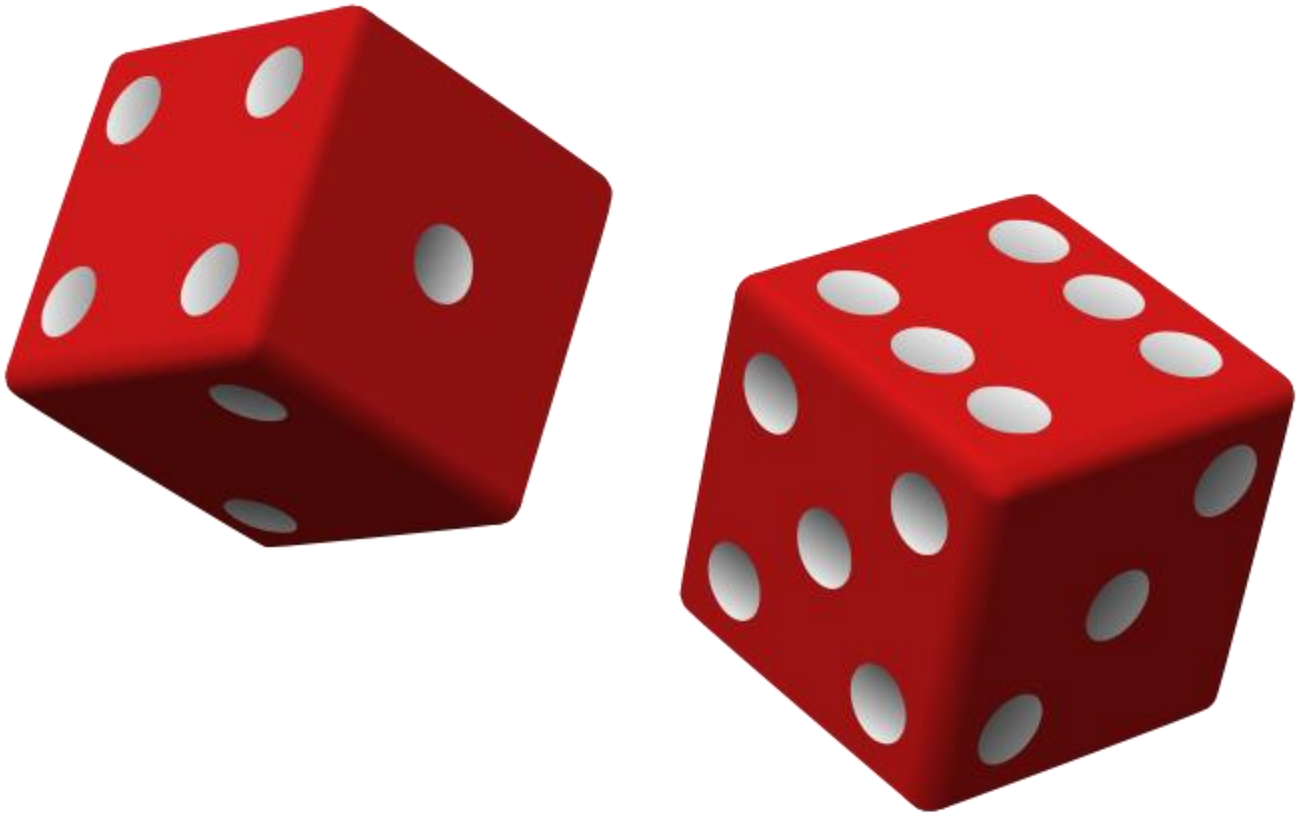


Introdução à Probabilidade

Probabilidade

- Probabilidade é um framework matemático para raciocinar sobre incerteza
- Modelo probabilístico é composto por:
  - Os possíveis resultados (espaço amostral); e
  - E uma lei (regra) que nos dita a probabilidade de cada evento
- Espaço amostral Ω (exaustivo e mutualmente exclusivo)
  - O conjunto de todos os resultados possíveis
- Evento
  - Qualquer conjunto de resultados possíveis

Problema de Probabilidade

- Considerando um dado não viciado, quais são as probabilidades de pois de um lançamento de ocorrer:
  a) Um número par, ou seja, A={2, 4, 6}
  b) Um número menor que 3, ou seja, B={1, 2}
  c) O número seis, ou seja, C={6}
  d) Um número maior que seis, ou seja, B={ }

Modelo de Probabilidade do exercício c)

Resultado	Sim (número 6)	Não (outro número)
Probabilidade	1/6	5/6

Axiomas da Probabilidade

- As probabilidades são valores entre
  - 0 e 1
- A soma das probabilidades de todos os eventos possíveis do experimento deve ser
  - Igual a 1
    - A probabilidade de você lançar um dado e ocorrer um dos 6 números é 1
      - Ou seja, certeza

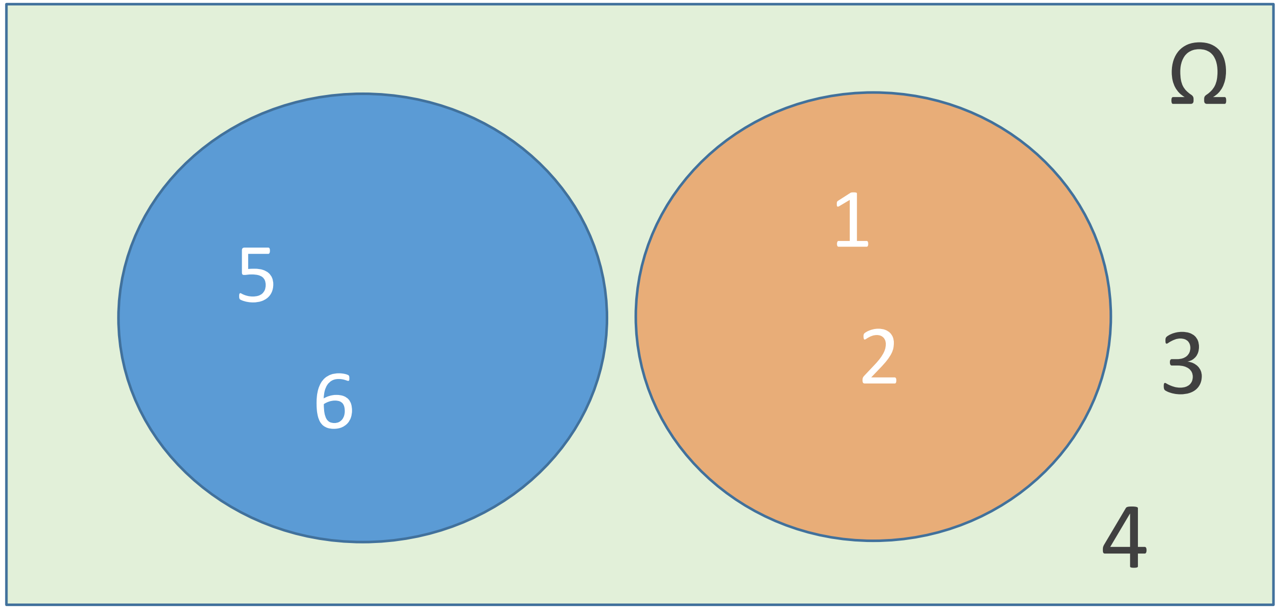
$$P(A) \geq 0 \qquad P(\Omega) \geq 0$$

Axiomas da Probabilidade

- Aditividade

$$Se\ A \cap B = \emptyset, então\ P(A \cup B) = P(A) + P(B)$$

- Exemplo: resultados de um lançamento de dado
  - maior que 4 ou menor que 3



Princípio da Equiprobabilidade

- Quando as características do experimento sugerem N possíveis resultados
  - E todos os resultados com igual probabilidade
    - A probabilidade de ocorrer um evento A, contendo n resultados, é:

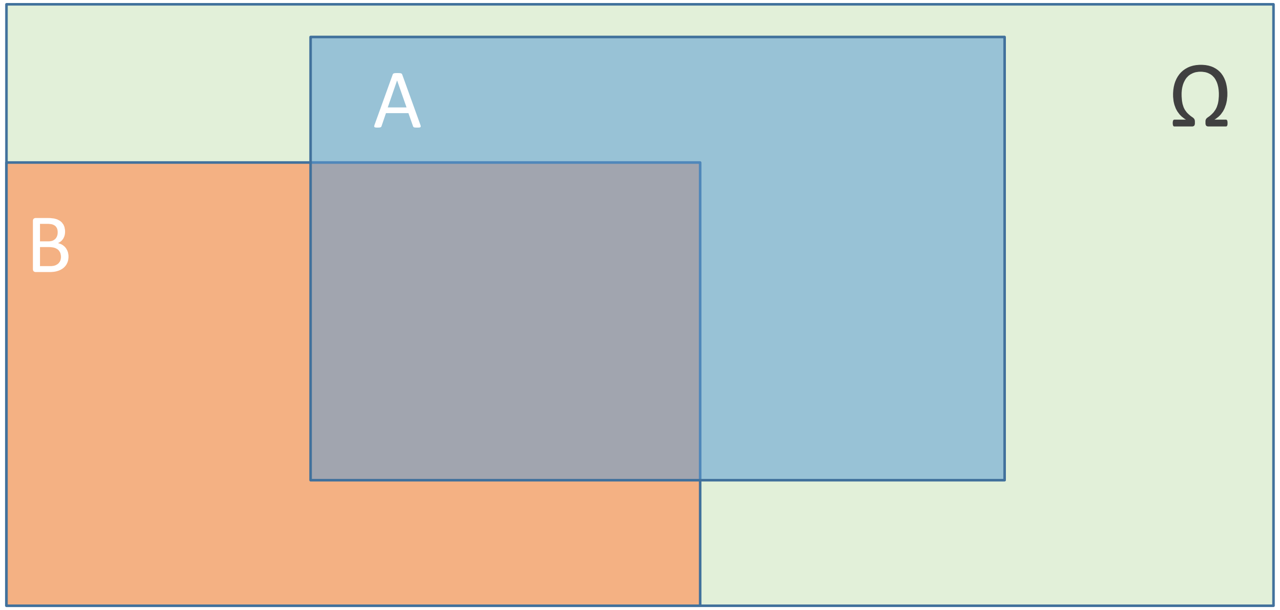
$$P(A) = \frac{n}{N} = \frac{número\ de\ resultados\ de\ A}{número\ total\ de\ resultados}$$

Probabilidade Condicional

- É uma proporção entre a probabilidade P(B)
  - E a probabilidade P(A ∩ B)

$$Def1: \; P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Mudança de contexto
  - Recontagem no novo espaço amostral



Eventos independentes

- O primeiro lançamento de uma moeda
  - Não adiciona informação (conhecimento) sobre o resultado do 2º lançamento.

- Definição: $Def2\text{: } P(A|B) = P(A)$
  - Ou seja, em eventos independentes, a ocorrência de B não altera a probabilidade da ocorrência de A
- Juntando com a definição Def1: $P(A \cap B) = P(A|B).P(B)$
- Obtemos

$$P(A \cap B) \overset{\text{Def1}}{=} P(A|B).P(B) \overset{\text{Def2}}{=} P(A).P(B)$$

Não compreender quando eventos são independentes

- Numa roleta (cassino), se a bolinha parou no preto 5 vezes seguidas
  - Qual a probabilidade de na próxima, cair no vermelho?
    - Permanece inalterada. São eventos independentes.
- Se você lançar uma moeda 1 milhão de vezes e obtiver 1 milhão de caras
  - A probabilidade do próximo lançamento ser coroa é 1/2

Eventos independentes: Morte súbita infantil (SMSI)

- É um fenômeno (raro) no qual um bebê perfeitamente saudável
  - Morre no berço (no UK conhecido como “Morte do berço”)
- Atraía mais atenção a medida que outras causas tornavam-se menos comuns
  - Por não se compreender o as causas, as mortes despertavam suspeitas
    - Um exame pós morte não diferencia SMSI de maus tratos
- Promotores e cortes britânicas convenceram-se de que seria possível separa maus-tratos de SMSI
  - Focando nas famílias que apresentassem múltiplas mortes por SMSI

Eventos independentes: Morte súbita infantil (SMSI)

- Sir Roy Meadow (um pediatra e perito judicial)
  - Era convocado a depor nesses casos
- A revista The Economist cunhou o termo “Lei de Meadow”
  - “A morte de 1 bebê é uma tragédia, 2 é uma suspeita e 3 é assassinato”
    - Duas ou mais ocorrências na mesma família são tão improváveis,
      - Que é praticamente impossível serem fruto do acaso.
- Meadow: a probabilidade de duas mortes na mesma família
  - $P(SMSI) = 1/8500$
    - $P(SMSI_1, SMSI_2) = (1/8500)^2 = 1/(73 \text{ milhões})$ => “cheira maus tratos”

Eventos independentes: Morte súbita infantil (SMSI)

- Muitos pais foram para a prisão com base nesse argumento
  - As vezes, sem qualquer outra evidência médica que corroborasse com esta conclusão
- Em alguns casos
  - Bebês foram retirados dos seus pais por causa da morte do irmão
- Entrou em cena a Royal Statistical Society
  - Esse cálculo seria apropriado para eventos independentes
    - Ou seja, mortes totalmente aleatórias sem estarem ligadas por algum outro fator desconhecido.
      - Fatores genéticos, dormir de bruços/de costas, etc.
- O governo britânico anunciou que reveria 258 julgamentos

Aglomerações acontecem

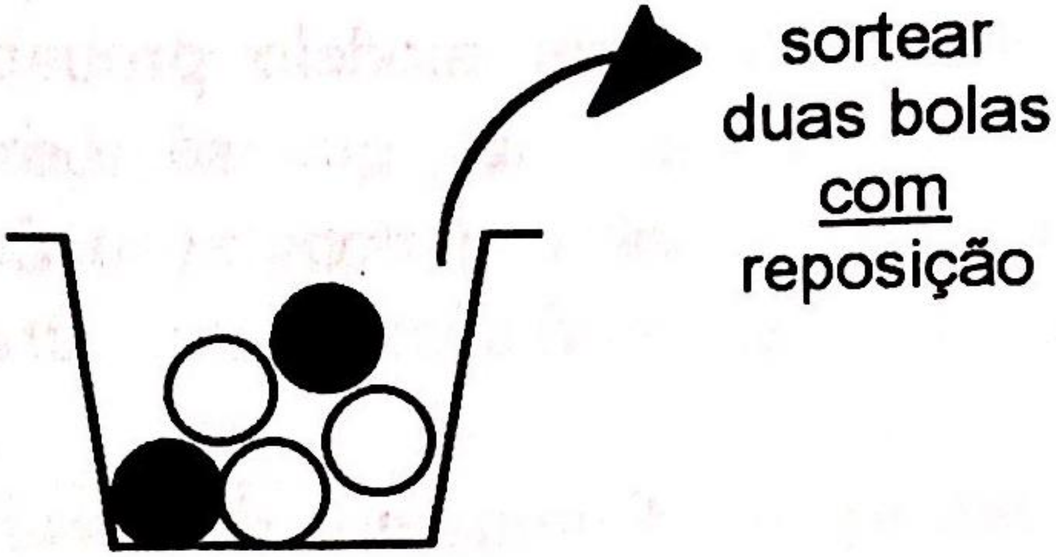
- Eventos raros podem se aglomerar
  - Desde que a amostra seja suficientemente grande
- Exemplo: 5 pessoas contraírem uma forma rara de leucemia
  - Na mesma escola, igreja ou local de trabalho
    - Pode ter uma probabilidade rara (Exemplo: 1 em 1 milhão)
  - Mas há milhões de escolas, igrejas e locais de trabalho
    - Apenas não estamos contabilizando todas as escolas, igrejas e locais de trabalho onde isso não ocorreu
- Vários lançamentos de moeda

Variável aleatória

- Variável aleatória é uma função
  - Associa um elemento do espaço amostral, Ω, a um número real
    - uma probabilidade.
- Funções são nomenclaturas genéricas
  - Úteis para substituir tabelas quando se quer uma comunicação concisa
    - Por exemplo, numa equação

Variável aleatória: exemplo 1

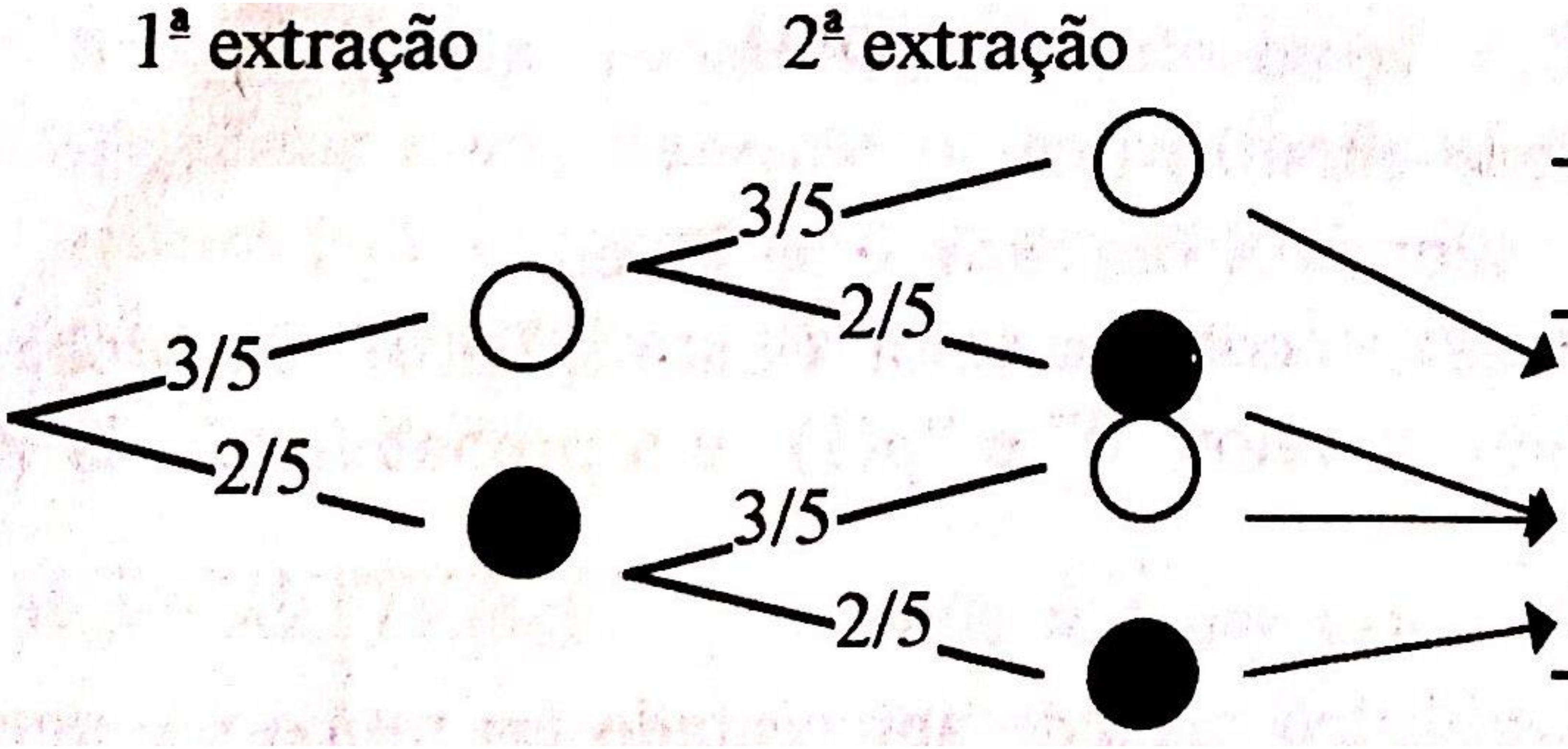
EXEMPLO 7.5 Seja uma urna com três bolas brancas e duas pretas. Extrair aleatoriamente duas bolas, sendo uma após a outra, tal que repomos na urna a primeira bola antes de extrairmos a segunda - *amostragem com reposição*. Queremos a distribuição de probabilidades da variável **X** = *número de bolas pretas extraídas na amostra*.



Variável aleatória: exemplo 1

- Variável **X**: número de bolas pretas extraídas
  - Eventos independentes

Distribuição de probabilidade de X



Distribuição de X

x	p(x)
0	9/25 (ou 0,36)
1	12/25 (ou 0,48)
2	4/25 (ou 0,16)

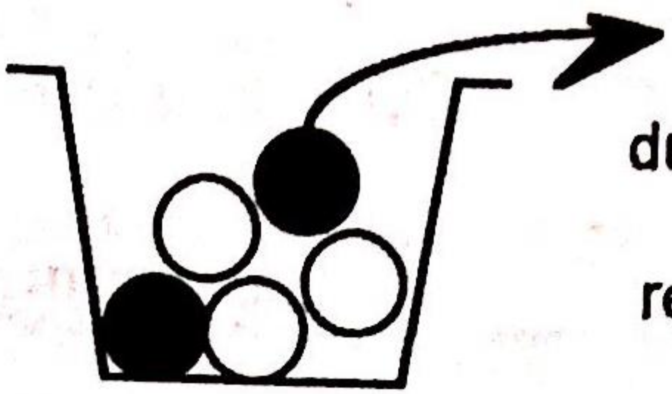
Tabela de probabilidades

TABELA II - Distribuição Binomial

n	x	π 0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
1	0	0,9500	0,9000	0,8500	0,8000	0,7500	0,7000	0,6500	0,6000	0,5500	0,5000
	1	0,0500	0,1000	0,1500	0,2000	0,2500	0,3000	0,3500	0,4000	0,4500	0,5000
2	0	0,9025	0,8100	0,7225	0,6400	0,5625	0,4900	0,4225	0,3600	0,3025	0,2500
	1	0,0950	0,1800	0,2550	0,3200	0,3750	0,4200	0,4550	0,4800	0,4950	0,5000
	2	0,0025	0,0100	0,0225	0,0400	0,0625	0,0900	0,1225	0,1600	0,2025	0,2500
3	0	0,8574	0,7290	0,6141	0,5120	0,4219	0,3430	0,2746	0,2160	0,1664	0,1250
	1	0,1354	0,2430	0,3251	0,3840	0,4219	0,4410	0,4436	0,4320	0,4084	0,3750
	2	0,0071	0,0270	0,0574	0,0960	0,1406	0,1890	0,2389	0,2880	0,3341	0,3750
	3	0,0001	0,0010	0,0034	0,0080	0,0156	0,0270	0,0429	0,0640	0,0911	0,1250

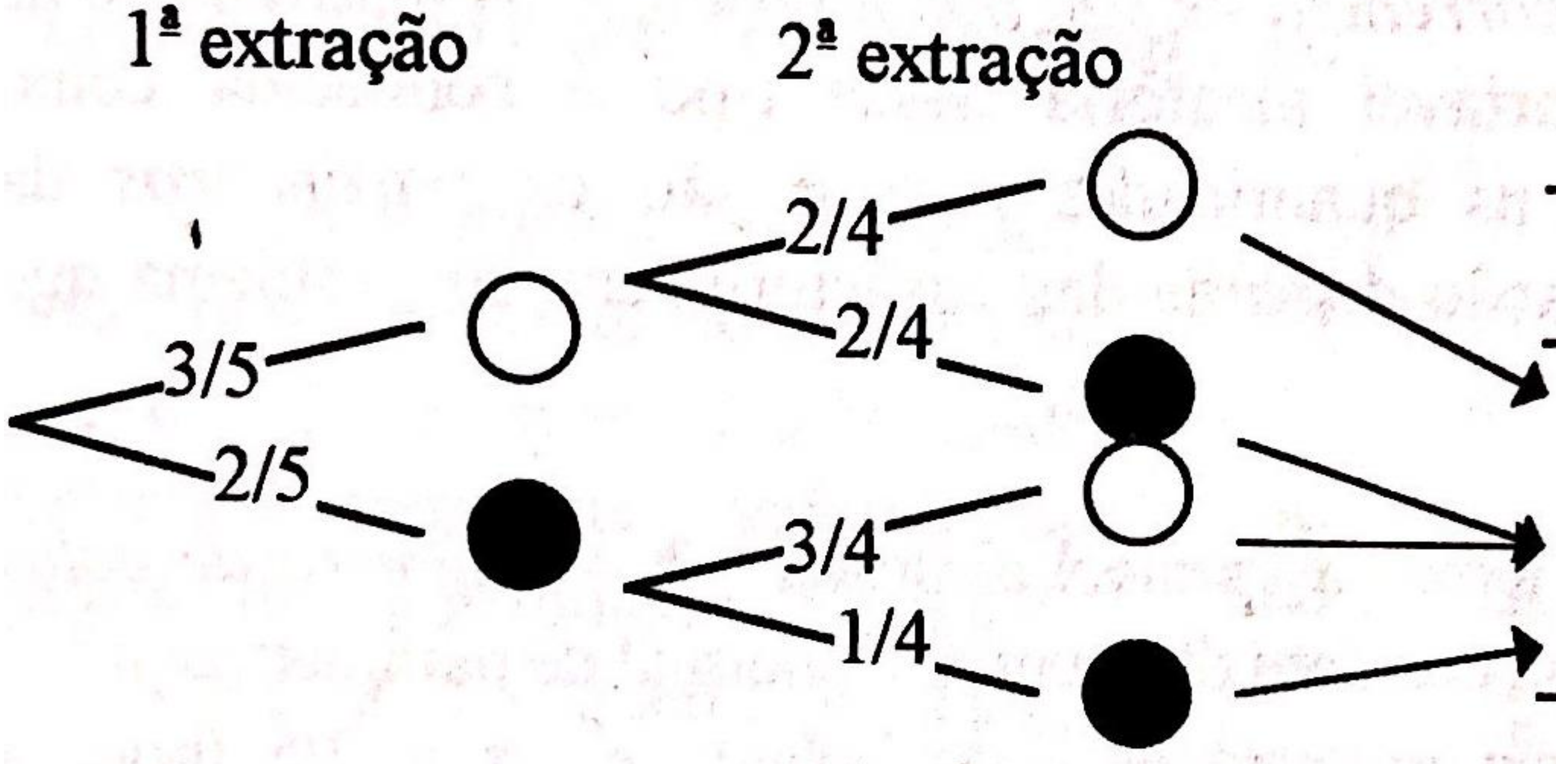
Variável aleatória: exemplo 2

EXEMPLO 7.6 Idem ao exemplo anterior, mas sem repor a primeira bola na segunda extração - *amostragem sem reposição*.



Variável aleatória: exemplo 2

- Variável **X**: número de bolas pretas extraídas
  - Eventos dependentes



Distribuição de X

x	p(x)
0	6/20 (ou 0,30)
1	12/20 (ou 0,60)
2	2/20 (ou 0,10)

Resultado importante

- Nos exemplos apresentados, a reposição causa uma diferença significativa nas probabilidades dos resultados
- Mas se houver um grande número de bolinhas (2000 brancas e 3000 pretas)
  - A distribuição de probabilidades da variável X será praticamente a mesma
    - Com ou sem reposição

$$P(0) = \frac{3000}{5000} \cdot \frac{3000}{5000} = \frac{9}{25} \qquad P(0) = \frac{3000}{5000} \cdot \frac{2999}{4999} \cong \frac{9}{25}$$

- Então, em grandes populações
  - Pode-se supor independência entre as extrações (ensaios)
    - Mesmo que a amostragem seja feita sem reposição

Experimento Binomial

1. Consiste de n ensaios (experimentos ou extrações)
2. Cada ensaio tem apenas dois resultados (Exemplo: sim ou não)
3. Os ensaios são independentes entre si, com probabilidade π de ocorrer sim, sendo π uma constante entre 0 e 1.

- A distribuição de probabilidades de uma variável aleatória desse tipo é conhecida como **distribuição binomial**

Tabela de probabilidades

TABELA II, continuação

n	x	π = 0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95
9	0	0,0008	0,0003	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	1	0,0083	0,0035	0,0013	0,0004	0,0001	0,0000	0,0000	0,0000	0,0000
	2	0,0407	0,0212	0,0098	0,0039	0,0012	0,0003	0,0000	0,0000	0,0000
	3	0,1160	0,0743	0,0424	0,0210	0,0087	0,0028	0,0006	0,0001	0,0000
	4	0,2128	0,1672	0,1181	0,0735	0,0389	0,0165	0,0050	0,0008	0,0000
	5	0,2600	0,2508	0,2194	0,1715	0,1168	0,0661	0,0283	0,0074	0,0006
	6	0,2119	0,2508	0,2716	0,2668	0,2336	0,1762	0,1069	0,0446	0,0077
	7	0,1110	0,1612	0,2162	0,2668	0,3003	0,3020	0,2597	0,1722	0,0629
	8	0,0339	0,0605	0,1004	0,1556	0,2253	0,3020	0,3679	0,3874	0,2986
	9	0,0046	0,0101	0,0207	0,0404	0,0751	0,1342	0,2316	0,3874	0,6302
10	0	0,0003	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	1	0,0042	0,0016	0,0005	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000
	2	0,0229	0,0106	0,0043	0,0014	0,0004	0,0001	0,0000	0,0000	0,0000
	3	0,0746	0,0425	0,0212	0,0090	0,0031	0,0008	0,0001	0,0000	0,0000
	4	0,1596	0,1115	0,0689	0,0368	0,0162	0,0055	0,0012	0,0001	0,0000
	5	0,2340	0,2007	0,1536	0,1029	0,0584	0,0264	0,0085	0,0015	0,0001
	6	0,2384	0,2508	0,2377	0,2001	0,1460	0,0881	0,0401	0,0112	0,0010
	7	0,1665	0,2150	0,2522	0,2668	0,2503	0,2013	0,1298	0,0574	0,0105
	8	0,0763	0,1209	0,1757	0,2335	0,2816	0,3020	0,2759	0,1937	0,0746
	9	0,0207	0,0403	0,0725	0,1211	0,1877	0,2684	0,3474	0,3874	0,3151
	10	0,0025	0,0060	0,0135	0,0282	0,0563	0,1074	0,1969	0,3487	0,5987

Experimento Binomial: exemplo 1

EXEMPLO 7.9 Seja a população de pessoas de um município, onde 70% são favoráveis a um certo projeto municipal (Exemplo 7.7(b)). Qual a probabilidade de que, numa amostra aleatória simples de 10 pessoas desta população, a maioria seja favorável ao projeto?

Experimento Binomial: exemplo 1

n	x	π 0,70
10	0	0,0000
	1	0,0001
	2	0,0014
	3	0,0090
	4	0,0368
	5	0,1029
	6	0,2001
	7	0,2668
	8	0,2335
	9	0,1211
	10	0,0282

$\boxed{X > 5}$

$$P(X > 5) =$$
$$= p(6) + p(7) + p(8) + p(9) + p(10) =$$
$$= 0{,}2001 + 0{,}2668 + 0{,}2335 + 0{,}1211 + 0{,}0282 =$$
$$= 0{,}8497.$$

Distribuição binomial graficamente

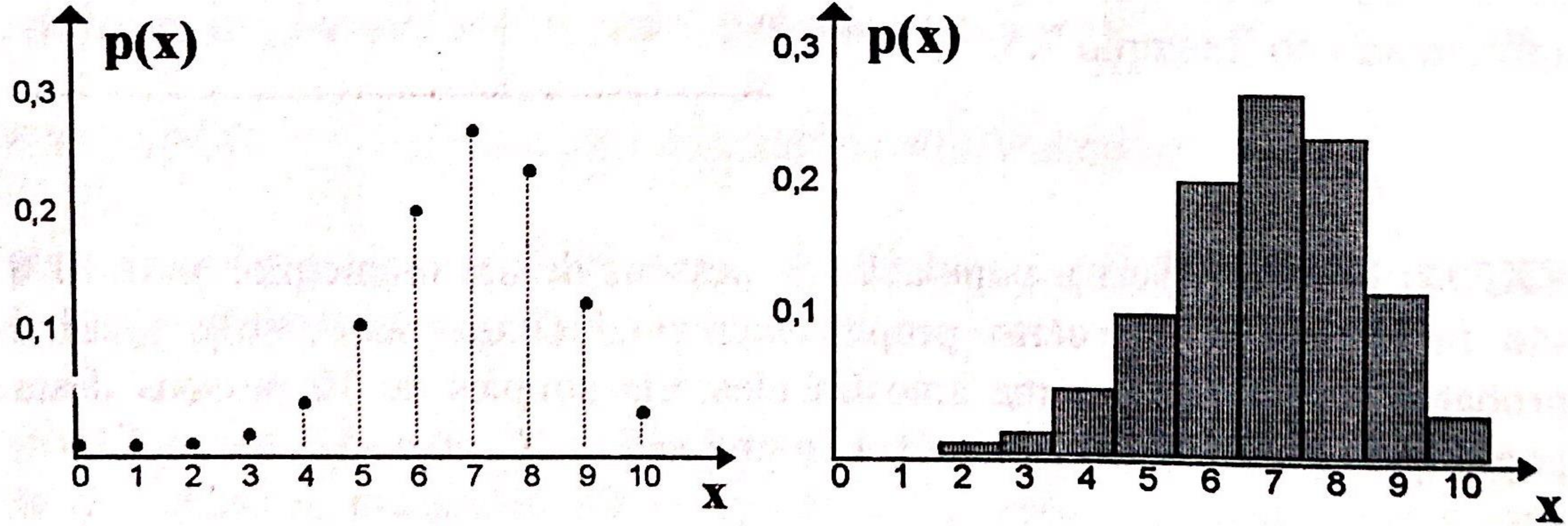
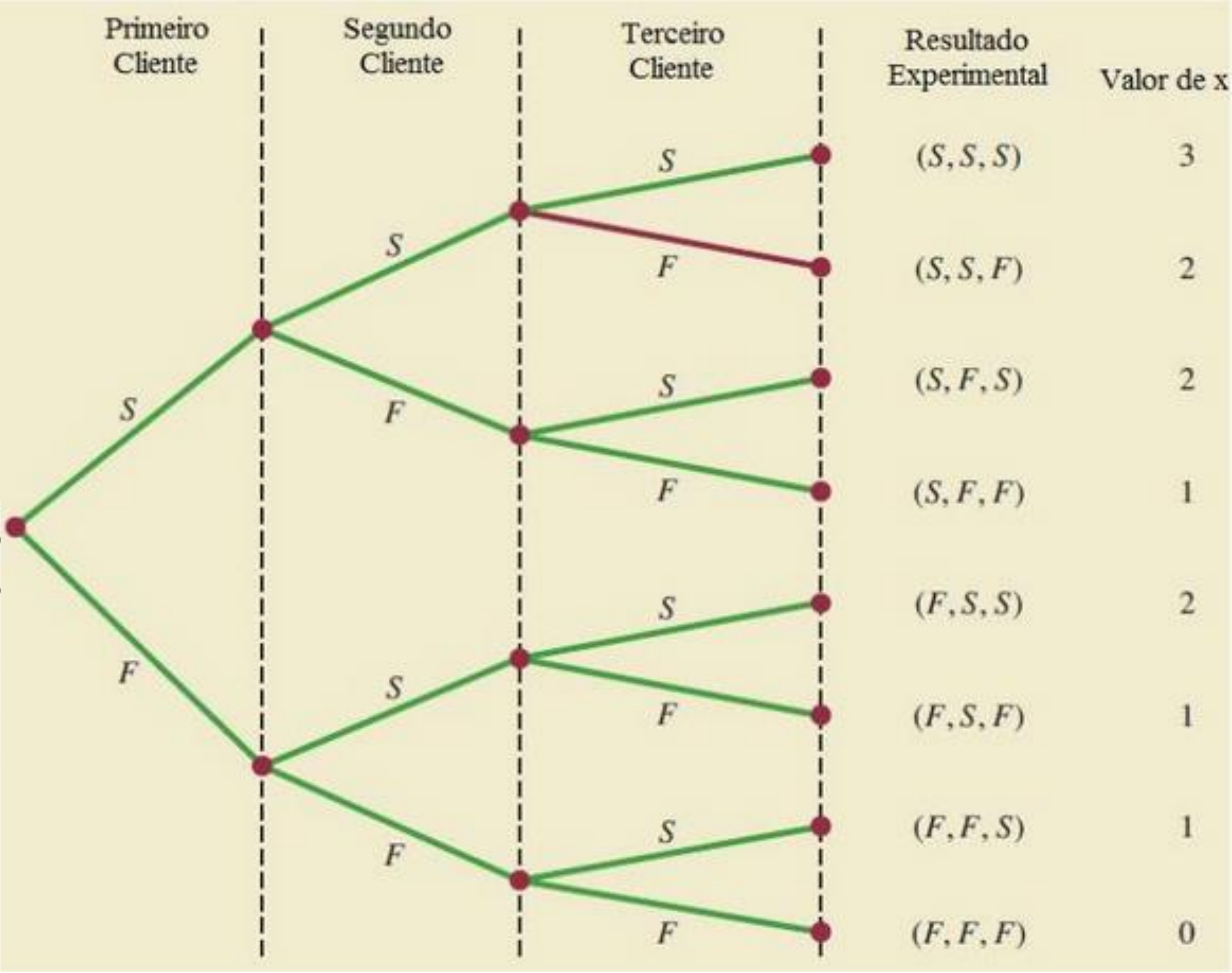


FIG. 7.2 Representações gráficas da distribuição binomial com n = 10 e $\pi = 0{,}7$ (Exemplo 7.7.b).

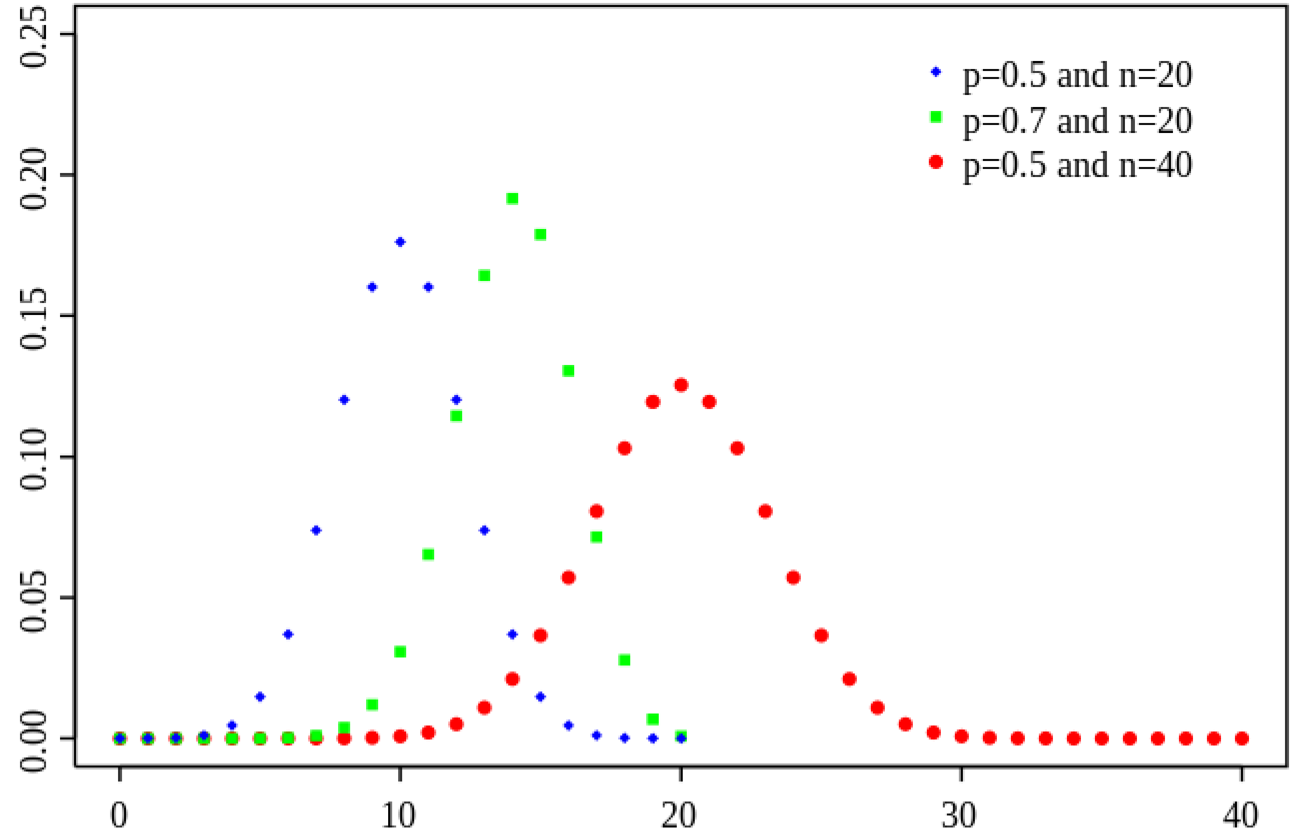
Experimento Binomial: exemplo 2

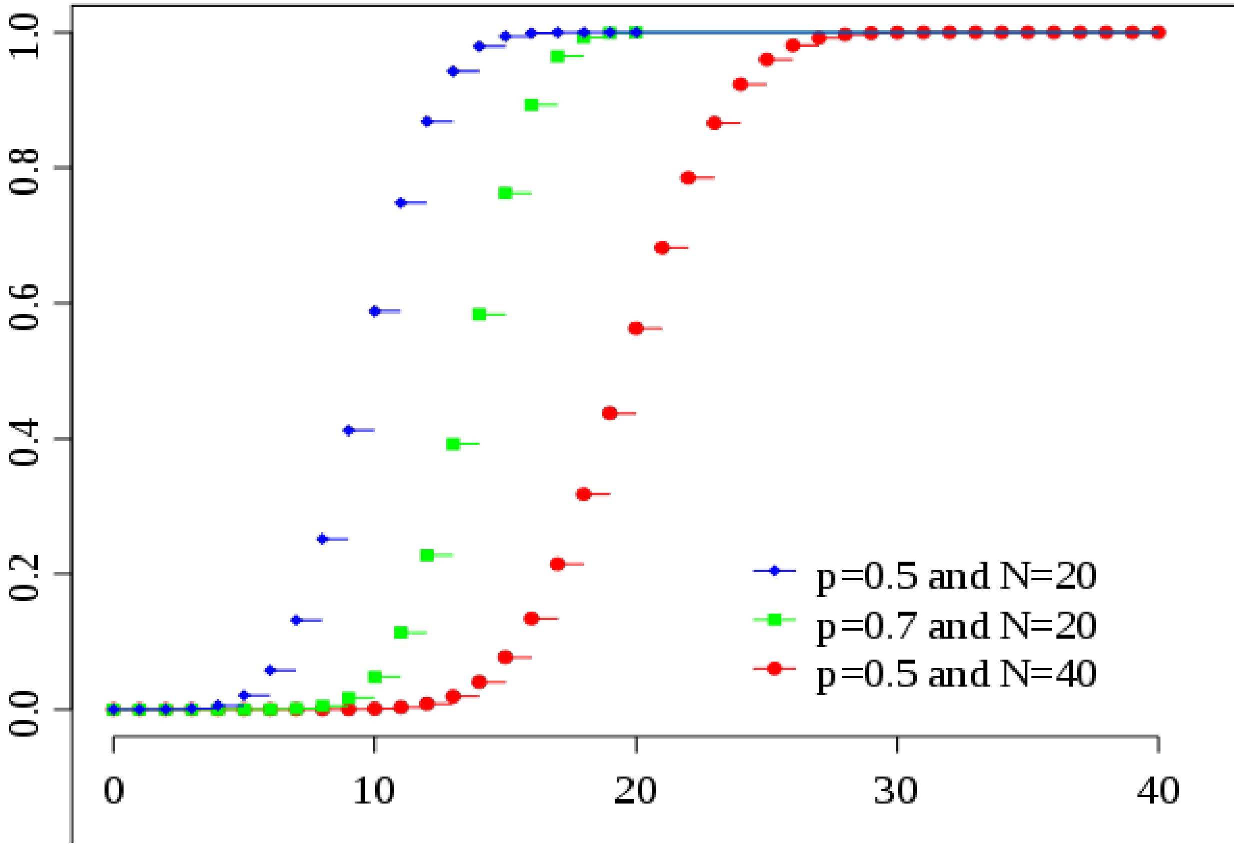
- Probabilidade de 3 clientes
  - Comprarem um produto
    - Considerando que os eventos sejam independentes
  - $\pi = 0.3$

n	x	π 0,05	0,10	0,15	0,20	0,25	0,30
1	0	0,9500	0,9000	0,8500	0,8000	0,7500	0,7000
	1	0,0500	0,1000	0,1500	0,2000	0,2500	0,3000
2	0	0,9025	0,8100	0,7225	0,6400	0,5625	0,4900
	1	0,0950	0,1800	0,2550	0,3200	0,3750	0,4200
	2	0,0025	0,0100	0,0225	0,0400	0,0625	0,0900
3	0	0,8574	0,7290	0,6141	0,5120	0,4219	0,3430
	1	0,1354	0,2430	0,3251	0,3840	0,4219	0,4410
	2	0,0071	0,0270	0,0574	0,0960	0,1406	0,1890
	3	0,0001	0,0010	0,0034	0,0080	0,0156	0,0270



Distribuição binomial graficamente

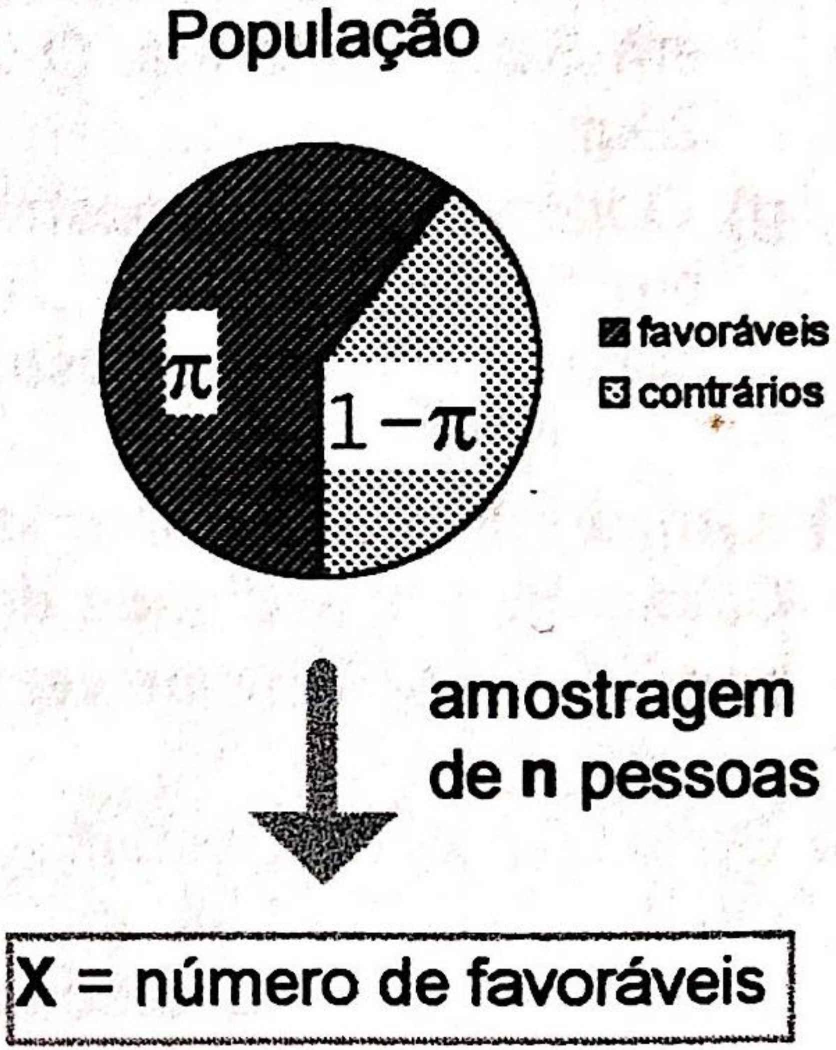




Formulação Matemática do Modelo Binomial

Considere o seguinte experimento: seja **X** o número de pessoas favoráveis a um certo projeto municipal, numa amostra aleatória simples de **n** pessoas, extraída de uma população, onde a proporção de favoráveis é igual a π, como ilustra o esquema ao lado.

Admitindo que o tamanho da população seja bastante superior ao tamanho da amostra, podemos supor que a variável aleatória **X** tenha distribuição binomial, com parâmetros **n** e π.



Formulação Matemática do Modelo Binomial

Para cada um dos indivíduos indagados a respeito do projeto, vamos representar por **S** a resposta *sim* (*favorável*) e por **N** a resposta *não* (*contrário*). A Figura 7.3 apresenta as possíveis combinações de respostas **S** e **N**, numa amostra de n = 4 pessoas. Esta figura também mostra os valores da variável aleatória **X** e suas respectivas probabilidades.

Formulação Matemática do Modelo Binomial

			SSNN		
			SNSN		
		SNNN	SNNS	SSSN	
		NSNN	NSSN	SSNS	
		NNSN	NSNS	SNSS	
	NNNN	NNNS	NNSS	NSSS	SSSS
Valores de X:	0	1	2	3	4
Probabilidades:	$(1-\pi)^4$	$4\pi(1-\pi)^3$	$6\pi^2(1-\pi)^2$	$4\pi^3(1-\pi)$	π^4

FIG. 7.3 Possíveis seqüências de respostas e construção de uma distribuição binomial de probabilidades com $n = 4$ e π genérico.

Binômio de Newton

- Como construir essa pirâmide?

- Qual é a equação do binômio de newton da 6ª linha ?

1

1 1

1 2 1

1 3 3 1

1 4 6 4 1

1 5 10 10 5 1

1 6 15 20 15 16 1

Como calcular os coeficientes do Binômio de Newton

- Combinações

$$\binom{n}{x} = \frac{n!}{(n-x)!\,x!}$$

		SSNN		
		SNSN		
	SNNN	SNNS	SSSN	
	NSNN	NSSN	SSNS	
	NNSN	NSNS	SNSS	
NNNN	NNNS	NNSS	NSSS	SSSS

Binômio de Newton

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 16 & & 1
\end{array}
\longleftrightarrow
\begin{array}{ccccccc}
 & & & \binom{0}{0} & & & \\
 & & \binom{1}{0} & & \binom{1}{1} & & \\
 & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & \\
\binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3}
\end{array}
$$

Expressão geral da distribuição binomial

$$p(x) = \binom{n}{x} . \pi^{x} . (1 - \pi)^{n-x}$$

- **Exemplo**: Seja a população de pessoas de um município, onde **70%** são favoráveis a um certo projeto municipal. Qual a probabilidade de, numa amostra aleatória simples de **quatro** pessoas desta população, encontrarmos exatamente **3** pessoas favoráveis ao projeto ?

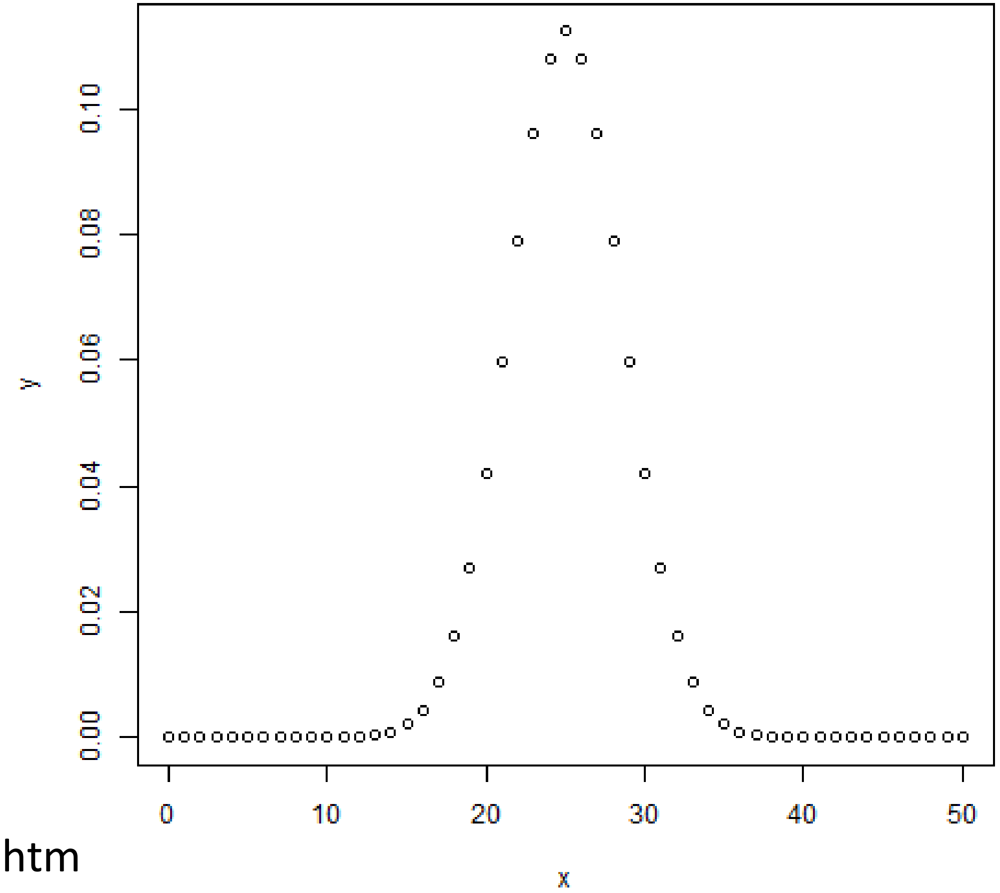
$$p(3) = \binom{4}{3} . 0{,}7^{3} . (0{,}3)^{1} = \frac{4!}{(4-3)!\,3!} . 0{,}7^{3} . (0{,}3) = 0{,}4116$$

Funções do R para o cálculo binomial

- dbinom()
  - Calcula a distribuição de probabilidade em cada ponto

```
# Cria uma amostra de 50 números
# sequenciais a partir de 1
x <- seq(0,50,by = 1)

# Cria a distribuição binomial
y <- dbinom(x,50,0.5)
```



Fonte: https://www.tutorialspoint.com/r/r_binomial_distribution.htm

Funções do R para o cálculo binomial

- pbinom()
  - Calcula a probabilidade acumulada de um evento

```
# Calcular a probabilidade de obter 26 ou menos caras a partir de 50
# lançamentos de moedas
x <- pbinom(26,51,0.5)

[1] 0.610116
```

Fonte: https://www.tutorialspoint.com/r/r_binomial_distribution.htm

Exercício de Avaliação - 1

Uma companhia de seguros vendeu apólices a **5** pessoas. Todas da mesma idade e com boa saúde. De acordo com as tábuas atuariais, a probabilidade de que uma pessoa daquela idade esteja viva daqui 30 anos é de **2/3**. Calcular a probabilidade de que, daqui a 30 anos:

a) Exatamente duas pessoas estejam vivas; e
b) Todas as pessoas estejam vivas.

Instruções:

1) Entregar a solução com a aplicação da fórmula do binômio de newton (equivalente ao exemplo anterior).
2) Usar somente calculadora.
3) Tirar uma foto e encaminhar por e-mail para o professor.

Exercício de Avaliação – 2 (no R)

Sobre o exercício anterior, calcule no R:

a) A probabilidade de que pelo menos 3 pessoas estejam vivas; e
b) A probabilidade de exatamente 4 pessoas estarem vivas.

Instruções:

1) Crie um novo Jupyter notebook para este exercício.
2) Encaminhe por e-mail para o arquivo do Jupyter notebook com a solução (você pode exportar o arquivo na opção File -> Download as -> Notebook).

Exercício de Avaliação – 3 (no R)

Plote um gráfico do tipo boxplot (gráfico de caixas) no R com algum dado real (não inventado por você).

Instruções:

1) Crie um novo Jupyter notebook para este exercício.
2) Encaminhe por e-mail para o arquivo do Jupyter notebook com a solução (você pode exportar o arquivo na opção File -> Download as -> Notebook).

Exercício de Avaliação – 4

Suponha o lançamento de um dado não viciado. O espaço amostral é $S = \{1, 2, 3, 4, 5, 6\}$. Seja A = face é 2 ou 3, e B = face é par (2, 4, 6). Calcular $P(A|B)$. Dica: use a formula da probabilidade condicional.

Instruções:

1) Entregar a solução com a aplicação da fórmula do binômio de newton (equivalente ao exemplo do slide anterior).
2) Tirar uma foto e encaminhar por e-mail para o professor.