- FUNÇÕES DE VARIÁVEIS ALEATÓRIAS

* SEJA X UMA V.A. CONTÍNUA COM FUNÇÃO DENSIDADE

$$f_X(x)$$

E SEJA $y = g(x)$ UMA FUNÇÃO BIJETIVA, COM INVERSA

$x = g^{-1}(y) \equiv h(y)$.

$\Rightarrow$ A FUNÇÃO DENSIDADE DA V.A.

$$Y = g(X)$$

É DADA POR

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = \frac{f_X(x)}{\left|\frac{dy}{dx}\right|} = f_X(x)\left|\frac{dy}{dx}\right|^{-1}$$

* SEJAM X E Y DUAS V.As. CONTÍNUAS COM DENSIDADE CONJUNTA

$$f_{XY}(x,y)$$

E SEJA

$$\begin{cases} z = g(x,y) \\ w = h(x,y) \end{cases}$$

UMA TRANSFORMAÇÃO BIJETIVA, COM INVERSA

$$\begin{cases} x = q(z,w) \\ y = r(z,w) \end{cases}$$

$\Rightarrow$ A FUNÇÃO DENSIDADE CONJUNTA DAS VARIÁVEIS

$$Z = g(X,Y)$$
$$W = g(X,Y)$$

É DADA POR $f_{ZW}(z,w) = f_{XY}(x,y)\,|J(x,y)|^{-1}$

ONDE $J(x,y)$ É O JACOBIANO DA TRANSFORMAÇÃO

$$\begin{cases} z = g(x,y) \\ w = h(x,y) \end{cases}$$

$$\Rightarrow J(x,y) = \begin{vmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{vmatrix}$$

* NO CASO DE TERMOS n FUNÇÕES DE n VARIÁVEIS:

$$Y_1 = g_1(X_1, \dots, X_n)$$
$$Y_2 = g_2(X_1, \dots, X_n)$$
$$\vdots \qquad \vdots$$
$$Y_n = g_n(X_1, \dots, X_n)$$

A FUNÇÃO DENSIDADE CONJUNTA DE $Y_1, Y_2, \dots, Y_n$ É DADA POR

$$f_{Y_1 \dots Y_n}(y_1, \dots, y_n) = f_{X_1, \dots, X_n}(x_1, \dots, x_n) \, |J(x_1, \dots, x_n)|^{-1}$$

ONDE

$$J(x_1, \dots, x_n) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \\ \frac{\partial g_n}{\partial x_1} & \dots & \frac{\partial g_n}{\partial x_n} \end{vmatrix}$$

É O JACOBIANO DA TRANSFORMAÇÃO

$$y_1 = g_1(x_1, \dots, x_n)$$
$$y_2 = g_2(x_1, \dots, x_n)$$
$$\vdots \qquad \vdots$$
$$y_n = g_n(x_1, \dots, x_n)$$

OBSERVAÇÕES:

i) NO CASO UNIVARIADO, SE A TRANSFORMAÇÃO $y = g(x)$ NÃO É BIJETIVA, HÁ VÁRIOS VALORES x_k, $k = 1, 2, \ldots$ QUE SATISFAZEM

$$y = g(x_k)$$

PARA UM DADO y

$\Rightarrow$ NESTE CASO, A FUNÇÃO DENSIDADE DE $Y = g(X)$ SE TORNA

$$f_Y(y) = \sum_k f_X(x_k) \, |g'(x_k)|^{-1}$$

ii) NO CASO MULTIVARIADO, SE A TRANSFORMAÇÃO $\tilde{y} = g(\tilde{x})$ TEM VÁRIAS SOLUÇÕES $\tilde{x}_k$ PARA O MESMO $\tilde{y}$, A FUNÇÃO DENSIDADE CONJUNTA PARA $\tilde{Y} = (Y_1, Y_2, \ldots Y_m)^T$ SE TORNA

$$f_{\tilde{Y}}(\tilde{y}) = \sum_k f_{\tilde{X}}(\tilde{x}_k) \, |J(\tilde{x}_k)|^{-1}$$

iii) A TRANSFORMAÇÃO DE VARIÁVEIS ALEATÓRIAS É ÚTIL PARA A GERAÇÃO DE AMOSTRAS DE CERTAS DISTRIBUIÇÕES DE PROBABILIDADE UTILIZANDO O COMPUTADOR.

iv) COMO O INVERSO DO JACOBIANO DE UMA TRANSFORMAÇÃO É O JACOBIANO DA TRANSFORMADA INVERSA, PODEMOS USAR, POR EXEMPLO,

$$|J(x, y)|^{-1} = |\bar{J}(z, w)|$$

ONDE

$$\bar{J}(z, w) = \begin{vmatrix} \frac{\partial q}{\partial z} & \frac{\partial q}{\partial w} \\ \frac{\partial r}{\partial z} & \frac{\partial r}{\partial w} \end{vmatrix} \equiv \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix}$$

v) NO CASO DISCRETO, SE $Y = g(X)$, TEMOS

$$f_Y(y) = \sum_{x \in g^{-1}(y)} f_X(x)$$

- GERAÇÃO COMPUTACIONAL DE VARIÁVEIS ALEATÓRIAS

1. GERAÇÃO DE VARIÁVEIS UNIFORMES

MÉTODO CONGRUENCIAL LINEAR:

⇒ GERA UMA SEQUÊNCIA DE INTEIROS z_i, COMO

$$z_i = (a z_{i-1} + c) \bmod m \,, \quad i = 1, 2, \ldots$$

DADOS

$z_0 \equiv$ SEMENTE (RAIZ)

$a \equiv$ MULTIPLICADOR

$c \equiv$ INCREMENTO

QDO. $c = 0$: MÉTODO MULTIPLICATIVO

QDO. $c \neq 0$: MÉTODO MISTO

$m \equiv$ MÓDULO

OPERAÇÃO MOD $\equiv$ RESTO DA DIVISÃO INTEIRA

e.g., $15 \bmod 6 = 3$

A PARTIR DE z_i, UMA SEQUÊNCIA UNIFORME EM $[0,1]$ É OBTIDA COMO

$$R_i = \frac{z_i}{m} \qquad R \sim U(0,1)$$

A SEQUÊNCIA GERADA ASSUME VALORES NO CONJUNTO

$$\left\{0, \frac{1}{m}, \frac{2}{m}, \ldots, \frac{m-1}{m}\right\}$$

NO MATLAB:

$$m = 2^{31} - 1 = 2147483647$$

$$c = 0$$

$$a = 7^5 = 16807$$

OBSERVAÇÕES:

i) AO SE INICIAR A GERAÇÃO COM A MESMA SEMENTE, OBTÉM-SE A MESMA SEQUÊNCIA UNIFORME.

ii) OS NÚMEROS GERADOS SÃO RACIONAIS, PORTANTO DEVE-SE UTILIZAR UM VALOR ALTO PARA m.

iii) O VALOR DE m TAMBÉM DETERMINA O MÁXIMO PERÍODO PARA A SEQUÊNCIA. A OBTENÇÃO DESTE MÁXIMO DEPENDE DA ESCOLHA ADEQUADA DOS VALORES DE a E c.

iv) PARA GERAR SEQUÊNCIAS $\sim U(A,B)$, EMPREGAMOS A TRANSFORMAÇÃO

$$U(A,B) = (B-A)\,U(0,1) + A$$

2. GERAÇÃO DE VARIÁVEIS ALEATÓRIAS CONTÍNUAS:

– MÉTODO DA TRANSFORMAÇÃO INVERSA

QUEREMOS GERAR AMOSTRAS DA DISTRIBUIÇÃO

$$F(x) = \Pr[X \leq x]$$

⇒ SE A INVERSA DE F É FÁCIL DE OBTER, SEGUIMOS O PROCEDIMENTO ABAIXO:

i) GERAMOS UM NÚMERO ALEATÓRIO $U \sim U(0,1)$

ii) OBTEMOS X COMO

$$X = F^{-1}(U)$$

⇒ A AMOSTRAGEM DESEJADA É OBTIDA PORQUE

$$\Pr[X \leq x] = \Pr[F^{-1}(U) \leq x] = \Pr[U \leq F(x)] = F(x)$$

ONDE NOS DOIS ÚLTIMOS PASSOS USAMOS A MONOTONICIDADE DE $F(x)$, E PORTANTO DE $F^{-1}(x)$, E O FATO DE QUE

SE U É UMA VARIÁVEL ALEATÓRIA $U(0,1)$

$$\Pr[U \leq \alpha] = \alpha$$

EXEMPLOS DE APLICAÇÃO:

i) VARIÁVEL EXPONENCIAL NEGATIVA

$$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0$$

⇒ INVERSA: $X = F^{-1}(U) = -\frac{1}{\lambda}\log(1-U)$

COMO $U \sim U(0,1)$ E $(1-U)$ SÃO A MESMA VARIÁVEL, AMOSTRAS DA DISTRIBUIÇÃO EXPONENCIAL PODEM SER OBTIDAS COMO

$$X = -\frac{1}{\lambda}\log(U)$$

ii) VARIÁVEL ALEATÓRIA DE CAUCHY

$$F(x) = \frac{1}{\pi}\left[\operatorname{ArcTg}(x) - \frac{\pi}{2}\right]$$

$\Rightarrow$ INVERSA: $X = F^{-1}(U) = Tg\left[\pi\left(U + \frac{1}{2}\right)\right]$

$$= Tg\left[\pi\left(U - \frac{1}{2}\right)\right]$$

iii) VARIÁVEL ALEATÓRIA DE RAYLEIGH

$$f(x) = x\, e^{-\frac{x^2}{2}}, \quad x \geq 0$$

$$\Rightarrow F(x) = \int_0^x v\, e^{-\frac{v^2}{2}}\, dv = 1 - e^{-\frac{x^2}{2}}$$

$\Rightarrow$ INVERSA:

$$X = F^{-1}(U) = \sqrt{-2\log(1-U)}$$

$$= \sqrt{-2\log U}$$

– MÉTODO DE ACEITAÇÃO-REJEIÇÃO

⇒ GERAR AMOSTRAS DA FUNÇÃO DENSIDADE $f(x)$, QUANDO O MÉTODO DA TRANSFORMAÇÃO INVERSA NÃO É ADEQUADO

SUPOMOS CONHECER UMA FUNÇÃO $t(x)$ TAL QUE

$$t(x) \geq f(x), \quad \forall x \in I$$

ONDE I É O SUPORTE DE $f(x)$.

⇒ SE $t(x)$ NÃO É UMA FUNÇÃO DENSIDADE, TRABALHAMOS COM

$$r(x) = \frac{t(x)}{c}$$

ONDE

$$c = \int_I t(x)\,dx$$

e.g., QUANDO O SUPORTE I É LIMITADO, PODEMOS USAR PARA $r(x)$ A DENSIDADE UNIFORME EM I, COM

$$t(x) = \max_{x \in I} f(x)$$

O MÉTODO: SE A DENSIDADE $r(x)$ É FÁCIL DE SIMULAR, O PROCEDIMENTO ABAIXO GERA AMOSTRAS X DA DENSIDADE $f(x)$:

i) GERE $Y \sim r$

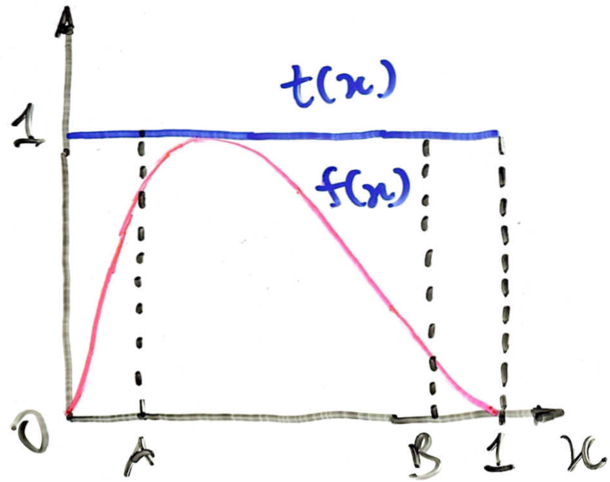
ii) GERE $U \sim U(0,1)$, INDEPENDENTE DE Y

iii) SE $U \leq \dfrac{f(Y)}{t(Y)}$, FAÇA $X = Y$

SE NÃO, REPITA O PROCEDIMENTO

OBS.: MOSTRA-SE QUE O NÚMERO DE ITERAÇÕES NECESSÁRIAS PARA FINALIZAR O PROCEDIMENTO COM UM VALOR ACEITO É c.

ILUSTRAÇÃO GRÁFICA:



USAMOS $Y \sim U(0,1)$

$\Rightarrow$ AS SUAS AMOSTRAS DISTRIBUEM-SE UNIFORMEMENTE NO INTERVALO UNITÁRIO

$\Rightarrow$ SE O PONTO A É GERADO, COMO $f(A) \simeq t(A)$, O PONTO SERÁ PROVAVELMENTE ACEITO, POIS $\dfrac{f(A)}{t(A)} \simeq 1$.

SE O PONTO B É GERADO, COMO $f(B)/t(B)$ É PEQUENO, A AMOSTRA SERÁ PROVAVELMENTE REJEITADA.

$\Rightarrow$ ESTE RESULTADO MOSTRA-SE DE ACORDO COM UMA AMOSTRAGEM DE $f(x)$.

3. GERAÇÃO DE VARIÁVEIS DISCRETAS:

- VARIÁVEIS COM CONTRADOMÍNIO FINITO

ASSUMINDO

$$P_R[X = x_j] = p_j, \quad j = 1, 2, \ldots, n$$

PODEMOS SIMULAR X UTILIZANDO UMA VARIÁVEL UNIFORME $U \sim U(0,1)$:

$$X = \begin{cases} x_1, & \text{se } U < p_1 \\ x_2, & \text{se } p_1 \leq U < p_1 + p_2 \\ \vdots & \\ x_j, & \text{se } \sum_{k=1}^{j-1} p_k \leq U < \sum_{k=1}^{j} p_k \\ \vdots & \end{cases}$$

$\Rightarrow$ O INTERVALO $[0,1]$ É DIVIDIDO EM n SUBINTERVALOS COM COMPRIMENTOS $p_1, p_2, \ldots, p_n$, E ESCOLHEMOS X DE ACORDO COM O INTERVALO EM QUE A VARIÁVEL U É GERADA.

- VARIÁVEL DE POISSON

$\Rightarrow$ V.A. DISCRETA COM CONTRADOMÍNIO INFINITO

$$P_R[X = k] = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

$\Rightarrow$ UTILIZAMOS O SEGUINTE RESULTADO:
PARA UMA SEQUÊNCIA i.i.d. DE VARIÁVEIS $U_i \sim U(0,1)$, MOSTRA-SE QUE

$$P_R[U_1 \geq p, U_1U_2 \geq p, \ldots, U_1U_2\cdots U_k \geq p, U_1U_2\cdots U_{k+1} < p] =$$

$$= \frac{p}{k!}(-1)^k[\log p]^k, \quad 0 \leq p \leq 1$$
$$k = 1, 2, \ldots$$

USANDO

$$p = e^{-\lambda}$$

NA EXPRESSÃO ANTERIOR, OBTEMOS A DENSIDADE DE POISSON:

$$Pr[U_1 \geq p, U_1U_2 \geq p, \ldots, U_1U_2\ldots U_k \geq p, U_1U_2\ldots U_{k+1} < p] =$$

$$= \frac{e^{-\lambda}\lambda^k}{k!}, \quad \text{PARA } p = e^{-\lambda}$$

$\Rightarrow$ PODEMOS EMPREGAR O SEGUINTE PROCEDIMENTO PARA GERAR AMOSTRAS X DA DENSIDADE DE POISSON:

$$\text{SE } U_1 < e^{-\lambda}, \quad X = 0$$

$$\left.\begin{array}{l} \text{SE } U_1 \geq e^{-\lambda}, \text{ E} \\ U_1U_2 < e^{-\lambda} \end{array}\right\} X = 1$$

$$\vdots$$

$$\left.\begin{array}{l} \text{SE } U_1 \geq e^{-\lambda}, \\ U_1U_2 \geq e^{-\lambda} \\ \vdots \\ U_1U_2\ldots U_k \geq e^{-\lambda}, \text{ E} \\ U_1U_2\ldots U_{k+1} < e^{-\lambda} \end{array}\right\} X = k$$

4. GERAÇÃO DE VARIÁVEIS GAUSSIANAS

- MÉTODO DE BOX-MULLER:

DUAS VARIÁVEIS X E Y, AMBAS $\sim N(0,1)$ PODEM SER GERADAS A PARTIR DE UMA VARIÁVEL DE RAYLEIGH

$$\Rightarrow \text{DENSIDADE } f(\rho) = \rho e^{-\rho^2/2}, \quad \rho \geq 0$$

E UMA VARIÁVEL UNIFORME EM $[0, 2\pi]$

$$\Rightarrow \text{DENSIDADE } g(\theta) = \frac{1}{2\pi}, \quad 0 \leq \theta < 2\pi$$

UTILIZANDO-SE A TRANSFORMAÇÃO

$$\begin{cases} X = \rho \cos\theta \\ Y = \rho \operatorname{sen}\theta \end{cases}$$

AS VARIÁVEIS ρ E θ PODEM SER GERADAS A PARTIR DE DUAS VARIÁVEIS UNIFORMES INDEPENDENTES EM $[0,1]$, COMO

$$\rho = \sqrt{-2 \log U_1}$$

$$\theta = 2\pi U_2$$

PORTANTO,

$$X = \sqrt{-2 \log U_1} \cos(2\pi U_2)$$

$$Y = \sqrt{-2 \log U_1} \operatorname{sen}(2\pi U_2)$$

5. SIMULAÇÃO DE UM VETOR ALEATÓRIO GAUSSIANO

⇒ QUEREMOS OBTER UM VETOR ALEATÓRIO GAUSSIANO $\tilde{X}$ COM UMA CERTA MATRIZ DE COVARIÂNCIAS

$$K_x = COV(\tilde{X}, \tilde{X}^T)$$

SEJA $\tilde{Z} = (Z_1, \ldots, Z_n)^T$ UM VETOR DE n VARIÁVEIS GAUSSIANAS INDEPENDENTES COM MÉDIA ZERO E VARIÂNCIA UNITÁRIA

$$\Rightarrow \tilde{Z} \sim N(\tilde{0}, I)$$

i.e., $E[\tilde{Z}] = \tilde{0}$

$COV(\tilde{Z}, \tilde{Z}^T) = I$ (MATRIZ IDENTIDADE)

- ABORDAGEM VIA DECOMPOSIÇÃO DE CHOLESKY

FAZEMOS

$$\tilde{X} = L\tilde{Z} \Rightarrow COV(\tilde{X}, \tilde{X}^T) = L I L^T = L L^T$$

⇒ COMO NÓS REQUEREMOS

$$COV(\tilde{X}, \tilde{X}^T) = K_x \Rightarrow L L^T = K_x$$

FATORAÇÃO DE CHOLESKY:

SE A MATRIZ K_x É DEFINIDA POSITIVA, i.e.,

$$\tilde{Y}^T K_x \tilde{Y} > 0, \quad \forall \tilde{Y} \neq 0$$

A MATRIZ L EXISTE, É TRIANGULAR INFERIOR E ÚNICA.

UMA MATRIZ DE COVARIÂNCIAS É SEMPRE DEFINIDA POSITIVA, A MENOS QUE UMA DAS VARIÁVEIS SEJA UMA COMBINAÇÃO LINEAR EXATA DAS DEMAIS.

EXEMPLO BIDIMENSIONAL:

$$K_x = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} l_{11} & l_{12} \\ 0 & l_{22} \end{pmatrix} \qquad l_{12} = l_{21}$$

$$l_{11}^2 = \sigma_1^2 \therefore l_{11} = \sigma_1$$

$$l_{11} l_{12} = \rho_{12}\sigma_1\sigma_2 \therefore l_{12} = l_{21} = \rho_{12}\sigma_2$$

$$l_{12}^2 + l_{22}^2 = \sigma_2^2 \therefore l_{22} = \sigma_2 \sqrt{1-\rho_{12}^2}$$

$$\therefore L = \begin{pmatrix} \sigma_1 & 0 \\ \rho_{12}\sigma_2 & \sqrt{1-\rho_{12}^2}\,\sigma_2 \end{pmatrix}$$

$\Rightarrow$ UM VETOR ALEATÓRIO GAUSSIANO COM MATRIZ DE COVARIÂNCIAS K_x E VETOR DE MÉDIAS $\tilde{\mu}$ PODE SER OBTIDO COMO

$$\tilde{X} = L\tilde{z} + \tilde{\mu}$$