MPEI 2020-2021

Resolução de exercícios
(consolidação)

Problema #1

- Considere o seguinte jogo de cartas entre si e um amigo:
  - Apenas se usam as figuras (Rei, Dama e Valete), o Ás e o Joker de um dos naipes
  - As cartas são baralhadas honestamente e extraídas uma a uma
  - Sempre que se extrai uma figura ganha 1 Euro, cada Ás obriga-o a pagar 1 Euro ao seu amigo, cada Joker implica pagar 2 Euros ao seu amigo.
  - Volta a colocar-se a carta no baralho e baralha-se

- Questão: **Qual o seu ganho médio ao fim de uma longa sequência de jogadas ?**

Fonte: “O ACASO”, J. M. Sá, Gradiva

Exemplo

- Para perceber melhor, analisemos um possível jogo:

Saiu	**A**	**R**	**A**	**R**	**A**	**J**	**V**	**D**	**V**	**…**
Seu ganho	-1	1	-1	1	-1	-2	1	1	1	
Ganho acumulado	-1	0	-1	0	-1	-3	-2	-1	0	…

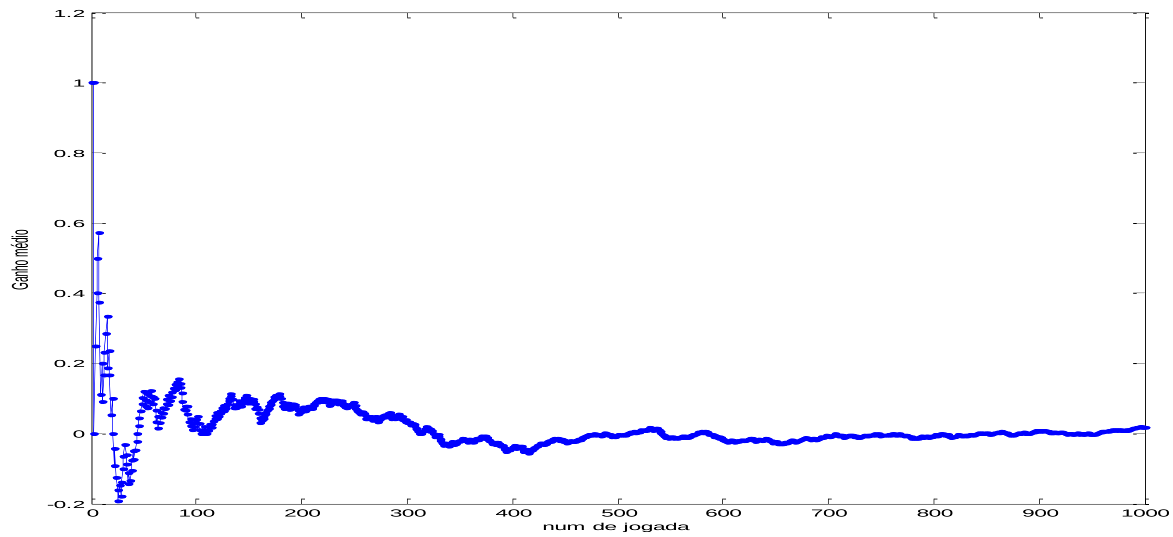
Simulemos ...

- Para simplificar, consideremos a seguinte codificação:
  Rei=1, Dama=2, Valete=3, Ás=4, Joker=5

```
% simular N extracções
e= floor (rand(1,N)*5)+1;   % inteiros de 1 a 5 (equiprováveis)

%  calcular ganho acumulado ao longo do jogo
g=e;  g(e==1 | e==2 | e=3)=1;      g(e==4 )=-1;       g(e==5 )=-2;
gacum= cumsum(g)

%  ganho médio ao longo do tempo
n=1:N;
gmed= gacum ./ n;

Plot(n,gmed, '.')
```

Exemplo de resultado

- 1000 jogadas



- É apenas uma de um número muito grande de sequências possíveis (são possíveis 5^{1000})

Resolução …

- Tentemos agora resolver sem simulação e usando o que já sabemos …

- Sugestões ?

Variável aleatória Ganho

- O Ganho em cada extracção de uma carta (experiência aleatória) pode ser considerada uma variável aleatória (G)
- Pode assumir os valores {-2,-1,1}

- Função de probabilidade ?

$$p_G(G = -2) = ?$$
$$p_G(G = -1) = ?$$
$$p_G(G = 1) = ?$$

Função de probabilidade de G

- $p_G(G = -2) = 1/5$
- Equivalente a P("sair um Joker")

- $p_G(G = -1) = 1/5$

- $p_G(G = 1) = 3/5$

Ganho médio …

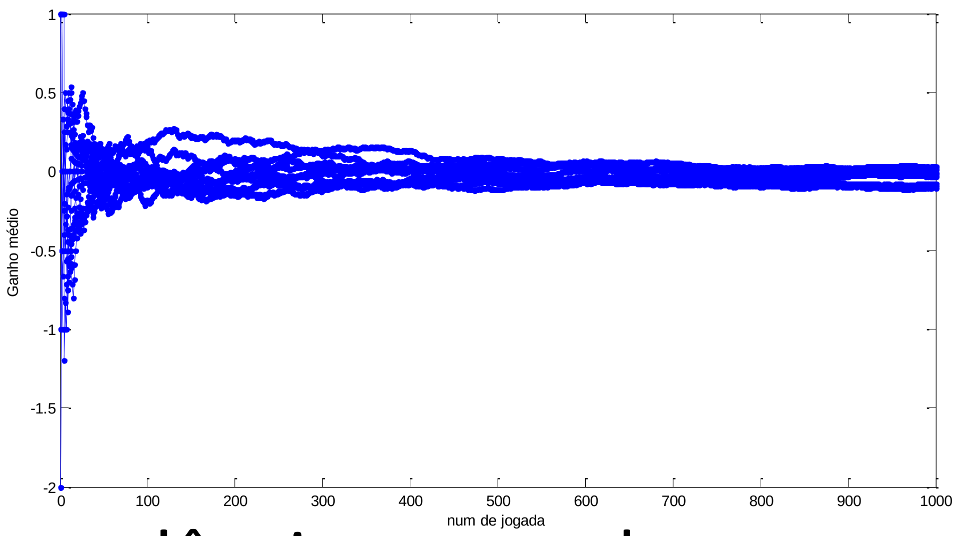
- O "Ganho médio esperado" ao fim de um grande número de jogadas é designado por **esperança matemática da variável Ganho**

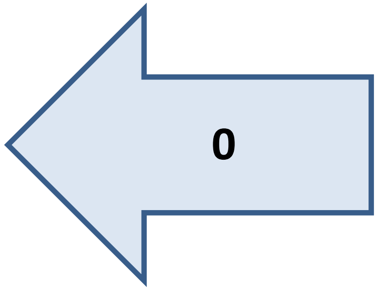
- $\mathrm{E}[G] = \sum x_i p(X = x_i)$

- Aplicando ao nosso caso…
- $\mathrm{E}[G] = 1 \times \frac{3}{5} + (-1) \times \frac{1}{5} + (-2) \times \frac{1}{5} = 0$

Longa sequência

- É altura de esclarecer um pouco mais o que se entende por “longa sequência”...
- Repetindo o jogo (ou melhor simulando..), temos





- Existe uma tendência para todas as curvas estabilizarem em torno da esperança (0)

Problema #2 (aniversários)

- Faz parte do Guião PL02

Consideremos uma festa em que há um certo número de pessoas, digamos n.

Qual deve ser o valor de n para que a **probabilidade de duas (ou mais) pessoas terem a mesma data de aniversário** (mês e dia) seja superior a 50 % ?

e para ser superior a 90% ?

Simulação

- Usar inteiros de 1 a 365 para representar os possíveis aniversários
- Gerar N conjuntos de n aniversários
  - Por exemplo, matriz de N colunas e n linhas

- Determinar o número de colunas que têm pelo menos 1 número repetido
  - Dá os casos favoráveis

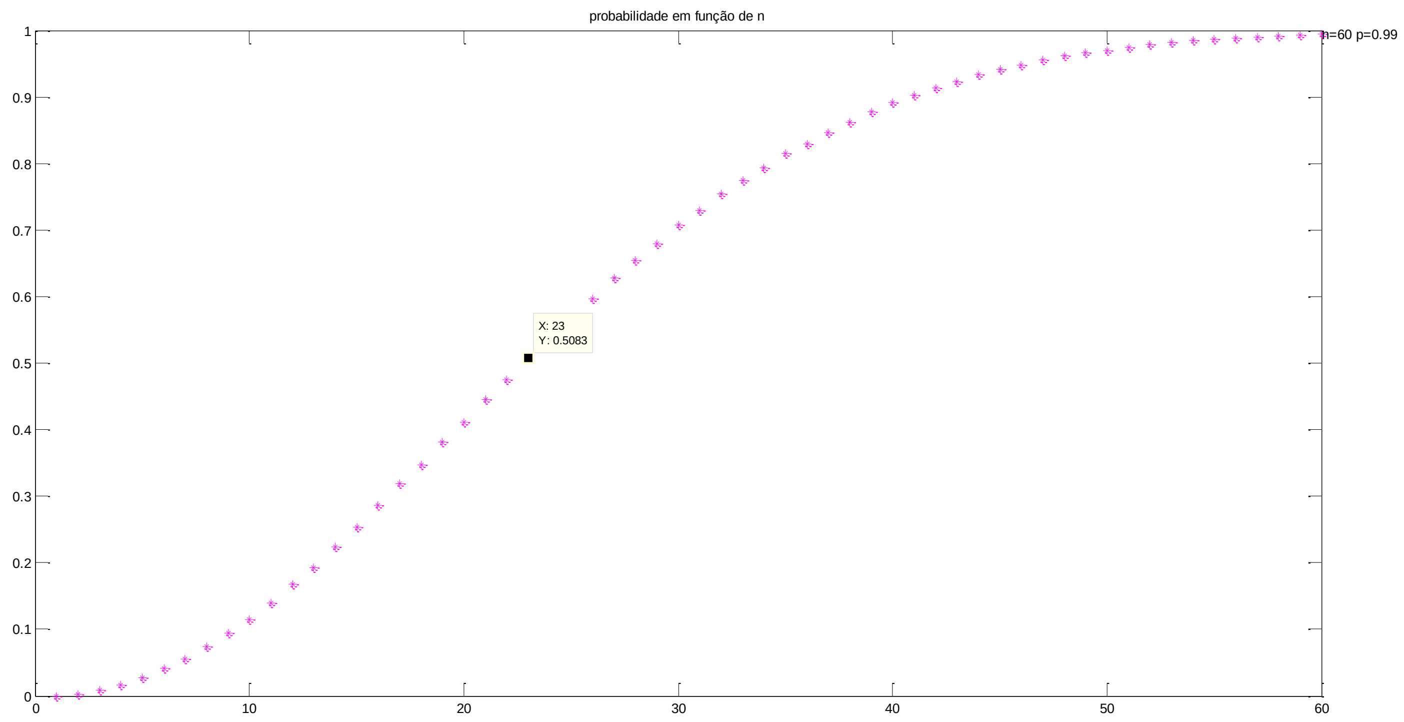
- Calcular a probabilidade (casos favoráveis / N)

Simulação (exemplo de código)

```
N=10e5;
diferentes=zeros(1,N);

for n=1:60
   M= randi(365, n, N);  %  inteiros de 1 a 365

   % verificar se houve repetição (1 ou mais)
   for i=1:N
      diferentes(i)= length(unique(M(:,i)));
   end
   colisoes=sum(diferentes~= n);

   pColisao=colisoes/N;
   fprintf(1,'n= %d, p=%.3f\n',n,pColisao);
end
```

Resultados da simulação

- Probabilidade em função de n:



Resposta às questões concretas

- superior a 50 % ?

  n=23

- superior a 90 % ?

  n=

Resolução "teórica"

- Mais simples resolver probabilidade de ninguém de n pessoas fazer anos no mesmo dia
  - Depois usa-se 1-p
- Os aniversários das n pessoas são independentes
- Para não haver nenhum dia com mais do que um aniversário, a primeira pessoa pode fazer anos em 365 dias dos 365, E a segunda em 364, E a terceira em 363 ..
- Ou seja $\frac{365}{365} \times \frac{364}{365} \times \frac{365-2}{365} \times \cdots \times \frac{365-n+1}{365}$
- Que é igual a $\frac{365\,!}{(365-n)!\,365^n}$
- Para n=23 temos 1-p=0.5073
- Para n=41 temos 1-p=0.9032

Problema #3 (n dardos e m alvos)

- Também parte de PL02

Considere o seguinte ''jogo'': lançamento com os olhos vendados de n dardos, um de cada vez, para m alvos garantindo-se que cada dardo atinge sempre um alvo (e apenas 1).

a) Qual a probabilidade de nenhum alvo ter sido atingido 2 ou mais vezes ?

b) Qual a probabilidade de pelo menos 1 alvo ter sido atingido 2 vezes ?
NOTA: Apesar de não ser 100% claro vamos considerar 2 vezes OU MAIS objectivo inicial do exercício

- …

Simulação

- Basta uma simples adaptação/extensão do problema anterior
  - Em vez de 365 dias temos m alvos
  - As n pessoas passam a ser os n dardos

Resolução teórica

- Seguindo ideia análoga aos aniversários …

- P(nenhum alvo ter sido atingido 2 ou mais vezes) = $\frac{m}{m} \times \frac{m-1}{m} \times \frac{m-2}{m} \times \cdots \times \frac{m-n+1}{m}$
- Que é igual a $\frac{m\,!}{(m-n)!\,m^n}$

Problema #4 (sequências de palavras)

Integra PL02

Considere uma linguagem com apenas 3 palavras "um", "dois'", "três" e que permite sequências de 2 palavras. Se todas as frases forem equiprováveis e as duas palavras na frase puderem ser iguais:

…

(d) Qual o valor de P[``sequência incluir a palavra um'' | ``sequência inclui palavra dois''] ?

(e**) Resolva a questão anterior para o caso de termos 10 palavras diferentes e sequências de 5 palavras com ajuda de um programa que calcule exaustivamente todas as possibilidades.**

Sugestão: use números de 1 a 10 para representar as palavras e use Matlab.

Em Matlab…

- A solução mais simples é usar 5 ciclos for, mas não é generalizável

- Possível alternativa (por Tomás Oliveira e Silva)
  - Geral
  - Elegante

```
npw = 5; % number of possible words
nws = 2; % number of words in a sequence

seq = zeros(nws,npw^nws); % all possible sequences will be kept here
seq(:,1) = 1;          % first sequence
nc = 1;                % number of columns of the block that will be replicated
for i=1:nws
  % replicate a block npw-1 times, and change the data of its i-th row
  for j=2:npw
    seq( : ,(j-1)*nc+1:j*nc ) = seq(:,1:nc); % copy a block
    seq( i ,(j-1)*nc+1:j*nc ) = j;         % change its i-th row
  end
  nc = nc*npw; % increase the block size
end

has_one = sum(reshape(seq == 1,size(seq))) >= 1;
has_two = sum(reshape(seq == 2,size(seq))) >= 1;

P= ⋯
```

Problema #5

1. (4 valores) Considere uma variável aleatória discreta X cujos valores possíveis estão no vetor x e cujas probabilidades de cada valor estão no vetor y. Considere o seguinte código MATLAB:

```
x = 1:4;
y = [1/2 1/4 1/8 1/8];
a = y.*x;
b = sum(a)
c = sqrt(sum(y.*x.^2)-b^2)
```

- Depois de executado o código, indique, justificando, o que representam as variáveis b e c
- Quais os valores de b e c ?

Problema #6

3. (2 valores) Considere duas variáveis X e Y com a função massa de probabilidade conjunta seguinte:

$X \setminus Y$	0	1
0	$\frac{1}{9}$	$\frac{2}{9}$
1	$\frac{2}{9}$	$\frac{4}{9}$

a) Calcule as probabilidades marginais de X e de Y.

b) Calcule a média e a variância de X.

Problema #7

4. (8 valores) Um sistema de transmissão de dados envia informação em pacotes de 100 bits. A probabilidade do receptor descodificar um pacote com erros (1 ou mais erros) é de %1.
Para a transmissão e recepção de 1600 pacotes de dados:

a) Indique o nome da distribuição (e os respectivos parâmetros) que descreve a probabilidade do número de pacotes descodificados com erros.

b) Calcule o número médio de pacotes descodificados com erro.

c) Indique a expressão exata para a probabilidade de obter 20 (e só 20) pacotes com erros.

d) Indique uma aproximação, usando a distribuição de Poisson, para a probabilidade de obter 20 (e só 20) pacotes com erros.

e) Calcule uma aproximação, usando a distribuição normal, para a probabilidade de obter 20 ou mais pacotes com erros.
Sugestão: deverá usar uma variável aleatória normal com média e variâncias adequadas.

Problema #8

3. (5 valores) Considere três programadores - André, Bernardo e Carlos - e que a distribuição do número de erros que cometem nos seus programas é a seguinte:

	Núm. erros (x_i)	0	1	2	3	4 ou mais
André	$p_A(X = x_i)$	0.7	0.1	0.08	0.05	0.07
Bernardo	$p_B(X = x_i)$	0.5	0.4	0.05	0.03	0.02
Carlos	$p_C(X = x_i)$	0.85	0.1	0.03	0.01	0.01

a) Escolhe-se aleatoriamente um programa de entre um conjunto de programas em que os do André são tantos como o conjunto correspondente ao Carlos mais o Bernardo, com estes dois a contribuirem com o mesmo número. O programa escolhido tem 3 ou mais erros. Qual a probabilidade de ter sido o André o autor do programa escolhido ?

(b) Quem é mais provável ter sido o autor do programa escolhido?

Problema #9

4. (3 valores) Considere que um programador P comete, em média, 10 erros em cada 1000 linhas de código que escreve. Se esse programador criar código para resolver 3 pequenos problemas e daí resultarem 100, 200 e 500 linhas de código, respectivamente, qual a probabilidade de nenhum dos 3 programas conter qualquer erro ?

   Assuma distribuição de Poisson ($p_X(k) = \frac{\lambda^k}{k!}e^{-\lambda}$).