequena
p = 0.4
u = runif(n)
u > 0.6  # visualizar o vetor lógico
x = as.integer(u > 0.6)
x
mean(x)  # teórico: mu = p = 0,4
var(x)  # teórico: sigma**2 = p(1-p) = 0,4*0,6 = 0,24

hist(x, breaks=c(-0.5, 0.5, 1.5), prob=T)
# para gerar o gráfico da função do R
y = rbinom(n, size=1, prob=p)
hist(y, breaks = c(-0.5, 0.5, 1.5), prob = T)
# amostra maior
n = 1000
```

Método da transformação inversa - caso discreto

Ex. 4: Considere a seguinte distribuição de probabilidade discreta:

x	1	2	3	4	5
$f(x)$	0,20	0,30	0,10	0,05	0,35

Método da transformação inversa - caso discreto

Ex. 4: Considere a seguinte distribuição de probabilidade discreta:

x	1	2	3	4	5
$f(x)$	0,20	0,30	0,10	0,05	0,35

```
n = 1000
p = c(0.2, 0.3, 0.1, 0.05, 0.35)
Fx = cumsum(p)
u = runif(n)
x = 1:n
for(i in 1:n){
  indice = 1
  while(u[i] > Fx[indice]){
    indice = indice + 1
  }
  x[i] = indice
}
hist(x, breaks = seq(0.5, 5.5))
hist(x, breaks = seq(0.5, 5.5), prob = T)
```

Método da aceitação/rejeição

Método da aceitação/rejeição

- Como vimos, obter uma fórmula explícita para F^{-1} a partir da função de distribuição de onde queremos gerar realizações de v.a.s nem sempre é possível
- Mesmo se for possível, podem existir alguns métodos mais eficientes para isso
- Um desses métodos é o da aceitação/rejeição ou, simplesmente, método da rejeição

Método da aceitação/rejeição

- Começamos assumindo que a F que queremos simular tem f.d.p. f (caso contínuo, o discreto é similar)
- A ideia é encontrar uma distribuição alternativa G, com função de densidade g que:
  - já tenhamos algum método eficiente de geração de v.a.s de g
  - g deve ser próxima de f
- Geralmente, assume-se que a razão $f_X(x)/g_X(x)$ é limitada por uma constante $c > 0$, ou seja,

$$\sup_x \{f(x)/g(x) \leq c\}$$

- Na prática, queremos c o mais próximo de 1 possível

Método da aceitação/rejeição

Algoritmo:

1. Gerar uma realização y de G
2. Gerar u de $U(0,1)$
3. Se
$$u \leq \frac{f(y)}{cg(y)},$$
aceitar y e fazer $x = y$, senão rejeitar y e repetir os passos

Método da aceitação/rejeição

- Assim,

$$P(\text{aceitar}|Y) = P\left(U \leq \frac{f(y)}{cg(y)} | Y\right) = \frac{f(y)}{cg(y)}.$$

  A última igualdade vale porque estamos avaliando a função distribuição de U, que é uma $U(0,1)$.

- Dessa forma,

$$0 < \frac{f(y)}{cg(y)} \leq 1.$$

Método da aceitação/rejeição

- A probabilidade total de aceitação para qualquer iteração é

$$\int P(\text{aceitar}|y)P(Y=y) = \int \frac{f(y)}{cg(y)} g(y) = \frac{1}{c}.$$

- O número de vezes N que os passos devem ser repetidos (número de iterações para gerar X) é uma v.a. que segue a distribuição geométrica com probabilidade de sucesso $p = \dfrac{1}{c}$
- Como $E(N) = 1/p$, em média, cada valor amostral de X requer c iterações

Método da aceitação/rejeição

Resumindo:

- Devemos escolher uma função g que minimize a constante c
- A melhor função seria $g(x) = f(x)$, que não é o que queremos, pois procuramos uma função g alternativa a f que seja fácil de simular
- Requer habilidade encontrar a função g

Método da aceitação/rejeição

Ex. 5: Considere o caso da distribuição beta. Em média, quantos números deverão ser simulados para gerar 1000 realizações da *beta*$(\alpha = 2, \beta = 2)$, usando o método da aceitação/rejeição?

Método da aceitação/rejeição

Ex. 5: Considere o caso da distribuição beta. Em média, quantos números deverão ser simulados para gerar 1000 realizações da $beta(\alpha = 2, \beta = 2)$, usando o método da aceitação/rejeição?

Resp.: Sabemos que esse valor dependerá da constante c que, por sua vez, depende da escolha de g. A f.d.p. da $beta(\alpha = 2, \beta = 2)$ é

$$f(x) = 6x(1-x), \qquad 0 < x < 1.$$

Método da aceitação/rejeição

Ex. 5 (cont.): Para visualizar o gráfico da f.d.p. $beta(2, 2)$:

```
y = seq(0, 1, length=100)
plot(y, 6*y*(1 - y), type='l')

# ou usando a própria densidade beta do R
curve(dbeta(x, 2, 2), col='red')
```

Método da aceitação/rejeição

Ex. 5 (cont.): Para visualizar o gráfico da f.d.p. *beta*$(2, 2)$:

```
y = seq(0, 1, length=100)
plot(y, 6*y*(1 - y), type='l')

# ou usando a própria densidade beta do R
curve(dbeta(x, 2, 2), col='red')
```

Vamos considerar g como sendo a $U(0, 1)$ e $c = 1,5$. Então y será aceita se

$$u \leq \frac{f(y)}{cg(y)} = \frac{6y(1-y)}{1,5 \cdot 1} = 4y(1-y)$$

Método da aceitação/rejeição

Ex. 5 (cont.): Então, na média, $c * n = 1500$ iterações (e 3000 números aleatórios) serão necessárias para gerar uma amostra de tamanho 1000.

```
# gerar n observações da B(2,2) usando o método da rejeição
# função g é a U(0,1) e c = 1,5
n = 1000
k = 0      # contador para aceitação
j = 0      # número de iterações
x = numeric(n)
while (k < n) {
  u = runif(1)
  j = j + 1
  y = runif(1)  # v.a. de g
  if (u <= 4*y*(1-y)) {    # aceitamos y
    k = k + 1
    x[k] = y
  }
}
j  # iterações necessárias para n
hist(x, prob = T)
```

Método da aceitação/rejeição

```
# densidade obtida e a amostra gerada pelo R
lines(density(x), col='blue')
lines(density(rbeta(n,2,2)), col='red')

# comparar quantis empíricos e teóricos
p = seq(.1, .9, .1)
Qhat = quantile(x, p)   # quantis amostrais
Q = qbeta(p, 2, 2)      # quantis teóricos
round(rbind(Qhat, Q), 3)
```

Método da aceitação/rejeição

O p-ésimo percentil $(0 < p < 1)$ de uma amostra de tamanho n é assintoticamente distribuído como uma normal com média igual ao p-ésimo percentil da distribuição de origem e com variância igual a

$$V[X_p] = \frac{p(1-p)}{n \cdot f(Q_p)^2}.$$

Assim, o erro padrão da estimativa do percentil é

$$EP[X_p] = \frac{\sqrt{p(1-p)}}{\sqrt{n} \cdot f(Q_p)}.$$

Método da aceitação/rejeição

```
# incluir o erro padrão das estimativas dos quantis
ep = sqrt(p * (1 - p)) / (sqrt(n) * dbeta(Q, 2, 2))
round(rbind(Qhat, Q, ep), 3)
```

Repetir para $n = 10000$ e ver que as estimativas ficam mais precisas.

Exercício

Suponha que queremos gerar uma v.a. gama(α, β). A função de distribuição da gama é

$$F_X(x; \alpha, \beta) = \int_0^x \frac{\beta e^{-\beta t}(\beta t)^{\alpha-1}}{(\alpha-1)!} dt.$$

Não é possível obter uma expressão fechada para sua inversa. Mas podemos usar o seguinte resultado:

Uma v.a. $X \sim$ gama(α, β) $=$ soma de α v.a.s exponenciais independentes, exp(β).

O que precisamos fazer é:

1. Gerar α números aleatórios $u_1, u_2, \ldots, u_\alpha$
2. Fazer $x = -\frac{1}{\beta} \ln u_1$ - ... $-\frac{1}{\beta} \ln u_\alpha = -\frac{1}{\beta} \ln(u_1 \ldots u_\alpha)$, pois $\sum^n \ln x_i = \ln(x_i \ldots x_n)$

Então, usando esses passos, gere uma amostra de tamanho $n = 1000$ da gama($\alpha = 2$, $\beta = 2$), gere o histograma, a estimação kernel e compare com a distribuição teórica (gráfico, média e variância).