# GERAÇÃO EFICIENTE DE UMA VARIÁVEL ALEATÓRIA GAUSSIANA

**Prof. Dulte José de Barros (★)**

**Prof. Dr. Walter Godoy Júnior (★★)**

## RESUMO

O presente artigo descreve sucintamente um procedimento alternativo para a obtenção de uma variável aleatória de distribuição gaussiana. Posteriormente, é efetuada uma análise comparativa entre a velocidade de um método tradicional que utiliza convoluções e o exposto aqui, o qual é otimizado de forma a aproveitar-se, ao máximo, a eficiência do ambiente utilizado.

## ABSTRACT

This paper describes an alternative procedure to generate a normally distributed random variable. Moreover, a comparative analysis has been done between the velocity of a popular method that uses convolutions and the optimized method exposed here, which takes advantage of the working environment.

★ **Dulte José de Barros** é mestrando do Curso de Pós-Graduação em Informática Industrial (CPGII), na área de Comunicação Digital, no Centro Federal de Educação Tecnológica do Paraná (CEFET-PR).

★★ **Walter Godoy Júnior** é professor pesquisador do Curso de Pós-Graduação em Informática Industrial (CPGII), na área de Comunicação Digital, no Centro Federal de Educação Tecnológica do Paraná (CEFET-PR).

## 1. INTRODUÇÃO

Nos últimos anos, com o incremento substancial da velocidade de processamento dos sistemas computacionais, o desenvolvimento de programas para simulação tem-se tornado cada vez mais atraente. Um exemplo da aplicabilidade de tais programas pode ser visto em [BRA 84] e em [CHA 89]. O requisito básico em tais ambientes é a geração de variáveis aleatórias com distribuição gaussiana. Esta talvez seja a mais importante distribuição contínua não uniforme utilizada. Também chamada de normal, é definida pela seguinte equação:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \qquad (1)$$

onde a variável aleatória X terá média nula e desvio padrão unitário. Esta restrição pode ser facilmente eliminada, utilizando-se uma variável auxiliar Y, a qual terá média   e desvio padrão  :

$$Y = \mu + \sigma X \qquad (2)$$

Segundo [GOD 90], um método de geração de uma variável aleatória comumente utilizado na prática está fundamentado no Teorema do Limite Central, o qual afirma que a convolução de infinitas variáveis aleatórias independentes e identicamente distribuídas terá distribuição gaussiana. Na prática, devido à impossibilidade de se utilizar a convolução com infinitos termos, obtêm-se excelentes resultados com 12 a 48 convoluções. Particularmente, em [CHA 89], foram utilizadas 19 convoluções. Deve-se ter em mente o compromisso entre rapidez e a precisão obtida, pois um número elevado de convoluções aumenta a precisão dos resultados, diminuindo, em contrapartida, a velocidade de execução.

No próximo item, apresenta-se o método polar descrito em [KNU 81], desenvolvido por G.E.P. Box, M.E. Muller e G. Marsaglia, no final da década de 50.

## 2. MÉTODO POLAR PARA GERAÇÃO DE DISTRIBUIÇÕES GAUSSIANAS

Por meio deste algoritmo poder-se-á calcular até duas variáveis aleatórias independentes e de distribuições gaussianas, $X_1$ e $X_2$, conforme detalhado a seguir:

**Passo 1:** Inicialmente, geram-se duas variáveis aleatórias independentes, $U_1$ e $U_2$, uniformemente distribuídas entre 0 e 1. Então, criam-se duas variáveis auxiliares $V_1$ e $V_2$, uniformemente distribuídas entre $-1$ e $+1$:

$$V_1 = 2\,U_1 - 1 \qquad (3)$$

$$V_2 = 2\,U_2 - 1 \qquad (4)$$

Na maioria dos computadores, neste passo, seria conveniente ter-se $V_1$ e $V_2$ representadas no formato de ponto flutuante.

**Passo 2:** Determina-se o valor de S:

$$S = V_1^2 + V_2^2 \qquad (5)$$

**Passo 3:** Se $S \geqslant 1$, retorna-se ao passo 1 (Pode-se demonstrar que os passos 1, 2 e 3 serão executados em média 1.27 vezes, com um desvio padrão de 0.587).

**Passo 4:** Atribui-se a $X_1$ e $X_2$ os seguintes valores:

$$X_1 = V_1 \sqrt{\frac{-2 \ln S}{S}} \qquad (6)$$

$$X_2 = V_2 \sqrt{\frac{-2 \ln S}{S}} \qquad (7)$$

As variáveis obtidas $X_1$ e $X_2$ possuem a distribuição gaussiana desejada.

No próximo item é efetuada a prova de que tal procedimento gera as distribuições normais.

## 3. DEMONSTRAÇÃO DA VALIDADE DO MÉTODO POLAR

Neste item, utilizam-se as propriedades elementares da Geometria Analítica e do Cálculo Integral e Diferencial.

Se $S < 1$ no passo 3, o ponto no plano cartesiano com coordenadas $(V_1, V_2)$ será um ponto aleatório uniformemente distribuído dentro de um círculo de raio unitário. Transformando-se para coordenadas polares, obtém-se:

$$V_1 = R \cos \Theta \qquad (8)$$

$$V_2 = R \operatorname{sen} \Theta \qquad (9)$$

e portanto:

$$S = R^2 \qquad (10)$$

$$X_1 = \sqrt{-2 \ln S}\, \cos \Theta \qquad (11)$$

$$X_2 = \sqrt{-2 \ln S}\, \operatorname{sen} \Theta \qquad (12)$$

Efetuando-se uma mudança de variáveis:

$$X_1 = R' \cos \Theta' \qquad (13)$$

$$X_2 = R' \operatorname{sen} \Theta' \qquad (14)$$

onde:

$$R' = \sqrt{-2 \ln S} \qquad (15)$$

$$\Theta' = \Theta \qquad (16)$$

Sabendo-se que as variáveis R e $\Theta$ são independentes dentro do círculo de raio unitário, pode-se afirmar que R' e $\Theta'$ também o são. Além disso, $\Theta$ e, por conseguinte, $\Theta'$ estão uniformemente distribuídas entre 0 e $2\pi$.

A probabilidade de que $R' \leqslant r$ será a probabilidade de que $-2 \ln S \leqslant r^2$, isto é, a probabilidade de que $S \geqslant e^{-r^2/2}$. Isto será igual a

$1 - e^{-r^2/2}$, visto que $S = R^2$ estará uniformemente distribuído entre 0 e 1.

A probabilidade de que R' esteja entre r e r + dr será, portanto, a diferencial de $1 - e^{-r^2}/2$, ou seja, $r\, e^{-r^2}/2\, dr$. Analogamente, a probabilidade de que $\Theta'$ esteja entre $\theta$ e $\theta + d\theta$ será $1/(2\pi)\, d\theta$.

Conseqüentemente, a probabilidade conjunta de que $X_1 \leqslant x_1$ e de que $X_2 \leqslant x_2$ poderá ser escrita como:

$$\int \{(r, \theta) \mid r \cos\theta \leqslant x_1,\ r \operatorname{sen}\theta \leqslant x_2\} \frac{1}{2\pi} e^{-r^2/2}\, r\, dr\, d\theta =$$

$$= \frac{1}{2\pi} \int \{(x, y) \mid x \leqslant x_1,\ y \leqslant x_2\} e^{-(x^2 + y^2)/2}\, dx\, dy =$$

$$= \left[\sqrt{\frac{1}{2\pi}} \int_{-\infty}^{x_1^1} e^{-x^2/2}\, dx\right] \left[\sqrt{\frac{1}{2\pi}} \int_{-\infty}^{x_2} e^{-y^2/2}\, dy\right] \qquad (17)$$

Isto prova que $X_1$ e $X_2$ são variáveis gaussianas independentes, conforme é desejado.

## 4. OTIMIZAÇÃO COMPUTACIONAL DO MÉTODO POLAR

O método polar, na forma como se descreve no item 2, é bastante eficiente. Entretanto, uma otimização poderá ser conseguida efetuando-se uma mudança de escala, a fim de suprimir todas as operações redundantes dentro da rotina que determinará a função gaussiana. Neste trabalho, uma otimização adicional foi efetuada, considerando a implementação da rotina no ambiente de programação TURBO C, versão 2.0. A geração das variáveis randômicas $U_1$ e $U_2$ está descrita detalhadamente em [MIL 88]. Na seqüência, é apresentada a forma final da rotina:

```
/★★★ Método Polar Otimizado. ★★★/
/★ Autor: Dulte José de Barros                    data: 18/08/90 ★/
/★ Includes & defines ★/
  include < math.h >
#define M   2147483647
#define A1  48271
#define A2  69621
#define Q1  44488
#define Q2  30845
#define R1  3399
#define R2  23902
#define KP  1.07374182E9
/★★★ Variáveis globais estáticas: ★★★/
static long seed1 = 1;
static long seed2 = 1;
/★★★ gauss (med, desv) Função que gera um número randômico com distribuição gaussiana de média "med" e desvio padrão "desv", a partir de duas variáveis randômicas independentes e uniformemente distribuídas pelo método polar otimizado. ★★★/
float gauss (float med, float desv)

      float v1,v2,s;
      while (v1 = ((seed1 = seed1 % Q1 ★ A1 — seed1 / Q1
      ★ R1) <= 0 ? seed1 += M : seed1) — KP,
      v2 = ((seed2 = seed2 % Q2 ★ A2 — seed2 / Q2
      ★ R2) <= 0 ? seed 2 += M : seed2) — KP,
      s = v1 ★ v1 + v2 ★ v2, s = 1.15292150E18);
      return med + desv ★ v1 ★ sqrt ((83.1776617 — 2 ★
      log(s)) /s);
```

## 5. ANÁLISE DA EFICIÊNCIA PROPORCIONADA PELO MÉTODO POLAR

A fim de efetuar-se uma comparação real entre os métodos das convoluções (fundamentado no Teorema do Limite Central), e o método apresentado acima, um teste simulado foi executado, obtendo-se os seguintes resultados:

**Teste de otimização da função Gauss.**

AMBIENTE: PC-AT 12 MHz com turbo e sem o co-processador 80287 (velocidade 6.9 vezes maior em relação ao IBM-PC padrão, de 4.77 MHz).

LINGUAGEM: TURBO C versão 2.0 (As opções de otimização do TURBO C foram deixadas no modo default).

Tempo de execução para 10000 iterações:

| | |
|---|---|
| 1) Método das convoluções com n_conv = 48: | 329.9s |
| 2) Método otim. das conv. com n_conv = 48: | 259.5s |
| 3) Método das convoluções com n_conv = 36: | 250.9s |
| 4) Método otim. das conv. com n_conv = 36: | 198.0s |
| 5) Método das convoluções com n_conv = 12: | 93.5s |
| 6) Método otim. das conv. com n_conv = 12: | 75.5s |
| 7) Método polar otimizado | 51.2s |

## 6. CONCLUSÃO

Com base nos resultados do item anterior, nota-se uma sensível melhora na velocidade de geração da variável aleatória gaussiana. Observa-se um ganho superior a 400% em relação ao método fundamentado no Teorema do Limite Central, otimizado para 48 convoluções. Vale ressaltar que este último é, por natureza, um método aproximado, enquanto o método polar possui um limite de precisão imposto apenas pelo ambiente de programação, sendo conseqüentemente mais preciso.

O programa apresentado no item 4 implementa apenas uma variável aleatória gaussiana. Porém, conforme visto no item 2, o método polar consegue gerar simultaneamente até duas variáveis aleatórias de distribuição gaussiana, o que facilita a simulação para códigos multiníveis. Assim, caso seja necessário, pode-se alterar o programa anterior para que, com um aumento mínimo de «overhead», forneça como saída estes dois valores.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS


[BRA 84] Brandão, J.C. Análise e Simulação de Transmissão Digital em Microcomputador, CETUC, PUC-RIO, 1984.

[CHA 89] Chamon, M.A. Contribuição ao Estudo de Sistemas de Modulação Codificada em Códigos Concatenados. Dissertação de Mestrado, CTA, S.J. dos Campos, 1989.

[GOD 90] Godoy, W.Jr. Introdução à Modulação Codificada com Códigos de Blocos BCM. Apostila do CPGII, Curitiba, maio 1990.

[KNU 81] Knuth D.E. The Art of Computer Programming 2nd Ed. Addison-Wesley, Reading Mass., 1981.

[MIL 88] Miller K.W. e Park S.K. Random Number Generators: Good Ones Are Hard to Find, Communications of ACM 10, vol 31, out 1988, pag. 1192-1201.