

Teoria da Probabilidade e Variável Aleatória

Prof. Waldecir Perrella

OBJETIVOS: Neste experimento vamos estudar:

- Propriedades de variáveis aleatórias, discretas e contínuas;
- Variáveis aleatórias uniformes e gaussianas em termos de suas funções densidades de probabilidades (PDF) e distribuição cumulativas (CDF);
- Medidas estatísticas da média, variância e potência média quadrática;
- Geração de variáveis aleatórias por computador.

INTRODUÇÃO TEÓRICA:

1. Teoria da Probabilidade

A teoria da probabilidade estuda fenômenos aleatórios que são caracterizados por seus resultados ou saídas que não são conhecidos antes de sua realização.

Para se definir probabilidade é conveniente utilizar-se a teoria de conjuntos. Assim, define-se espaço amostral S como o conjunto cujos elementos são todas as possíveis saídas de um experimento. Todos os outros conjuntos com formados com as saídas de um experimento são subconjuntos ou eventos do espaço amostral S .

Podemos definir a probabilidade usando a frequência relativa. A probabilidade de um evento A , denotada por $P(A)$, pode ser definida em termos da frequência relativa da ocorrência do evento A em n realizações do experimento:

$$P(A) = \lim_{n \rightarrow \infty} \left(\frac{n_A}{n} \right) \quad (1)$$

onde n_A é o número de vezes que o evento A ocorreu em n realizações.

Na prática, n deve ser um número razoavelmente para termos uma boa aproximação.

Da definição de probabilidade, equação (1), o valor de $P(A)$ para qualquer evento A tem a propriedade:

$$0 \leq P(A) \leq 1 \quad (2)$$

onde $P(A) = 0$ quando A é o evento nulo Φ , isto é, o evento nunca ocorre, e $P(A) = 1$ quando A é o evento certo U , isto é, o evento sempre ocorre.

A probabilidade de um evento conjunto AB , ($AB = A \cap B$), é definida por:

$$P(AB) = \lim_{n \rightarrow \infty} \left(\frac{n_{AB}}{n} \right) \quad (3)$$

onde n_{AB} é o número de vezes que o evento AB ocorreu em n realizações.

Dois eventos, A e B , são chamados de mutuamente exclusivos se AB é o evento nulo, que implica que $P(AB) = 0$.

A probabilidade da união de dois eventos pode ser calculada usando-se o evento resultante ou por meio das probabilidades dos eventos simples e usando-se o seguinte teorema:

Seja $E = A + B$, então:

$$P(E) = P(A + B) = P(A) + P(B) - P(AB) \quad (4)$$

Prova: Considere que n_1 é o número de vezes que ocorreu somente o evento A , em n realizações, n_2 é o número de vezes que ocorreu somente o evento B , em n realizações, e n_{AB} é o número que o evento AB ocorreu. Assim

$$P(A + B) = \lim_{n \rightarrow \infty} \left(\frac{n_1 + n_2 + n_{AB}}{n} \right) = \lim_{n \rightarrow \infty} \left(\frac{n_1 + n_{AB}}{n} \right) + \lim_{n \rightarrow \infty} \left(\frac{n_2 + n_{AB}}{n} \right) - \lim_{n \rightarrow \infty} \left(\frac{n_{AB}}{n} \right)$$

que é idêntico à equação (4) pois:

$$P(A) = \lim_{n \rightarrow \infty} \left(\frac{n_1 + n_{AB}}{n} \right), P(B) = \lim_{n \rightarrow \infty} \left(\frac{n_2 + n_{AB}}{n} \right) \text{ e } P(AB) = \lim_{n \rightarrow \infty} \left(\frac{n_{AB}}{n} \right)$$

A probabilidade condicional é a probabilidade de um evento A ocorrer, dado que um evento B já ocorreu. Denotamos esta probabilidade por $P(A | B)$ e é definida por:

$$P(A | B) = \lim_{n_B \rightarrow \infty} \left(\frac{n_{AB}}{n_B} \right) \quad (5)$$

Teorema de Bayes: Considerando $E = AB$, então:

$$P(E) = P(A)P(B | A) = P(B)P(A | B) \quad (6)$$

Prova:

$$P(AB) = \lim_{n \rightarrow \infty} \left(\frac{n_{AB}}{n} \right) = \lim_{\substack{n \rightarrow \infty \\ n \rightarrow \text{grande}}} \left(\frac{n_{AB}}{n_A} \frac{n_A}{n} \right) = P(B | A)P(A) \quad (7)$$

Dois eventos A e B , são independentes se:

$$P(B | A) = P(B) \text{ ou } P(A | B) = P(A) \quad (8)$$

Usando as equações (4) e (6), podemos mostrar que para um conjunto de eventos independentes, A_1, A_2, \dots, A_n , a condição necessária é:

$$P(A_1 A_2 \dots A_n) = P(A_1) P(A_2) \dots P(A_n) \quad (9)$$

2. Variáveis aleatórias

Uma variável aleatória é uma função definida sobre os eventos (elementos) de um sistema de probabilidades. Até agora, nós definimos probabilidades em termos dos eventos A, B, C de um espaço amostral S . é mais conveniente descrever estes conjuntos por valores numéricos de tal forma que equações podem ser obtidas como uma função de valores numéricos ao invés de funções de parâmetros alfanuméricos.

Como exemplo, considere o experimento do lançamento de uma moeda. Os possíveis eventos são cara (A) ou coroa (B) e mutuamente exclusivos, isto é $P(S) = P(A) + P(B) = 1$. Cada um desses eventos pode ser mapeado a valores conveniente de uma variável aleatória X . Os valores atribuídos para X podem ser positivos, negativos, fracionário, inteiro ou real. Na Tabela 1 temos o mapeamento escolhido:

Tabela 1

Evento [.]	Valor da Variável Aleatória $x[.]$	Probabilidade do evento $P(x)$
A	0	0.5
B	1	0.5
		Total = 1

Na Figura 1 temos: o diagrama de Venn mostrando os eventos, o mapeamento da variável aleatória e a função de probabilidade $P(x)$.

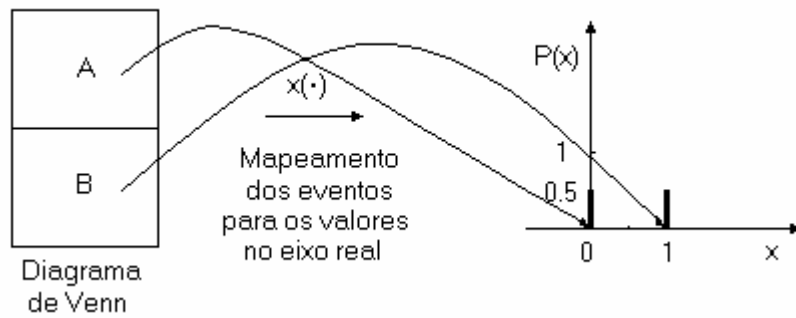


Figura 1. Mapeamento da variável aleatória X .

A função distribuição cumulativa (CDF) de uma variável aleatória X é denotada por $F(a)$, e é dada por:

$$F(a) = P(x \leq a) \equiv \lim_{n \rightarrow \infty} \left(\frac{n_{x \leq a}}{n} \right) \quad (10)$$

onde $F(a)$ é uma função sem unidade.

A função densidade de probabilidade (PDF) de uma variável aleatória é denotada por $f(x)$, e é dada por:

$$f(x) = \left| \frac{dF(a)}{da} \right|_{a=x} = \left| \frac{dP(x \leq a)}{da} \right|_{a=x} = \lim_{\substack{n \rightarrow \infty \\ \Delta x \rightarrow 0}} \left(\frac{1}{\Delta x} \left(\frac{n_{\Delta x}}{n} \right) \right) \quad (11)$$

onde $f(x)$ é uma função com unidade de $1/x$.

A CDF para o exemplo acima, ilustrada na Figura 2-a foi obtida usando-se a equação (8). Note que a CDF começa com o valor zero para $(a \rightarrow -\infty)$ e tende para a unidade para $(a \rightarrow \infty)$.

A PDF, ilustrada na Figura 2-b, é obtida usando-se a equação (9). A PDF para este exemplo consiste de impulsos (funções delta de Dirac) localizados nos valores (discretos) da variável aleatória e tendo seus pesos iguais a probabilidades dos eventos associados.

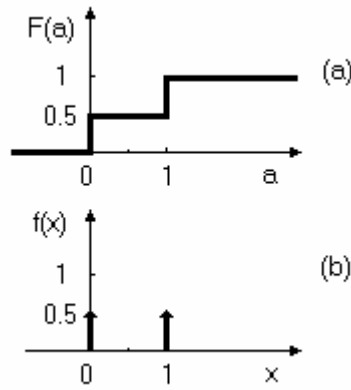


Figura 2 CDF e PDF do exemplo acima

Propriedades da CDF e PDF:

1. $F(a)$ é uma função não decrescente.
2. $F(a)$ é uma função contínua à direita, isto é:

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} F(a + \varepsilon) = F(a) \quad (12)$$

$$F(a) = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon > 0}} \int_{-\infty}^{a+\varepsilon} f(x) dx \quad (13)$$

3. $0 \leq F(a) \leq 1$.
4. $F(-\infty) = 0$.
5. $F(+\infty) = 1$.
6. $F(x) \geq 0$, isto é, $f(x)$ é sempre não negativa.
7. $\int_{-\infty}^{+\infty} f(x) dx = F(+\infty) = 1$.
8. $F(b) - F(a) = P(x \leq b) - P(x \leq a) = P(a < x \leq b) = \int_a^b f(x) dx$
onde $b > a$ e a distribuição é contínua.

No exemplo acima temos uma distribuição discreta, isto é, a variável aleatória tem M valores discretos x_1 e x_2 . Consequentemente, a CDF cresce somente em pulos ($F(a)$ é descontínua) quando a cresce, e a PDF consiste de impulsos (funções delta de Dirac), localizados nos valores discretos da variável aleatória. Diferentemente deste exemplo de variável aleatória com distribuição discreta existem distribuições contínuas, como a considerada a seguir. Uma variável aleatória é distribuída continuamente se ela pode assumir qualquer valor dentro de algum intervalo. Como exemplo considerando que uma variável aleatória representa as voltagens de uma pilha de 1,5 V. Se o número de pilhas fosse infinito, o número dos diferentes valores de voltagem (eventos) que seria também seria infinito, de tal forma que as distribuições, CDF e PDF, seriam funções contínuas. Suponha que, por meio de experimentação, a CDF é determinada primeiro usando

$F(a) = P(x \leq a) = \lim_{n \rightarrow \infty} \left(\frac{n_{x \leq a}}{n} \right)$, onde n é o número de baterias da coleção e $n_{x \leq a}$ é o número de baterias da coleção que tem voltagens menor ou igual à a volts, é um parâmetro. A CDF que obtida

obter, é ilustrada na figura 3-a. A PDF associada é obtida tomando-se a derivada da CDF, como mostra a Figura 3-b.

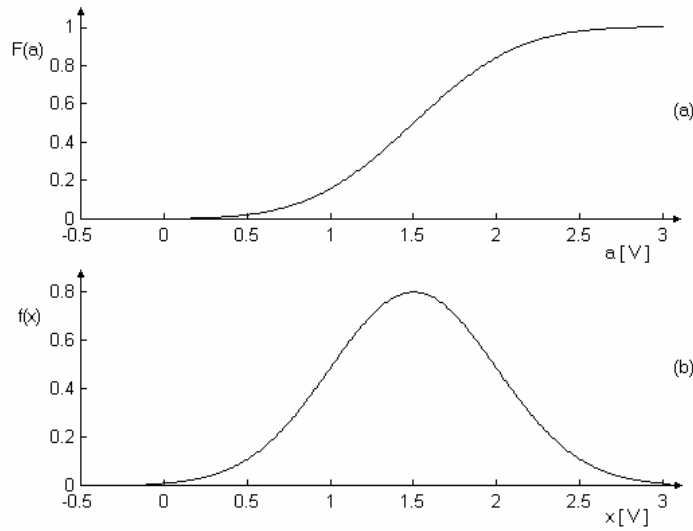


Figura 3 Distribuição Contínua; (a) CDF e (b) PDF

Quando a distribuição for discretamente distribuída a PDF é dada por:

$$f(x) = \sum_{i=1}^M P(x_i) \delta(x - x_i) \quad (14)$$

onde M é o número de eventos discretos e $P(x_i)$ é a probabilidade de ocorrer o evento discreto x_i . A CDF, neste caso é dada por:

$$F(a) = \sum_{i=1}^L P(x_i) \quad (15)$$

onde L é o maior inteiro tal que $x_L \leq a$ e $L \leq M$.

3. Transformação de Variáveis Aleatórias

Em muitos problemas de engenharia, dado um modelo matemático de um sistema e um sinal de entrada, deseja-se determinar as características do sinal de saída. Como um simples exemplo, considere o problema de achar a PDF Y da saída de um retificador com função de transferência $y=x^2$, sabendo que a PDF X da entrada é gaussiana com média zero e variância σ_x^2 .

Teorema: Seja X uma VA contínua com PDF $f_X(x)$, e $y=g(x)$ uma dada transformação de X . Para determinar a PDF de Y , resolve-se a equação $y=g(x)$ em termos de Y . Se x_1, x_2, \dots, x_n são todas as soluções reais de $x_i=g^{-1}(y)$ e $g'(x)$ é a derivada da transformação, então:

$$f_Y(y) = \sum_{i=1}^n \frac{f_X(x_i)}{|g'(x_i)|} \quad (16)$$

Para o exemplo acima $y=x^2$, tem-se $n=2$ e assim: $x_1=\sqrt{y}$ e $x_2=-\sqrt{y}$, $g'(x)=2x$, a PDF de X é dada por $f_x = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{x^2}{2\sigma_x^2}\right)$ e aplicando-se o teorema tem-se:

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\left(\frac{y}{2\sigma_x^2}\right)} \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\left(\frac{y}{2\sigma_x^2}\right)} \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi y}\sigma_x} e^{-\left(\frac{y}{2\sigma_x^2}\right)} \quad p/y \geq 0 \quad (17)$$

Para uma transformação linear $y=ax+b$ e X a mesma VA definida no exemplo anterior, tem-se $n=1$ e assim $x_1=(y-b)/a$ e $g'(x)=a$, resulta:

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{((y-b)/a)^2}{2\sigma_x^2}} \frac{1}{|a|} = \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}} \quad (18)$$

onde $\mu_Y=b$, $\sigma_Y=|a|/\sigma_x$ e assim $\sigma_Y^2 = a^2 \sigma_x^2$.

4. Médias Estatísticas e Momentos (“Ensemble Average”)

Um dos usos primários da teoria de probabilidades é determinar o valor médio de uma variável aleatória ou de alguma função dessa variável aleatória, denotada por $y = h(x)$.

Define-se valor médio ou valor esperado (“ensemble average”) de $y = h(x)$ por:

$$E[Y] = \int_{-\infty}^{\infty} [h(x)] f(x) dx \quad (19)$$

Se X for uma variável aleatória discreta podemos usar a definição (13) e assim, obtemos:

$$E[Y] = \sum_{i=1}^M P(x_i) h(x_i) \quad (20)$$

Momentos são definidos como os valores esperados de algumas específicas funções de X . Por exemplo, o r -ésimo momento é definido fazendo $y = h(x) = (x - x_0)^r$.

Momento de uma variável aleatória em relação ao ponto $x = x_0$ é dado por:

$$E[(X - x_0)^r] = \int_{-\infty}^{\infty} (x - x_0)^r f(x) dx \quad (21)$$

O valor médio m_x é definido pelo primeiro momento em relação a origem, isto é, $x_0=0$.

$$m_x = E[X] = \int_{-\infty}^{\infty} x f(x) dx \quad (22)$$

A variância σ^2 é definida pelo segundo momento em relação a média:

$$\sigma_x^2 = E[(X - m_x)^2] = \int_{-\infty}^{\infty} (x - m_x)^2 f(x) dx \quad (23)$$

O desvio padrão σ_x é a raiz quadrada da variância. Assim:

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{\int_{-\infty}^{\infty} (x - m_x)^2 f(x) dx} \quad (24)$$

O valor médio quadrático é definido pelo segundo momento em relação a origem:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx \quad (25)$$

A seguir mostra-se a seguinte relação $\sigma^2 = E[X^2] - m^2$

Partindo-se da definição de variância tem-se:

$$\begin{aligned} \sigma_x^2 &= E[(x - m_x)^2] = E[x^2 - 2x m_x + m_x^2] = \\ &= E[x^2] - 2m_x E[x] + m_x^2 = \\ &= E[x^2] - m_x^2 \end{aligned} \quad (26)$$

5. Principais distribuições

Densidade de probabilidade uniforme é definida pela PDF:

$$f(x) = \begin{cases} 0 & x < a \\ 1 & a \leq x \leq b \\ \frac{b-a}{b-a} & \\ 0 & x > b \end{cases} \quad (27)$$

onde $b > a$.

Densidade de probabilidade gaussiana é definida pela PDF:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (28)$$

onde m é a média e σ^2 é a variância.

Para calcular a CDF de variáveis gaussianas é necessário utilizarmos as seguintes funções que podem ser calculadas numericamente:

A função de erro, denotada por **erf(u)**, é dada pela seguinte expressão:

$$\text{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-z^2} dz \quad (29)$$

A função de erro complementar é definida por:

$$\text{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^{\infty} e^{-z^2} dz \quad (30)$$

Elas são relacionadas como segue:

$$\text{erfc}(u) = 1 - \text{erf}(u) \quad (31)$$

Outra maneira de calcular a CDF de gaussianas é utilizando a função $Q(v)$ definida por:

$$Q(v) = \frac{1}{\sqrt{2\pi}} \int_v^{\infty} e^{-\left(\frac{x^2}{2}\right)} dx \quad (32)$$

Pode-se mostrar que a função $Q(v)$ é relacionada com a função erro complementar como segue:

$$Q(v) = \frac{1}{2} \text{erfc}\left(\frac{v}{\sqrt{2}}\right) \quad \text{ou} \quad \text{erfc}(u) = 2Q(\sqrt{2}u) \quad (33)$$

Em comunicação digital freqüentemente é necessário calcular a probabilidade de um evento ($x > a$), denotado por $P(X > a)$, quando X é uma variável aleatória gaussiana com média m e variância σ^2 . Assim tem-se:

$$P(x > a) = \int_a^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{(x-m)^2}{2\sigma^2}\right)} dx \quad (34)$$

Fazendo-se a seguinte mudança de variável, $z = \frac{x-m}{\sigma}$, tem-se:

$$P(x > a) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-m}{\sigma}}^{\infty} e^{-\left(\frac{z^2}{2}\right)} dz \quad (35)$$

Utilizando-se as equações (32) e (33) tem-se:

$$P(X > a) = Q\left(\frac{a-m}{\sigma}\right) = \frac{1}{2} \text{erfc}\left(\frac{a-m}{\sqrt{2}\sigma}\right) \quad (36)$$

No Matlab tem-se $\text{erf}(x)$ e $\text{erfc}(x)$.

4. Geração de números pseudo aleatórios

A geração de números pseudo aleatórios consiste na obtenção de uma sequência de números, que embora obtidos deterministicamente, tem a aparência de serem uma variável aleatória independente e distribuído uniformemente entre 0 e 1.

5.1 Método Congruencial multiplicativo

Neste método da geração de números a partir de uma semente x_0 utilizando-se a operação módulo m , como mostra a equação:

$$x_{n+1} = (a \bullet x_n) \bmod m \quad (37)$$

Para a obtenção de números entre 0 e 1 divide-se x_n por m .
Como exemplo, considere $a=3$, $m=7$ e $x_0=2$ temos a Tabela 2

Tabela 2

$\mathbf{X_{n+1}=[a \cdot X_n] \bmod m}$	$\mathbf{X_{n+1}}$	$\mathbf{X_{n+1}/m}$
X_0	2	
$X_1=[3 \cdot 2] \bmod 7$	6	$6/7=0,857$
$X_2=[3 \cdot 6] \bmod 7$	4	$4/7=0,571$
$X_3=[3 \cdot 4] \bmod 7$	5	$5/7=0,714$
$X_4=[3 \cdot 5] \bmod 7$	1	$1/7=0,143$
$X_5=[3 \cdot 1] \bmod 7$	3	$3/7=0,429$
$X_6=[3 \cdot 3] \bmod 7$	2	$2/7=0,285$
$X_7=[3 \cdot 2] \bmod 7$	6	$6/7=0,857$
$X_8=[3 \cdot 6] \bmod 7$	4	$4/7=0,571$

Podemos observar que a seqüência é periódica, com período igual a $m-1=6$. No Matlab podemos utilizar as instrução $x(n+1)=\text{mod}(a \cdot x(n),m)$.

5.2 Método Congruencial mixto

Neste método da geração de números a partir de uma semente x_0 utilizando-se a operação módulo m , como mostra a equação:

$$x_{n+1} = (a \bullet x_n + c) \bmod m \quad (38)$$

Em geral as constantes nos dois métodos precisam satisfazer 3 critérios:

1. Para cada semente inicial, a seqüência resultante deve ter a aparência de uma variável aleatória independente e uniforme entre 0 e 1.
2. Para qualquer semente inicial, o número de variáveis aleatórias que podem ser geradas antes da repetição deve ser muito grande.
3. O valores devem ser gerados eficientemente em um computador.

Para atender esses critérios m precisa ser escolhido como o maior número primo que pode ser representado num computador. Exemplo, num computador de 32 bits, onde o primeiro bit é de sinal, temos $m=2^{31}-1$ e $a=7^5=16807$, resultando numa seqüência com período igual a $2,1 \cdot 10^9$.

No Matlab, gera-se esses números aleatórios através das seguintes funções:

rand(N) gera uma matriz ($N \times N$)

rand(M,N) gera uma matriz ($M \times N$)

rand(N,1) gera um vetor coluna ($M \times 1$)

rand gera um escalar

Este gerador tem um período de 2^{1492} .

O Matlab inicializa o estado do gerador toda vez que o Matlab é inicializado e assim gera sempre a mesma seqüência de números.

Para se saber o estado do gerador usa-se a instrução $s=\text{rand('state')}$ e para forçar o estado do gerador usam-se as seguintes instruções:

rand('state', s) inicializa para o estado s ;

rand('state', 0) inicializa para o estado inicial;

rand('state', j) inicializa para o estado j;
*rand('state', sun(100*clock))* inicializa para diferentes estados conforme o relógio do computador;

6. Geração de Variáveis Aleatórias

6.1 Geração de Variáveis Aleatórias Discretas.

- **Método Inverso**

Deseja-se gerar uma variável aleatória X com probabilidade de massa dada por:

$$P[X=x_j] = p_j \text{ para } j = 0, 1, \dots \text{ e } \sum_{j=0}^{\infty} p_j = 1 \quad (39)$$

Para isso gere uma variável aleatória uniforme U no intervalo $[0,1]$ e a determine os valores de X usando o seguinte teste:

$$X = \begin{cases} x_0 & \text{se } u < p_0 \\ x_1 & \text{se } p_0 \leq u < p_0 + p_1 \\ \vdots & \\ x_j & \text{se } \sum_{i=1}^{j-1} p_i \leq u < \sum_{i=1}^j p_i \end{cases} \quad (40)$$

Isto pode ser mostrado pois para $0 < a < b < 1$ tem-se que $P[a \leq U < b] = b - a$ e assim resulta:

$$P[X = x_j] = P\left[\sum_{i=1}^{j-1} p_i \leq U \leq \sum_{i=1}^j p_i\right] = p_j \quad (41)$$

6.2 Geração de Variáveis Aleatórias Contínuas Gaussianas.

- **Teorema do Limite Central**

Em resumo, o teorema do limite central mostra que somando um grande número de variáveis aleatórias com densidade de probabilidade qualquer tem-se uma variável aleatória com densidade de probabilidade gaussiana. Assim partindo-se da variável aleatória uniforme U definida no intervalo $[0,1]$ e somando-se 12 dessas variáveis tem-se uma variável aleatória gaussiana X dada por:

$$x = \sum_{i=1}^{12} u_i \quad (42)$$

onde a média de X é dada por:

$$\mu_x = E[X] = E\left[\sum_{i=1}^{12} U_i\right] = \sum_{i=1}^{12} E[U_i] = 0,5 * 12 = 6 \quad (43)$$

E a variância é dada por:

$$\sigma_x^2 = E[(X - \mu_x)^2] = 1 \quad (44)$$

Para obter-se uma variável gaussiana Y com média zero e variância um usa-se a seguinte transformação:

$$y = x - 6 \quad (45)$$

Neste método de simulação de variável aleatória gaussiana que os valores possíveis de cada realização de Y pertencem a faixa $[-6.0, +6.0]$ e uma variável gaussiana tem valores na faixa $[-\infty, +\infty]$. Assim, na simulação de eventos raros é necessário usar outros métodos.

- **Método de Box Müller**

É conhecido que se x e y são variáveis aleatórias gaussianas com médias zero e independentes, então $r = (x^2 + y^2)^{1/2}$ e $\theta = \tan^{-1}(y/x)$ tem PDFs de Rayleigh e uniforme, respectivamente. Este fato pode ser usado para gerar duas amostras de uma variável gaussiana usando a transformação de um par de variáveis Rayleigh e uniforme como segue:

Gera-se duas variáveis independentes, U_1 e U_2 uniformemente distribuídas no intervalo $[0, 1]$ e então gera-se as variáveis independentes X e Y , gaussianas com média zero e variância 1, utilizando-se as seguintes expressões:

$$x = [-2 \ln(u_1)]^{1/2} \cos(2\pi u_2) \quad (46)$$

$$y = [-2 \ln(u_1)]^{1/2} \sin(2\pi u_2) \quad (47)$$

Para finalizar para obter-se uma variável aleatória gaussiana Z com média μ_z e variância σ_z^2 usa-se a seguinte transformação:

$$Z = \sigma_z X + \mu_z \quad (48)$$