

Geração de números aleatórios

Matheus Victal Cerqueira

15/05/2021

Números pseudoaleatórios

Um processo pseudoaleatório se trata de um processo que parece ser aleatório mas na realidade é determinístico. Sequências pseudoaleatórias em geral apresentam *aleatoriedade estatística*, quando na realidade, o processo que as gerou é inteiramente determinístico. Tais números são de mais fácil geração do que números genuinamente aleatórios e também possuem a vantagem de poderem ser repetidos para inspeção de software a partir de uma *seed*.

Geração de amostras aleatórias

A geração de amostras aleatórias de uma população com uma distribuição de interesse por vias computacionais é de grande interesse para simulações computacionais. Por exemplo, podemos querer gerar um grafo cuja distribuição de grau respeite um modelo exponencial, ou um modelo normal, dependendo da aplicação. Extremamente importante em simulações estocásticas, poder gerar amostras de uma população alvo aumenta muito a possibilidade de investigação estatística de forma empírica. Quando soluções analíticas são de difícil obtenção para o problema em questão, a geração de amostras aleatórias é uma ferramenta poderosa na compreensão de mecanismos probabilísticos em problemas reais.

Entre os métodos de geração de amostras aleatórias mais utilizados, tem-se o *método da inversão* e o *método da rejeição*, os quais serão tratados aqui com mais detalhes.

Método da Inversão (ou Método da Transformação):

O método da inversão consiste na utilização da função de densidade acumulada (f.d.a.) da variável aleatória de interesse para a obtenção de uma amostra aleatória. Primeiramente, a f.d.a. $F(x)$ da variável aleatória alvo tem de ser conhecida e a obtenção de sua função inversa generalizada deve ser possível. A função inversa generalizada de $F(x)$ é dada por:

$$F^{-}(x) = \inf\{t \in \mathbb{R} : F(t) \geq x\}, \text{ para } x \in [0, 1]$$

Obs.: $F(x)$ contínua e estritamente crescente $\Rightarrow F^{-}(x) = F^{-1}(x), \forall x \in [0, 1]$

Com os conceitos acima, pode-se gerar amostras aleatórias de uma variável aleatória de interesse X a partir do seguinte resultado:

Resultado: Se U_1, \dots, U_n é uma a.a. da distro. $U(0, 1)$, então $X_1 = F^{-}(u_1), \dots, X_n = F^{-}(u_n)$ é uma a.a. da v.a. X cuja f.d.a. é $F(x)$.

Dessa forma, é possível gerar a.a. de qualquer distribuição com $F^{-}(x)$ obtível e conhecida utilizando-se de um gerador de amostras aleatórias de uma distribuição uniforme implementado. A esse procedimento, dá-se o nome de método da inversão.

Método da Rejeição

O método da aceitação-rejeição foi proposto por von Neumann em 1951 e é extretamente utilizado no contexto de aplicações simulacionais até hoje. Muitas vezes, o método da inversão não é aplicável pelo simples fato de não conhecermos $F^-(x)$ para a variável aleatória de interesse, ou ela não ser obtível. Digamos que o objetivo seja obter uma a.a. de uma variável aleatória X com distribuição $f(x)$ a qual se enquadre na situação descrita, o método da rejeição tem seu conceito base na obtenção de uma função de distribuição $g(y)$ de uma v.a. Y que respeite o seguintes critério:

- $f(x) \leq Mg(x), \forall x$ no domínio de f e g (repare que tais funções devem ter o mesmo domínio) em que $1 < M < \infty$.

Assim, dizemos que $Mg(x)$ envelope $f(x)$. O algoritmo para o método da rejeição segue os seguintes passos:

1. Gerar uma observação y de Y ;
2. Gerar uma observação u de $U(0, 1)$;
3. Se $u \leq \frac{f(x)}{Mg(x)}$, y é tomada como uma observação x da v.a. X , c.c., rejeita-se y e retorna-se ao passo 1;
4. Repetir os passos de 1 a 4 até a obtenção de n observações de x , que irão compor a amostra aleatória de X .

Recomenda-se tomar $M = \max \frac{f(x)}{g(x)}$, para algum x . Qualquer outro $M' > M$ poderia se utilizado.

Propriedade: O número de tentativas até obter um valor da distribuição de X tem distribuição geométrica com média igual a M . A probabilidade p de aceitação é $1/M$.

O método baseia-se na propriedade de que a distribuição da variável Y , condicionada no evento de aceitação descrito no algoritmo, coincide com a distribuição de X .

Abaixo tem-se um exemplo em R para a geração de a.a. de uma distribuição normal utilizando-se do método da rejeição.