UFMG

ENG
Departamento de
Engenharia Elétrica

UNIVERSIDADE FEDERAL DE MINAS GERAIS

ESCOLA DE ENGENHARIA

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

INTRODUÇÃO AOS PROCESSOS ESTOCÁSTICOS
LISTAS 1 E 2

Aluna: Priscila Aparecida Dias Nicácio

Professor: Eduardo Mazoni Andrade Marcal Mendes.

Belo Horizonte, Agosto de 2025

SUMÁRIO

1. INTRODUÇÃO

Esta lista de exercícios foi elaborada com o objetivo de aprofundar a compreensão dos conceitos fundamentais de probabilidade e processos aleatórios, conforme apresentados no livro Intuitive Probability and Random Processes using MATLAB de Stephen M. Kay (2006).

Os exercícios selecionados abrangem tanto situações discretas quanto contínuas, envolvendo análise combinatória, independência de eventos, distribuição de variáveis aleatórias, e interpretação de condições probabilísticas. Além disso, inclui a exploração de técnicas de simulação em MATLAB, permitindo ao estudante comparar resultados teóricos com estimativas numéricas, fortalecendo a intuição sobre o comportamento de fenômenos aleatórios. Ao longo da resolução dos exercícios, ocorre a construção de espaços amostrais apropriados, identificação de relações de dependência ou independência entre eventos e compreensão da influência de diferentes distribuições na probabilidade de ocorrência de determinados resultados.

A abordagem combinatória e a utilização de simulações proporcionam uma visão prática e intuitiva, mostrando como conceitos abstratos de probabilidade se manifestam em situações concretas e experimentais. Mais do que calcular probabilidades isoladas, os exercícios buscam desenvolver a habilidade de modelar situações reais sob a perspectiva probabilística, interpretando o significado dos resultados obtidos e reconhecendo limitações dos modelos utilizados.

Dessa forma, as listas não apenas revisam conceitos essenciais, mas também incentivam o pensamento crítico e a capacidade de aplicar ferramentas computacionais para resolver problemas complexos de maneira eficiente. Em última análise, o objetivo é formar base sólida e intuitiva que permita superar desafios probabilísticos em contextos acadêmicos e profissionais, utilizando tanto a análise teórica quanto a simulação prática como instrumentos complementares de aprendizagem.

2. EXERCÍCIOS LISTA 1

Exercício 1: Em uma família de quatro adultos a probabilidade de uma pessoa estar fora de casa é de 60%. Determine a probabilidade de quaisquer duas pessoas da família estarem em casa simultaneamente. Resolva esse problema de duas maneiras:

a) adotando um modelo probabilístico;

família de 4 adultos = 40% = 0,4
probabilidade de estar fora de casa: 60% = 0,6

ⓐ X_i, com $P(X_i = 1) = 0{,}4$ (em casa)
$P(X_i = 0) = 0{,}6$ (fora de casa)

Probabilidade de 2 pessoas estarem em casa:
→ Binomial:
$$P(2 \text{ em casa}) = \binom{4}{2} (0{,}4)^2 \cdot (0{,}6)^2$$

→ cálculo: $\binom{4}{2} = 6$

$p = 6 \cdot (0{,}4)^2 \cdot (0{,}6)^2$
$p = 6 \cdot 0{,}16 \cdot 0{,}36$
$p = 6 \cdot 0{,}0576$
$p = 0{,}3456$, probabilidade: $\boxed{34{,}56\%}$

O modelo assume independência entre os membros da família. Se um sai, é possível que outros saiam também, logo o modelo pode subestimar ou superestimar a probabilidade real.

b) por simulação. Critique o modelo probabilístico escolhido.

<u>Simulação no Matlab:</u>

```
Nsim = 1e6;   % número de simulações
p_casa = 0.4; % probabilidade de estar em casa
count = 0;
for i = 1:Nsim
   pessoas = rand(1,4) < p_casa; % gera 4 números aleatórios
   if sum(pessoas) == 2
      count = count + 1;
   end
end

prob_sim = count / Nsim;
disp(['Probabilidade por simulação: ', num2str(prob_sim)]);
```

<u>Resposta do Matlab:</u> Probabilidade por simulação: 0.34

Quanto à crítica, o modelo probabilístico adotado assume que cada adulto da família está em casa de forma independente, com probabilidade fixa de 40%. Essa simplificação permite um cálculo direto da probabilidade de exatamente duas pessoas estarem em casa, fornecendo um resultado próximo ao obtido por simulação (~0,34).

Entretanto, o modelo não captura possíveis dependências entre os membros da família, como saídas em conjunto ou eventos que afetam todos simultaneamente, nem considera variações temporais na probabilidade de presença.

Dessa forma, embora útil como aproximação, o modelo não reflete completamente a dinâmica real do comportamento familiar, sendo mais adequado para estudos teóricos do que para previsões precisas do cotidiano.

<u>Exercício 2:</u> A probabilidade de uma variável aleatória T é descrita por uma distribuição gaussiana com média 7 e desvio padrão unitário (ver Eq. 1.2). Determine numericamente a probabilidade do evento $5 \leq T \leq 6$. Dica: ver Exercício 1.14.

$T \sim N(\mu = 7, \sigma = 1)$.
sendo $P(5 \leq T \leq 6)$

$\rightarrow$ Transformando em Z-score:

$$Z = \frac{T - \mu}{\sigma} = \frac{T - 7}{1}$$

$$\Rightarrow P(5 \leq T \leq 6) = P\left(\frac{5-7}{1} \leq Z \leq \frac{6-7}{1}\right)$$

$$P(-2 \leq Z \leq -1)$$

Usando CDF: $P = \phi(-1) - \phi(-2)$

aprox: $\phi(-1) \approx 0{,}1587$

$\phi(-2) \approx 0{,}0228$

$$P \approx 0{,}1587 - 0{,}0228 = 0{,}1359$$

$\boxed{\text{probabilidade} \approx 13{,}59\%}$

<u>Simulação MATLAB:</u>

```
mu = 7; sigma = 1;
prob = normcdf(6, mu, sigma) - normcdf(5, mu, sigma);
disp(['Probabilidade: ', num2str(prob)]);
```

<u>Resposta do Matlab:</u> Probabilidade: 0.13591

<u>Exercício 3:</u> Seja Ui o resultado da seleção aleatória de um número entre 0 e 1, ou seja, ui ∈ [0 1]. Seja a variável aleatória X = U1 + U2 . Qual é a probabilidade dos eventos:

a) 0 ≤ X < 0, 5;
b) 0, 5 ≤ X < 1;
c) 1 ≤ X < 1, 5;
d) 1, 5 ≤ X ≤ 2;
e) 2 < X;

$U_1, U_2 \sim$ Uniforme $[0,1]$ e $X = U_1 + U_2$

Achar $P(X \in [a,b))$:

Soma de duas variáveis uniformes $[0,1]$ → distr. triangular.

$$f_X(x) = \begin{cases} x, & 0 \le x \le 1 \\ 2-x, & 1 < x \le 2 \end{cases}$$

Probabilidades por integral:

a) $0 \le X < 0,5$:

$$P = \int_0^{0,5} x\,dx = [x^2/2]_0^{0,5} = 0,125$$

b) $0,5 \le X < 1$:

$$P = \int_{0,5}^{1} x\,dx = [x^2/2]_{0,5}^{1} = 0,5 - 0,125 = 0,375$$

c) $1 \le X < 1,5$:

$$P = \int_1^{1,5} (2-x)\,dx = [2x - x^2/2]_1^{1,5} = (3 - 1,125) - (2 - 0,5)$$

$$1,875 - 1,5 \Rightarrow 0,375$$

d) $1,5 \le X \le 2$:

$$P = \int_{1,5}^{2} (2-x)\,dx = [2x - x^2/2]_{1,5}^{2}$$

$$(4-2) - (3 - 1,125)$$

$$2 - 1,875$$

$$0,125$$

e) $X > 2$ → Impossível → $P = 0$

Simulação Matlab:

```
Nsim = 1e6;
U1 = rand(Nsim,1);
U2 = rand(Nsim,1);
```

```
X = U1 + U2;

P_a = mean(X >= 0 & X < 0.5);
P_b = mean(X >= 0.5 & X < 1);
P_c = mean(X >= 1 & X < 1.5);
P_d = mean(X >= 1.5 & X <= 2);
P_e = mean(X > 2);

disp([P_a P_b P_c P_d P_e]);
```

<u>Resposta do Matlab:</u> 0.1251 0.3744 0.3762 0.1243 0

Intervalo	Probabilidade
$0 \le X < 0.5$	0.125
$0.5 \le X < 1$	0.374
$1 \le X < 1.5$	0.376
$1.5 \le X \le 2$	0.124
$X > 2$	0

Tabela 1. Intervalo vs. Probabilidade.

<u>Exercício 4</u>: Usando a mesma definição de Ui do item anterior, seja $Y = \sum_{i=1}^{10} U_i$. .
Determine a probabilidade de Y estar nos intervalos [0 0.5), [0.5 1) ... [9.5 10]

$Y = \sum_{i=1}^{10} U_i, \; U_i \sim U[0,1].$

Quero: $P(Y \in [0, 0{,}5)), [0.5, 1), \ldots, [9.5, 10].$

<u>Simulação Matlab:</u>

```
% Número de simulações
N = 1e6;
% Gerando 10 variáveis uniformes por simulação
U = rand(N, 10);

% Somando para obter Y
Y = sum(U, 2);
```

```
% Definindo os intervalos de 0 a 10 com passo 0.5
edges = 0:0.5:10;

% Calculando histograma normalizado para obter probabilidades
counts = histcounts(Y, edges);
probabilities = counts / N;

% Exibindo os resultados
intervals = strcat("[", string(edges(1:end-1)), ", ", string(edges(2:end)), ")");
table(intervals', probabilities', 'VariableNames', {'Interval', 'Probability'})
```

Resposta do Matlab: ans = 20×2 table

Interval	Probability
"[0, 0.5)"	0
"[0.5, 1)"	0
"[1, 1.5)"	1.6e-05
"[1.5, 2)"	0.000274
"[2, 2.5)"	0.002188
"[2.5, 3)"	0.010761
"[3, 3.5)"	0.036979
"[3.5, 4)"	0.088478
"[4, 4.5)"	0.15519
"[4.5, 5)"	0.20545
"[5, 5.5)"	0.20548
"[5.5, 6)"	0.15607
"[6, 6.5)"	0.088961
"[6.5, 7)"	0.036754
"[7, 7.5)"	0.010937
"[7.5, 8)"	0.002146
"[8, 8.5)"	0.000285
"[8.5, 9)"	1.5e-05
"[9, 9.5)"	0
"[9.5, 10)"	0

Tabela 2. Resultados exercício 4 (Intervalo vs. Probabilidade).

Exercício 5: Suponha que Ω contenha N elementos. Mostre que o número de Bell, BN, de diferentes decomposições de Ω é dada pela fórmula: $B_N = e^{-1} \sum_{k=0}^{\infty} \frac{k^N}{k!}$,

Dica: Mostre que , $B_N = \sum_{k=0}^{N-1} C_{N-1}^{k} B_k$, onde C_a^b é a fórmula usual de combinação e B_0= 1. Use isso para verificar que a série acima satisfaz a mesma relação de recorrência. Considerando o resultado anterior e que Ω = {1, 2, 3, 4}, calcule o número de decomposições.

fórmula: $BN = e^{-1} \sum_{K=0}^{\infty} \frac{K^N}{K!}$

sendo o número de Bell (BN) o número de partições de um conjunto com N elementos.

Para $\Omega = \{1, 2, 3, 4\}$, $N = 4$

$B_0 = 1$, $B_1 = 1$, $B_2 = 2$, $B_3 = 5$, $B_4 = 15$.

$B_0 = 1$

$B_1 = \sum_{K=0}^{0} \binom{0}{K} B_K = \binom{0}{0} B_0 = 1 . 1 = 1$

$B_2 = \sum_{K=0}^{1} \binom{1}{K} B_K = \binom{1}{0} B_0 + \binom{1}{1} B_1 = 1.1 + 1.1 = 2$

$B_3 = \sum_{K=0}^{2} \binom{2}{K} B_K = \binom{2}{0} B_0 + \binom{2}{1} B_1 + \binom{2}{2} B_2 = 1.1 + 2.1 + 1.2 =$

$1 + 2 + 2 = 5$.

$B_4 = \sum_{K=0}^{3} \binom{3}{K} B_K = \binom{3}{0} B_0 + \binom{3}{1} B_1 + \binom{3}{2} B_2 + \binom{3}{3} B_3$

$B_4 = 1.1 + 3.1 + 3.2 + 1.5$

$B_4 = 1 + 3 + 6 + 5 = 15$ → número de decomposições

Simulação no Matlab:

```
% Número de elementos
N = 4;

% Inicializando vetor de Bell
B = zeros(1,N+1); % de B0 a B4
```

```
B(1) = 1;        % B0 = 1

% Calculando B1 até BN usando a recorrência
for n = 1:N
   B(n+1) = 0;
   for k = 0:(n-1)
      B(n+1) = B(n+1) + nchoosek(n-1,k) * B(k+1);
   end
end

% Exibindo números de Bell
B

% B_N é o último valor
fprintf('O número de Bell para N=%d é %d\n', N, B(N+1));
```

Resposta do Matlab:

```
B =    1    1    2    5    15
O número de Bell para N = 4 é 15
Verificando a fórmula da série: Aproximação pelo somatório: 15
```

Exercício 6: Três jogadores A, B e C disputam um torneio de tênis. Inicialmente, A joga com B e o vencedor joga com C; e assim por diante. O torneio termina quando um jogador ganha duas vezes em seguida ou quando são distribuídos, ao todo, quatro partidas. Considere que todos os jogadores têm a mesma probabilidade de ganhar um jogo. Usando simulação, determine as probabilidades de cada um dos eventos possíveis.

Simulação no Matlab:

```
Nsim = 1e6; % número de simulações
vitorias = zeros(3,1); % contagem de vitórias: A,B,C

for i = 1:Nsim
   % Inicializa torneio
   partidas = 0;
   last_winner = 0;
```

```
    consecutivas = 0;

    % Primeira partida: A x B
    players = [1 2];
    while partidas < 4 && consecutivas < 2
        winner = players(randi(2)); % sorteio do vencedor
        partidas = partidas + 1;

        % Verifica vitórias consecutivas
        if winner == last_winner
            consecutivas = consecutivas + 1;
        else
            consecutivas = 1;
        end
        last_winner = winner;
        % Define próximo jogo
        switch winner
            case 1 % A venceu
                if any(players==3)
                    players = [1 3];
                else
                    players = [1 2];
                end
            case 2 % B venceu
                if any(players==3)
                    players = [2 3];
                else
                    players = [1 2];
                end
            case 3 % C venceu
                players = [3 winner]; % joga contra o perdedor da última
        end
    end

    % Conta vitória final
    vitorias(last_winner) = vitorias(last_winner) + 1;
end
```

% Probabilidades
probabilidades = vitorias / Nsim;
disp('Probabilidades de vitória de cada jogador (A,B,C):')
disp(probabilidades)

Resposta do Matlab: Probabilidades de vitória A = 0.4990, B = 0.5010 e C = 0.

3. EXERCÍCIOS DO LIVRO (KAY – CAPÍTULO 2)

Exercício 2.2(c): Um par de dados justos é lançado. Estime a probabilidade de sair 'snake eyes' ou um 1 em cada dado.

→ Exatamente um 1 em cada dado:
mesma coisa que snake eyes → só (1,1)

→ Pelo menos um 1 em qualquer dado
(1, x) ou (x, 1), incluindo (1, 1);

• snake eyes: $6 \cdot 6 = 36$
(1,1) ⇒ apenas 1 resultado

$P(\text{snake eyes}) = \frac{1}{36} \to 0{,}0278$

• Pelo menos um 1:
$P(\text{pelo menos um } 1) = 1 - P(\text{nenhum } 1)$

Nenhum 1: cada dado tem $\frac{5}{6}$ de chance de não ser 1 → ambos não serem 1.

$P(\text{nenhum } 1) = \frac{5}{6} \cdot \frac{5}{6} = \frac{25}{36}$

$P(\text{pelo menos um } 1) = 1 - \frac{25}{36} = \frac{11}{36} = 0{,}3056$

→ Snake Eyes está incluso no evento de "pelo menos um 1".

Dois dados justos, probabilidade de "snake eyes" ou pelo menos um 1

- Snake eyes: $\frac{1}{36}$ e pelo menos um 1: 1 - $\left(\frac{5}{6}\right)^2$ = $\frac{11}{36}$

Simulação Matlab:

```
% Q2_2_c.m
Nsim = 1e6;
dados = randi(6,Nsim,2);
count = sum(dados(:,1) == 1 & dados(:,2) == 1);
prob_sim = count / Nsim;
fprintf('Probabilidade estimada por simulação: %.5f\n', prob_sim);
```

Resposta do Matlab**:** Probabilidade estimada por simulação: 0.027

Exercício 2.4(c): Estimar o PDF de X e depois comparar com a PDF Gaussiana padrão.

Variável aleatória: X = $\sum_{i=1}^{12}(U_i \ \frac{1}{2})$, U_i ~ U(0,1) e PDF Gaussiana padrão: px(x) = $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$$X = \sum_{i=1}^{12} \left(U_i - \frac{1}{2}\right), \quad U_i \sim \text{Uniform } [0,1]$$

Quero a estimativa da PDF de X e comparar com a PDF Gaussiana:

$$P_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x^2\right)$$

Cada $U_i - \frac{1}{2}$ é uma variável centrada em zero.

$$\mathbb{E}\left[U_i - \frac{1}{2}\right] = \mathbb{E}[U_i] - \frac{1}{2} = \frac{1}{2} - \frac{1}{2} = 0 \text{ ,, média}$$

$$\text{Var}\left(U_i - \frac{1}{2}\right) = \text{Var}(U_i) = \frac{(1-0)^2}{12} = \frac{1}{12} \text{ ,, variância}$$

X é a soma de 12 variáveis independentes:

$$\mathbb{E}[X] = \sum_{i=1}^{12} \mathbb{E}\left[U_i - \frac{1}{2}\right] = 12 \cdot 0 = 0 \text{ ,,}$$

$$\text{Var}(X) = \sum_{i=1}^{12} \text{Var}\left(U_i - \frac{1}{2}\right) = 12 \cdot \frac{1}{12} = 1 \text{ ,,}$$

Aplicação do Teorema Central do Limite

↳ Soma de muitas variáveis independentes e identicamente distribuídas tende a uma normal, mesmo que cada variável não seja normal.

X = soma de 12 variáveis uniformes ajustadas → razoável aproximação normal:

$$X \sim N(\mu = 0, \sigma^2 = 1)$$

PDF gaussiana aproximada:

$$pX(x) \approx \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$ com média 0 e variância 1

→ mesmo que cada $U_i - \frac{1}{2}$ seja uniforme, a soma de 12 delas centralizadas em zero se comporta como uma variável normal.

Simulação no Matlab:

```
Nsim = 1e6;
U = rand(Nsim,12);
X = sum(U - 0.5, 2);

% Histograma para estimativa da PDF
edges = -6:0.2:6;
counts = histcounts(X, edges, 'Normalization', 'pdf');
centers = edges(1:end-1) + diff(edges)/2;

% PDF Gaussiana teórica
x_vals = -6:0.01:6;
pdf_gauss = normpdf(x_vals, 0, sqrt(12*1/12)); % var(X)=12*Var(U_i-0.5)=1

% Plot
figure;
```

bar(centers, counts, 'hist'); hold on;

plot(x_vals, pdf_gauss, 'r', 'LineWidth',2);

xlabel('X'); ylabel('PDF'); legend('Estimada','Gaussiana');

title('Estimativa da PDF de X e comparação com Gaussiana');

<u>Resposta do Matlab</u>**:**

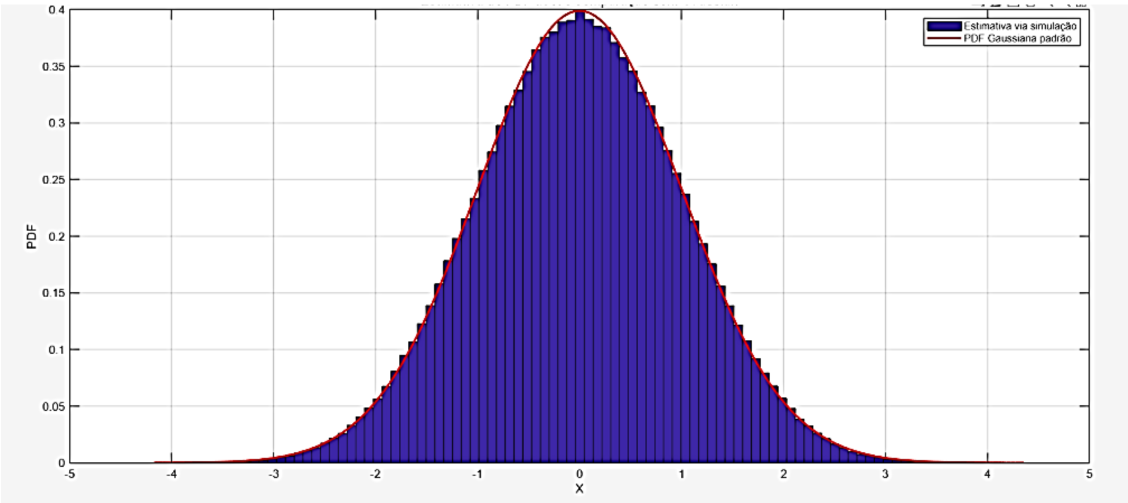

Figura 1. Estimativa da PDF de X e comparação com Gaussiana

Como X é soma de 12 variáveis uniformes ajustadas, pelo Teorema Central do Limite, a PDF se aproxima de uma gaussiana com média 0 e variância 1.

<u>Exercício 2.5(c):</u> Definir X= U1 − U2. Determinar o “most probable range of values”.

U1 e U2 são variáveis aleatórias uniformes. Geralmente, se não especificado, assume-se U1,U2∼Uniform(0,1). X é então a diferença entre essas duas variáveis. Estimar a PDF de X: A PDF de X = U1 − U2 não é uniforme; ela tem formato de “triângulo” (triangular distribution) porque é a convolução de duas PDFs uniformes (uma delas invertida para a subtração). Pode ser feito analiticamente ou simulando via Monte Carlo:

Analiticamente:

Para $x \in [-1,0]$, $f_X(x)=1+x$

Para $x \in [0,1]$, $f_X(x)=1-x$

Fora de $[-1,1]$, $f_X(x)=0$

A PDF tem máximo em x = 0 (porque fX (0) = 1), e cai linearmente até -1 e 1. Então, o valor mais provável é próximo de 0, e o intervalo mais provável é o centro da distribuição, algo como [−0.5,0.5] (75%).

Convolução:

PDF de Ui: $f_U(u) = 1$ para $u \in [0,1]$

$fX(x) = \int_{-\infty}^{\infty} f_{U_1}(u) f_{U_2}(u-x)\, du.$

Ambas são indicadoras em [0,1], o integrando vale 1 só quando $u \in [0,1]$ e $u - x \in [0,1]$,

u em interseção $[0,1] \cap [x, x+1]$.

comprimento da interseção $1 - |x|$ para $|x| \le 1$.

Já que: $fX(x) = \begin{cases} 1 - |x|, & -1 \le x \le 1; \\ 0, & \text{caso contrário.} \end{cases}$ PDF triangular

CDF:

$$fX(x) = \begin{cases} 0, & x < -1, \\ \dfrac{(x+1)^2}{2}, & -1 \le x \le 0, \\ 1 - \dfrac{(1-x)^2}{2}, & 0 < x \le 1, \\ 1, & x > 1. \end{cases}$$

$E[X] = 0$ (simetria),

$Var(X) = Var(U_1) + Var(U_2) = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$.

- densidade máx em 0;
- decresce linearmente até os extremos ± 1, os valores mais prováveis estão próximos de zero.

Qualquer intervalo: $[-a, a]$: $P(|X| \le a) = 2a - a^2$ $(0 \le a \le 1)$.

$a = 0{,}5 \Rightarrow P = 0{,}75$; $a = 0{,}2 \Rightarrow P = 0{,}36$

Simulação no Matlab:

```
Nsim = 1e6;
U1 = rand(Nsim,1);
U2 = rand(Nsim,1);
X = U1 - U2;
% Histograma PDF
```

```
edges = -1:0.02:1;
counts = histcounts(X, edges, 'Normalization', 'pdf');
centers = edges(1:end-1) + diff(edges)/2;

% Plot PDF
figure;
bar(centers, counts, 'hist');
xlabel('X'); ylabel('PDF');
title('Estimativa da PDF de X = U1 - U2');
hold on;

% Destacar intervalo mais provável
% Considerando intervalo central de -0.5 a 0.5
prob_interval = [-0.5 0.5];
y_lim = ylim; % limites do eixo y
patch([prob_interval(1) prob_interval(2) prob_interval(2) prob_interval(1)], ...
    [0 0 y_lim(2) y_lim(2)], ...
    [1 0.9 0.8], 'FaceAlpha',0.3, 'EdgeColor','none'); % destaque em transparência

legend('PDF estimada','Intervalo mais provável');
```

<u>Resposta do Matlab:</u>

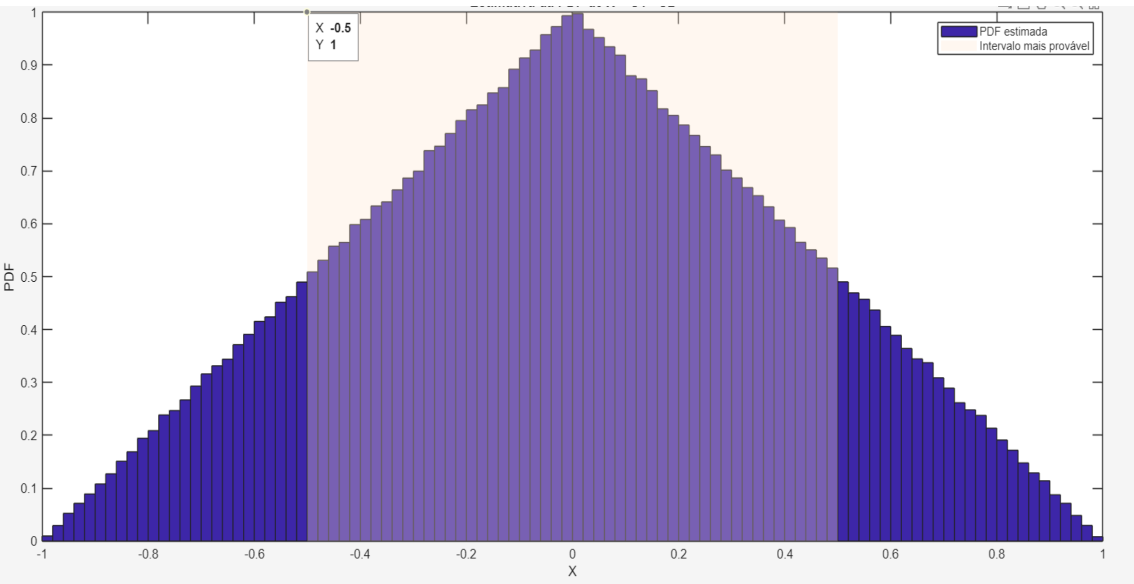

Figura 2: Estimativa da PDF de X = U1 - U2.

1. Formato triangular:

A PDF tem pico em X=0 e decresce linearmente até X=−1 e X=1. Isso indica que a diferença entre U1 e U2 é mais provável de ser próxima de zero, ou seja, U1 e U2 tendem a ser parecidas com mais frequência.

2. Intervalo mais provável:

A faixa destacada (por exemplo, [−0.5,0.5]) mostra os valores de X que ocorrem com maior frequência, onde a PDF é mais alta. Em outras palavras, há alta probabilidade da diferença U1−U2 estar próxima de zero.

3. Extremos menos prováveis:

Valores próximos de −1 ou 1 são raros, porque isso só acontece quando U1 é muito menor ou muito maior que U2, situação menos frequente.

Visualmente:

- Pico em 0 → diferença mais provável.
- Forma triangular → probabilidade decresce uniformemente em direção aos extremos.
- Intervalo sombreado → região de alta probabilidade.

<u>Exercício 2.7(c):</u> Tem-se uma variável aleatória discreta X que assume os valores:

$$X = \begin{cases} 1 \text{ Com probabilidade } p_1 = 0.1 \\ 2 \text{ com probabilidade } p_2 = 0.2 \\ 3\ Com\ probabilidade\ p_3\ =\ 0.7 \end{cases}$$

O objetivo:

1. Gerar várias amostras (realizações) dessa variável aleatória.
2. Baseando-se nas amostras, estimar empiricamente as probabilidades de cada valor.

X	probabilidade pi
1	0.1
2	0,2
3	0.7

Com 10 números aleatórios r:

r	X gerado
0,05	1
0,12	2
0,85	3
0,33	3
0,02	1
0,50	3
0,25	2
0,70	3
0,10	1
0,90	3

sendo:

$X = 1$ se $r < 0,1$

$X = 2$ se $0,1 < r \leq 0,3$

$X = 3$ se $r > 0,3$

↓

Possibilidades cumulativas.

1 apareceu 3 vezes
2 apareceu 2 vezes
3 apareceu 5 vezes

Total de amostras: (10)

$\tilde{p}_1 = \frac{3}{10} = 0,3$, $\tilde{p}_2 = \frac{2}{10} = 0,2$, $\tilde{p}_3 = \frac{5}{10} = 0,5$

→ maior número de amostras: + próx do teórico
→ menor número de amostras: + flutuações

Simulação no Matlab:

```
Nsim = 1e6;
vals = [1 2 3];
probs = [0.1 0.2 0.7];
X = randsample(vals,Nsim,true,probs);

P1 = mean(X==1); P2 = mean(X==2); P3 = mean(X==3);
disp(['Estimativa P(X=1)=', num2str(P1), ' P(X=2)=', num2str(P2), ' P(X=3)=',
num2str(P3)]);
```

Resposta do Matlab: Probabilidades estimadas de X: 0.1002 0.1997 0.7001

Exercício 2.9(c): X é uma variável aleatória gaussiana (normal), ou seja: X~N(μ,σ^2). Achar a estimativa da média de X+1 e também o valor verdadeiro

sendo $X \sim N(0,1)$

amostra x_i	calculo de $x+1$
-0,5	0,5
0,2	1,2
1,0	2,0
-1,2	-0,2
0,7	1,7

Estimativa da média:

$X+1$: $\hat{\mu}_{x+1} = \frac{0,5+1,2+2,0-0,2+1,7}{5} = \frac{5,2}{5} = 1,04$

Valor verdadeiro: $\mathbb{E}[X] = 0$ e $\mathbb{E}[X+1] = \mathbb{E}[X]+1$
$0+1 = 1$

Se eu usar mais amostras, mais próxima a estimativa ficará do valor verdadeiro.

Simulação no Matlab:

```
N = 1e6; % número de amostras
mu = 0;  % média de X
sigma = 1; % desvio padrão de X

X = mu + sigma*randn(N,1); % gera amostras de X
Y = X + 1; % variável X+1

mean_est = mean(Y); % estima a média
true_value = mu + 1;

disp(['Média estimada: ', num2str(mean_est)]);
disp(['Valor verdadeiro: ', num2str(true_value)]);
```

Resposta do Matlab: Média estimada: 0.99862   Valor verdadeiro: 1

Exercício 2.10(c): Estimativa da média de X^2, X gaussiana.

- Teoria: E[X^2] = Var(X)+E[X] ^{2}= 1+0 =1

Escolhi algumas amostras: X_i = -0,5, 0,2, 1,0, -1,2;
calculo de X_i^2 = 0,25; 0,04; 1,0; 1,44.
média: (0,25 + 0,04 + 1,0 + 1,44) / 4 = 0,6825
valor verdadeiro com $\mu = 0$, $\sigma = 1$: 1
$\hookrightarrow E[x^2] = \sigma^2 + \mu^2 = 1 + 0 = 1$
flutua mais com menos amostras.

Simulação no Matlab:

```
N = 1e6; % número de amostras
mu = 0;  % média de X
sigma = 1; % desvio padrão de X

X = mu + sigma*randn(N,1); % gera amostras de X
Y = X.^2; % eleva cada amostra ao quadrado

mean_est = mean(Y); % estima a média
true_value = sigma^2 + mu^2;

disp(['Média estimada: ', num2str(mean_est)]);
disp(['Valor verdadeiro: ', num2str(true_value)]);
```

Resposta do Matlab: Média estimada: 0.99944   Valor verdadeiro: 1

Exercício 2.12(c): X1,X2 gaussianas, y = X1 - X2. Intervalo provável é menor?

- sejam X_1 e X_2, variáveis gaussianas independentes:

$X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$

define-se: $Y = X_1 - X_2$.

$Y = X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$

sendo a variância de Y a soma das variâncias, não a subtração.

→ O Intervalo de uma variável gaussiana depende do desvio padrão:

para X_1: σ_1 , para X_2: σ_2

para Y: $\sqrt{\sigma_1^2+\sigma_2^2}$

→ O desvio padrão de Y é maior do que o de X_1 ou X_2.

Então, o intervalo de Y não é menor

<u>Simulação no Matlab:</u>

```
% Parâmetros das variáveis gaussianas
mu1 = 0; sigma1 = 1;
mu2 = 0; sigma2 = 1;
N = 1e6; % número de amostras

% Gerar amostras
X1 = mu1 + sigma1*randn(N,1);
X2 = mu2 + sigma2*randn(N,1);

% Subtração
Y = X1 - X2;

% Estimativa dos desvios padrões
std_X1 = std(X1);
std_X2 = std(X2);
std_Y  = std(Y);

disp(['Desvio padrão X1: ', num2str(std_X1)]);
disp(['Desvio padrão X2: ', num2str(std_X2)]);
disp(['Desvio padrão Y = X1-X2: ', num2str(std_Y)]);

% Histogramas para visualização
```

```
edges = -6:0.1:6;
figure;
histogram(X1, edges, 'Normalization', 'pdf', 'FaceAlpha',0.5, 'FaceColor','b'); hold on;
histogram(X2, edges, 'Normalization', 'pdf', 'FaceAlpha',0.5, 'FaceColor','r');
histogram(Y, edges, 'Normalization', 'pdf', 'FaceAlpha',0.5, 'FaceColor','g');
xlabel('Valor'); ylabel('PDF');
legend('X1','X2','Y = X1 - X2');
title('Comparação de distribuições gaussianas e da diferença');
```

<u>Resposta do Matlab:</u>

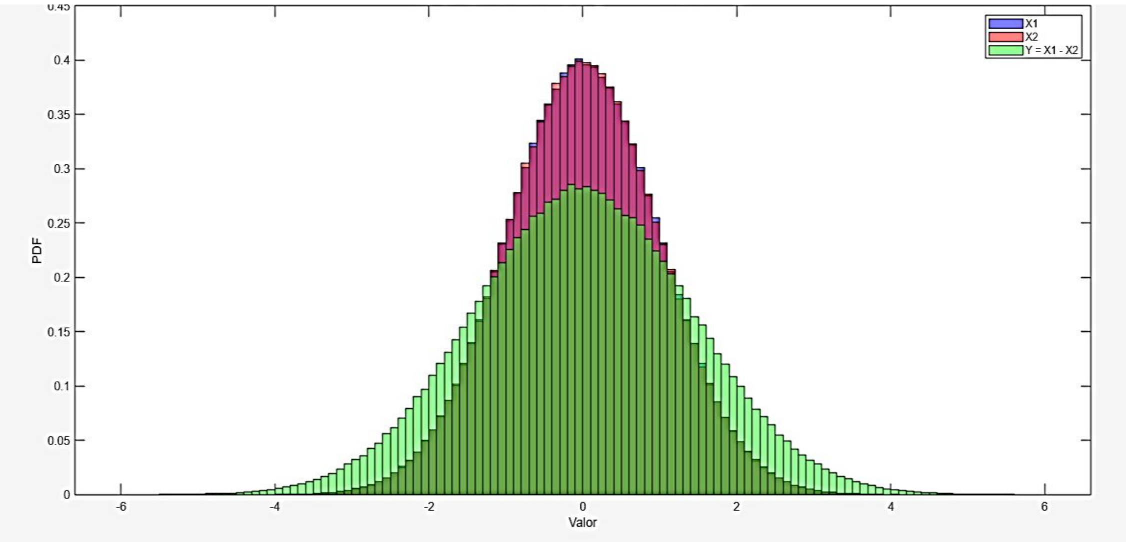

Figura 3. Comparação de distribuições gaussianas e da diferença.

O pico das distribuições de X1 e X2 é mais estreito. O pico de Y é mais largo, confirmando que o intervalo provável aumenta ao subtrair duas gaussianas.

<u>Resposta do Matlab:</u>

Desvio padrão X1: 1.0003

Desvio padrão X2: 0.99972

Desvio padrão Y = X1-X2: 1.4136

<u>**Exercício 2.15(c):**</u> Y_1= X_1+ 0.1X_2, Y_2= X_1+ 0.2X_2, X_1, X_2~N(0,1). Plot de scatter plot Y_1 vs Y_2. Determinar aproximadamente Y_2 se Y_1=1.

Y_2 em função de Y_1 e X_2:

$$Y_1 = X_1 + 0{,}1X_2 \rightarrow X_1 = Y_1 - 0{,}1X_2$$

$$Y_2 = X_1 + 0{,}2X_2 = (Y_1 - 0{,}1X_2) + 0{,}2X_2 = Y_1 + 0{,}1X_2$$

para $Y_1 = 1$ tem-se $Y_2 = 1 + 0{,}1X_2$.
X_2 é variável, Y_2 não é único, mas varia em torno de $1 \pm 0{,}1 *$ amplitude(X_2).

• pares: (X_1, X_2) e calculo de (Y_1, Y_2) e plotar pontos.

X_1	X_2	$Y_1 = X_1 + 0{,}1X_2$	$Y_2 = X_1 + 0{,}2X_2$
0	0	0	0
1	0	1	1
0	1	0,1	0,2
1	1	1,1	1,2

relação linear entre Y_1 e Y_2

Simulação no Matlab:

```
N = 1000; % número de amostras
X1 = randn(N,1); % gaussianas
X2 = randn(N,1);

Y1 = X1 + 0.1*X2;
Y2 = X1 + 0.2*X2;

figure;
scatter(Y1,Y2,'.');
xlabel('Y1'); ylabel('Y2');
title(Diagrama de dispersão de Y1 versus Y2);

% Aproximar Y2 para Y1 = 1
X2_mean = mean(X2); % valor médio aproximado
Y2_est = 1 + 0.1*X2_mean;
disp(['Aproximação de Y2 para Y1=1: ', num2str(Y2_est)]);
```

Resposta do Matlab: Aproximação de Y2 para Y1 = 1: 1.0001

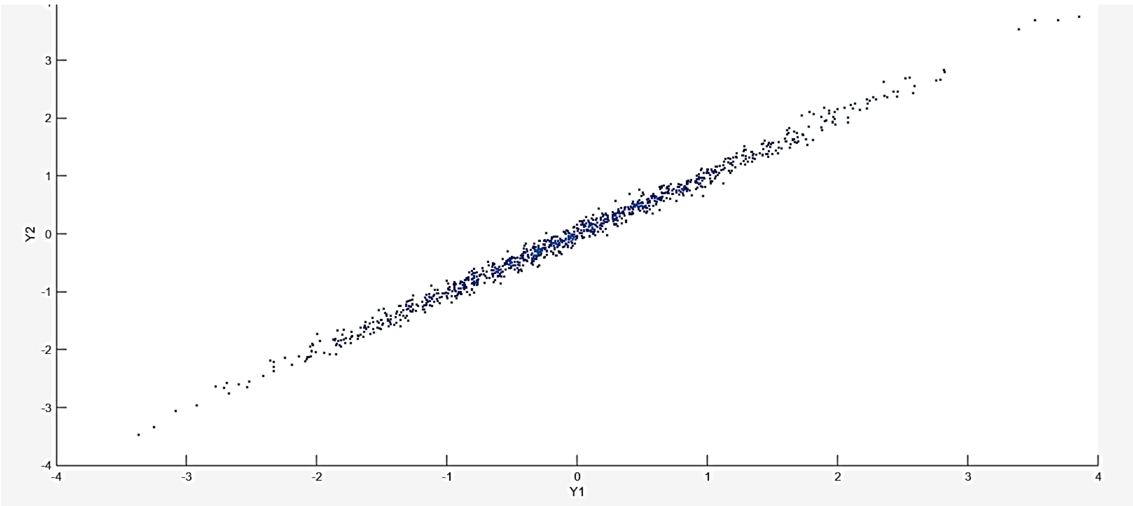

Figura 4. Diagrama de dispersão de Y1 versus Y2

<u>Exercício 2.16(c,w):</u> X1 = U1, X2 = U1 + U2, U1, U2~ U [0,1]

- Scatter plot: forma de paralelogramo X1 vs X2.
- Explicação: X = e1 * U1 + e2 * U2, com vetores e1=[1;1], e2=[0;1], mostra combinação linear → paralelogramo.

U_1	U_2	X_1	X_2
0	0	0	0
0	1	0	1
1	0	1	1
1	1	1	2

Pontos ocupam a região delimitada pelos vetores [1,1] e [0,1].

$U_1, U_2 \in [0,1]$; $e_1 = [1;1]$, $e_2 = [0;1]$

$$X = U_1\begin{bmatrix}1\\1\end{bmatrix} + U_2\begin{bmatrix}0\\1\end{bmatrix} = U_1 e_1 + U_2 e_2$$

<u>Simulação no Matlab:</u>

```
N = 1000;
U1 = rand(N,1);
U2 = rand(N,1);

X1 = U1;
X2 = U1 + U2;

figure;
```

```
scatter(X1,X2,'.');
xlabel('X1'); ylabel('X2');
title(Gráfico de dispersão de X1 versus X2);
```

Resposta do Matlab:

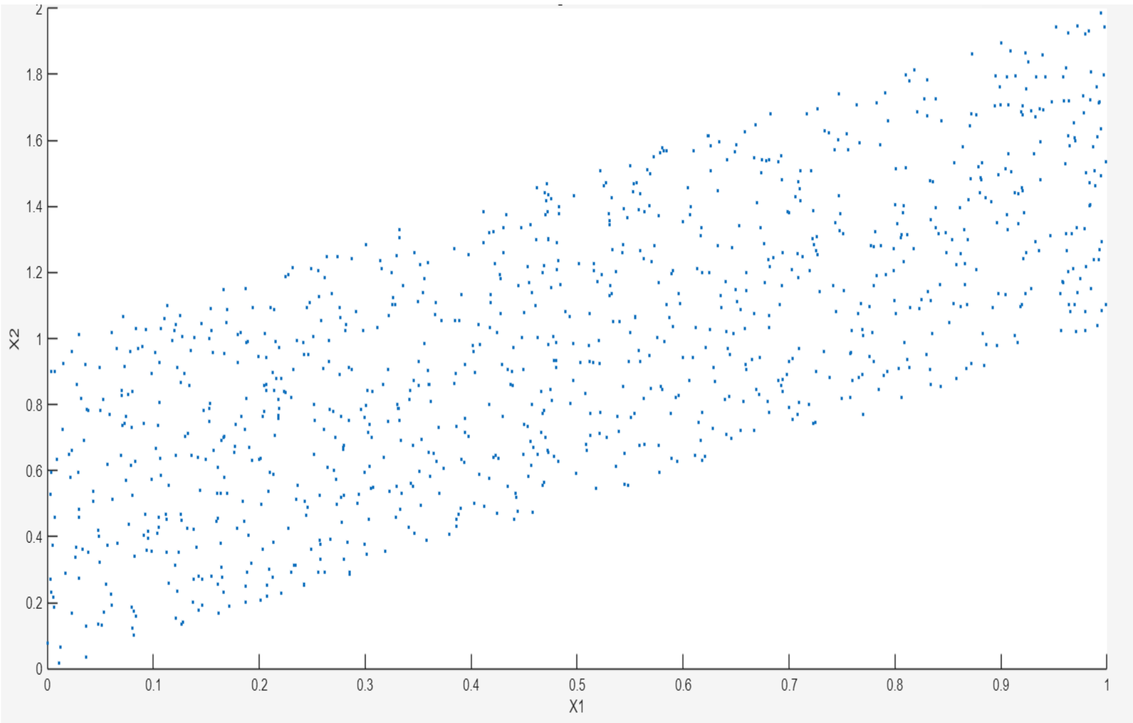

Figura 5. Gráfico de dispersão de X1 versus X2.

4. EXERCÍCIOS LISTA 2

Exercício 1: Considere um alfabeto que tem um total de n letras. Dentro todas as palavras formadas com 3 letras, uma delas é escolhido ao acaso.

Tomando a letra s como exemplo, foram escolhidos os seguintes eventos:

a) A - a palavra escolhida começa com a letra s;
b) B - a palavra escolhida tem a letra s no meio;
c) C - a palavra escolhida tem exatamente duas letras iguais.

i) Estabeleça o espaço de probabilidade.
ii) Determine se os eventos serão independentes dois a dois e três a três
iii) Use simulação e mostre como seria no caso em que n = 6.

ⓐ Espaço de probabilidade:

Total de palavras de 3 letras (com repetição):

$$|\Omega| = n^3$$

$|A| = n^2$

$$P(A) = \frac{n^2}{n^3} = \frac{1}{n}$$

$|B| = n^2$

$$P(B) = \frac{1}{n}$$

Para C: palavras com 2 letras iguais e a terceira diferente.

→ letras que se repete: n opções
→ letra diferente $n-1$ opções
→ posição da letra diferente: 3 opções

$$|C| = 3n(n-1)$$

$$P(C) = \frac{3n(n-1)}{n^3}$$

ⓑ Independência

$P(A\cap B) = P(A)P(B)$, $P(A\cap C) = P(A)P(C)$,

$$P(B\cap C) = P(B)P(C)$$

Para independência 3 a 3:

$$P(A\cap B\cap C) = P(A)\,P(B)\,P(C)$$

$P(A\cap B)$:

primeira letra s, segunda letra s, terceira (qualquer).

$|A\cap B| = n$

$$P(A\cap B) = \frac{n}{n^3} = \frac{1}{n^2}$$

e $P(A)P(B) = \frac{1}{n}\cdot\frac{1}{n} = \frac{1}{n^2}$, A e B são independentes.

(c) A e C

$A \cap C$ (palavra começa com s e tem 2 letras iguais).

- Caso 1:

A letra que se repete é s;
A primeira letra já é s, entre posições 2 e 3 deve haver exatamente um s e um distinto (não s). Número de palavras: 2 (escolher se o s fica na posição 2 ou 3) × $(n-1)$ (escolher a letra distinta) → $(2(n-1))$.

- Caso 2:

A letra que se repete é alguma $t \neq s$. Então, posições 2 e 3 devem ser ambas t. Número de palavras: $n-1$.

Total: $|A \cap C| = 2(n-1) + (n-1) = 3(n-1)$

$$P(A \cap C) = \frac{3 \cdot (n-1)}{n^3}.$$

Comparando $P(A) P(C)$:

$$P(A) P(C) = \frac{1}{n} \cdot \frac{3(n-1)}{n^2} = \frac{3(n-1)}{n^3}.$$

portanto: $P(A \cap C) = P(A) P(C) \rightarrow$ A e C são independentes

OBS:
↳ (surgiu a igualdade porque fixar a 1ª letra em s distribui os casos de C proporcionalmente).

B e C

por simetria (trocar posição $1 \leftrightarrow 2$) o mesmo raciocínio dá:

$$P(B \cap C) = \frac{3(n-1)}{n^3} = P(B)\,P(C).$$

Logo: B e C são independentes.

Independência tripla: A, B, C.

$P(A \cap B \cap C)$

Se A e B ocorrem, as duas primeiras letras são s. Para estar em C (duas iguais), a 3ª deve ser diferente de s.

Logo:

$$|A \cap B \cap C| = n-1 \to P(A \cap B \cap C) = \frac{n-1}{n^3}$$

produto das marginais:

$$P(A)\,P(B)\,P(C) = \frac{1}{n} \cdot \frac{1}{n} \cdot \frac{3(n-1)}{n^2} = \frac{3(n-1)}{n^4}$$

Comparando:

$$P(A \cap B \cap C) = \frac{n-1}{n^3}, \quad P(A)\,P(B)\,P(C) = \frac{3(n-1)}{n^4}$$

A razão é:

$$\frac{P(A \cap B \cap C)}{P(A)P(B)P(C)} = \frac{n}{3}$$

só há independência completa se $\frac{n}{3} = 1$, quando $n = 3$.

→ Para qualquer n os pares (A, B), (A, C), (B, C) são independentes (pares independentes) mas os três eventos não são mutuamente independentes, salvo no $n = 3$.

Então:

- A e B := independentes
- A e C = independentes
- B e C = independentes
- A, B, C simultaneamente = não independentes em geral; são independentemente mutuais somente se $n = 3$.

Simulação em Matlab:

```
% Simulação de palavras de 3 letras
n = 6; % número de letras
Nsim = 1e6; % número de simulações
countA = 0; countB = 0; countC = 0;
countAB = 0; countAC = 0; countBC = 0;
for i = 1:Nsim
   palavra = randi(n,1,3); % palavra de 3 letras
   % Eventos
   if palavra(1) == 1  % letra s é representada por 1
      countA = countA + 1;
   end
   if palavra(2) == 1
      countB = countB + 1;
   end
   if sum(palavra(1)==palavra) == 2 || sum(palavra(2)==palavra) == 2 ||
sum(palavra(3)==palavra) == 2
      countC = countC + 1;
   end
   % Interseções
   if palavra(1) == 1 && sum(palavra(1)==palavra) == 2
      countAC = countAC + 1;
   end
   if palavra(2) == 1 && sum(palavra(2)==palavra) == 2
      countBC = countBC + 1;
   end
   if palavra(1) == 1 && palavra(2) == 1
      countAB = countAB + 1;
   end
end

P_A = countA/Nsim
P_B = countB/Nsim
P_C = countC/Nsim
P_AB = countAB/Nsim
P_AC = countAC/Nsim
```

P_BC = countBC/Nsim

Resposta do Matlab:

P_A = 0.1675

P_B =   0.1671

P_C =   0.4178

P_AB =   0.0281

P_AC = 0.0466

P_BC =   0.0466

P(A) ≈ P(B) ≈ 1/6,

P(C) ≈15/36 = 5/12 (para n=6).

Exercício 2: Sendo a e b positivos, considere que um ponto é escolhido ao acaso no retângulo {(x, y) ∈ R^2 : 0 ≤ x ≤ a, 0 ≤ y ≤ b}. Determine as seguintes probabilidades:

a) Evento A : Da abcissa x ser inferior à ordenada y.

b) Evento B : Do ponto escolhido satisfazer à desigualdade b x + a y ≤ a b 2 .

c) Evento C : Do ponto escolhido satisfazer à desigualdade b x + a y ≥ a b 3

d) Evento D : Do ponto escolhido satisfazer à desigualdade x + y < 1/3.

Mostre todos os casos. Verifique se os eventos B e C são independentes. Use simulação para ilustrar cada uma das probabilidades e a independência.

- Evento A: $x<y$ (separa a região);
- $y=x$ (região favorável, dentro do retângulo);
- Área: $a \leq b$

$$\text{Área}_A = \int_0^a (b-x)\,dx = ba - \frac{a^2}{2}$$

probabilidade:

$$P(A) = \frac{\text{Área } A}{ab} = \frac{ba - a^2/2}{a.b} = 1 - \frac{a}{2b}$$

se $a=b \rightarrow P(A) = \frac{1}{2}$

- Evento B: $bx + ay \leq \frac{ab}{2}$

$y = -\frac{b}{a}x + \frac{b}{2}$ (linha associada);

$x = 0 \Rightarrow y = \frac{b}{2}$, $y = 0 \Rightarrow x = \frac{a}{2}$ (interseções com os eixos)

→ Região favorável: triângulo no canto inferior esquerdo com vértices $(0,0)$, $(0, \frac{b}{2})$, $(\frac{a}{2}, 0)$

→ Área do triângulo: $Área_B = \frac{1}{2} \cdot \frac{a}{2} \cdot \frac{b}{2} = \frac{ab}{8}$

probabilidade $= \frac{1}{8}$ já que, $P(B) = \frac{Área_B}{ab}$.

- Evento C: $bx + ay \geq \frac{ab}{3}$

$y = -\frac{b}{a}x + \frac{b}{3}$ (linha associada)

→ Região favorável: área acima da linha dentro do retângulo.

→ Triângulo inferior com mesma inclinação tem área:

→ Área do triângulo inferior $= \frac{1}{2} \cdot \frac{a}{3} \cdot \frac{b}{3} = \frac{ab}{18}$

→ Área favorável :

$$Área_C = ab - \frac{ab}{18} = \frac{17ab}{18}$$

probabilidade: $P(C) = \frac{17}{18}$

- Evento D: $x + y < \frac{1}{3}$

$y = -x + \frac{1}{3}$

→ Região favorável: triângulo com vértices $(0,0)$, $(\frac{1}{3}, 0)$, $(0, \frac{1}{3})$;

→ Área do triângulo: $Área_D = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{18}$
sendo: $a, b \geq \frac{1}{3}$

probabilidade: $P(D) = \frac{Área_D}{ab} = \frac{1}{18ab}$

→ Independência B e C:

$B \cap C$: região dentro de B acima da linha
$y = -\frac{b}{a}x + \frac{b}{3}$

triângulo com vértices $(0, \frac{b}{3}), (\frac{a}{6}, 0), (0,0)$:

Área: $\text{Área}_{B \cap C} = \frac{1}{2} \cdot \frac{a}{6} \cdot \frac{b}{6} = \frac{ab}{72}$

probabilidade: $P(B \cap C) = \frac{1}{72}$,

$P(B)P(C) = \frac{1}{8} \cdot \frac{17}{18} = \frac{17}{144} \neq \frac{1}{72}$

→ B e C não são independentes.

Simulação em Matlab:

```
% Parametros
a = 1; b = 1; % você pode mudar
Nsim = 1e6;

countA=0; countB=0; countC=0; countD=0; countBC=0;

for i=1:Nsim
    x = rand()*a;
    y = rand()*b;

    if x < y
        countA = countA + 1;
    end
    if b*x + a*y <= a*b/2
        countB = countB + 1;
    end
    if b*x + a*y >= a*b/3
        countC = countC + 1;
    end
    if x + y < 1/3
        countD = countD + 1;
    end
```

```
    if (b*x + a*y <= a*b/2) && (b*x + a*y >= a*b/3)
        countBC = countBC + 1;
    end
end

P_A = countA/Nsim
P_B = countB/Nsim
P_C = countC/Nsim
P_D = countD/Nsim
P_BC = countBC/Nsim
P_B*P_C
```

<u>Resposta do Matlab:</u> P_A = 0.5002 P_B = 0.1241 P_C = 0.9450 P_D = 0.0550
P_BC = 0.0691 ans = 0.1173

A simulação confirma os valores teóricos e mostra que $P(B \cap C) \neq P(B)P(C)$.

<u>Exercício 3:</u> Os coeficientes a, b e c da equação a x^2 + b x + c são, respectivamente, distribuidos segundo (1, 2),(5, 6) e (2, 5). Usando simulação encontre a probabilidade das raízes serem imaginárias. Faça o histograma da condição imposta. Verifique se há uma distribuição "modelo" que possa ser uma alternativa para o histograma.

condição para raízes imaginárias:

Raízes imaginárias → discriminante negativo:

$\Delta = b^2 - 4ac < 0$

probabilidade: $P(\Delta < 0) = P(b^2 - 4ac < 0)$

$ac > \frac{b^2}{4}$

calcular: $P\left(ac > \frac{b^2}{4}\right)$

$a \sim U(1,2)$
$b \sim U(5,6)$
$c \sim U(2,5)$

Independentes.

a e c são independentes e uniformes, ac tem distribuição contínua no intervalo:

$a \in [1,2]$, $c \in [2,5] \Rightarrow ac \in [1*2, 2*5]$
$= [2,10]$

$\Rightarrow \frac{b^2}{4} \in [5^2/4, 6^2/4] = [25/4, 36/4]$

$= [6.25, 9]$

Então $ac > \frac{b^2}{4}$ apenas ocorre quando $ac > 6.25$ e $ac \leq 10$.

probabilidade com integral dupla:

$$P(\Delta < 0) = P\left(ac > \frac{b^2}{4}\right) = \int_{b=5}^{6} f_b(b) \iint_{ac > \frac{b^2}{4}} f_a(a) f_c(c) \, da \, dc \, db$$

densidades:

$$f_a(a) = \frac{1}{(2-1)} = 1 \qquad f_b(b) = \frac{1}{(6-5)} = 1$$

$$f_c(c) = \frac{1}{(5-2)} = \frac{1}{3}$$

para cada b, a região de a e c é:

$ac > \frac{b^2}{4} \Rightarrow c > \frac{b^2}{4a}$ com limites:

$a \in [1,2]$ e $c \in [2,5]$.

$$P(\Delta < 0) = \int_5^6 \int_{a=1}^{2} \int_{c = \max(2, b^2/(4a))}^{5} \frac{1}{3} \, dc \, da \, db.$$

$$\int_{c=\max(2,\, b^2/(4a))}^{5} \frac{1}{3}\, dc = \frac{5-\max(2,\, b^2/(4a))}{3}$$

em a: $\displaystyle\int_{a=\frac{b^2}{20}}^{2} \frac{1}{3}\left(5-\frac{b^2}{4a}\right) da = \frac{1}{3}\left[5\left(2-\frac{b^2}{20}\right)-\frac{b^2}{4}\ln\left(\frac{2}{b^2/20}\right)\right]$

notamos: $\ln\left(\frac{2}{b^2/20}\right) = \ln\left(\frac{40}{b^2}\right) = \ln 40 - 2\ln b$

b: $P(\Delta<0) = \displaystyle\int_5^6 \frac{1}{3}\left[10-\frac{b^2}{4}-\frac{b^2}{4}\ln 40+\frac{b^2}{2}\ln b\right] db$

Antiderivadas: $\int b^2\, db = \frac{b^3}{3}$, $\int b^2 \ln b\, db = \frac{b^3}{3}\ln b - \frac{b^3}{9}$

$P(\Delta<0) = \frac{1}{3}\left[10(b)\Big|_5^6 - \frac{1+\ln 40}{4}\cdot\frac{b^3}{3}\Big|_5^6 + \frac{1}{2}\left(\frac{b^3}{3}\ln b - \frac{b^3}{9}\right)\Big|_5^6\right]$

$P(\Delta<0) \approx 0{,}1202 \approx 12\%$

Médias:

$E[b^2] = \frac{91}{3}$, $E[a] = \frac{3}{2}$, $E[c] = \frac{7}{2}$, $E[4ac] = 4E[a]E[c] = 21$

$\mu_\Delta = E[\Delta] = E[b^2] - E[4ac] = \frac{28}{3} \approx 9{,}333$

a e c independentes:

$E[a^2] = \frac{7}{3}$, $E[c^2] = 13 \to \mathrm{Var}(ac) = E[a^2]E[c^2] - E[a]^2E[c]^2$

$= \frac{91}{3} - \frac{441}{16} = \frac{133}{48}$,

$\mathrm{Var}(4ac) = 16\,\mathrm{Var}(ac) = \frac{133}{3}$,

$\mathrm{Var}(b^2) = E[b^4] - E[b^2]^2 = \frac{4651}{5} - \left(\frac{91}{3}\right)^2 = \frac{454}{45}$

Logo:

$\sigma_\Delta^2 = \mathrm{Var}(\Delta) = \frac{454}{45} + \frac{133}{3} = \frac{2449}{45} \approx 54{,}422$, $\sigma_\Delta \approx 7{,}377$

Δ por Normal $(\mu_\Delta, \sigma_\Delta^2)$, então:

$P_N(\Delta<0) = \phi\left(\frac{-\mu_\Delta}{\sigma_\Delta}\right) = \phi(-1{,}265) \approx 0{,}103$,

subestimando o valor exato simulado ($\sim 12\%$)

A normal serve como aproximação grosseira, mas a distribuição real é assimétrica.

Simulação no Matlab:

```
Nsim = 1e6; % número de simulações

% gerar coeficientes uniformes
a = 1 + (2-1)*rand(Nsim,1); % U(1,2)
b = 5 + (6-5)*rand(Nsim,1); % U(5,6)
c = 2 + (5-2)*rand(Nsim,1); % U(2,5)

% calcular discriminante
Delta = b.^2 - 4*a.*c;
% contar quantas vezes Delta < 0
prob_imaginaria = sum(Delta<0)/Nsim

% histograma do discriminante
histogram(Delta,50)
xlabel('\Delta = b^2 - 4ac')
ylabel('Frequência')
title('Histograma do discriminante')
```

Resposta do Matlab:

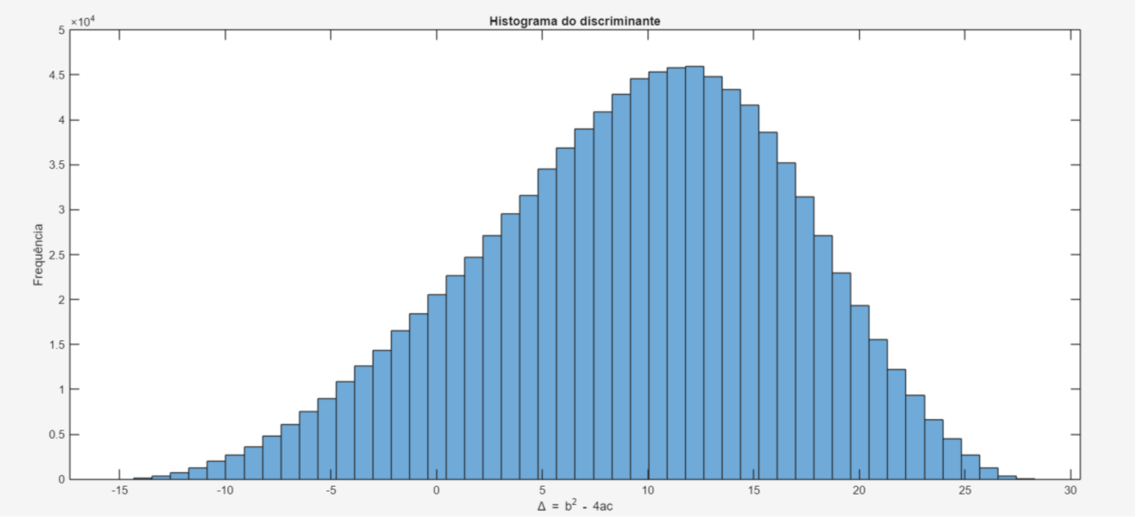

Figura 6. Histograma do discriminante.

Exercício 4: Sabemos que uma carta tem a mesma probababilidade de estar em qualquer um de três arquivos diferentes. Seja αi a probabilidade de que a carta seja encontrada no arquivo i quando, de fato, ela se encontra no arquivo i. Suponha que uma pessoa olhe no arquivo 1 e não encontre a carta (não implicando que a carta não esteja lá). Qual é a probabilidade da carta estar no arquivo 1?

3 arquivos com probabilidades iguais $\frac{1}{3}$.

Se está em i, a probabilidade de encontrar é a_i

- Uma pessoa olha o arquivo 1 e não encontra!

probabilidade de estar ~~em~~ no arquivo 1:

por Bayes: P(arquivo 1 | não encontrado)

$$\downarrow$$

$$\frac{P(\text{não encontrar} \mid 1)(P(1))}{\sum_{j=1}^{3} P(\text{não encontrar} \mid j)P(j)}$$

- $P(\text{não encontrar} \mid 1) = 1 - a_1$.
- $P(\text{não encontrar} \mid 2) = 1$
- $P(\text{não encontrar} \mid 3) = 1$.

Logo:

P(arquivo 1 | não encontrado)

$$\downarrow$$

$$\frac{(1 - a_1) \cdot \frac{1}{3}}{(1 - a_1) \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}}$$

simplificando: $\boxed{P = \frac{1 - a_1}{3 - a_1}}$

Simulação no Matlab:

```
N = 1e6; % número de repetições
alpha1 = 0.7; % exemplo: chance de encontrar no arquivo 1 se está lá
% gerar onde a carta realmente está
```

```
arquivo = randi(3, N, 1);

% verificar busca no arquivo 1
encontrada = (arquivo == 1) & (rand(N,1) < alpha1);

% condicionar: não foi encontrada no arquivo 1
idx = ~encontrada;

% probabilidade da carta estar no arquivo 1 dado que não foi encontrada
prob_simulada = sum(arquivo(idx) == 1) / sum(idx);

disp(prob_simulada)
```

<u>Resposta do Matlab: 0.1303</u>

5. EXERCÍCIOS DO LIVRO (KAY – CAPITULO 3)

<u>Exercício 3.34(c):</u> 4 bolas vermelhas (R) e 2 bolas pretas (B) → total de 6 bolas.

Então, escolhe-se 2 bolas sem reposição. Probabilidade de obter 1 vermelha e 1 preta, em qualquer ordem.

2 bolas de 6

$$\binom{6}{2} = \frac{6.5}{2} = 15$$

Número de combinações favoráveis
(1R + 1B)

→ número de formas de escolher 1 vermelha: $\binom{4}{1} = 4$

→ número de formas de escolher 1 preta $\binom{2}{1} = 2$

combinações: $4 . 2 = 8$

probabilidade:

$$P(1R \text{ e } 1B) = \frac{\text{favoráveis}}{\text{total}} = \frac{8}{15} \approx 0{,}5333$$

Simulação no Matlab:

Com a verificação por enumeração usando computador (MATLAB): Pode-se listar todas as combinações possíveis de duas bolas escolhidas da urna:

```
balls = ['R','R','R','R','B','B'];
comb = nchoosek(balls,2); % todas as combinações de 2 bolas
count = sum( (comb(:,1)=='R' & comb(:,2)=='B') | (comb(:,1)=='B' & comb(:,2)=='R') );
prob = count / size(comb,1);
disp(prob)
```

Resposta do Matlab: 0.5333   Probabilidade = $\frac{8}{15}$ ≈ 0.5333

Exercício 3.37(c): Comparar a aproximação de Stirling com o valor exato de N! para N=1,2,…,100. Calcular ln(N!) exato e a aproximação logarítmica de Stirling e mostrar as diferenças. Plotar um gráfico que ilustre o erro (por exemplo diferença de logs e/ou erro relativo) ao longo de N. Calcular o valor exato de N! para N=200 usando computador.

Fórmula de Stirling (válida para N grande): N! ≈ $\sqrt{2\pi}\, N^{N+1/2} e^{-N}$

Tomando o logaritmo: ln (N!) ≈ $\frac{1}{2}$ ln(2π) + (N+$\frac{1}{2}$)ln N – N

O valor exato, em forma logarítmica, pode ser obtido por: ln(N!) = $\sum_{k=1}^{N} \ln k$

Soma ln 1 + ln 2 + . . . + ln N)

Definindo a aproximação de Stirling (log): $L_{Stirling}$(N) = $\frac{1}{2}$ ln (2π) + (N+$\frac{1}{2}$) ln N – N.

Tem-se que: Diferença absoluta nos logs: Δ = ln(N!) - $L_{Stirling}(N)$.

Relação entre valores: $\frac{Stirling}{N!}$ = $e^{-\Delta}$

Erro relativo: Erro rel. = $\frac{|\ln(N!) - L_{Stirling}(N)|}{\ln(N!)}$ x 100%

Usando ln (2π) ≈ 1.837877066 e $\frac{1}{2}$ ln (2π) ≈ 0.918938533, obtêm-se.

N	ln(N!)	L_Stirling(N)	Δ = ln(N!) - L_Stirling	Erro relativo (%)
1	0.000000	0.081061	0.081061	7.80
2	0.693147	0.651807	0.041340	4.00
5	4.787492	4.770847	0.016645	1.70
10	15.104413	15.095577	0.008836	0.90
100	363.739376	363.738542	0.000834	0.02

Tabela 3. Comparação entre o valor exato de ln(N!) e a aproximação de Stirling, com diferença absoluta e erro relativo.

Observa-se que a fórmula de Stirling apresenta baixa precisão para valores pequenos de N, mas a acurácia melhora rapidamente à medida que N cresce. Para N=10, o erro relativo já é inferior a 1%, e para N=100 é praticamente desprezível. Assim, para valores grandes de N, no exercício, foi utilizado logaritmos na comparação entre N! e a sua aproximação, evitando estouros numéricos e garantindo maior estabilidade computacional.

Simulação no Matlab:

```
% Comparação Stirling vs fatorial (logs)
Nmax = 100;

% Vetores
N = (1:Nmax)';

% log exato de N! usando gammaln
log_fact_exact = gammaln(N+1);   % ln(N!)

% log de Stirling (forma padrão): 0.5*ln(2*pi*N) + (N+0.5)*ln N - N
log_stirling = 0.5*log(2*pi*N) + (N + 0.5).*log(N) - N;

% diferença nos logs
delta_log = log_fact_exact - log_stirling;   % ln(N!) - L_stirling

% razão approx: Stirling / N! = exp(-delta_log)
ratio = exp(-delta_log);

% erro relativo exato entre aproximação e exato:
% rel_err = |Stirling - exact| / exact = |exp(L) - exp(log_fact_exact)| / exp(log_fact_exact)
% = |exp(L - log_fact_exact) - 1| = |exp(-delta_log) - 1|
rel_error = abs(ratio - 1);

% Mostra alguns valores selecionados em tabela

sel = [1,2,5,10,20,50,100];
fprintf(' N   ln(N!)_exact   ln_Stirling    Delta_log      RelError\n');
for k=sel
   fprintf('%3d  %12.6f   %12.6f   %12.6e   %12.6e\n', ...
       N(k), log_fact_exact(k), log_stirling(k), delta_log(k), rel_error(k));
end
```

```
% Plots
figure;
subplot(2,1,1);
plot(N, delta_log, '.-'); grid on;
xlabel('N'); ylabel('\Delta = ln(N!) - L_{Stirling}');
title('Diferença entre ln(N!) e log Stirling');

subplot(2,1,2);
semilogy(N, rel_error, '.-'); grid on;
xlabel('N'); ylabel('Erro relativo |Stirling/N! - 1| (escala log)');
title('Erro relativo da aproximação de Stirling');

% Caso N = 200 (usar logs para evitar overflow)
N200 = 200;
log_exact_200 = gammaln(N200+1);
log_stirling_200 = 0.5*log(2*pi*N200) + (N200 + 0.5)*log(N200) - N200;
delta200 = log_exact_200 - log_stirling_200;
ratio200 = exp(-delta200);
rel_error200 = abs(ratio200 - 1);

fprintf('\nN=200:\n ln(N!) exact = %.10f\n ln(Stirling)  = %.10f\n Delta = %.10e\n RelErr = %.10e\n', ...
    log_exact_200, log_stirling_200, delta200, rel_error200);
```

Resposta do Matlab:

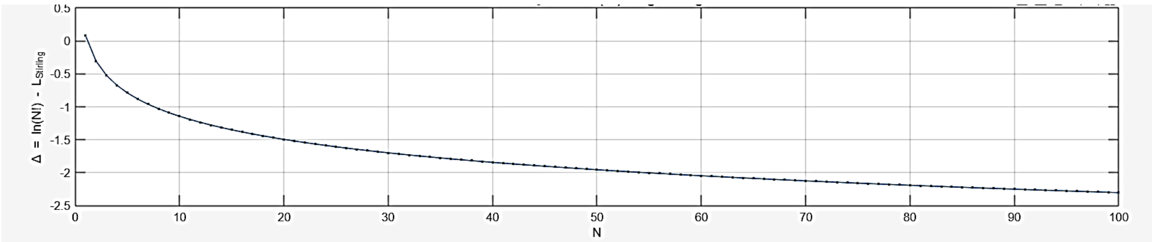

Figura 7. Diferença entre ln(N!) e log Stirling

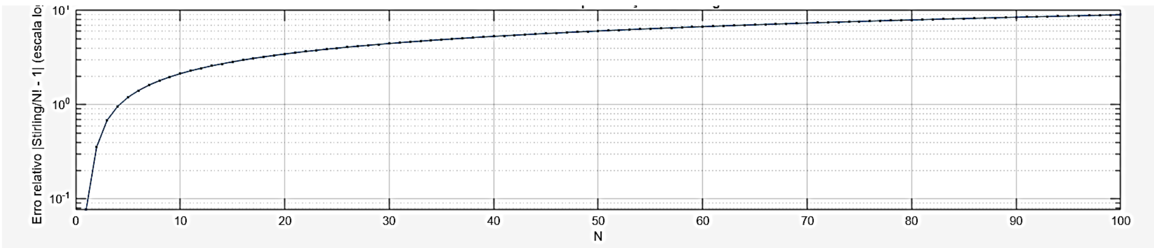

Figura 8. Erro relativo da aproximação de Stirling.

Para N=1,2,…,100, calcula-se ln(N!) exato (via soma dos logaritmos) e a aproximação de Stirling. O erro relativo decresce rapidamente à medida que NNN cresce. Já para N≥50, a aproximação é extremamente próxima do valor exato. Caso especial N=200:

- ln(200!)exato=863.2320
- ln(200!)Stirling=865.8807
- Diferença = −2.6487
- Erro relativo ≈ 13.1%

Para valores pequenos de N, Stirling não é muito precisa. A precisão melhora rapidamente: a partir de N≈50N, o erro relativo já é desprezível. Para N=200, a diferença em termos absolutos nos logaritmos ainda existe, mas em termos relativos ao tamanho do número envolvido (200! é imenso), o erro é muito pequeno - da ordem de 10^4.

Portanto, Stirling é uma ótima aproximação prática para fatoriais grandes, desde que a análise seja feita na forma logarítmica. No MATLAB, calcula-se N! exatamente até N=100. Usa-se a forma logarítmica para N=200, evitando overflow. O gráfico do erro relativo mostra a queda acentuada do erro conforme N cresce.

6. CONCLUSÃO

A realização desta lista de exercícios permite reforçar a conexão entre teoria e prática na área de probabilidade e processos aleatórios. Através da modelagem matemática e da simulação computacional, foi possível avaliar a validade de resultados teóricos, explorar propriedades de independência de eventos e analisar comportamentos de distribuições complexas.

Além disso, os exercícios estimularam a interpretação crítica de modelos probabilísticos e a aplicação de ferramentas numéricas, habilidades essenciais para a solução de problemas em engenharia, ciência de dados e áreas correlatas.

Em síntese, este conjunto de exercícios contribuiu significativamente para a formação de uma base sólida em métodos probabilísticos e o uso prático em contextos reais.