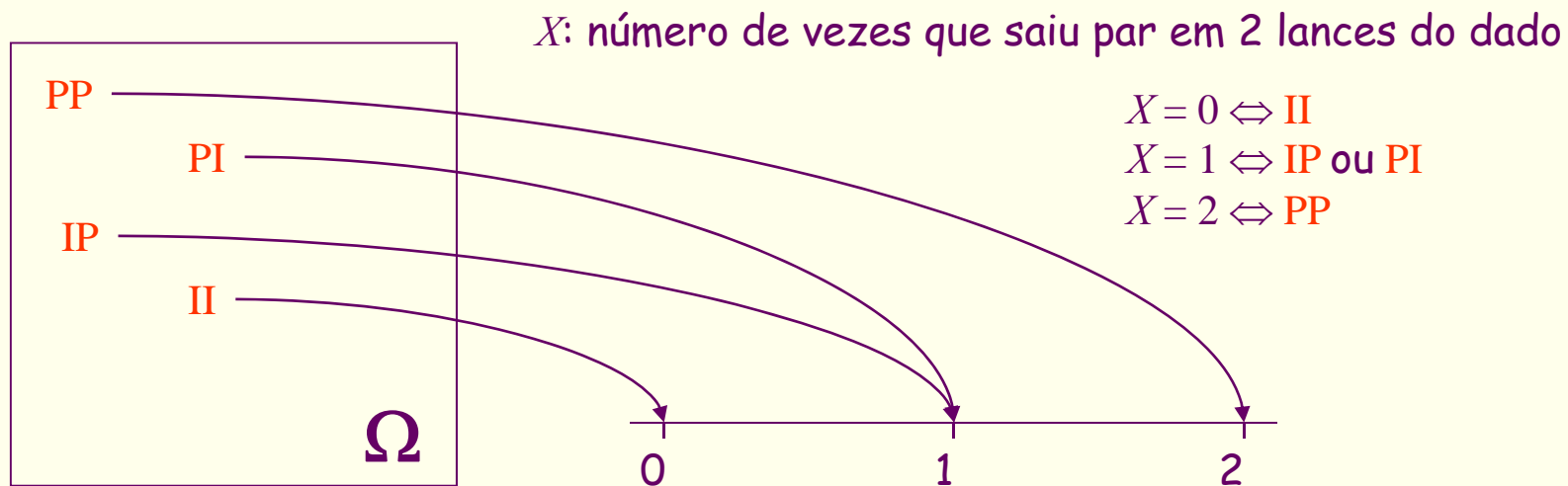


VARIÁVEIS ALEATÓRIAS e DISTRIBUIÇÃO BINOMIAL

Variável Aleatória

Uma função X que associa a cada elemento ω do espaço amostral Ω um valor $x \in \mathbf{R}$ é denominada uma *variável aleatória*.

Experimento: jogar 1 dado duas vezes e observar o resultado
(P = par e I = ímpar)



Variável Aleatória

Uma variável aleatória pode ser classificada em:

- Variável aleatória discreta
- Variável aleatória contínua

Exemplos:

1) Observa-se o **sexo** (característica) das crianças em famílias com três filhos (M : masculino e F : feminino).

Espaço amostral:

$$\Omega = \{(\underbrace{MMM}_{\omega_1}), (\underbrace{MMF}_{\omega_2}), (\underbrace{MFM}_{\omega_3}), (\underbrace{FMM}_{\omega_4}), (\underbrace{MFF}_{\omega_5}), (\underbrace{FMF}_{\omega_6}), (\underbrace{FFM}_{\omega_7}), (\underbrace{FFF}_{\omega_8})\}$$

Defina X : n°. de crianças do sexo masculino (M).

Ω	MMM	MMF	MFM	FMM	MFF	FMF	FFM	FFF
X	3	2	2	2	1	1	1	0

→ Então X assume valores no conjunto $\{0, 1, 2, 3\}$, logo é uma variável aleatória discreta.

Variável Aleatória

Uma variável aleatória pode ser classificada em:

- Variável aleatória discreta

Uma v.a. é discreta quando o conjunto de valores possíveis que ela assume for finito ou infinito enumerável.

Exemplos:

1) No mesmo experimento...

Espaço amostral:

$$\Omega = \{(\underbrace{MMM}_{\omega_1}), (\underbrace{MMF}_{\omega_2}), (\underbrace{MFM}_{\omega_3}), (\underbrace{FMM}_{\omega_4}), (\underbrace{MFF}_{\omega_5}), (\underbrace{FMF}_{\omega_6}), (\underbrace{FFM}_{\omega_7}), (\underbrace{FFF}_{\omega_8})\}$$

Podemos definir agora

Y : n°. de crianças do sexo feminino (F).

Ω	MMM	MMF	MFM	FMM	MFF	FMF	FFM	FFF
Y	0	1	1	1	2	2	2	3

→ Então Y também assume valores no conjunto $\{0, 1, 2, 3\}$, porém, para outros valores de Ω .

Exemplos:

2) Observar o tempo de vida, em horas, de lâmpadas produzidas por uma fábrica.



Defina T : tempo de vida, em horas, da lâmpada escolhida, ao acaso, da fábrica.

→ Então, T é uma variável aleatória contínua que assume qualquer valor real não negativo.

Variável Aleatória

Uma variável aleatória pode ser classificada em:

- Variável aleatória contínua

Uma v.a. é contínua quando o conjunto de valores possíveis que ela assume for não enumerável.

VARIÁVEL ALEATÓRIA DISCRETA

Caracterização

Função de probabilidade: É a função que atribui a cada valor x_i da v. a. discreta X sua probabilidade de ocorrência e pode ser representada pela tabela:

x	x_1	x_2	\dots	x_n
$P(X=x)$	$P(X=x_1)$	$P(X=x_2)$	\dots	$P(X=x_n)$

Uma função de probabilidade deve satisfazer:

$$0 \leq P(X = x_i) \leq 1 \quad \text{e} \quad \sum_{i=1}^n P(X = x_i) = 1$$

Exemplo 1:

O Departamento de Estatística é formado por 35 professores, sendo 21 homens e 14 mulheres. Uma comissão de 3 professores será constituída sorteando, ao acaso, três membros do departamento. Qual é a probabilidade da comissão ser formada por *pelo menos duas mulheres*?

Vamos definir a v.a.

X : nº. de mulheres na comissão.

Quais são os possíveis valores que X pode assumir?

Espaço amostral	Probabilidade	X
(HHH)	$\frac{21}{35} \times \frac{20}{34} \times \frac{19}{33} = 0,203$	0
(HHM)	$\frac{21}{35} \times \frac{20}{34} \times \frac{14}{33} = 0,150$	1
(HMH)	$\frac{21}{35} \times \frac{14}{34} \times \frac{20}{33} = 0,150$	1
(MHH)	$\frac{14}{35} \times \frac{21}{34} \times \frac{20}{33} = 0,150$	1
(HMM)	$\frac{21}{35} \times \frac{14}{34} \times \frac{13}{33} = 0,097$	2
(MHM)	$\frac{14}{35} \times \frac{21}{34} \times \frac{13}{33} = 0,097$	2
(MMH)	$\frac{14}{35} \times \frac{13}{34} \times \frac{21}{33} = 0,097$	2
(MMM)	$\frac{14}{35} \times \frac{13}{34} \times \frac{12}{33} = 0,056$	3

x	0	1	2	3
$P(X = x)$	0,203	0,450	0,291	0,056

Assim, $P(X \geq 2) = P(X=2) + P(X=3) = 0,291 + 0,056 = 0,347$.

Exemplo 2: Um dado é lançado duas vezes, de forma independente. Qual é a probabilidade da soma dos pontos nos dois lançamentos ser menor do que 6?

$$\Omega = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), \\ (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), \\ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), \\ (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), \\ (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), \\ (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}.$$

Qual é a probabilidade de cada ponto w_i de Ω ?

Admitindo-se que o dado seja perfeitamente homogêneo e sendo os lançamentos independentes,

$$P(w_i) = 1/36, \text{ qualquer } w_i \in \Omega.$$

Defina X : soma dos pontos nos dois lançamentos do dado.

Função de probabilidade de X :



x	2	3	4	5	6	7	8	9	10	11	12
$P(X=x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Então,

$$\begin{aligned}P(X < 6) &= P(X=5) + P(X=4) + P(X=3) + P(X=2) \\&= 4/36 + 3/36 + 2/36 + 1/36 \\&= 10/36 = 0,278\end{aligned}$$

X : Soma dos pontos nos dois lançamentos

		2 ^o . lançamento					
		1	2	3	4	5	6
1 ^o . Lan ça- men- to	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12



Y : valor máximo obtido dentre os dois lançamentos.

		2 ^o . lançamento					
		1	2	3	4	5	6
1 ^o . Lan ça- men- to	1	1	2	3	4	5	6
	2	2	2	3	4	5	6
	3	3	3	3	4	5	6
	4	4	4	4	4	5	6
	5	5	5	5	5	5	6
	6	6	6	6	6	6 ¹⁴	6

Podemos estar interessados em outras variáveis aleatórias definidas para o mesmo espaço amostral.

Y : valor máximo obtido dentre os dois lançamentos.

y	1	2	3	4	5	6
$P(Y = y)$	1/36	3/36	5/36	7/36	9/36	11/36



Z : diferença entre os pontos do 2º. e do 1º. lançamento.

z	-5	-4	-3	-2	-1	0	1	2	3	4	5
$P(Z = z)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

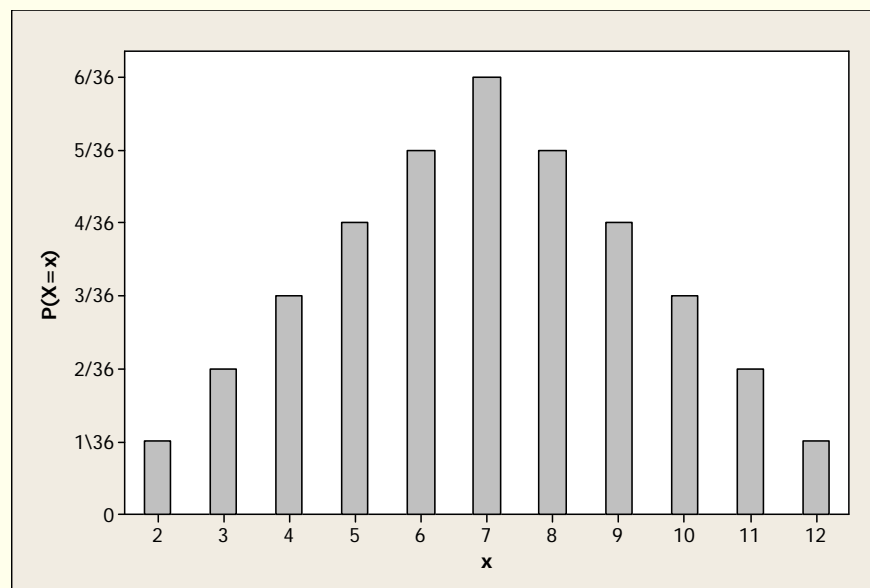
VALOR ESPERADO E VARIÂNCIA

Qual é o valor médio da soma dos pontos (X) no lançamento de dois dados?

$\Omega = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6),$
 $(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),$
 $(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),$
 $(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),$
 $(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),$
 $(6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}$

\Rightarrow 36 pontos igualmente prováveis

x	$P(X = x)$
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36



VALOR ESPERADO E VARIÂNCIA

Valor Esperado (“média”): Dada a v.a. X , assumindo os valores x_1, x_2, \dots, x_n , chamamos de *valor médio*, ou *valor esperado*, ou *esperança matemática* da distribuição de X o valor

$$E(X) = x_1 \times P(X = x_1) + \dots + x_n \times P(X = x_n) = \sum_{i=1}^n x_i \times P(X = x_i)$$

Notação: $\mu = E(X)$

No exemplo, para média de X (soma de pontos), temos:

$$\begin{aligned} E(X) &= 2 \times (1/36) + 3 \times (2/36) + \dots + 11 \times (2/36) + 12 \times (1/36) \\ &= 252/36 = 7, \end{aligned}$$

ou seja, em média, a soma dos pontos no lançamento dos dois dados é igual a 7.

Variância: É o valor esperado da v.a. $(X - E(X))^2$, ou seja, se X assume os valores x_1, x_2, \dots, x_n , então

$$\text{Var}(X) = \sum_{i=1}^n [x_i - E(X)]^2 \times P(X = x_i)$$

Notação: $\sigma^2 = \text{Var}(X)$.

Da relação acima, segue que

$$\text{Var}(X) = E(X^2) - [E(X)]^2.$$

Desvio Padrão: É definido como a raiz quadrada positiva da variância, isto é,

$$\text{DP}(X) = \sqrt{\text{Var}(X)}.$$

Notação: $\sigma = \text{DP}(X)$.

No exemplo,

x	2	3	4	5	6	7	8	9	10	11	12
$P(X=x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

$$\begin{aligned}\text{Var}(X) &= (2-7)^2 \times \frac{1}{36} + (3-7)^2 \times \frac{2}{36} + \dots + (11-7)^2 \times \frac{2}{36} + (12-7)^2 \times \frac{1}{36} \\ &= \frac{210}{36} = 5,83.\end{aligned}$$

Alternativamente, poderíamos calcular

$$\begin{aligned}E(X^2) &= 2^2 \times \frac{1}{36} + 3^2 \times \frac{2}{36} + \dots + 11^2 \times \frac{2}{36} + 12^2 \times \frac{1}{36} \\ &= \frac{1974}{36} = 54,83\end{aligned}$$

e, portanto, $\text{Var}(X) = 54,83 - 7^2 = 5,83$.

Propriedades:

1) Se $P(X = a) = 1$, então

$$E(X) = a \quad \text{e} \quad \text{Var}(X) = 0.$$

2) Se $Y = aX + b$, em que a e b são constantes, então

$$E(Y) = E(aX + b) = aE(X) + b$$

e

$$\text{Var}(Y) = \text{Var}(aX + b) = a^2 \text{Var}(X).$$

- MODELOS PROBABILÍSTICOS DISCRETOS -

Modelo de Bernoulli ou Binário

Na prática, existem muitos experimentos que admitem apenas dois resultados.

Exemplos:

- uma peça é classificada como boa ou defeituosa;
- o resultado de um exame médico para detecção de uma doença é positivo ou negativo;
- um paciente submetido a um tratamento, durante um período de tempo fixo, cura-se ou não da doença;
- um entrevistado concorda ou não com a afirmação feita;
- no lançamento de um dado ocorre ou não a face “5”.

Situações com alternativas *dicotômicas* podem ser representadas, genericamente, por respostas do tipo sucesso-fracasso.

Esses experimentos recebem o nome de *Ensaaios de Bernoulli* e originam uma v.a. com distribuição de Bernoulli.

Variável aleatória de Bernoulli: É uma v.a. que assume apenas dois valores:

- 1 se ocorrer *sucesso*,
- 0 se ocorrer *fracasso*.

Geralmente, a probabilidade de sucesso é representada por p , $0 < p < 1$.

“ $X \sim \text{Bernoulli}(p)$ ” indica uma v.a. com *distribuição de Bernoulli* com parâmetro p , isto é,

$$X = \begin{cases} 1, & \text{se ocorrer “sucesso”} \\ 0, & \text{se ocorrer “fracasso”} \end{cases}$$

e sua função de probabilidade pode ser representada pela tabela

X	1	0
$P(X=x)$	p	$1 - p$

Segue que

$$\begin{aligned} E(X) &= p, \\ \text{Var}(X) &= p(1 - p). \end{aligned}$$

→ Repetições independentes de um ensaio de Bernoulli, com a mesma probabilidade de ocorrência de “sucesso”, dão origem ao *modelo de probabilidade binomial*.

Modelo Binomial

Exemplo: Um dado equilibrado é lançado 3 vezes.
Qual é a probabilidade de se obter a face 5 duas vezes?

Denotamos,

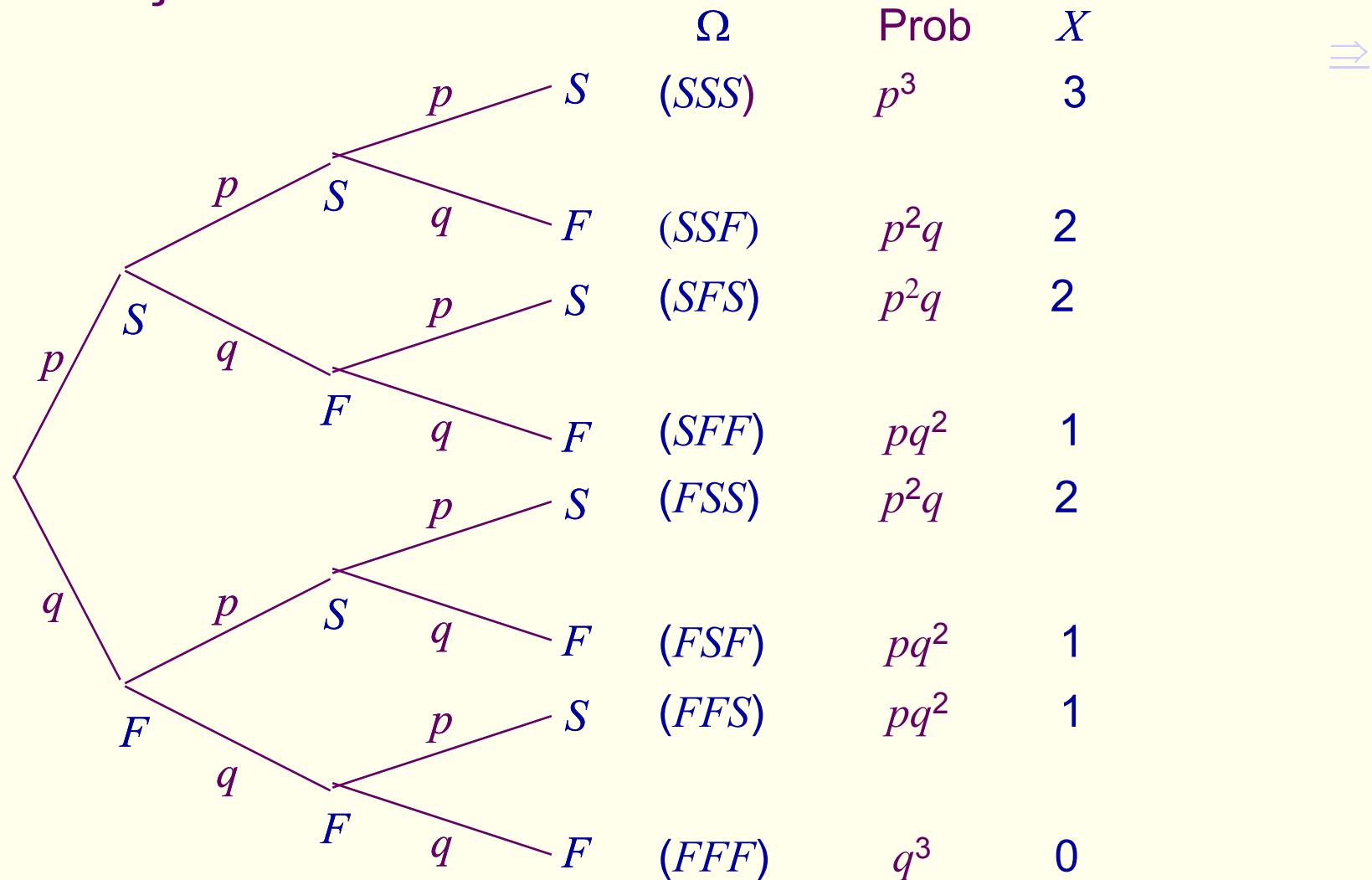
S : “sucesso”, ocorrer face 5;

F : “fracasso”, não ocorrer face 5.

É fácil ver que $p = P(\text{sucesso}) = 1/6$ e
 $q = 1 - p = P(\text{fracasso}) = 5/6$

$\Omega = \{SSS, SSF, SFS, FSS, SFF, FSF, FFS, FFF\}$

Estamos interessados no número total de sucessos que, no caso, é o número de vezes que a face 5 é observada nos 3 lançamentos do dado.



A função de probabilidade de X é dada por:

Probabilidades binomiais para $n = 3$ e $P(S) = p$

nº. de sucessos	probabilidades	$p = 1/6$
0	q^3	125/216=0,5787
1	$3pq^2$	75/216=0,3472
2	$3p^2q$	15/216=0,0694
3	p^3	1/216=0,0046

Podemos escrever essa função como

$$P(X = k) = \binom{3}{k} p^k q^{3-k}, \quad k = 0, 1, 2, 3.$$

No exemplo, para $n = 3$ e $p = 1/6$, $P(X = 2) = 0,0694$.

Distribuição binomial:

A v.a. X correspondente ao número de sucessos em n ensaios de Bernoulli independentes e com mesma probabilidade p de sucesso, tem distribuição binomial com parâmetros n e p .

Sua função de probabilidade é dada por

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n.$$

Notação: $X \sim b(n; p)$.

Resultado: Se $X \sim b(n; p)$, então

valor esperado: $\mu = E(X) = n \times p$

variância: $\sigma^2 = \text{Var}(X) = n \times p \times (1-p)$

Exemplo utilizando o R:

Considere uma prova com 12 questões, cada uma com 4 alternativas. Suponha que o aluno escolha a resposta ao acaso. Qual é a probabilidade de que ele *acerte pelo menos 6 questões*?

X : n°. de questões que o aluno acertará

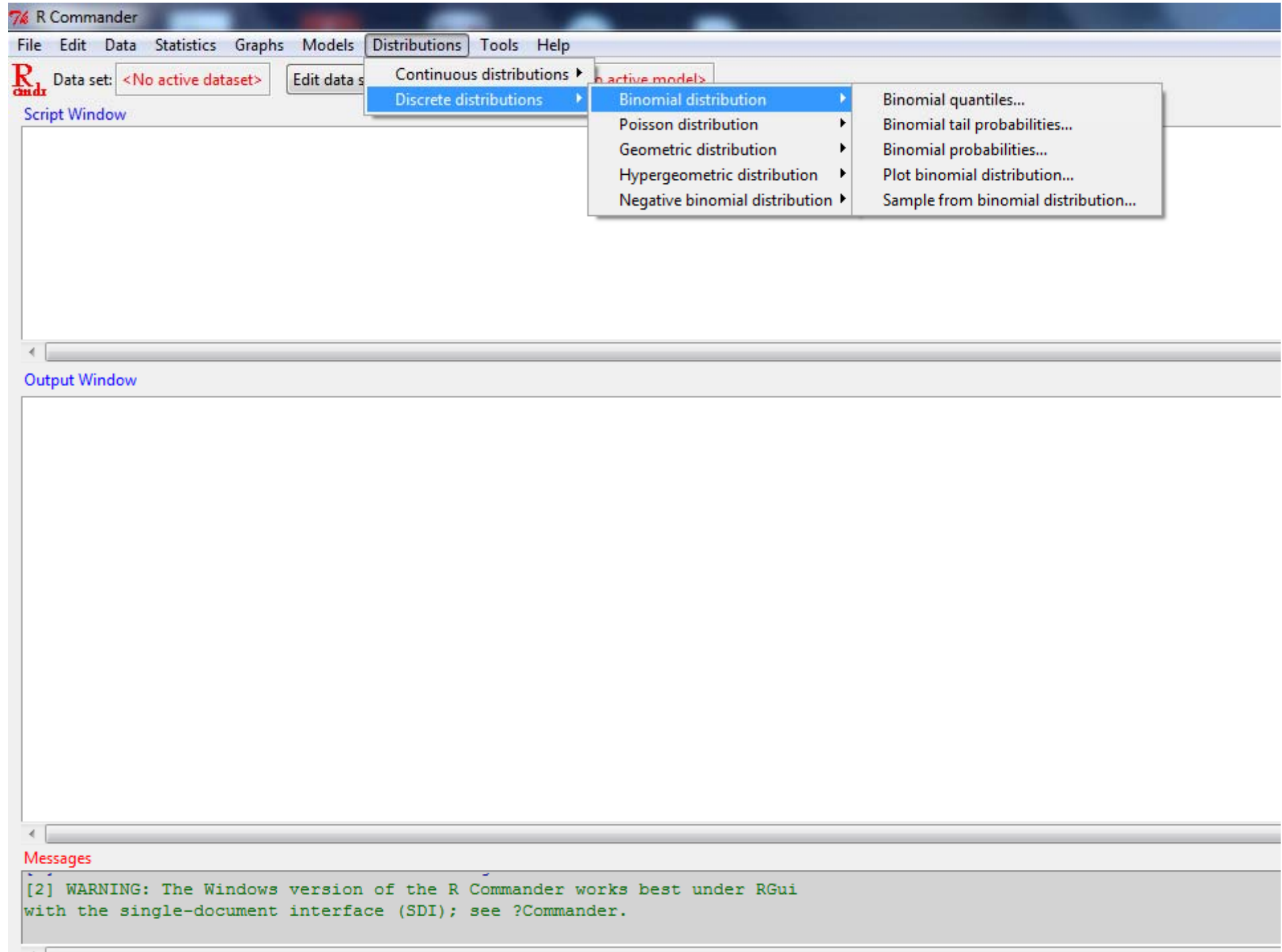
$$X = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 \text{ ou } 12\}$$

$$X \sim b(12; 0,25)$$

$$P(X = x) = \binom{12}{x} 0,25^x (1 - 0,25)^{12-x}$$



Uso do R para os cálculos!



Binomial Probabilities

Binomial trials: 12

Probability of success: 0.25

OK Cancel Help

Model: <No active model>

1=FALSE)
1=FALSE)
1=FALSE)
1=FALSE)
1=FALSE)

```
Table <- data.frame(r = rbinom(12, size=12, prob=0.25))  
rownames(.Table) <- 0:12  
.Table  
remove(.Table)
```

Output Window

```
> rownames(.Table) <- 0:12
```

```
> .Table
```

	Pr
0	3.167635e-02
1	1.267054e-01
2	2.322932e-01
3	2.581036e-01
4	1.935777e-01
5	1.032414e-01
6	4.014945e-02
7	1.147127e-02
8	2.389848e-03
9	3.540516e-04
10	3.540516e-05
11	2.145767e-06
12	5.960464e-08

```
> remove(.Table)
```

Messages

```
[2] WARNING: The Windows version of the R Commander works best under RGui  
with the single-document interface (SDI); see ?Commander.
```



Data set: <No active dataset>

Edit data set

View data set

Model: <No active model>

Script Window

```
rownames(.Table) <- 0:12
.Table
remove(.Table)
pbinom(c(6), size=12, prob=0.25, lower.tail=TRUE)
pbinom(c(6), size=12, prob=0.25, lower.tail=FALSE)
.Table <- data.frame(Pr=dbinom(0:12, size=12, prob=0.25))
rownames(.Table) <- 0:12
.Table
remove(.Table)
```

Output Window

```
> rownames(.Table) <- 0:12
```

```
> .Table
```

	Pr
0	3.167635e-02
1	1.267054e-01
2	2.322932e-01
3	2.581036e-01
4	1.935777e-01
5	1.032414e-01
6	4.014945e-02
7	1.147127e-02
8	2.389848e-03
9	3.540516e-04
10	3.540516e-05
11	2.145767e-06
12	5.960464e-08

```
> remove(.Table)
```

Messages

```
[2] WARNING: The Windows version of the R Commander works best under RGui
with the single-document interface (SDI); see ?Commander.
```



```
> .Table
```

	Pr
0	3.167635e-02
1	1.267054e-01
2	2.322932e-01
3	2.581036e-01
4	1.935777e-01
5	1.032414e-01
6	4.014945e-02
7	1.147127e-02
8	2.389848e-03
9	3.540516e-04
10	3.540516e-05
11	2.145767e-06
12	5.960464e-08

Portanto,

$$P(X \geq 6) = 0,0544.$$

Temos $E(X) = n \times p =$
 $12 \times 0,25 = 3,$

ou seja, *em média*, o aluno
que responder ao acaso
todas as questões acertará
3.

R Commander

File Edit Data Statistics Graphs Models Distributions Tools Help

Model: <No active model>

Binomial Probabilities

Variable value(s) 5

Binomial trials 12

Probability of success 0.25

Lower tail ☐

Upper tail ☒

OK Cancel Help

1=FALSE)
1=FALSE)
1=FALSE)
1=FALSE)
1=FALSE)
1=FALSE)

Output Window

```
> pbinom(c(5), size=12, prob=0.25, lower.tail=FALSE)
[1] 0.05440223
```

Messages

```
[2] WARNING: The Windows version of the R Commander works best under RGui
with the single-document interface (SDI); see ?Commander.
```