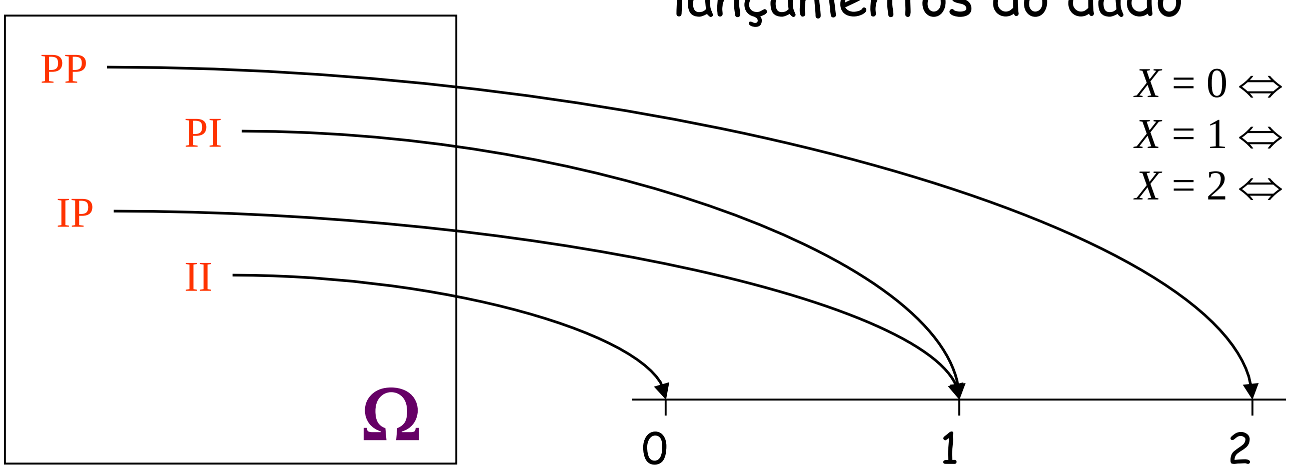
# VARIÁVEIS ALEATÓRIAS
# e
# DISTRIBUIÇÃO BINOMIAL

# Variável Aleatória

Uma **variável aleatória** é uma função $X$ que associa cada elemento $\omega$ do espaço amostral $\Omega$ a um valor $x \in \mathbb{R}$.

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

$X$: número de vezes que saiu par em 2 lançamentos do dado




$X = 0 \Leftrightarrow$ II
$X = 1 \Leftrightarrow$ IP ou PI
$X = 2 \Leftrightarrow$ PP

# Variável Aleatória

Uma variável aleatória pode ser classificada em:

- Variável aleatória **discreta**

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

- Variável aleatória **contínua**

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

# Exemplos:

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

Espaço amostral:

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

Defina $X$: número de crianças do gênero masculino ($M$)

| $\Omega$ | $MMM$ | $MMF$ | $MFM$ | $FMM$ | $MFF$ | $FMF$ | $FFM$ | $FFF$ |
|---|---|---|---|---|---|---|---|---|
| $X$ | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 0 |

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

# Exemplos:

2) No mesmo experimento...

Espaço amostral:

$$\Omega = \{(MMM), (MMF), (MFM), (FMM), (MFF), (FMF), (FFM), (FFF)\}$$

$$\omega_1 \quad \omega_2 \quad \omega_3 \quad \omega_4 \quad \omega_5 \quad \omega_6 \quad \omega_7 \quad \omega_8$$

Podemos definir agora uma outra variável aleatória

***Y***: número de crianças do sexo feminino (*F*)

| $\Omega$ | *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 elementos de **Ω**. Também é uma variável aleatória **discreta**.

# Exemplos:

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

Defina $T$: tempo de vida, em horas, de uma lâmpada escolhida ao acaso.

$\rightarrow$ Então, $T$ é uma variável aleatória **contínua** que assume qualquer valor real não negativo.

# VARIÁVEL ALEATÓRIA DISCRETA

## Caracterização

Função (ou distribuição) 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 por

| $x$ | $x_1$ | $x_2$ | ... | $x_n$ |
|---|---|---|---|---|
| $P(X=x)$ | $P(X=x_1)$ | $P(X=x_2)$ | ... | $P(X=x_n)$ |

Uma função de probabilidade deve satisfazer:

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

## 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-se, ao acaso, três membros do departamento.

Qual é a probabilidade da comissão ser formada por ***pelo menos duas mulheres***?

Vamos definir a v.a.

$X$: número 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 das faces 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 $\omega_i$ de $\Omega$ ?*

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

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

Defina $X$ : soma das faces nos dois lançamentos do dado.

Função de probabilidade de $X$: $\Rrightarrow$

| $x$ | 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,

$$P(X < 6) = P(X=5) + P(X=4) + P(X=3) + P(X=2)$$

$$= \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36}$$

$$= \frac{10}{36} = 0,278$$

Podemos considerar outras variáveis aleatórias definidas no mesmo espaço amostral.

$Y$: valor máximo obtido entre 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 |

$\Leftarrow$

$Z$: diferença entre as faces 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 |

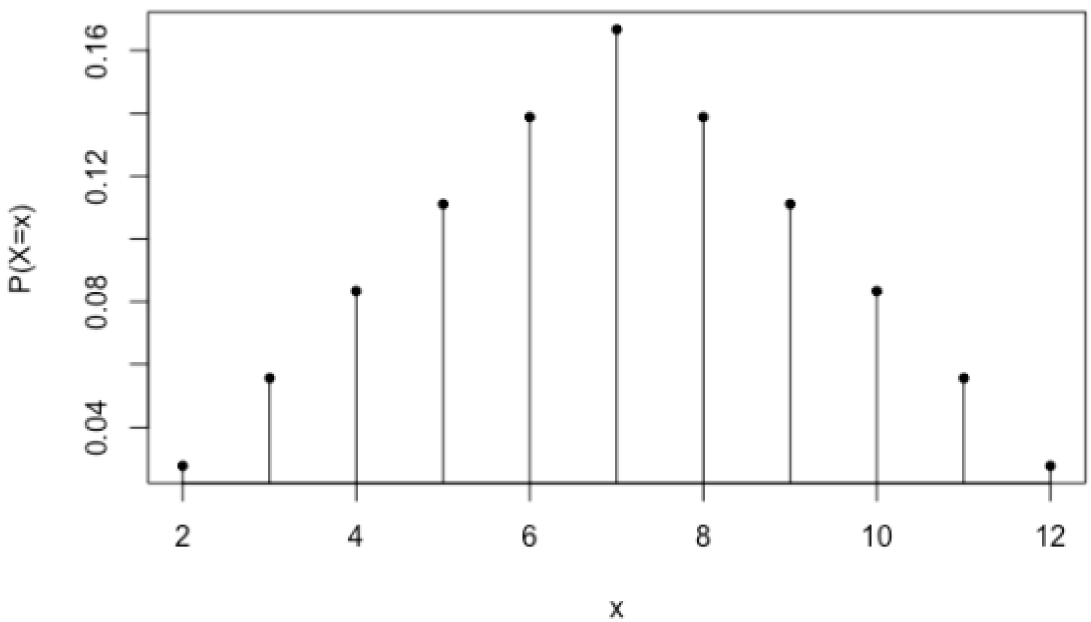
# VALOR ESPERADO E VARIÂNCIA

*Qual é o valor médio da soma das faces ($X$) 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 amostrais igualmente prováveis

| $x$ | $P(X = x)$ |
|---|---|
| 2 | 1/36=0,028 |
| 3 | 2/36=0,056 |
| 4 | 3/36=0,083 |
| 5 | 4/36=0,111 |
| 6 | 5/36=0,139 |
| 7 | 6/36=0,167 |
| 8 | 5/36=0,139 |
| 9 | 4/36=0,111 |
| 10 | 3/36=0,083 |
| 11 | 2/36=0,056 |
| 12 | 1/36=0,028 |



# VALOR ESPERADO E VARIÂNCIA

Valor Esperado ("média"): Dada uma v.a. $X$, assumindo os valores $x_1$, $x_2$, ..., $x_n$, chamamos de **valor médio**, ou **valor esperado**, ou **esperança matemática** da distribuicão de $X$ o valor

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

Notação: $\mu = E(X)$

No exemplo, para média de $X$ (soma das faces), temos:

$$E(X) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + \cdots + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} = \frac{252}{36} = 7$$

ou seja, em média, a soma das faces 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 $x_1, x_2, \ldots, x_n$, então

$$\begin{aligned}\mathrm{Var}(X) &= E\{[X - E(X)]^2\} \\ &= \sum_{i=1}^{n} [x_i - E(X)]^2 \times P(X = x_i)\end{aligned}$$

Desenvolvendo a fórmula acima, e lembrando que $E(X) = \mu$, obtemos a seguinte **fórmula alternativa**

$$\begin{aligned}\mathrm{Var}(X) &= E(X^2) - [E(X)]^2 = E(X^2) - \mu^2 \\ &= \sum_{i=1}^{n} x_i^2 \cdot P(X = x_i) - \mu^2\end{aligned}$$

A notação usual de variância é

$$\mathrm{Var}(X) = \sigma^2$$

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

$$\mathrm{DP}(X) = \sqrt{\mathrm{Var}(X)}$$

Notação: $\mathrm{DP}(X) = \sigma$

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 |

$$\mathrm{Var}(X) = (2-7)^2 \frac{1}{36} + (3-7)^2 \frac{2}{36} + \cdots + (11-7)^2 \frac{2}{36} + (12-7)^2 \frac{1}{36}$$

$$= \frac{210}{36} = 5{,}83.$$

Podemos também calcular pela fórmula alternativa

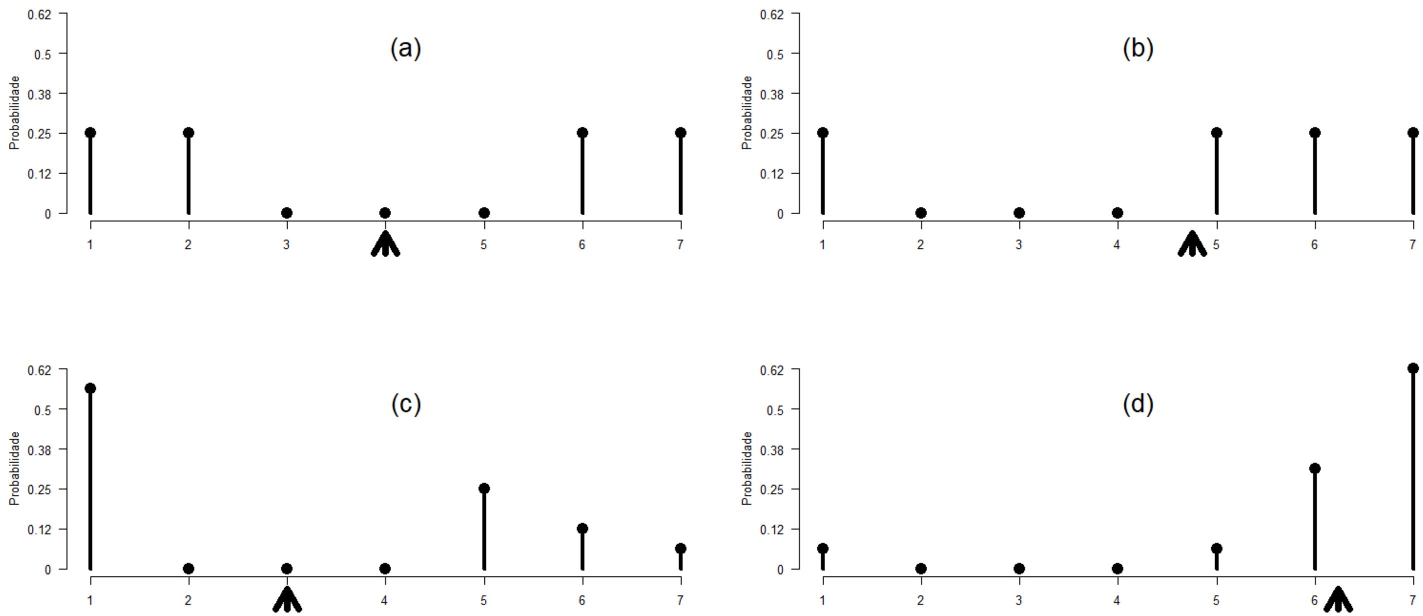
$$\mathrm{E}(X^2) = 2^2 \frac{1}{36} + 3^2 \frac{2}{36} + \cdots + 11^2 \frac{2}{36} + 12^2 \frac{1}{36}$$

$$= \frac{1974}{36} = 54{,}83$$

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

Suponha uma característica que pode assumir os valores:
1, 2, 3, 4, 5, 6, e 7, porém com diferentes funções de probabilidades dependendo o caso.

| P(X=x) | P(Y=y) | P(W=w) | P(Z=z) |
|---|---|---|---|
| 0,25 | 0,25 | 0,56 | 0,06 |
| 0,25 | 0,00 | 0,00 | 0,00 |
| 0,00 | 0,00 | 0,00 | 0,00 |
| 0,00 | 0,00 | 0,00 | 0,00 |
| 0,00 | 0,25 | 0,25 | 0,06 |
| 0,25 | 0,25 | 0,13 | 0,29 |
| 0,25 | 0,25 | 0,06 | 0,59 |
| E(X)=4 | E(Y)=4,75 | E(W)=3 | E(Y)=6,24 |
| Var(X)=3,25 | Var(Y)=3,52 | Var(W)=2,25 | Var(Z)=1,61 |




**Fonte:http://recologia.com.br/2017/05/valor-esperado-expected-value-a-famosa-media/**

# Propriedades:

1) Se $P(X = a) = 1$, então

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

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

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

e

$$Var(Y) = Var(aX + b) = a^2\, Var(X).$$

# - 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 é submetido a um tratamento: o tratamento é eficaz ou não;
- 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 *Ensaios de Bernoulli* e originam uma variável aleatória 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 $p$, $0 < p < 1$.

"***$X \sim$ Bernoulli ($p$)***" indica uma v.a. com ***distribuição de Bernoulli*** com parâmetro $p$, isto é,

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

e sua função de probabilidade pode ser representada pela tabela

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

Segue que

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

→ 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,

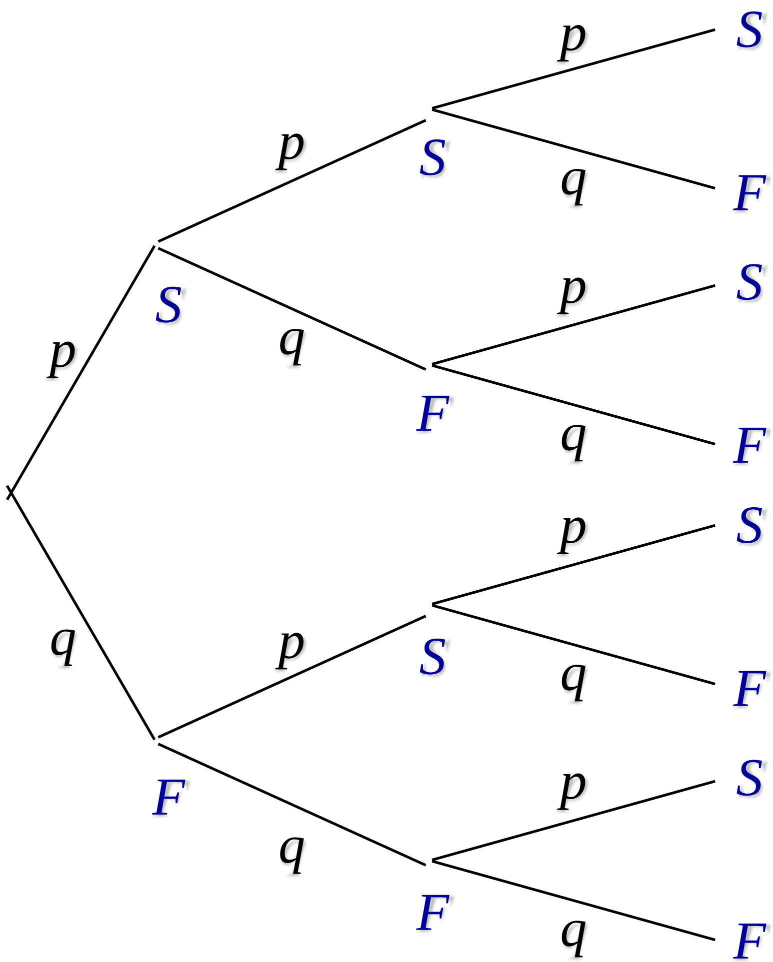
***S***: "sucesso", ocorrer face 5;
***F***: "fracasso", não ocorrer face 5.

É fácil ver que $\boldsymbol{p} = P(\text{sucesso}) = 1/6$ e

$$\boldsymbol{q} = 1 - \boldsymbol{p} = P(\text{fracasso}) = 5/6$$

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

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




| $\Omega$ | Prob | $X$ |
|---|---|---|
| $(SSS)$ | $p^3$ | 3 |
| $(SSF)$ | $p^2q$ | 2 |
| $(SFS)$ | $p^2q$ | 2 |
| $(SFF)$ | $pq^2$ | 1 |
| $(FSS)$ | $p^2q$ | 2 |
| $(FSF)$ | $pq^2$ | 1 |
| $(FFS)$ | $pq^2$ | 1 |
| $(FFF)$ | $q^3$ | 0 |

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

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

Assim, a probabilidade de obter a face 5 duas vezes é

$P(X=2) = 0,0694$

## 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

$$\boxed{P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{para } k = 0, 1, \ \ldots \ , n}$$

Notação: $X \sim b(n; p)$.

# Resultado:

Se $X \sim b(n;\, p)$, então

$$\text{valor esperado: } \mu = \mathrm{E}(X) = n\,p$$

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

# 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?***

$X$: nº. de questões que o aluno acertará

X ϵ {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,12}

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

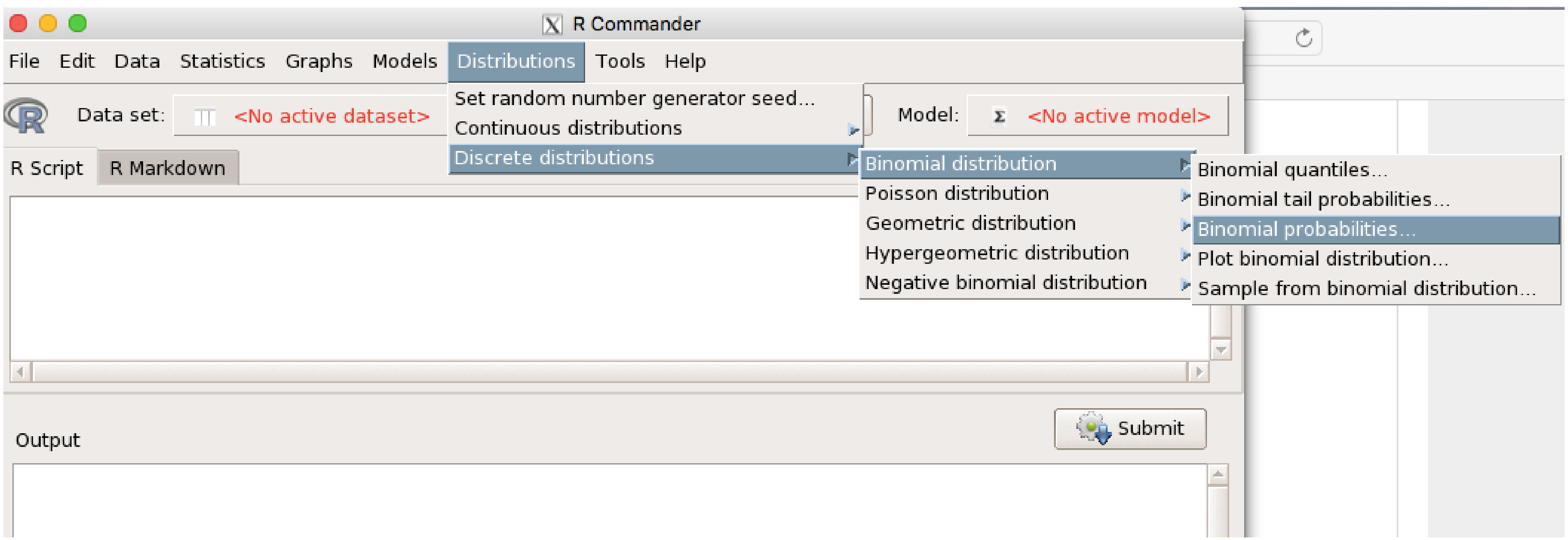
$$\boxed{P(X = x) = \binom{12}{x} 0{,}25^{x} (1 - 0{,}25)^{12-x}}$$

$\Downarrow$

Uso do R para os cálculos!

Para obter a distribuição de probabilidades de uma binomial no R Commander siga o menu

*Distribuições -> Distribuições Discretas -> Distribuição Binomial -> Probabilidades da binomial*



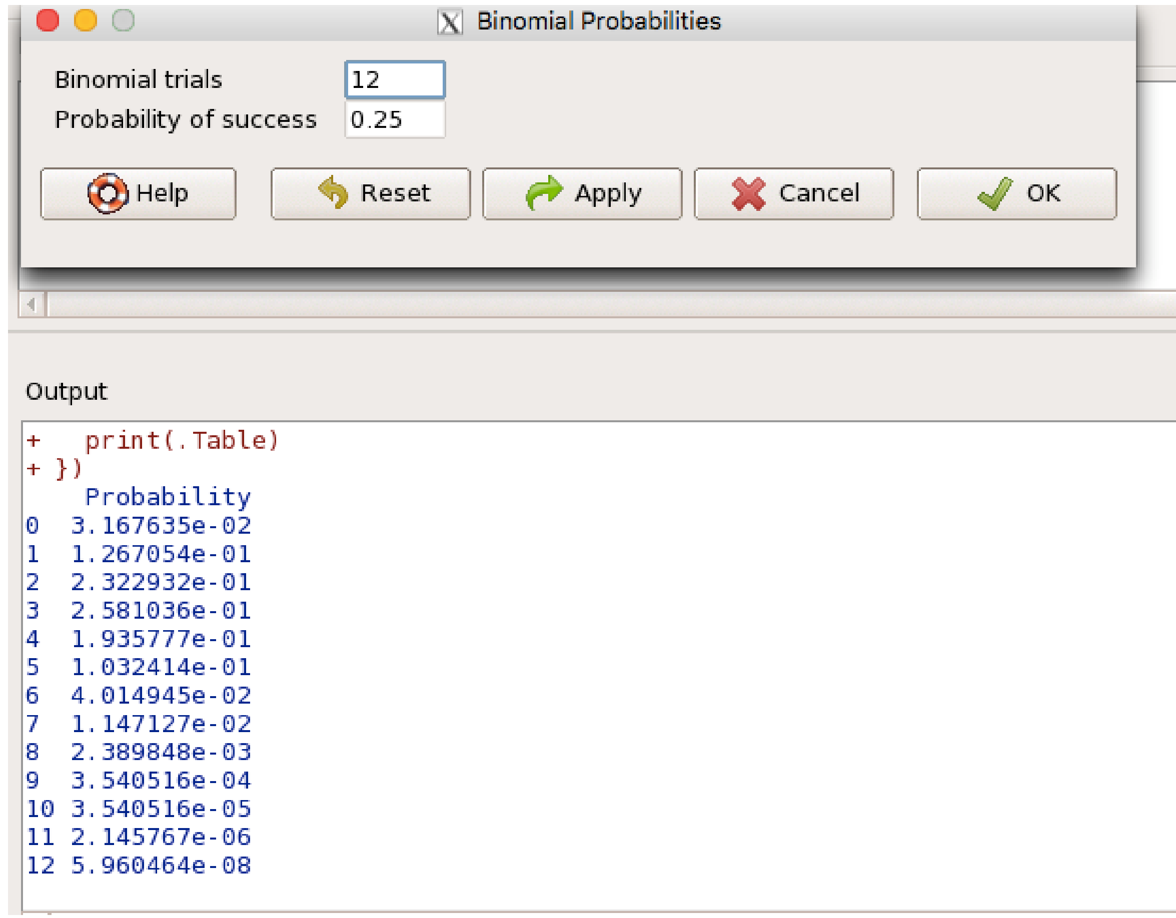
No exemplo, a variável aleatória

$X$: nº. de questões que o aluno acertará

$X \sim b(n = 12; p = 0{,}25)$

Observação:

- Em português usa-se "," para decimal.
- No R (em inglês ou em português) usa-se sempre "." para decimal.



```
> .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, a probabilidade de acertar pelo menos 6 questões é

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

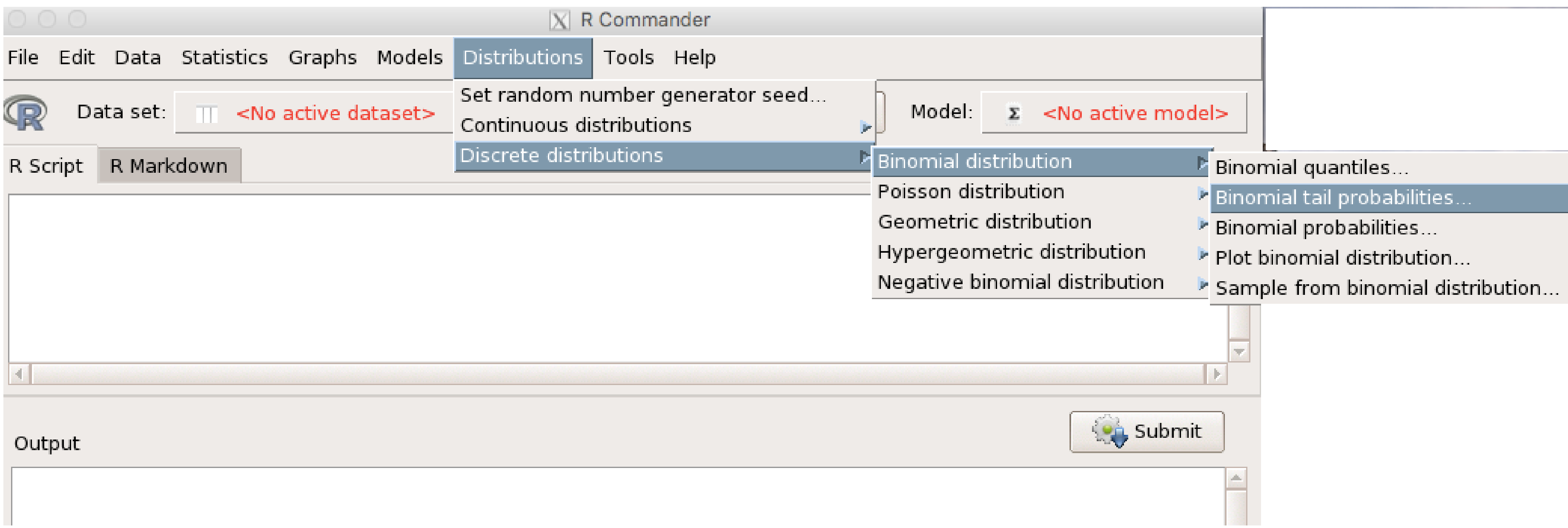
Além disso,

$$E(X) = n.p = 12\ (0{,}25) = 3.$$

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

Podemos também calcular probabilidades **caudais** através de

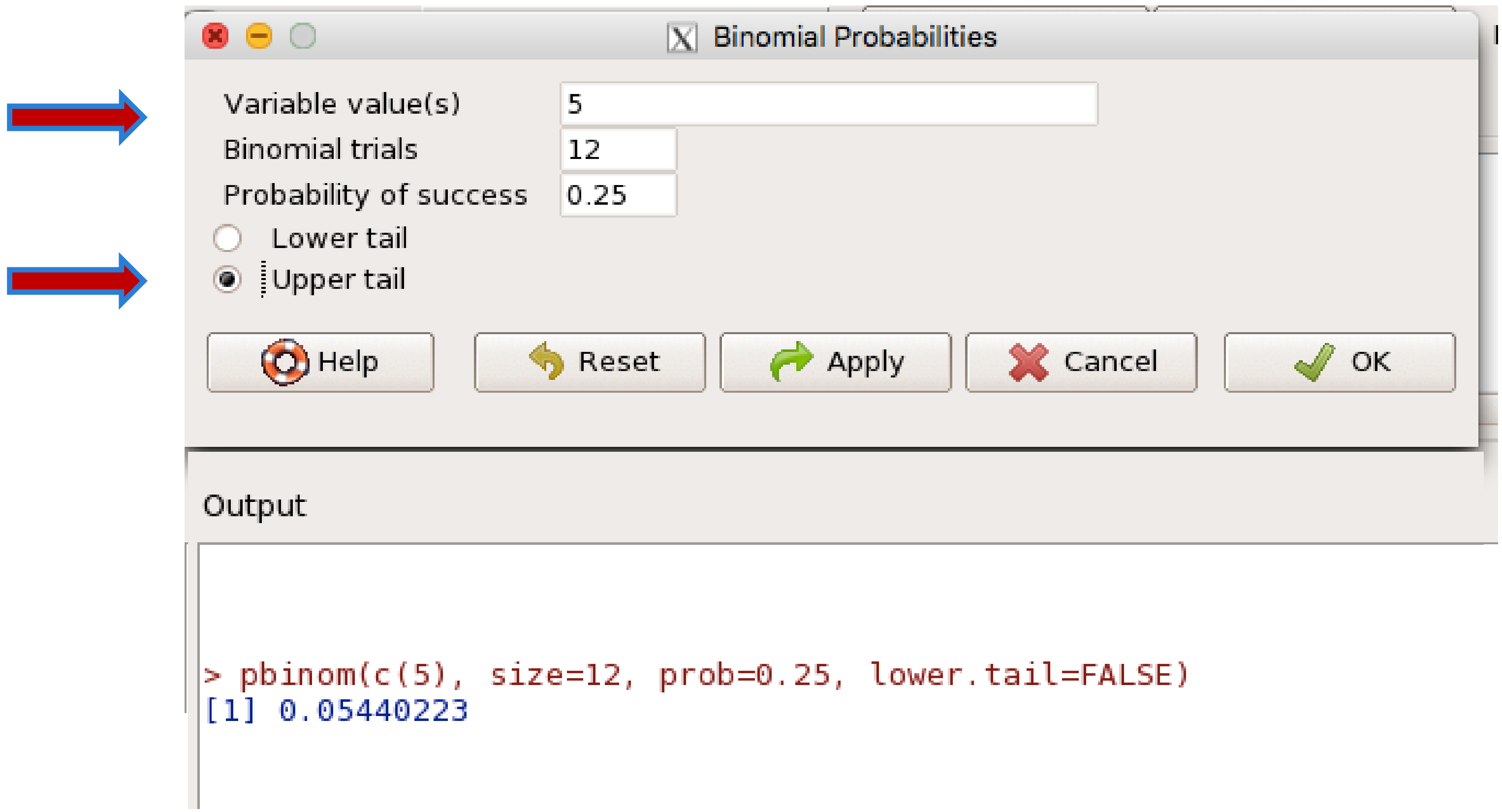
*Distribuições -> Distribuições Discretas -> Distribuição Binomial -> Probabilidades das caudas da binomial*



A probabilidade da cauda **inferior** a $b$ é $P(X \leq b)$.

A probabilidade da cauda **superior** a $b$ é $P(X > b)$.

Portanto, no exemplo, $P(X \geq 6) = P(X > 5)$




Portanto, $P(X \geq 6) = 0{,}0544$