Probabilidades e Aplicações

CC e MAT

2020/21

TP+PL

Emilia Athayde

mefqa@math.uminho.pt

Simulação, LGN e método de Monte Carlo

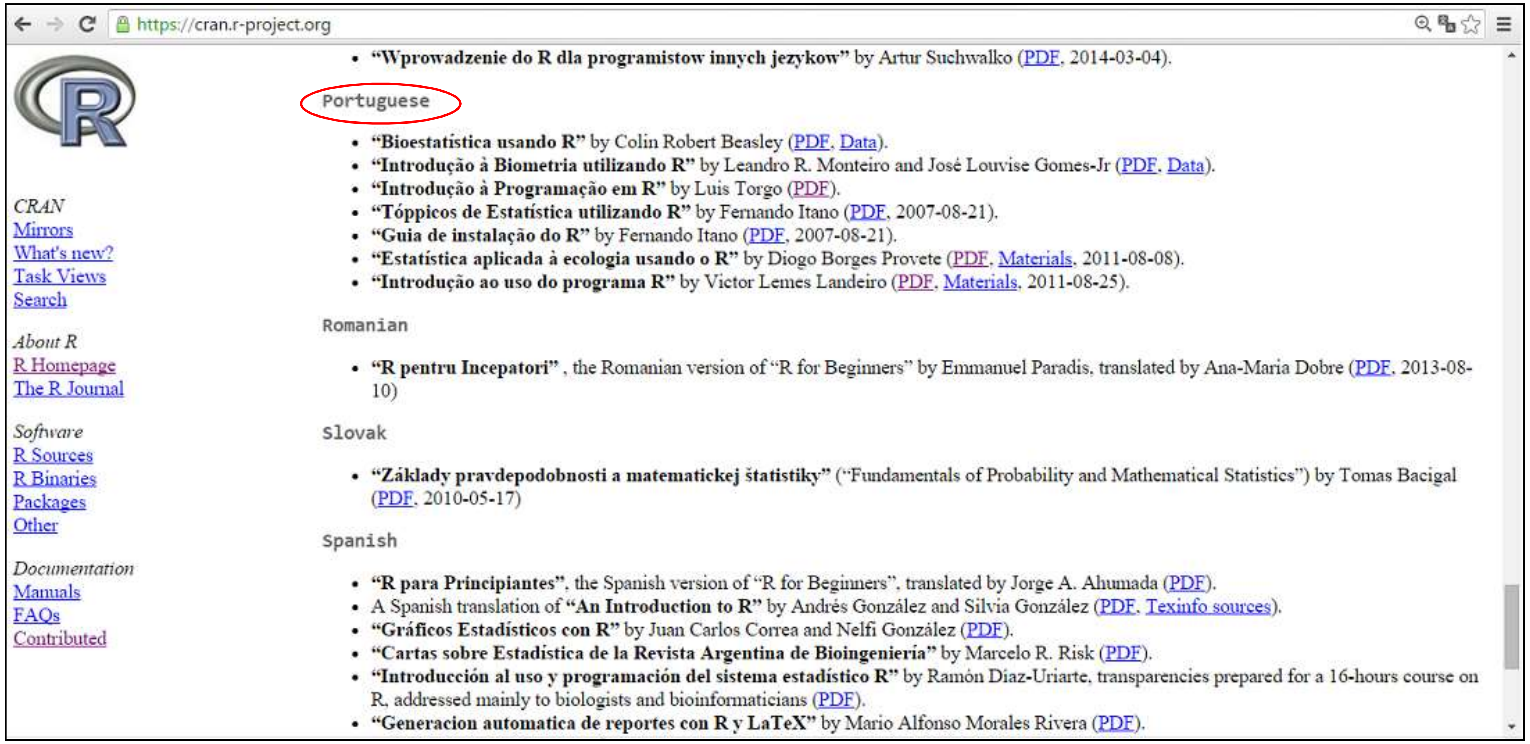
(no sistema R)

sistema computacional R https://cran.r-project.org/

Instalação (menu *Software*) para Linux / Mac OS X / Windows:

R Binaries > Windows > base > Download R 4.0.2 for Windows

Documentação (menu *Documentation*) em português, etc.: Contributed



https://cran.r-project.org

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed

- **"Wprowadzenie do R dla programistow innych jezykow"** by Artur Suchwalko (PDF, 2014-03-04).

Portuguese

- **"Bioestatística usando R"** by Colin Robert Beasley (PDF, Data).
- **"Introdução à Biometria utilizando R"** by Leandro R. Monteiro and José Louvise Gomes-Jr (PDF, Data).
- **"Introdução à Programação em R"** by Luis Torgo (PDF).
- **"Tóppicos de Estatística utilizando R"** by Fernando Itano (PDF, 2007-08-21).
- **"Guia de instalação do R"** by Fernando Itano (PDF, 2007-08-21).
- **"Estatística aplicada à ecologia usando o R"** by Diogo Borges Provete (PDF, Materials, 2011-08-08).
- **"Introdução ao uso do programa R"** by Victor Lemes Landeiro (PDF, Materials, 2011-08-25).

Romanian

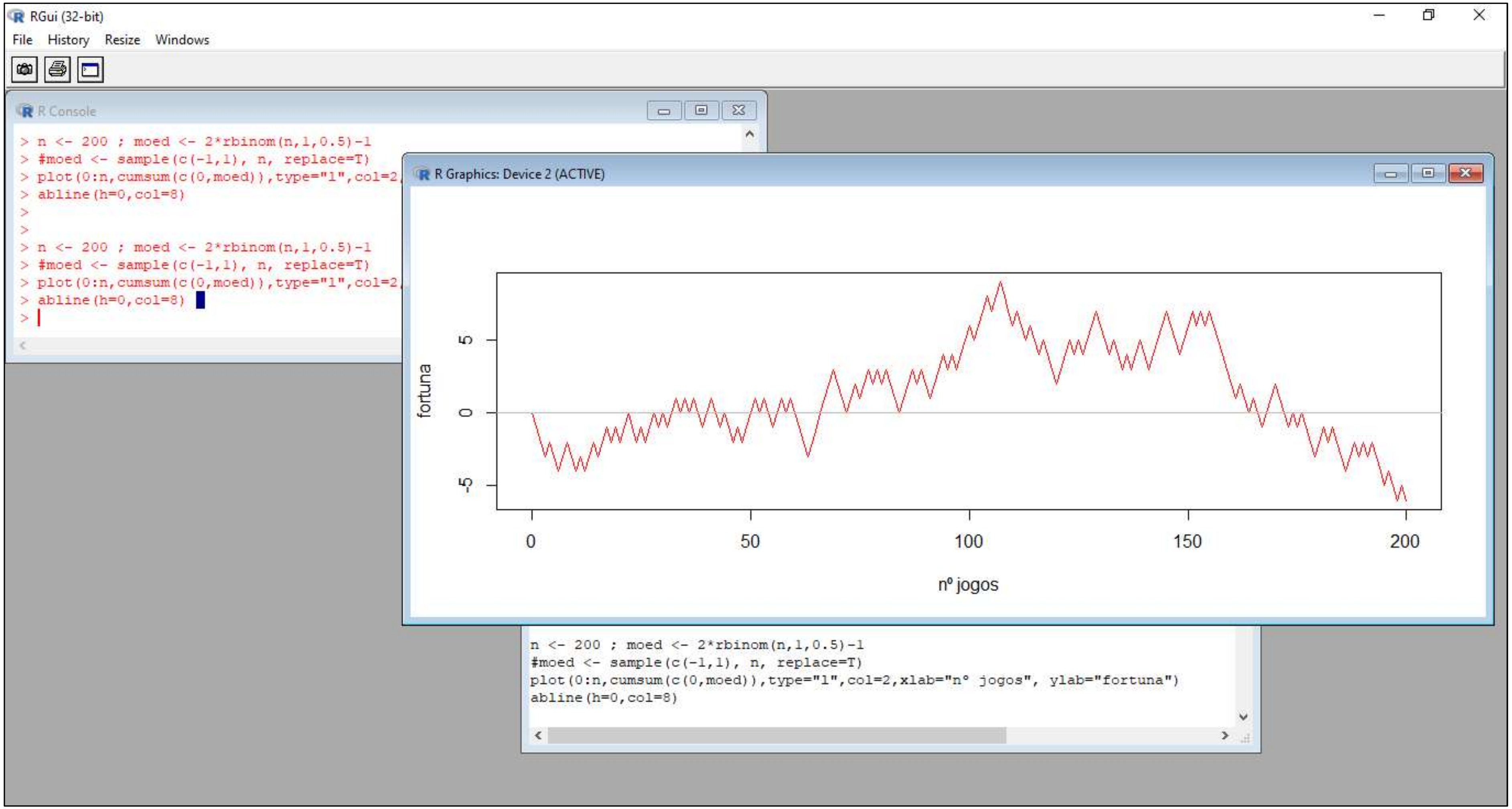
- **"R pentru Incepatori"** , the Romanian version of "R for Beginners" by Emmanuel Paradis, translated by Ana-Maria Dobre (PDF, 2013-08-10)

Slovak

- **"Základy pravdepodobnosti a matematickej štatistiky"** ("Fundamentals of Probability and Mathematical Statistics") by Tomas Bacigal (PDF, 2010-05-17)

Spanish

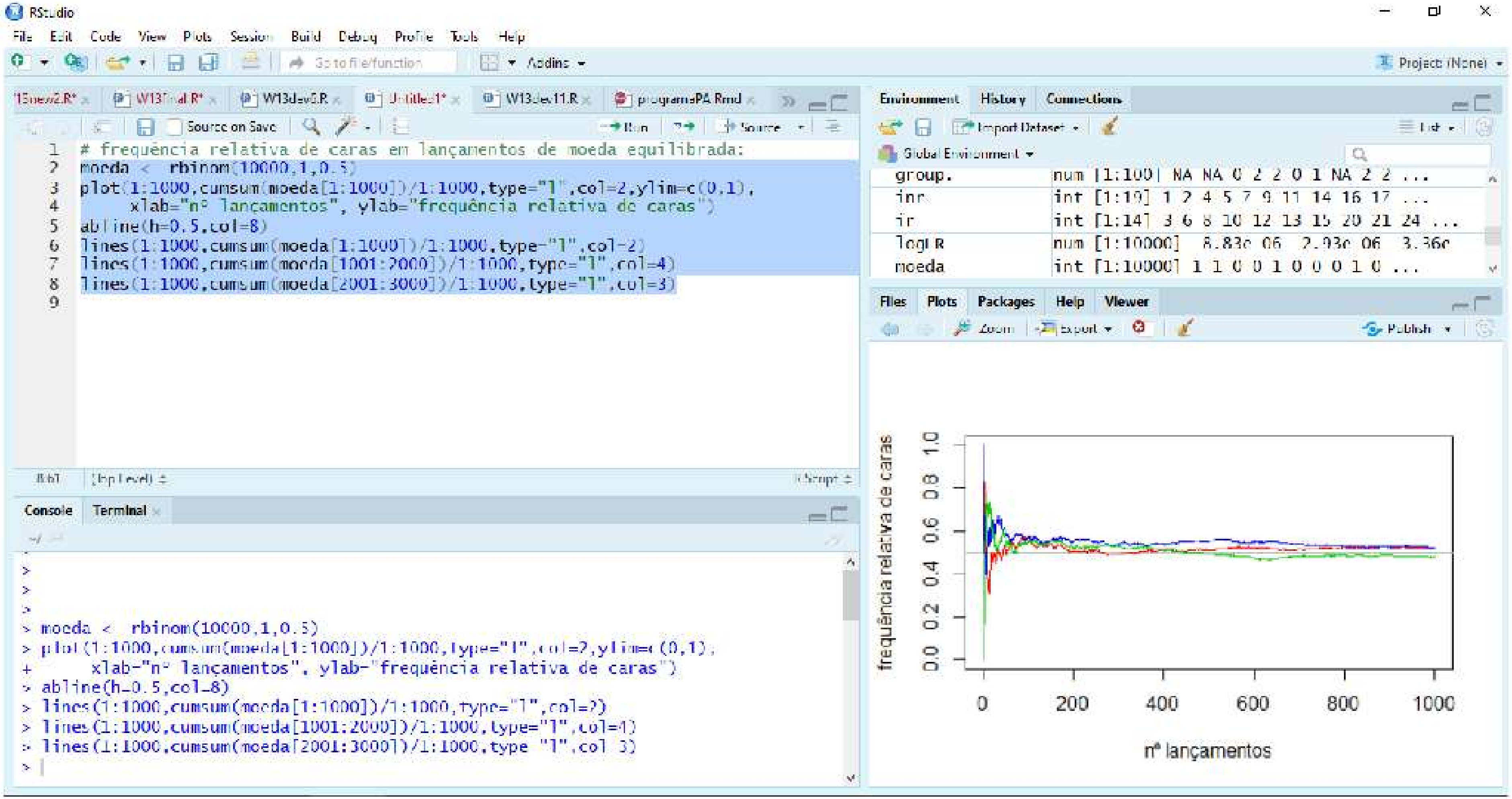
- **"R para Principiantes"**, the Spanish version of "R for Beginners", translated by Jorge A. Ahumada (PDF).
- A Spanish translation of **"An Introduction to R"** by Andrés González and Silvia González (PDF, Texinfo sources).
- **"Gráficos Estadísticos con R"** by Juan Carlos Correa and Nelfi González (PDF).
- **"Cartas sobre Estadística de la Revista Argentina de Bioingeniería"** by Marcelo R. Risk (PDF).
- **"Introducción al uso y programación del sistema estadístico R"** by Ramón Díaz-Uriarte, transparencies prepared for a 16-hours course on R, addressed mainly to biologists and bioinformaticians (PDF).
- **"Generacion automatica de reportes con R y LaTeX"** by Mario Alfonso Morales Rivera (PDF).


RGui (32-bit)
File History Resize Windows
R Console
> n <- 200 ; moed <- 2*rbinom(n,1,0.5)-1
> #moed <- sample(c(-1,1), n, replace=T)
> plot(0:n,cumsum(c(0,moed)),type="l",col=2,
> abline(h=0,col=8)
>
>
> n <- 200 ; moed <- 2*rbinom(n,1,0.5)-1
> #moed <- sample(c(-1,1), n, replace=T)
> plot(0:n,cumsum(c(0,moed)),type="l",col=2,
> abline(h=0,col=8)
>
R Graphics: Device 2 (ACTIVE)
fortuna
-5
0
5
0
50
100
150
200
nº jogos
n <- 200 ; moed <- 2*rbinom(n,1,0.5)-1
#moed <- sample(c(-1,1), n, replace=T)
plot(0:n,cumsum(c(0,moed)),type="l",col=2,xlab="nº jogos", ylab="fortuna")
abline(h=0,col=8)

RStudio

https://rstudio.com/

interface mais apelativa, com mais funcionalidades



R / RStudio

Inserção/importação de dados:

`c`, `scan`, `read.table`, `seq`, `rep`, `...`

`file > Import Dataset > ...`

Manipulação de dados. Seleção condicional.

Objetos básicos: **`vector`**, **`matrix`**, `data.frame`, `list`

Tipos de dados: numéricos, factores, lógicos, ...

Simulação de dados: **`sample`**, **`runif`**, **`rbinom`**, **`rnorm`**, **`r...`**

Gráficos: funções *high-level*, *low-level*, *interactive*.

`plot`, **`lines`**, **`abline`**, **`points`**, **`curve`**, **`...`**

Scripts. `File > New script` / `File > New file > R script`

Outras funções: **`function`**, `help`, `search()`, `...`

```
c(4,9,-7, 9)
[1]  4  9 -7  9
x <- 2:7
x
[1] 2 3 4 5 6 7
( x <- 3:8 )
[1] 3 4 5 6 7 8
x + 1:2
[1] 4  6  6  8  8 10
(y <- x^2)
[1]  9 16 25 36 49 64
y[1:3]
[1]  9 16 25
y>40
[1] FALSE FALSE FALSE
[4] FALSE TRUE  TRUE
y[y>40]
[1] 49 64
which(y>40)
[1] 5 6
```

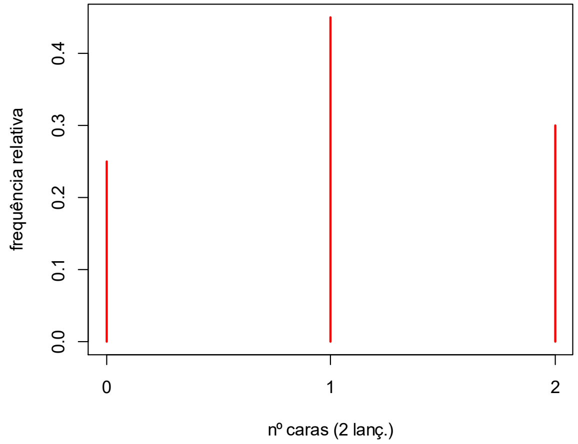
```
length(x)
[1]  6
class(x)
[1] "integer"
c("mat",3:2)
[1] "mat" "3"   "2"
class(c("mat",3:2))
[1] "character"
prod(1:6)
[1] 720
factorial(6)
[1] 720
choose(10,4)
[1] 210
choose(10,0:4)
[1]   1  10  45 120 210
(m <- matrix(1:6,nr=2))
     [,1] [,2] [,3]
[1,]    1    3    5
[2,]    2    4    6
```

```
apply(m,2,sum)
[1]  3  7 11
m[2,3] <- 10; m
 [,1] [,2] [,3]
[1,]    1    3    5
[2,]    2    4   10
m[2,]
[1]  2  4 10
dim(m)
[1] 2 3
# produto termo a termo:
m*m
     [,1] [,2] [,3]
[1,]    1    9   25
[2,]    4   16  100
# produto matricial:
m %*% t(m)
 [,1] [,2]
[1,]   35   64
[2,]   64  120
```

```
sample(x, size, replace = FALSE, prob = ...)
```

simula `size` extrações (por defeito, sem reposição; c.c. `replace=TRUE`) do universo `x` (vector dim. k) com probabilidades especificadas em `prob` (vector dim. k)

```
# totoloto (6 extrações sem reposição de urna com bolas de 1 a 49; vector 1:49):
(toto <- sample(1:49,6))
 [1]  8 46  6 20 24 48
# simulação de 10 lançamentos de moeda-½ com faces C e E:
(caras <- sample(c("C","E"),10,replace=T))
 [1] "E" "C" "E" "C" "E" "C" "E" "C" "C" "C"
# simulação de 20 lançamentos de moeda-½  (1-Cara; 0-Euro)
( moeda <- sample(0:1, 20, replace=T) )
[1] 1 0 1 0 0 0 0 0 0 1 1 1 0 1 1 1 0 0 1 0
# 20 simulações de nº caras em 2 lançamentos de moeda-½ :
( moed.2 <- sample(0:2, 20, T, prob=c(1,2,1)/4) )
 [1] 1 0 2 1 0 1 0 2 1 1 0 2 1 1 1 1 2 1 1 1
table(moed.2)
 0  1  2
 5  9  6
plot(table(moed.2)/20, col=2, ylab="frequência relativa",xlab="nº caras (2 lanç.)")
```



`rbinom(n,size,p)` – gera `n` réplicas do nº caras em `size` lançam. de moeda-`p`
`dbinom(x,size,p)` – calcula P(nº caras = `x`), em `size` lançamentos de moeda-`p`

```
# simular 20 lançam de moeda-0.7 (viciada a favor de "cara"), usando sample:
sample(0:1, 20, replace=T, p = c(0.3,0.7))
[1] 0 1 0 1 1 1 1 1 1 0 1 0 1 0 0 0 0 1 1 1
# o mesmo (nova simulação), usando rbinom:
rbinom(20, 1, 0.7)
[1] 1 1 1 1 0 1 0 1 1 1 1 1 1 0 1 1 0 1 1 0
# 20 simulações do n° de caras em 5 lançamentos de moeda:
rbinom(20, 5, 0.7)
[1] 2 3 3 4 3 4 4 2 3 0 4 5 3 3 2 2 3 4 3 3

# probab de "n° caras em 10 lançamentos ser 0,1,2,...,10"
dbinom(0:10,10,1/2)
# simul do n° caras acumulado ao fim dos sucessivos lançamentos:
( moed <- rbinom(20, 1, 0.5) )
[1] 1 0 1 1 1 1 1 0 0 0 1 0 1 1 1 1 1 1 0 0
cumsum(moed)
[1]  1  1  2  3  4  5  6  6  6  6  7  7  8  9 10 11 12 13 13 13
```

Cap.10 Simulação, LGN e método de Monte Carlo

- Simulação – números “pseudo-aleatórios” (NPA)

- Lei dos grandes números (LGN) –
  - o “teorema de ouro” de J. Bernoulli (1689)

- Método de Monte Carlo (Ulam & von Neuman, anos 40)

NPA

- Gerados/simulados de forma determinística (recursiva): métodos "middle square" 1949, congruenciais, "Mersenne twister" 1998)
- Comportam-se, na prática, como se fossem aleatórios
- A partir de um gerador de NPAs no intervalo (0,1) é possível gerar números aleatórios com outra distribuição (vários métodos)
- Onde? C++, EXCEL, Maple, Mathematica, Matlab, Octave, Python, **R**, SPSS, …

LGN

A LGN (Bernoulli) estabelece que a frequência relativa de um acontecimento *A* em *n* repetições (independentes) de uma experiência aleatória converge (em certo sentido…) para a probabilidade de *A*

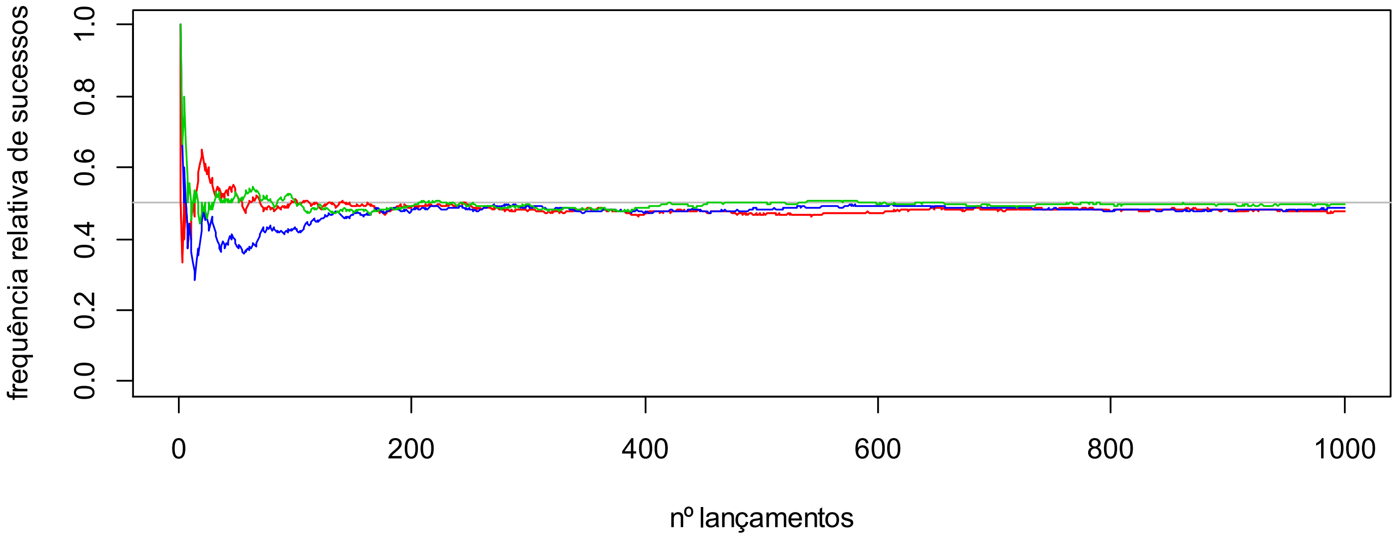
MMC

Pela LGN, a probabilidade de *A* pode ser "estimada" pela frequência relativa de *A* em *n* repetições da experiência (quanto maior for *n*, melhor <–> maior a "precisão" da estimativa)

Interpretação frequencista de Probabilidade

a probabilidade de um acontecimento A é o limite (quando $n \to \infty$) da frequência relativa de ocorrência de A em n realizações da experiência, que representamos por $f_n(A)$.

Esse limite existe? E caso exista, vai ser independente da sucessão particular de experiências? Veremos bem mais adiante que a LGN é consequência da axiomática... mas para já, vamos simular...



1000 lançamentos de moeda equilibrada; a frequência relativa de caras parece estabilizar à volta de ½
(3 trajetórias simuladas)

Nota: em n lançamentos de moeda equilibrada (n par), a LGN não implica que P("nº caras = nº de coroas") $\to$ 1 quando $n \to \infty$. Para que valor converge então? Calcule ou use o R para descobrir o resultado...

Exercício 1: (i) calcular (ii) usar o R – para obter o limite da probabilidade, p_n , de em $2n$ lançamentos (moeda equilibrada), o nº caras ser n (igual ao nº coroas).

Resolução:

(i) calcular a probabilidade (fácil) → $p_n = \binom{2n}{n} 2^{-2n}$

e o seu limite quando $n \to \infty$?

(hum… com a fórmula de Stirling é fácil) → $\lim p_n = \ldots = \lim \frac{2}{n\sqrt{2\pi}} = 0$

E sem esta fórmula?

```
n <- 1:100; plot(n,choose(2*n,n)/2^(2*n))
# arcaico... No Mathematica é de caras!
```

Limit[Binomial[2n,n]/2^{2n}, n->Infinity] dá zero!

Fórmula de Stirling:

$$\boxed{n! \sim \sqrt{2\pi}\, n^{n+1}\, e^{-n}}$$

significa que

$$\frac{n!}{\sqrt{2\pi}\, n^{n+1}\, e^{-n}} \xrightarrow[n\to\infty]{} 1$$

(ii) usar o R, para n =50, 500, 5×10^3, 5×10^4, … (fácil)

```
# usando dbinom(x,n,p)
# prob de 50 caras em 100, com p=0.5:
dbinom(50,100,0.5)
[1] 0.079589237
round(dbinom( 5*10^(1:7), 10*10^(1:7), 0.5),4)
[1] 0.0796 0.0252 0.0080 0.0025 0.0008 0.0003 0.0001
```

Exercício 1 (cont.): Sendo $p = P(A)$,
a LGN de Bernoulli estabelece que

$$\boxed{\forall_{\varepsilon>0} \quad P(|\, f_n(A) - p \,| \le \varepsilon) \xrightarrow[n\to\infty]{} 1}$$

(esta convergência chama-se convergência em probabilidade; notação: $f_n(A) \xrightarrow{P} p$)

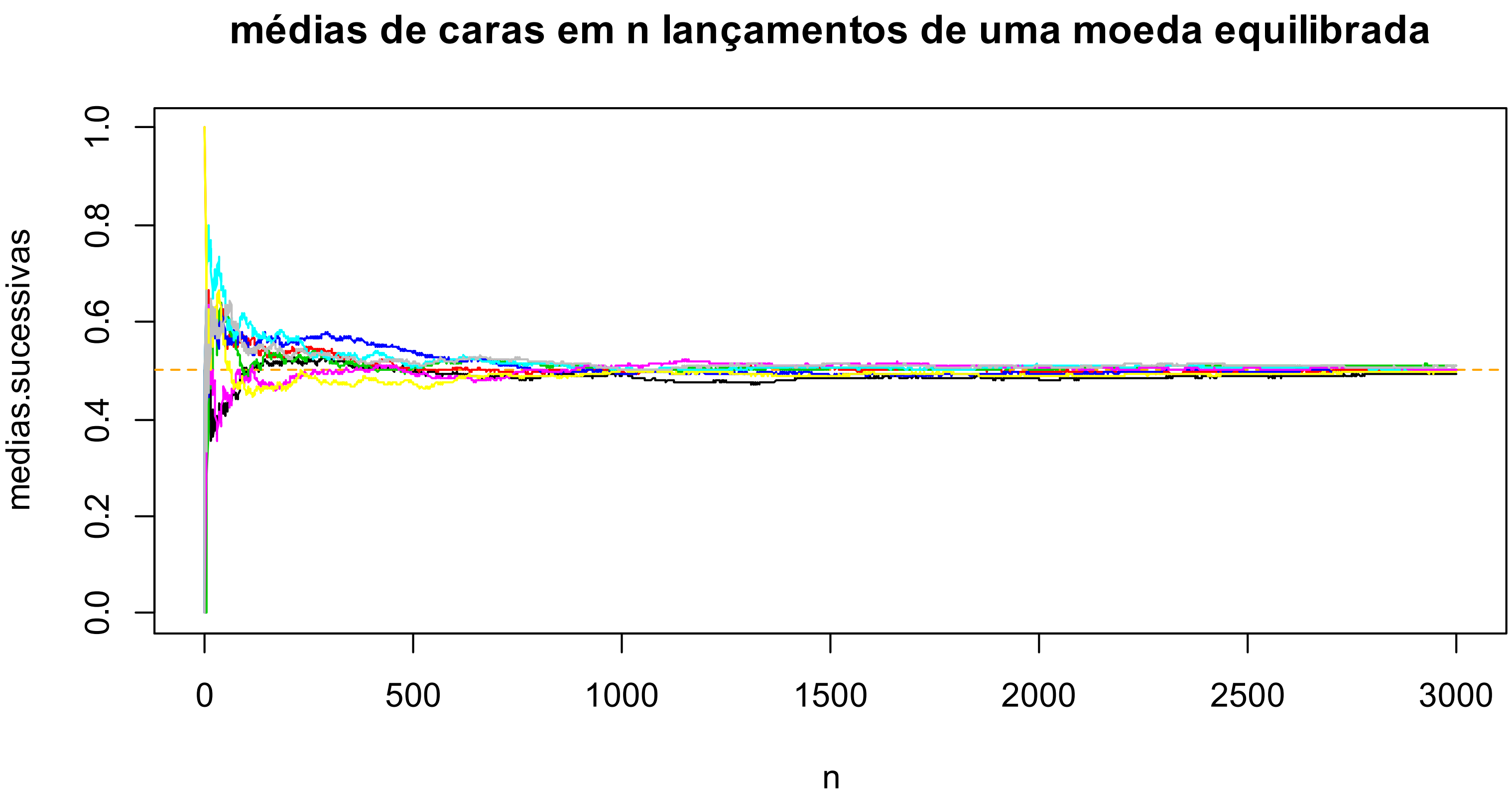
Ou seja, fixando $\varepsilon > 0$, por mais pequeno que seja, tem-se $P(|\, f_n(A) - p \,| \le \varepsilon) \xrightarrow[n\to\infty]{} 1$

No caso da moeda equilibrada, sendo Y o nº de caras nos n lançamentos, $p = ½$ e $\varepsilon = 0.01$, calcule-se então $P(|Y/n - ½\,| \le 0.01)$ ou seja $P(|Y - n/2\,| \le 0.01 \times n)$:

n	Probabilidade a calcular			Resultado
100	$P(\,\lvert Y\text{-}50\rvert \le 1\,)$	=	$P(\,49 \le Y \le 51\,)$	0.2356
1 000	$P(\,\lvert Y\text{-}500\rvert \le 10\,)$	=	$P(\,490 \le Y \le 510\,)$	0.4933
10 000	$P(\,\lvert Y\text{-}5000\rvert \le 100\,)$	=	$P(\,4900 \le Y \le 5100\,)$	0.9556
100 000	$P(\,\lvert Y\text{-}50000\rvert \le 1000\,)$	=	$P(\,49000 \le Y \le 51000\,)$	0.99999999975

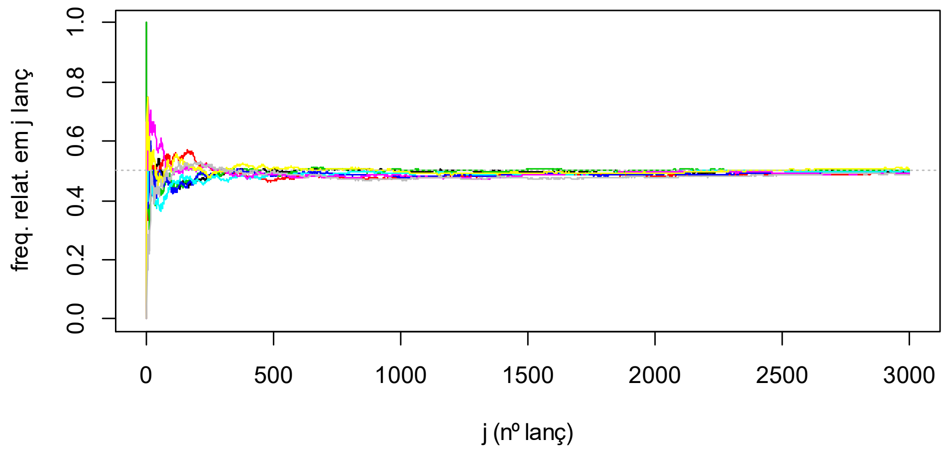
Nem por sombras é a probabilidade de “nºcaras = nºcoroas”

Exercício 2: Represente graficamente 8 trajectórias (cores 1:8) da frequência relativa de caras ao longo de n (de 1 a 3000) lançamentos de uma moeda equilibrada.



Resolução: versão 1) simular $n = 3000$ jogos de moeda-½; calcular a frequência relativa de caras ao fim de j jogos (j=1, 2, …, n); representar graficamente (`plot`) a trajetória correspondente; repetir usando `lines` (low-level) ou ciclo `for` e `lines`; versão 2) simular matriz com $3000{\times}8$ jogos e usar `matplot` para fazer o gráfico.

```
## (1) usando rbinom, for, plot, lines, abline:
moeda <- rbinom(3000,1,0.5)
plot(1:3000,cumsum(moeda)/1:3000,type="l",ylim=c(0,1),xlab="n° lançamento",
        ylab="freq. relativa de caras")
for (i in 2:8)
        {moeda <- rbinom(3000,1,0.5);
        lines(1:3000,cumsum(moeda)/1:3000,col=i)}
# linha horizontal cinzenta, a ponteado
abline(h=0.5,col=8,lty=3)
## (2) usando matriz 3000×8 e matplot:
# simulação em matriz
moeda8 <- matrix(rbinom(3000*8,1,0.5),nc=8)
# nova matriz (colunas com somas acumuladas)
z <- apply(moeda8,2,cumsum)
# gráfico com matplot:
matplot(1:3000,z/1:3000, type="l",col=1:8, lty=1,
         xlab="j (n° lanç)", ylab="freq. relat. em j lanç")
abline(h=0.5,col=8,lty=3)
```



Exercício 3: O problema (clássico) dos aniversários

Dadas n pessoas escolhidas ao acaso, qual o menor n para o qual a probabilidade p_n de haver pelo menos uma coincidência no dia de aniversário (supondo equiprováveis os 365 dias do ano) seja

> 0.5 ?

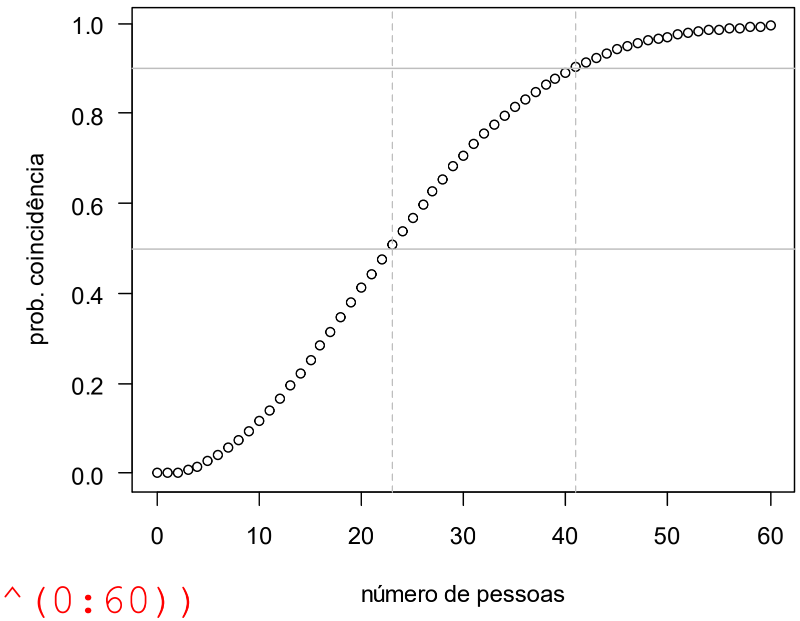
> 0.9 ?

> 0.99 ?

Pela regra de Laplace (usando o acontecimento complementar, "não haver coincidências"), temos

$$p_n = 1 - \frac{365 \times 364 \times \ldots \times (365-n+1)}{365^n} = 1 - \binom{365}{n} \frac{n!}{365^n}$$

Resolução: calcule-se a probabilidade p_n de haver pelo menos uma coincidência no dia de aniversário para obter a resposta à questão; represente-se graficamente em função de n.



$$p_n = 1 - \frac{365 \times 364 \times \ldots \times (365 - n + 1)}{365^n} = 1 - \binom{365}{n} \frac{n!}{365^n}$$

```
## básico:
plot(0:60,1-choose(365,0:60)*factorial(0:60)/365^(0:60))
## usando "function":
coinc <- function(n) 1-choose(365,n)*factorial(n)/365^n
plot(0:60,coinc(0:60),xlab="número de pessoas",ylab="prob. coincidência",las=1)
## para responder à pergunta:
min(which(coinc(1:50)>0.5))       # resposta: 23
min(which(coinc(1:50)>0.9))       # resposta: 41
min(which(coinc(1:60)>0.99))      # resposta: 57
## linhas adicionais:
abline(h=c(0.5,0.9),col=8); abline(v=c(23,41),col=8,lty=2)
```

O passeio aleatório

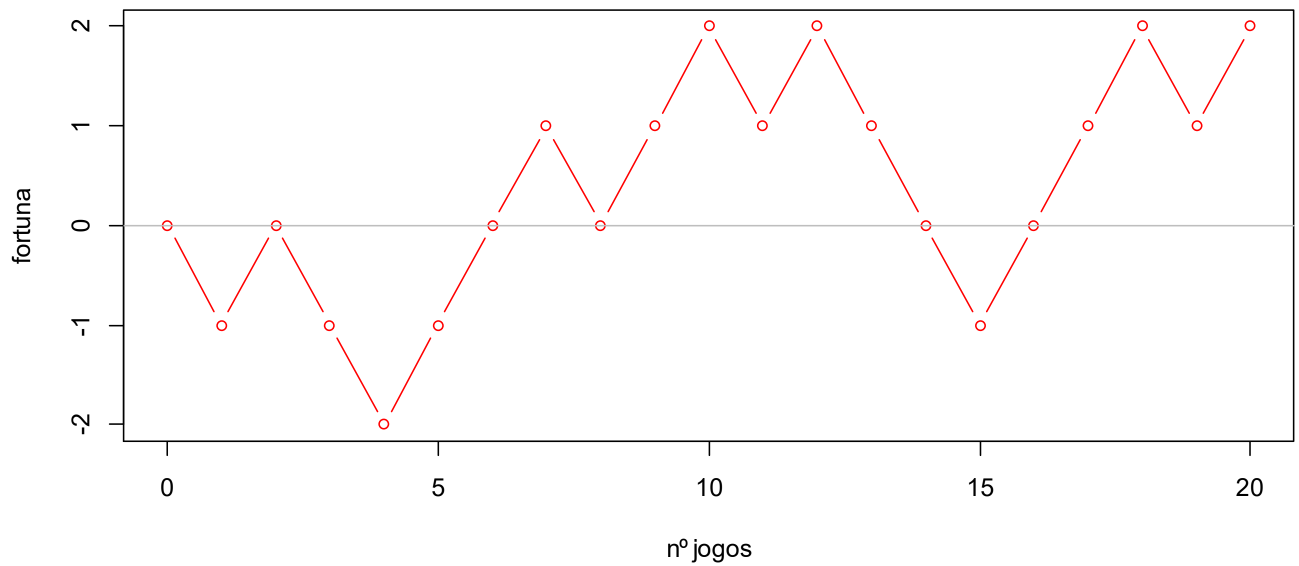
"We shall encounter theoretical conclusions which not only are unexpected but actually come as a shock to intuition and common sense"

"even the simple coin-tossing game leads to paradoxical results that contradict our intuition"

W. Feller, cap.III

Passeio aleatório

Um jogador em cada passo lança uma moeda e ganha ou perde $1€$ conforme sai cara ou coroa. Partindo de uma fortuna inicial $S_0 = 0$, representa-se a fortuna passo a passo por $S_1, S_2, \ldots, S_n, \ldots$ e a trajectória do processo num gráfico, unindo os pontos (i, S_i):



Nesta trajetória com 20 passos há 5 retornos à origem (zero), dos quais 3 são cruzamentos do eixo dos xx (correspondem a mudança de sinal na fortuna).

6 passos iniciais: fortuna ≤ 0,
8 passos seguintes: fortuna ≥ 0,
2 passos seguintes: fortuna ≤ 0,
restantes 4 passos: fortuna ≥ 0
(3 mudanças de sinal)

Note que um retorno à origem (i.e., nº caras = nº coroas) apenas pode ocorrer num instante par.

O passeio aleatório também pode ser visto como o movimento de uma partícula que se desloca no conjunto $\mathbb{Z}$ dos inteiros, transitando em cada instante (passo) para o ponto adjacente ($+1$ ou -1, com probabilidades p e $1-p$, resp.), partindo do 0 (a posição inicial da partícula é $S_0 = 0$). Representa-se a posição da partícula no tempo por $S_1, S_2, S_3, \ldots,$ etc.

O passeio aleatório diz-se simétrico se $p = ½$.

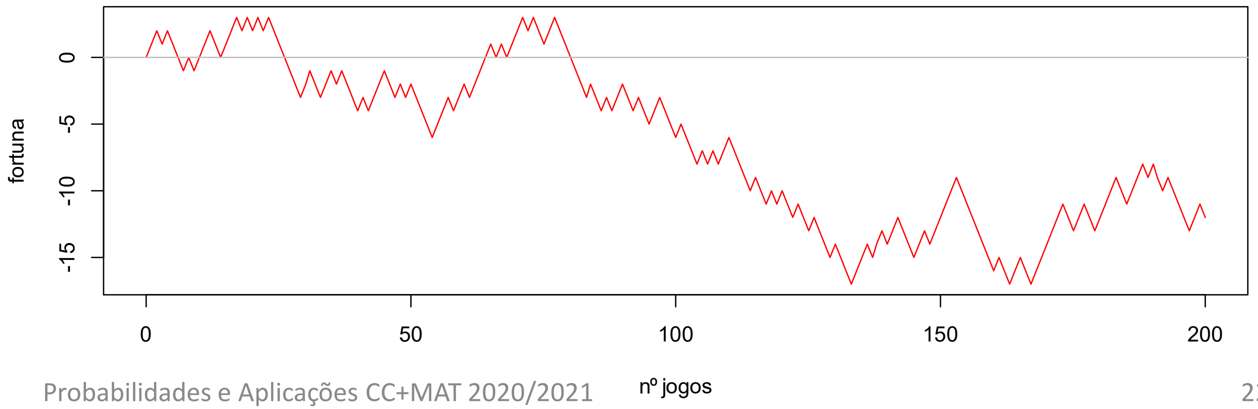
Este passeio aleatório (unidimensional, $d = 1$) generaliza-se a 2 ou 3 dimensões (em $\mathbb{Z}^2$ ou $\mathbb{Z}^3$). Diz-se simétrico se as probabilidades de transição para os pontos adjacentes (4 pontos no caso $d = 2$: para a *frente, trás, direita, esquerda*; e 6 no caso $d = 3$: idem, mais *cima* e *baixo*)

Qual será a probabilidade de retorno à origem (mais tarde ou mais cedo) nos casos $d = 1, 2, 3$?

Exercício 4: (nº4) Simular um passeio aleatório *simétrico* (i.e., moeda equilibrada) em n passos; representar graficamente a trajetória.

Resolução: Simula-se $n = 200$ jogos com ganho -1 ou $+1$ (equiprováveis) em cada jogo (usando `sample` ou `rbinom`) e calcula-se o ganho acumulado ao fim dos jogos sucessivos, $S_1, S_2, \dots, S_n$ (usando `cumsum`)

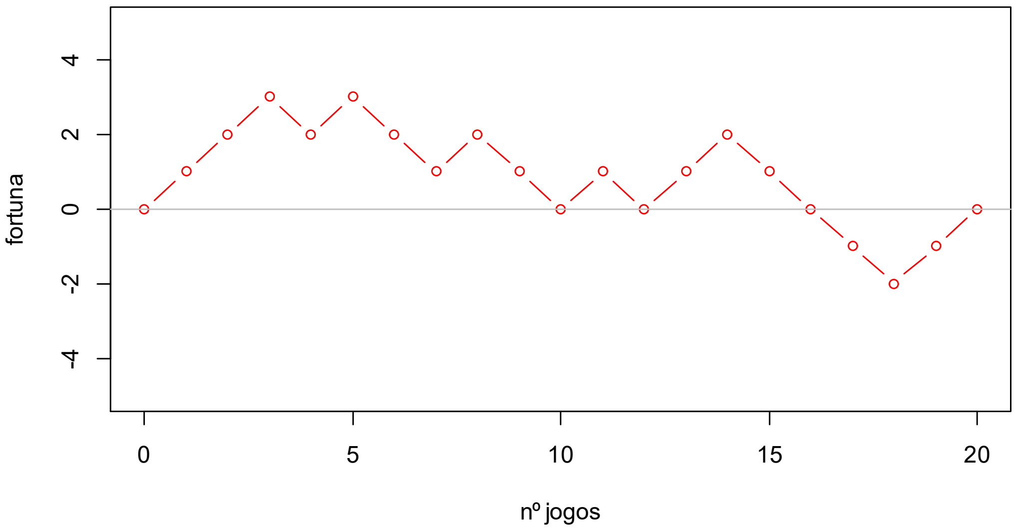
```
## passeio aleatório simétrico (cara:+1, coroa:-1) e trajectória:
n <- 200
moed <- sample(c(-1,1), n, replace=T)
moed <- 2*rbinom(n,1,0.5)-1
plot(0:n,cumsum(c(0,moed)),type="l",col=2,xlab="nº jogos", ylab="fortuna")
abline(h=0, col=8)
```

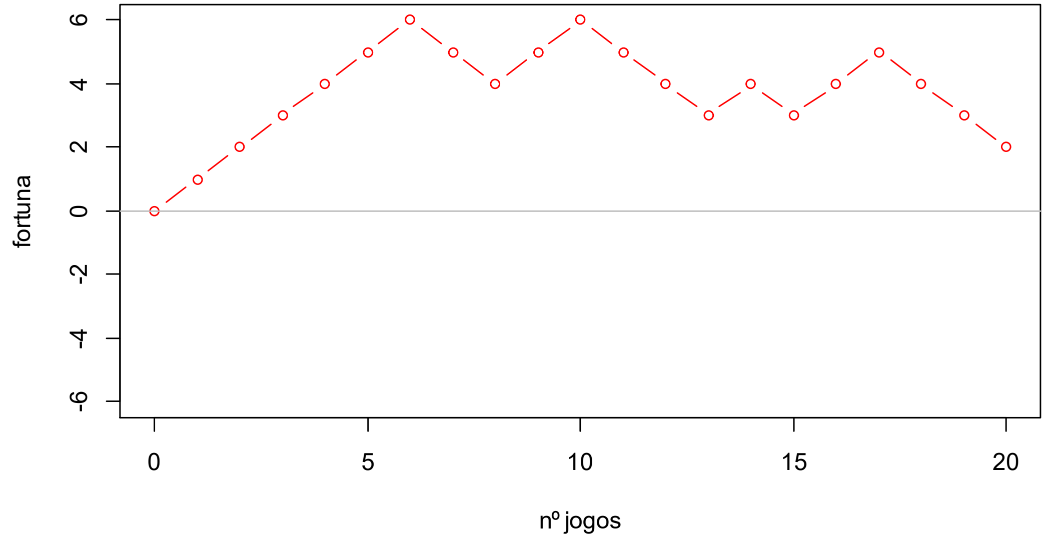


Destes dois acontecimentos, qual será mais provável (para n fixo)?

$S_{2n} = 0$ (retorno à origem no instante $2n$)	ou	$S_1 \neq 0, S_2 \neq 0, \ldots, S_{2n} \neq 0$ (trajetória sempre > 0 ou sempre < 0)

trajetórias com $n = 10$ (i.e., com 20 passos):

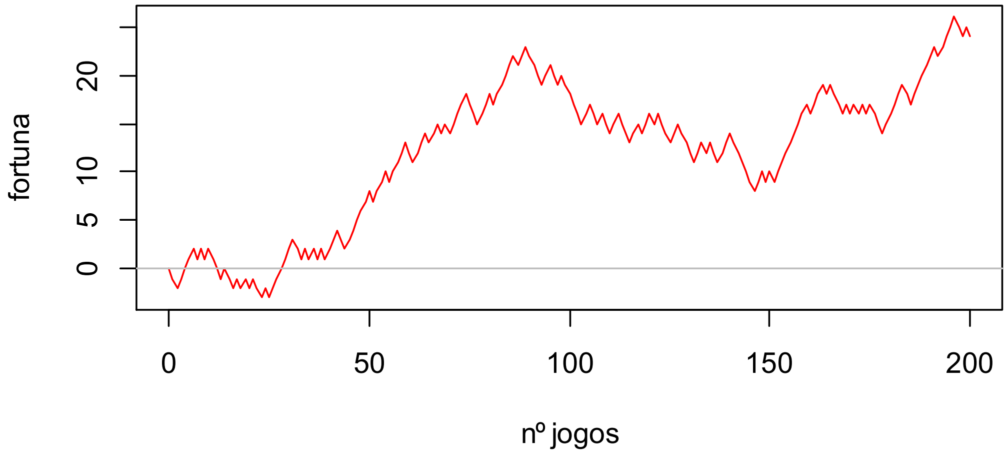


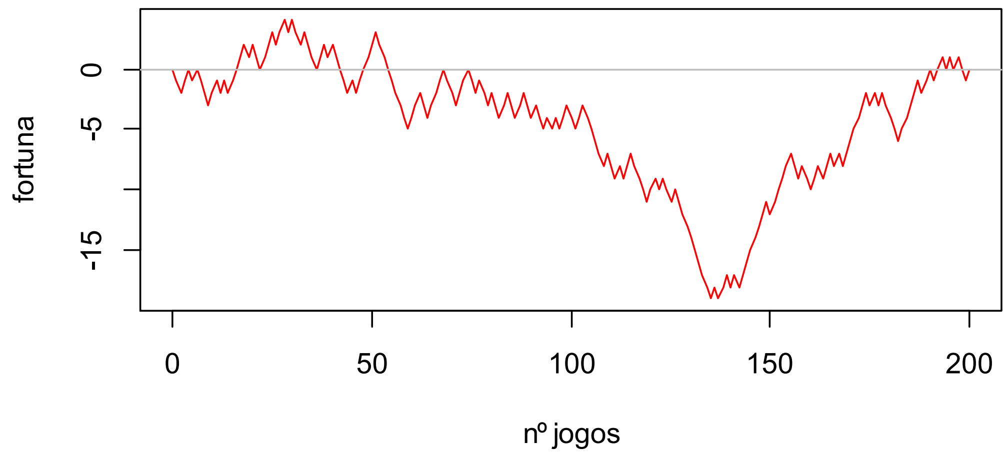


Destes dois acontecimentos, qual será mais provável (para n fixo)?

o último retorno a zero ocorrer na 1ª parte do jogo (antes do instante $n/2$)	ou	o último retorno a zero ocorrer na 2ª parte do jogo (depois do instante $n/2$)

trajetórias com $n = 200$ passos



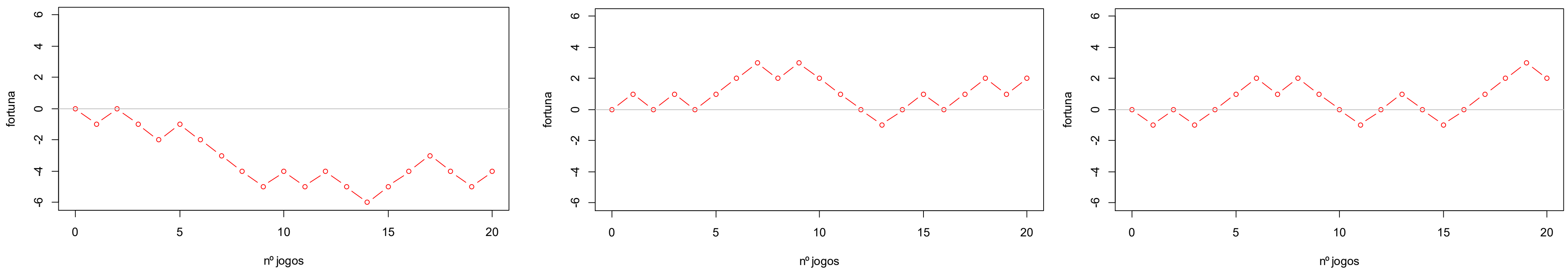


Seja $\xi_{r,n}$ a probabilidade de ser r o nº de mudanças de sinal (da fortuna) em n passos. Como varia $\xi_{r,n}$ em função de r (considerando n fixo)?

Note que o nº de mudanças de sinal, para n par, varia entre 0 e $n/2$

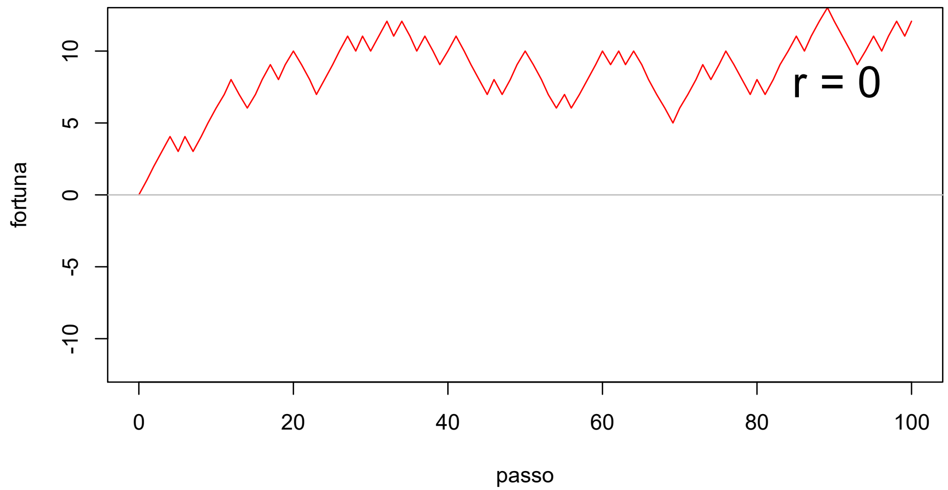
Por exemplo, para $n = 20$, qual é mais provável (r varia entre 0 e 10):

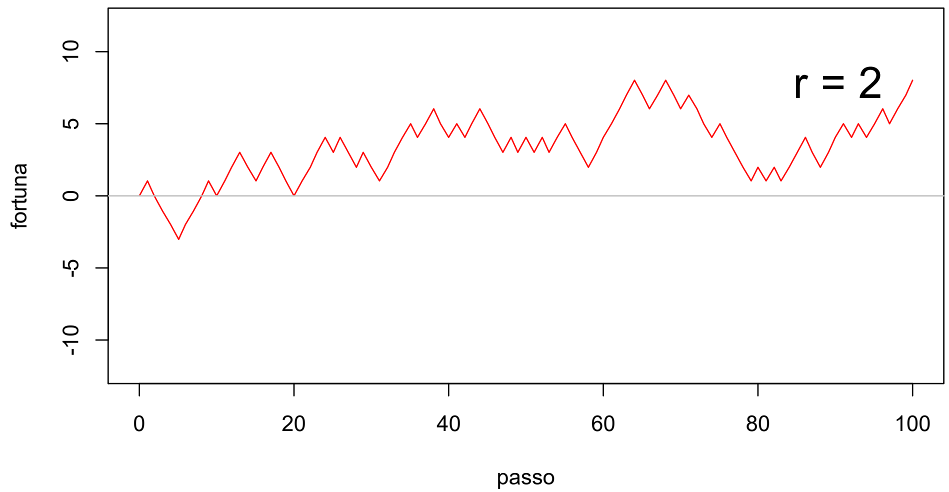
$r = 0$	$r = 2$	$r = 5$
Sem cruzamentos do eixo dos xx *(nenhuma mudança de sinal)*	Dois cruzamentos do eixo dos xx *(duas mudanças de sinal)*	Cinco cruzamentos do eixo dos xx *(cinco mudanças de sinal)*

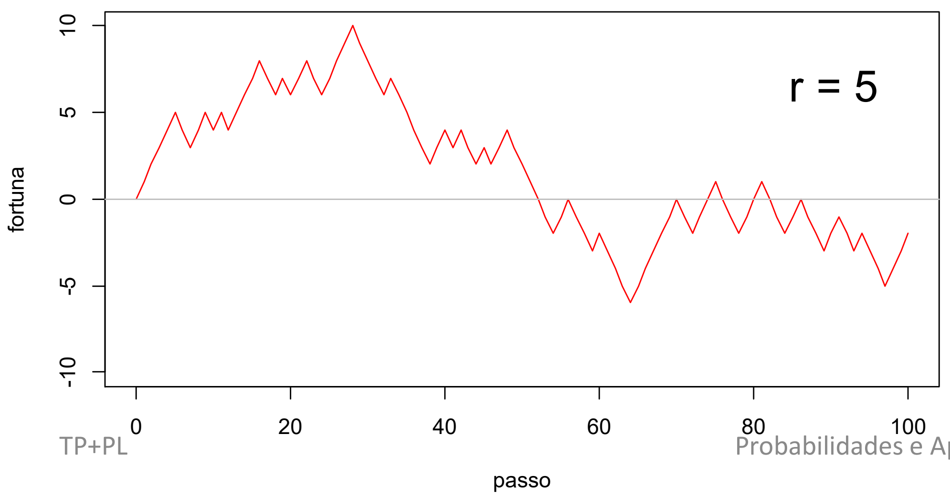


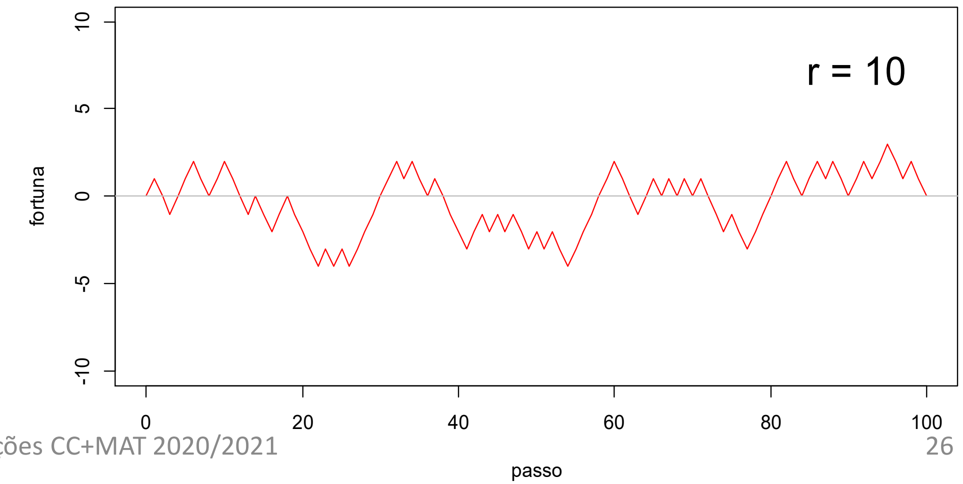
Por exemplo, para $n = 100$, qual é mais/menos provável (r varia entre 0 e 50):

Nenhum cruzamento do eixo dos xx, 2 , 5 ou 10 cruzamentos?





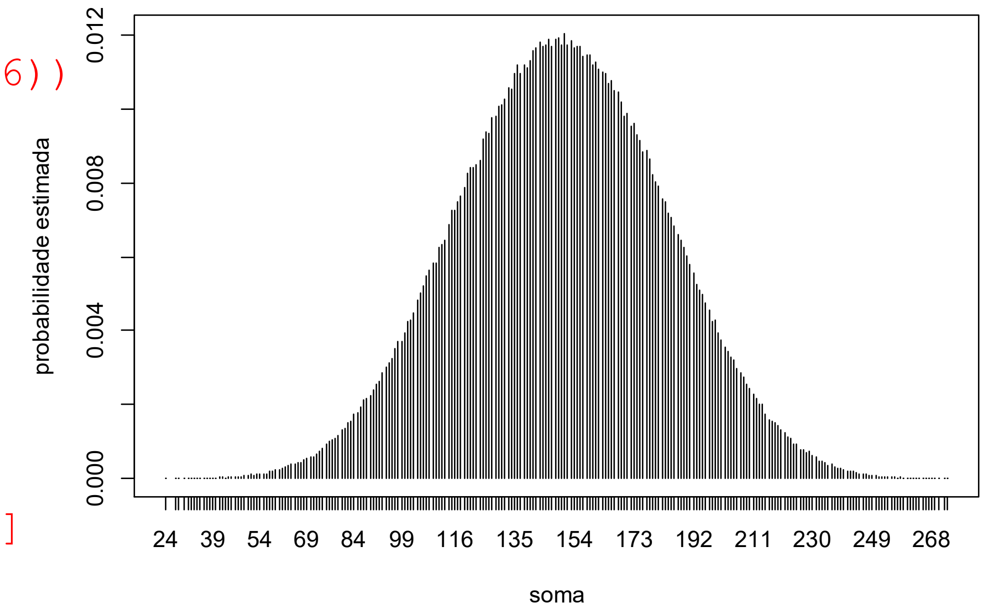




Exercício 5: **(nº2)** Num totoloto (bolas $1{:}49$) estime a probabilidade de que a soma dos 6 números extraídos seja < 50, por simulação (10^5 ou 10^6 réplicas). Melhor: estime a probabilidade p_j de essa soma ser j ($j = 21, 22, \ldots, 279$). Represente graficamente essas probabilidades estimadas em função de j.

Resolução:

```
soma <- 0 ; r <- 10^6
for (i in 1:r) soma[i] <- sum(sample(1:49,6))
# estimativa de P(soma<50):
sum(soma<50)/r
# ou então: length(soma[soma<50])/r
plot(table(soma)/r)
## variante com histograma:
hist(soma,40,freq=F)
# qual a prob de que a soma seja 21?
1/choose(49,6)
# e a estimativa? Executar table(soma)[1:5]
```



Exercício 6: (nº 14) Prove que se dois acontecimentos são independentes,

então (i) qualquer um deles e o complementar do outro também o são

(ii) os seus complementares também o são.

Resolução:

(i) Sejam A e B independentes, i.e., $P(A\cap B)=P(A)P(B)$. Então

$$P(A\cap \overline{B})=P(A-(A\cap B))=P(A)-P(A\cap B)=P(A)-P(A)P(B)=P(A)(1-P(B))=P(A)P(\overline{B})$$

donde A e $\overline{B}$ são independentes. Analogamente (trocando A com B) temos que $\overline{A}$ e B são independentes.

(ii) Aplicando (i) duas vezes, temos A e B indep $\Rightarrow$ $\overline{A}$ e B indep $\Rightarrow$ $\overline{A}$ e $\overline{B}$ indep

o complementar de um deles (A) e o outro (B) são independentes

o complementar de um deles (B) e o outro (A^C) são independentes

Exercício 7: **(nº 20)** Pretende-se ensinar um rato a virar à direita num labirinto. Introduz-se o rato num compartimento com duas saídas, uma à esquerda e outra à direita. O rato será recompensado se sair pela direita e castigado se sair pela esquerda. Supõe-se que na 1ª tentativa o rato sai ao acaso por qualquer das saídas, e que a seguir a uma recompensa [castigo] sai pela direita com probabilidade 0.6 [0.8]. Seja A_n o acontecimento "o rato sai pela direita na n-ésima tentativa" e $p_n = P(A_n)$. Calcule

(i) p_2 e p_3

(ii) p_{n+1} em função de p_n

(iii) $\lim\limits_{n\to\infty} p_n$

Resolução:

Temos $p = P(A_1) = 0.5,\ P(A_{n+1} \mid A_n) = 0.6,\ P(A_{n+1} \mid \overline{A_n}) = 0.8$

Resolução (cont.): (i) p_2 e p_3

Aplicando o TPT com partição $\{A_1, \overline{A_1}\}$, temos

$$P(A_2) = P(A_2 \mid A_1)P(A_1) + P(A_2 \mid \overline{A_1})P(\overline{A_1}) = 0.6 \times 0.5 + 0.8 \times 0.5 = 0.7$$

Analogamente $P(A_3) = \ldots = 0.66$

(ii) p_{n+1} em função de p_n

Aplica-se o TPT com a partição $\{A_n, \overline{A_n}\}$

$$p_{n+1} = P(A_{n+1}) = P(A_{n+1} \mid A_n)P(A_n) + P(A_{n+1} \mid \overline{A_n})P(\overline{A_n}) = 0.6 \times p_n + 0.8 \times (1 - p_n) = 0.8 - 0.2 \times p_n$$

(iii) $\lim\limits_{n\to\infty} p_n$

$$\lim_{n\to\infty} p_{n+1} = 0.8 - 0.2 \times \lim_{n\to\infty} p_n \quad \Rightarrow \quad p = 0.8 - 0.2 \times p \quad \Rightarrow \quad p = \frac{8}{12} = \frac{2}{3}$$

Exercício 8: (nº 21) Duelo triangular – A, B e C jogam por ordem alfabética, ciclicamente. Na sua vez de jogar, cada um escolhe um dos restantes para alvo e dispara (os tiros são mortais caso acertem no alvo). A, B e C acertam no alvo com prob 0.3, 1, 0.5, respectivamente. O 1º a jogar, A, tem 3 estratégias à escolha: Escolher B para alvo, escolher C, ou passar a vez a B. Qual a melhor estratégia para sobreviver ao duelo?

Resolução:

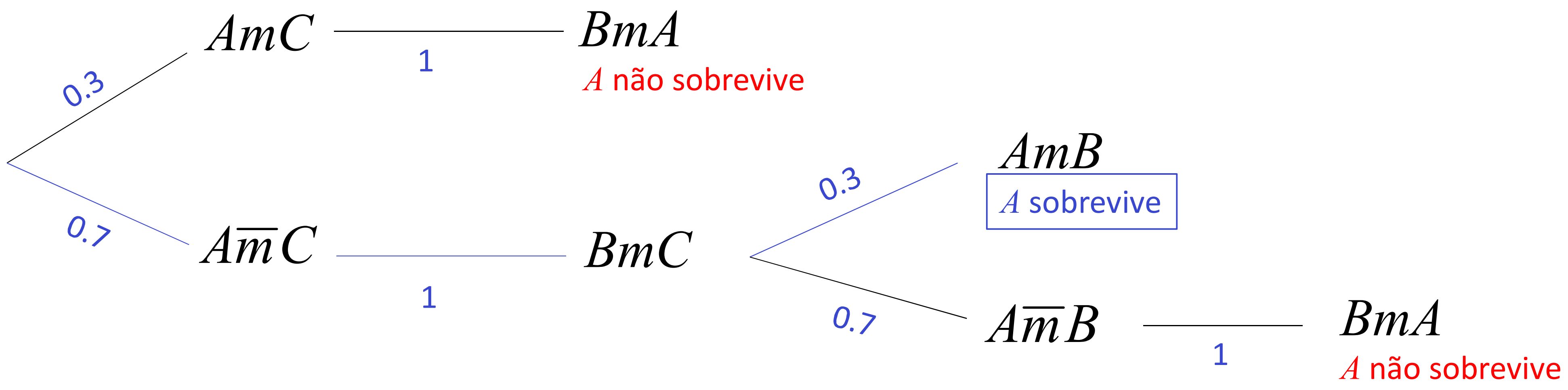
Calcula-se $P(A \text{ sobreviver})$ de acordo com cada estratégia.

- Estratégia 1 – A escolhe B para alvo
- Estratégia 2 – A escolhe C para alvo
- Estratégia 3 – A passa a vez a B

Escolhe-se a estratégia que maximiza $P(A \text{ sobreviver})$

Resolução:

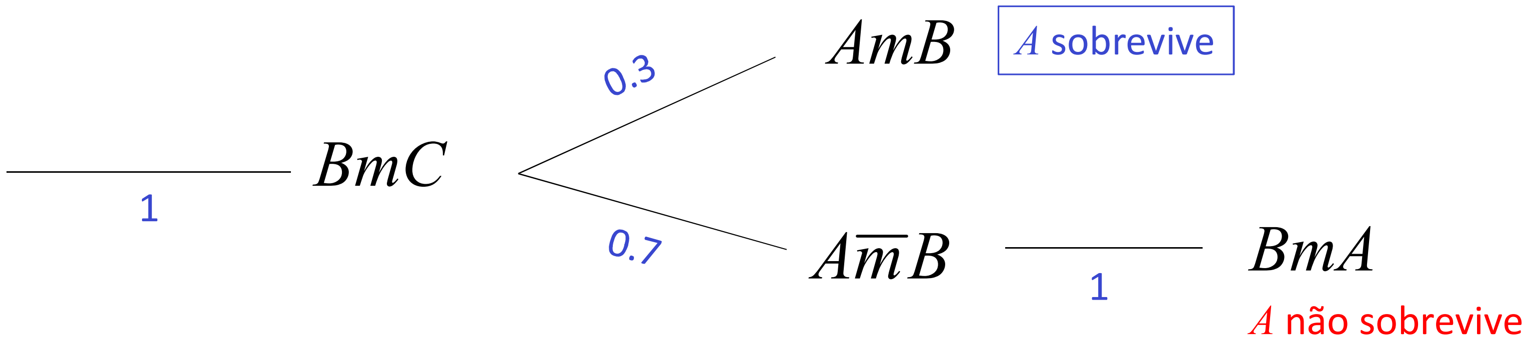
Estratégia 2 – A escolhe C; então AmC (A mata C) ou não; a seguir joga B; se AmC então BmA (A não sobrevive); se A~mC então BmC; depois joga A, que sobrevive sse AmB



$$P(A\ \ sobreviver) = 0.7 \times 1 \times 0.3 = 0.21$$

Resolução:

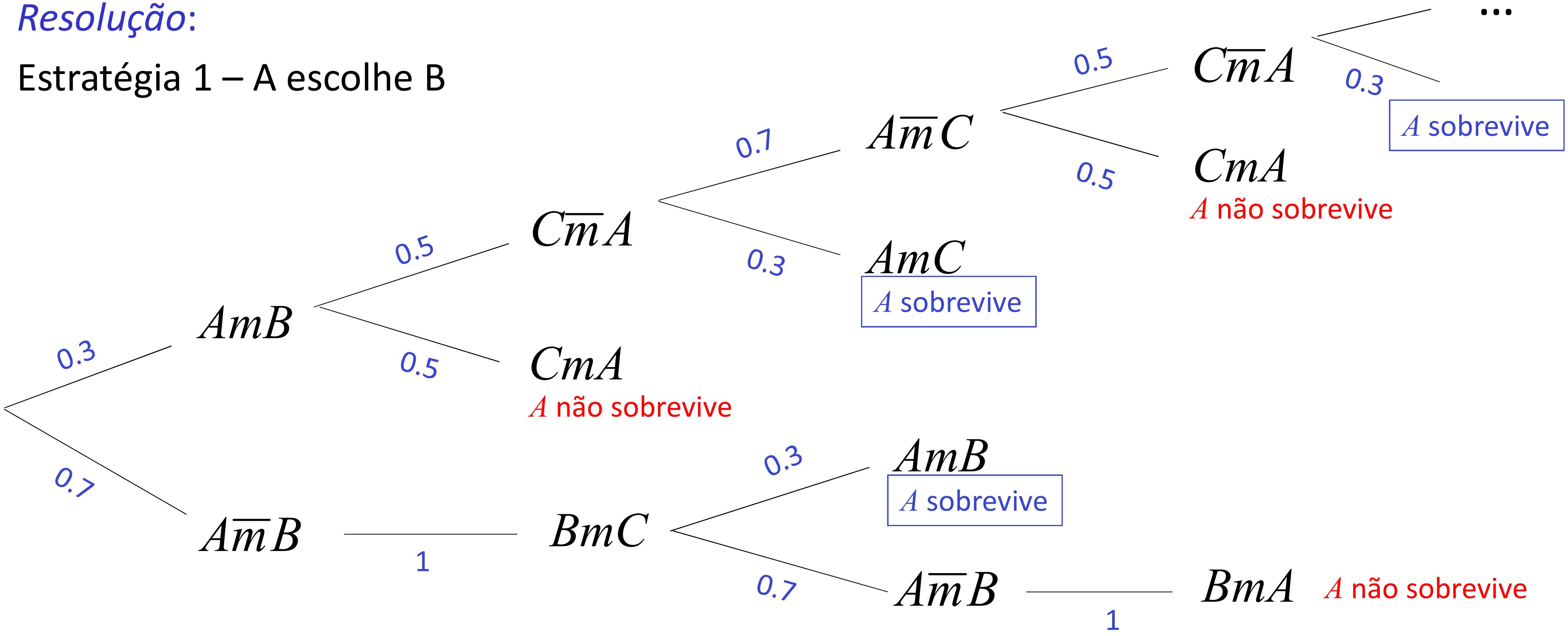
Estratégia 3 – A passa a vez; a seguir joga B; B escolhe C; BmC (B mata C); a seguir joga A; ou AmB (e então A sobrevive) ou não (A não sobrevive).



$$P(A\ \ sobreviver) = 1 \times 0.3 = 0.3$$

Resolução:

Estratégia 1 – A escolhe B



$$P(A\ \ sobreviver) = 0.7\times 1\times 0.3\ +\ 0.3\times 0.5\times 0.3\ +\ 0.3\times 0.5\times (0.7\times 0.5)\times 0.3 + \ldots =$$

$$= 0.21 + (0.3\times 0.5\times 0.3)\sum_{j\geq 0}(0.7\times 0.5)^j = 0.21 + 0.045\frac{1}{1-0.35} = 0.2792308$$

Exercício 9: (nº 42) Calcular a distribuição da soma de duas v.a.'s, X e Y, independentes, nos casos (i) $bi(m,p)$ e $bi(n,p)$ (ii) $Poisson(\lambda)$ e $Poisson(\mu)$

Resolução:

(i) Como $X+Y$ representa o nº de sucessos em $m+n$ provas de Bernoulli(p) independentes, temos imediatamente que $X+Y \frown bi(m+n,p)$.

(ii) Cálculo de $P(X+Y=n)$, escrevendo o acontecimento $\{X+Y=n\}$ como união disjunta dos acontecimentos $\{X=i, Y=n-i\}$, com $i=0, 1, 2, \ldots, n$

$$P(X+Y=n)=\sum_{i=0}^{n} P(X=i,Y=n-i)=\sum_{i=0}^{n} P(X=i)P(Y=n-i)=$$

$$=\sum_{i=0}^{n} e^{-\lambda}\frac{\lambda^i}{i!}\, e^{-\mu}\frac{\mu^{n-i}}{(n-i)!}=e^{-\lambda-\mu}\sum_{i=0}^{n}\tfrac{1}{n!}\binom{n}{i}\lambda^i\mu^{n-i}=\tfrac{1}{n!}e^{-(\lambda+\mu)}(\lambda+\mu)^n,\quad n=0,1,2,\ldots$$

Logo $X+Y \frown Poisson(\lambda+\mu)$.

Exercício 10: (nº 5) Num passeio aleatório simétrico, representando a fortuna do jogador ao fim de j passos por S_j , estime, por meio de simulação, a probabilidade de (i) $S_{2n} = 0$ (i.e.,haver retorno à origem ao fim de $2n$ passos)

(ii) $S_1 \neq 0,\ S_2 \neq 0\ ,\ldots,\ S_{2n} \neq 0$ (não haver retorno à origem durante os primeiros $2n$ passos), para $n = 10,\ 20\ ,50.$

Elabore um gráfico com as estimativas das probabilidades de o último empate (num jogo com $2n$ passos) ocorrer na jogada $2k$ ($k = 0,\ 1,\ 2,\ \ldots,\ 20$) e (iii) estime a probabilidade de que o último empate ocorra antes do passo n [depois do passo n], no caso $n = 20.$

Solução: ver resolução do Trabalho 1

Exercício 11: (nº 6) Num passeio aleatório simétrico em n passos, estime a probabilidade $\xi_{r,n}$ de ser r o nº de mudanças de sinal (da fortuna) e analise como varia em função de r (para n fixo). Simule para os casos $n = 25$ e $n = 100.$

Resultados teóricos – passeio aleatório (a confirmar por simulação):

- $P(S_1 \neq 0, S_2 \neq 0, \ldots, S_{2n} \neq 0) = P(S_{2n} = 0) = u_{2n} = \binom{2n}{n} 2^{-2n} \sim \frac{1}{\sqrt{n\pi}}$
- $\xi_{r,n}$ (prob de r mudanças de sinal em n passos) decresce com r:

$$\xi_{0,n} \geq \xi_{1,n} > \xi_{2,n} > \ldots$$

$$\xi_{r,2n+1} = \binom{2n+1}{r+n+1} \frac{1}{2^{2n}}$$

Exemplo: em 10 000 passos, a probabilidade de “o nº mudanças de sinal” ser < 9 é $\cong 0.14$; a de ser > 78 é $\cong 0.12$

- Em $2n$ passos, a probabilidade de o último retorno à origem (empate) ocorrer no instante $2k$ é $u_{2k}u_{2n-2k}$, $k = 0, 1, \ldots, n$ (note que é simétrica em relação a n). Logo a probabilidade de o último retorno ocorrer na 1ª metade do jogo (antes do passo n) é igual à de ocorrer na 2ª metade (depois de n).

Fórmulas úteis:

Binómio de Newton:

$$\boxed{(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j \, b^{n-j}}$$

donde

caso $a = p$ e $b = 1 - p$:

$$\boxed{1 = \sum_{j=0}^{n} \binom{n}{j} p^j \, (1-p)^{n-j}}$$

caso $a = 1$ e $b = 1$:

$$\boxed{2^n = \sum_{j=0}^{n} \binom{n}{j}}$$

Se $\# \, \Omega = n$, então $\# \, \mathcal{P}(\Omega) = 2^n$

Generalização do binómio (3 termos):

$$\boxed{(a+b+c)^n = \sum_{i=0}^{n} \sum_{j=0}^{n-i} \frac{n!}{i! \; j! \; (n\text{-}i\text{-}j)!} \; a^i \, b^j c^{n-j}}$$

Fórmula de Stirling:

$$\boxed{n! \sim \sqrt{2\pi} \; n^{n+1} \; e^{-n}}$$

significa que

$$\frac{n!}{\sqrt{2\pi} \; n^{n+1} \; e^{-n}} \xrightarrow[n\to\infty]{} 1$$

Fórmulas úteis (cont.):

$$\boxed{1+2+3+\ldots+n=\frac{n(n+1)}{2}}$$

$$\boxed{1^2+2^2+3^2+\ldots+n^2=\frac{n(n+1)(2n+1)}{6}}$$

Soma dos termos de uma progressão geométrica de razão x :

$$\boxed{1+x+x^2+x^3+\ldots+x^n=\frac{1-x^{n+1}}{1-x}}$$

$$\boxed{\sum_{j=0}^{\infty} x^j=\frac{1}{1-x}, \qquad |x|<1}$$

Desenvolvimento em série de e^x :

$$\boxed{e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots=\sum_{j\geq 0}\frac{x^j}{j!} \;, x\in\mathbb{R}}$$

$$\boxed{\lim_{x\to+\infty}\left(1+\frac{a}{x}\right)^x=e^a}$$

Desenvolvimento em série de $\log(1+x)$:

$$\boxed{\log(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots=\sum_{j\geq 1}(-1)^{j-1}\frac{x^j}{j} \;, \; |x|<1}$$