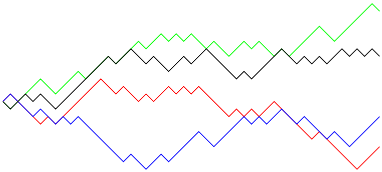
Baralhos e passeios aleatórios

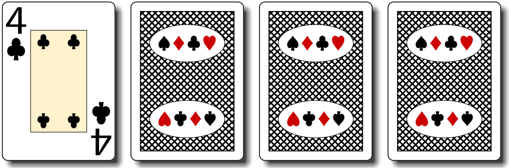


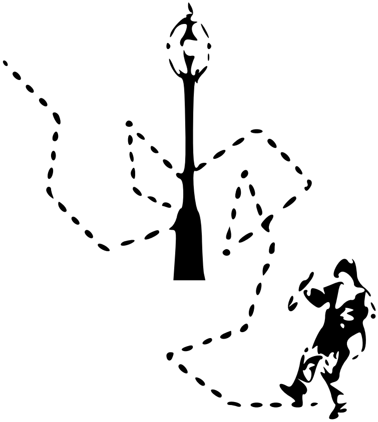
João V. A. Pimenta[♣]
Luiz F. S. Marques[◇]
Raphael A. Duarte[♠]
Sabrina Estácio[♡]

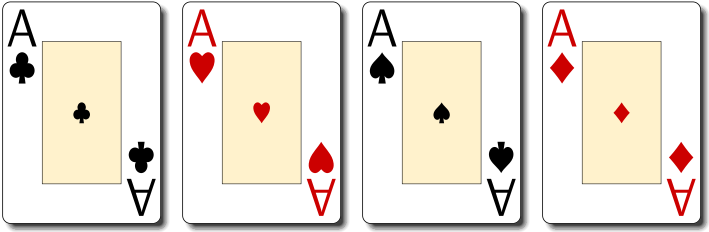
[♣]IFSC - USP, [◇]IME - USP, [♠]FFCLRP - USP, [♡]ICEX - UFMG

Sumário

Simulando um passeio aleatório






A
A
A
A

- Nós podemos.

- Nós podemos.
- É prático.

- Nós podemos.
- É prático.
- É legal.

Cor	Movimento atribuído
♣ ♠	$\uparrow$
♥ ♦	$\downarrow$

Cor	Movimento atribuído
♣♠	$\uparrow$
♥♦	$\downarrow$

- Diremos que esta é a **simulação tradicional**.

Cor	Movimento atribuído
♣♠	$\uparrow$
♥♦	$\downarrow$

- Diremos que esta é a **simulação tradicional**.
- Se retirarmos cartas de uma mesma cor seguidamente, o passeio começa a ficar enviesado.

Cor	Movimento atribuído
♣ ♠	$\uparrow$
♥ ♦	$\downarrow$

- Diremos que esta é a **simulação tradicional**.

Cor	Movimento atribuído
♣♠	$\uparrow$
♥♦	$\downarrow$

- Diremos que esta é a **simulação tradicional**.
- Intuitivamente, quanto mais cartas tem o baralho, melhor ele aproxima um passeio aleatório.

A simulação tradicional

Cor	Movimento atribuído
♣ ♠	$\uparrow$
♥ ♦	$\downarrow$

- Diremos que esta é a **simulação tradicional**.

Cor	Movimento atribuído
♣ ♠	$\uparrow$
♥ ♦	$\downarrow$

- Diremos que esta é a **simulação tradicional**.
- Se utilizarmos todas as cartas, voltaremos ao ponto inicial.

Naipe	Movimento atribuído
♣	$\uparrow\uparrow$
♥	$\uparrow\downarrow$
♠	$\downarrow\uparrow$
♦	$\downarrow\downarrow$

E se utilizarmos os naipes?

Naipe	Movimento atribuído
♣	$\uparrow\uparrow$
♥	$\uparrow\downarrow$
♠	$\downarrow\uparrow$
♦	$\downarrow\downarrow$

- Com o baralho dividido em $4 = 2^2$ grupos, damos 2 passos com cada carta.

E se utilizarmos os naipes?

Naipe	Movimento atribuído
♣	$\uparrow\uparrow$
♥	$\uparrow\downarrow$
♠	$\downarrow\uparrow$
♦	$\downarrow\downarrow$

- Com o baralho dividido em $4 = 2^2$ grupos, damos 2 passos com cada carta.
- De maneira mais geral, com o baralho dividido em 2^k grupos, obtemos uma simulação que dá k passos por cartas.

E se utilizarmos os naipes?

Naipe	Movimento atribuído
♣	$\uparrow\uparrow$
♥	$\uparrow\downarrow$
♠	$\downarrow\uparrow$
♦	$\downarrow\downarrow$

- Com o baralho dividido em $4 = 2^2$ grupos, damos 2 passos com cada carta.
- De maneira mais geral, com o baralho dividido em 2^k grupos, obtemos uma simulação que dá k passos por cartas.
- Temos o mesmo enviesamento que ocorre na simulação tradicional.

- Entender o quão bem a **simulação tradicional** simula um passeio aleatório simétrico.

- Entender o quão bem a **simulação tradicional** simula um passeio aleatório simétrico.
- Quantificar a relação entre o número de cartas e o número de passos necessários para aproximar bem um passeio aleatório.

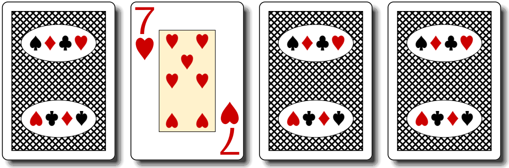
Nosso objetivo

- Entender o quão bem a **simulação tradicional** simula um passeio aleatório simétrico.
- Quantificar a relação entre o número de cartas e o número de passos necessários para aproximar bem um passeio aleatório.
- Para isso,
  (a) Definiremos precisamente os objetos em questão (baralho, embaralhamentos, simulação, ...);

Nosso objetivo

- Entender o quão bem a **simulação tradicional** simula um passeio aleatório simétrico.
- Quantificar a relação entre o número de cartas e o número de passos necessários para aproximar bem um passeio aleatório.
- Para isso,
  (a) Definiremos precisamente os objetos em questão (baralho, embaralhamentos, simulação, ...);
  (b) Estabeleceremos um meio preciso de avaliar a qualidade de uma simulação.

Definições e o resultado principal



- Sejam $p \in (0,1)$ e $(\xi_j)_{j\in\mathbb{N}}$ uma sequência de variáveis aleatórias i.i.d com

$$\mathbb{P}(\xi_j = 1) = 1 - \mathbb{P}(\xi_j = -1) = p.$$

- Sejam $p \in (0,1)$ e $(\xi_j)_{j\in\mathbb{N}}$ uma sequência de variáveis aleatórias i.i.d com

$$\mathbb{P}(\xi_j = 1) = 1 - \mathbb{P}(\xi_j = -1) = p.$$

Considere $X_0 := 0$ e, para cada $n \geq 1$,

$$X_n := \sum_{j=1}^{n} \xi_j.$$

- Sejam $p \in (0,1)$ e $(\xi_j)_{j\in\mathbb{N}}$ uma sequência de variáveis aleatórias i.i.d com

$$\mathbb{P}(\xi_j = 1) = 1 - \mathbb{P}(\xi_j = -1) = p.$$

Considere $X_0 := 0$ e, para cada $n \geq 1$,

$$X_n := \sum_{j=1}^{n} \xi_j.$$

- A sequência $(X_n)_{n\in\mathbb{N}_0}$ é dita o **passeio aleatório em** $\mathbb{Z}$.

- Se $p = 1/2$,

$$\mathbb{P}(\xi_j = 1) = \mathbb{P}(\xi_j = -1) = 1/2$$

e a sequência $(X_n)_{n\in\mathbb{N}_0}$ é dita o **passeio aleatório simétrico em $\mathbb{Z}$**.

- Se $p = 1/2$,

$$\mathbb{P}(\xi_j = 1) = \mathbb{P}(\xi_j = -1) = 1/2$$

  e a sequência $(X_n)_{n \in \mathbb{N}_0}$ é dita o **passeio aleatório simétrico em $\mathbb{Z}$**.
- É este o caso que desejamos simular!

- Se $p = 1/2$,

$$\mathbb{P}(\xi_j = 1) = \mathbb{P}(\xi_j = -1) = 1/2$$

e a sequência $(X_n)_{n\in\mathbb{N}_0}$ é dita o **passeio aleatório simétrico em $\mathbb{Z}$.**

- É este o caso que desejamos simular!
- Mas queremos simular um passeio finito...

Seja $T \in \mathbb{N}$.

- O vetor aleatório

$$(X_n)_{0 \leq n \leq T} = (X_0, X_1, \ldots, X_T)$$

é dito o **passeio aleatório de tamanho** T.

Seja $T \in \mathbb{N}$.

- O vetor aleatório

$$(X_n)_{0 \leq n \leq T} = (X_0, X_1, \ldots, X_T)$$

  é dito o **passeio aleatório de tamanho** T.
- $(X_n)_{0 \leq n \leq T}$ assume valores em

$$\Gamma_T := \{(x_0, \ldots, x_T) \in \mathbb{Z}^T : x_0 = 0, |x_j - x_{j-1}| = 1\}$$

  com lei uniforme.

Seja $T \in \mathbb{N}$.

- O vetor aleatório

$$(X_n)_{0 \le n \le T} = (X_0, X_1, \ldots, X_T)$$

é dito o **passeio aleatório de tamanho** T.

- $(X_n)_{0 \le n \le T}$ assume valores em

$$\Gamma_T := \{(x_0, \ldots, x_T) \in \mathbb{Z}^T : x_0 = 0, |x_j - x_{j-1}| = 1\}$$

com lei uniforme.

- Convém observar que

$$\Gamma_T \cong \{-1, 1\}^T =: \Omega_T.$$

Sejam $N = 2K$ e T naturais com $T < N$.

- Um **baralho** com N cartas é o conjunto

$$\Lambda_N := \{1, \ldots, N\}.$$

Sejam $N = 2K$ e T naturais com $T < N$.

- Um **baralho** com N cartas é o conjunto
$$\Lambda_N := \{1, \dots, N\}.$$
- Um **embaralhamento** de Λ_N é um elemento do conjunto
$$S_N := \{(\sigma(1), \dots, \sigma(N)) : \sigma \text{ é uma bijeção de } \Lambda_N \text{ em } \Lambda_N\}.$$

Sejam $N = 2K$ e T naturais com $T < N$.

- Um **baralho** com N cartas é o conjunto

$$\Lambda_N := \{1, \ldots, N\}.$$

- Um **embaralhamento** de Λ_N é um elemento do conjunto

$$S_N := \{(\sigma(1), \ldots, \sigma(N)) : \sigma \text{ é uma bijeção de } \Lambda_N \text{ em } \Lambda_N\}.$$

- Uma **simulação** é uma função

$$X : S_N \longrightarrow \Omega_T.$$

Sejam $N = 2K$ e T naturais com $T \leq K$.

- Seja $\varphi : \Lambda_N \longrightarrow \{-1, 1\}$ a função

$$\varphi(j) = \begin{cases} -1, & \text{se } j \text{ é par;} \\ 1, & \text{se } j \text{ é impar.} \end{cases}$$

Sejam $N = 2K$ e T naturais com $T \leq K$.

- Seja $\varphi : \Lambda_N \longrightarrow \{-1, 1\}$ a função

$$\varphi(j) = \begin{cases} -1, & \text{se } j \text{ é par;} \\ 1, & \text{se } j \text{ é impar.} \end{cases}$$

A **simulação tradicional** é a função $X : S_N \longrightarrow \Omega_T$ dada por

$$X(\sigma_1, \ldots, \sigma_N) = (\varphi(\sigma_1), \ldots, \varphi(\sigma_T)).$$

Sejam $N = 2K$ e T naturais com $T \leq K$.

- Seja $\varphi : \Lambda_N \longrightarrow \{-1, 1\}$ a função

$$\varphi(j) = \begin{cases} -1, & \text{se } j \text{ é par;} \\ 1, & \text{se } j \text{ é impar.} \end{cases}$$

A **simulação tradicional** é a função $X : S_N \longrightarrow \Omega_T$ dada por

$$X(\sigma_1, \ldots, \sigma_N) = (\varphi(\sigma_1), \ldots, \varphi(\sigma_T)).$$

- Por exemplo, com $N = 10$ e $T = 4$

$$X(4, 6, 7, 2, 3, 8, 5, 1, 10, 9) = (-1, -1, 1, -1)$$

- Sejam Ω um conjunto enumerável e $\mu, \nu : \mathscr{P}(\Omega) \longrightarrow [0,1]$ medidas de probabilidade.

- Sejam Ω um conjunto enumerável e μ, $\nu : \mathscr{P}(\Omega) \longrightarrow [0,1]$ medidas de probabilidade. A **distância de variação total** entre μ e ν é

$$d_{\mathrm{VT}}(\mu,\nu) := \frac{1}{2}\sum_{\omega\in\Omega}|\mu(\omega)-\nu(\omega)|$$

- Sejam Ω um conjunto enumerável e $\mu, \nu : \mathscr{P}(\Omega) \longrightarrow [0,1]$ medidas de probabilidade. A **distância de variação total** entre μ e ν é

$$d_{\mathrm{VT}}(\mu,\nu) := \frac{1}{2}\sum_{\omega\in\Omega}|\mu(\omega)-\nu(\omega)|$$

Proposição

$$d_{\mathrm{VT}}(\mu,\nu) = \sum_{\omega\,:\,\nu(\omega)\leq\mu(\omega)} \mu(\omega)-\nu(\omega).$$

Sejam $\mathbb{Q}_N$ e $\mathbb{P}_T$ as leis uniformes em S_N e Ω_T respectivamente.

- Cada simulação $X : S_N \longrightarrow \Omega_T$ determina uma medida de probabilidade em Ω_T dada por

$$\mu_X(\omega) := \mathbb{Q}_N(X = \omega),\ \omega \in \Omega_T.$$

Sejam $\mathbb{Q}_N$ e $\mathbb{P}_T$ as leis uniformes em S_N e Ω_T respectivamente.

- Cada simulação $X : S_N \longrightarrow \Omega_T$ determina uma medida de probabilidade em Ω_T dada por

$$\mu_X(\omega) := \mathbb{Q}_N(X = \omega),\ \omega \in \Omega_T.$$

- O passeio aleatório de tamanho T assume valores em Ω_T com lei unifome.

Sejam $\mathbb{Q}_N$ e $\mathbb{P}_T$ as leis uniformes em S_N e Ω_T respectivamente.

- Cada simulação $X : S_N \longrightarrow \Omega_T$ determina uma medida de probabilidade em Ω_T dada por

$$\mu_X(\omega) := \mathbb{Q}_N(X = \omega),\ \omega \in \Omega_T.$$

- O passeio aleatório de tamanho T assume valores em Ω_T com lei unifome.
- Assim, gostaríamos de estimar

$$d(N) := d_{\mathrm{VT}}(\mu_X, \mathbb{P}_T) = \frac{1}{2} \sum_{\omega \in \Omega_T} \left| \mu_X(\omega) - 2^{-T} \right|$$

Já temos ingredientes suficientes para enunciar o resultado principal:

Já temos ingredientes suficientes para enunciar o resultado principal:

Teorema

Para todo $c \geq 2$,

$$d(cT) = \operatorname{erf}\left(\frac{\sigma_c}{\sqrt{2(c-1)/c}}\right) - \operatorname{erf}\left(\frac{\sigma_c}{\sqrt{2}}\right) + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right),$$

Já temos ingredientes suficientes para enunciar o resultado principal:

Teorema

Para todo $c \geq 2$,

$$d(cT) = \operatorname{erf}\left(\frac{\sigma_c}{\sqrt{2(c-1)/c}}\right) - \operatorname{erf}\left(\frac{\sigma_c}{\sqrt{2}}\right) + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right),$$

em que

$$\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2}\, dt$$

Já temos ingredientes suficientes para enunciar o resultado principal:

Teorema

Para todo $c \geq 2$,

$$d(cT) = \operatorname{erf}\left(\frac{\sigma_c}{\sqrt{2(c-1)/c}}\right) - \operatorname{erf}\left(\frac{\sigma_c}{\sqrt{2}}\right) + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right),$$

em que

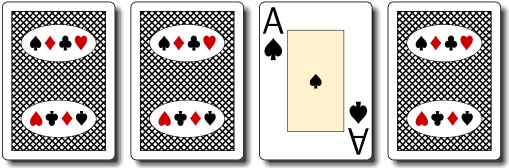
$$\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2}\, dt$$

e

$$\sigma_c = \sqrt{(c-1)\log\left(\frac{c}{c-1}\right)}$$

$d(cT)$	c
0.16	2
0.1	2.94
0.05	5.35
0.01	24.7
0.005	48.89
0.001	242.47

Uma ideia da prova



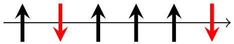
Queremos estimar a distância

$$d(N) = \frac{1}{2} \sum_{\omega \in \Omega_T} |\mu_X(\omega) - \mathbb{P}_T(\omega)|,$$

onde, lembrando,

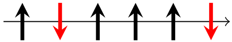
- $\mathbb{P}_T$ é a lei uniforme em Ω_T e
- μ_X é a lei da simulação $X : S_N \longrightarrow \Omega_T$.

O primeiro a se fazer é obter uma expressão para $\mu_X(\omega)$. Para isso, suponha que ω é a seguinte trajetória:



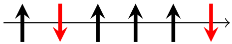
- N cartas no baralho.
- K cartas para cima no baralho (e K para baixo).

Qual é a lei de X?



$$\mu_X(\omega) = \frac{K}{2K} \cdot \frac{K}{2K-1} \cdot \frac{K-1}{2K-2} \cdot \frac{K-2}{2K-3} \cdot \frac{K-3}{2K-4} \cdot \frac{K-1}{2K-5}$$

Qual é a lei de X?



Podemos reordenar os termos multiplicados.

$$\mu_X(\omega) = \left(\frac{K}{2K} \cdot \frac{K-1}{2K-2} \cdot \frac{K-2}{2K-3} \cdot \frac{K-3}{2K-4}\right)\left(\frac{K}{2K-1} \cdot \frac{K-1}{2K-5}\right)$$

Reescrevendo $\mu_X(\omega)$ temos:

$$\mu_X(\omega) = \frac{\frac{K!}{(K-\lambda_\omega)!} \cdot \frac{K!}{(K-T+\lambda_\omega)!}}{\frac{(2K)!}{(2K-T)!}}$$

Se definirmos λ_ω a quantidade de $\uparrow$ do passeio ω, podemos chamar $T - \lambda_\omega$ a quantidade de $\downarrow$.

E ainda temos que duas trajetórias ω e γ vão ter a mesma probabilidade se $\lambda_\omega = \lambda_\gamma$.

$$\mu_X(\omega) = \frac{\frac{K!}{(K-\lambda_\omega)!} \cdot \frac{K!}{(K-T+\lambda_\omega)!}}{\frac{(2K)!}{(2K-T)!}}$$

Chame $L_\lambda = \{\omega \in \Omega_T \colon \lambda_\omega = \lambda\}$ (caminhos com λ $\uparrow$'s)

$$\mu_X(L_\lambda) = \frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}} \cdot \binom{T}{\lambda}$$

Qual é a lei de X?

Chame $L_\lambda = \{\omega \in \Omega_T \colon \lambda_\omega = \lambda\}$ (caminhos com λ $\uparrow$'s)

$$\mu_X(L_\lambda) = \frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}} \cdot \binom{T}{\lambda}$$

Visto que $\binom{T}{\lambda}$ é a quantidade de elementos em L_λ.

Temos então:

$$\mu_X(L_\lambda) = \frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}} \cdot \binom{T}{\lambda}$$

e

$$\mathbb{P}_T(L_\lambda) = \frac{1}{2^T} \cdot \binom{T}{\lambda}$$

Voltemos, agora, à distância,

$$d(N) = \frac{1}{2}\sum_{\lambda=0}^{T}\left|\frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}}\binom{T}{\lambda} - \frac{1}{2^T}\binom{T}{\lambda}\right|$$

Voltemos, agora, à distância,

$$d(N) = \frac{1}{2}\sum_{\lambda=0}^{T}\left|\frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}}\binom{T}{\lambda} - \frac{1}{2^T}\binom{T}{\lambda}\right|$$

Podemos então escrever a distância de variação total de duas formas:

1. $d(N) = \frac{1}{2}\sum_{\lambda=0}^{T}\left|\mu_H(\lambda) - \mu_B(\lambda)\right|$
2. $d(N) = \frac{1}{2}\sum_{\lambda=0}^{T}\frac{1}{2^T}\binom{T}{\lambda}\left|2^T\frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}} - 1\right|$

Nos atentando à primeira escrita da $d(N)$:

$$d(N) = \frac{1}{2} \sum_{\lambda=0}^{T} |\mu_H(\lambda) - \mu_B(\lambda)|$$

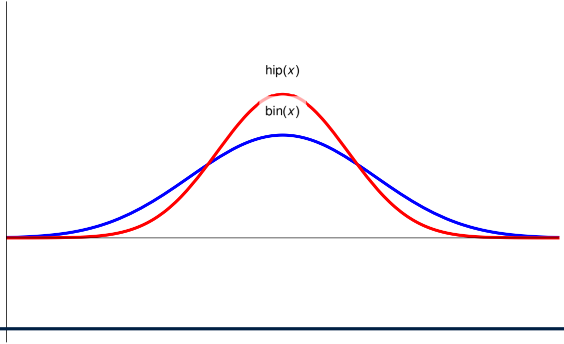
Estamos representando

- H como uma variável aleatória com distribuição hipergeométrica.
- B como uma variável aleatória com distribuição binomial.

Nos atentando à primeira escrita da $d(N)$:

$$d(N) = \frac{1}{2}\sum_{\lambda=0}^{T} |\mu_H(\lambda) - \mu_B(\lambda)|$$

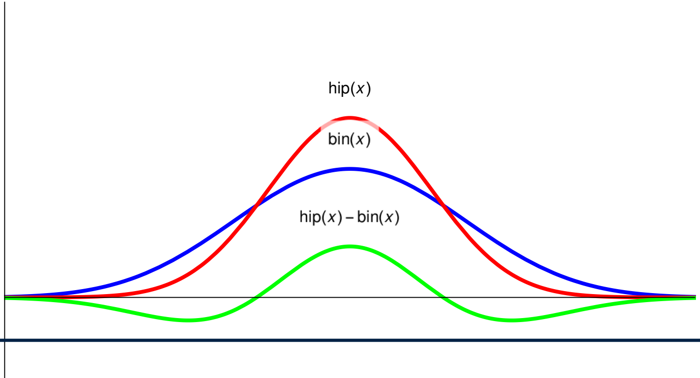
Os plots de $\mu_H(\lambda)$ e $\mu_B(\lambda)$ são os seguintes (com N e T escolhidos):



Nos atentando à primeira escrita da $d(N)$:

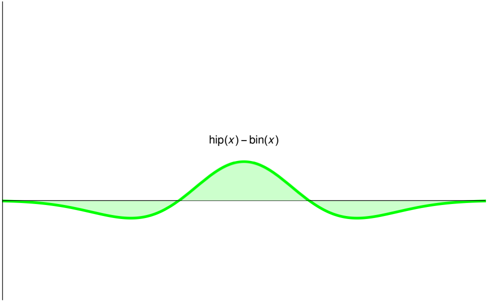
$$d(N) = \frac{1}{2} \sum_{\lambda=0}^{T} |\mu_H(\lambda) - \mu_B(\lambda)|$$

Avaliando a diferença entre a curva vermelha e azul obtemos a curva verde:



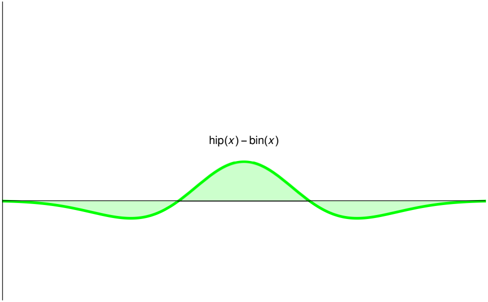
Sem o módulo, a soma resulta em 0.

$$\frac{1}{2}\sum_{\lambda=0}^{T}\mu_H(\lambda) - \mu_B(\lambda) = 0$$



Sem o módulo, a soma resulta em 0.

$$\frac{1}{2}\sum_{\lambda=0}^{T}\mu_H(\lambda) - \mu_B(\lambda) = 0$$



Vamos calcular a $d(N)$ na região positiva (sem módulo).

Existem 3 partes para a demonstração:

1. Encontrar uma maneira melhor de escrever a $d(N)$.
2. Encontrar uma estimativa da soma em uma região positiva.
3. Encontrar a estimativa da $d(N)$ em toda a região positiva.

Para escrever a $d(N)$ de maneira mais conveniente vamos reescrever:

$$d(N) = \frac{1}{2}\sum_{\lambda=0}^{T}\frac{1}{2^T}\binom{T}{\lambda}\left|2^T\frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}}-1\right|$$

Com algumas contas, mostramos que

Para escrever a $d(N)$ de maneira mais conveniente vamos reescrever:

$$d(N) = \frac{1}{2} \sum_{\lambda=0}^{T} \frac{1}{2^T} \binom{T}{\lambda} \left| 2^T \frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}} - 1 \right|$$

Com algumas contas, mostramos que

$$2^T \frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}} = \prod_{j=0}^{\lambda-1} \left(1 - \frac{j}{K}\right)\left(1 - \frac{j}{2K}\right)^{-1} \prod_{j=0}^{T-\lambda-1} \left(1 - \frac{j}{K}\right)\left(1 - \frac{j}{2K}\right)^{-1}.$$

$$2^T \frac{\binom{2K-T}{K-\lambda}}{\binom{2K}{K}} = \prod_{j=0}^{\lambda-1} \left(1 - \frac{j}{K}\right)\left(1 - \frac{j}{2K}\right)^{-1} \prod_{j=0}^{T-\lambda-1} \left(1 - \frac{j}{K}\right)\left(1 - \frac{j}{2K}\right)^{-1}.$$

Avaliando o log no produtório, e escrevendo N como cT, é possível mostrar que:

$$d(cT) = \frac{1}{2} \sum_{\lambda=0}^{T} \frac{1}{2^T} \binom{T}{\lambda} \left| \frac{1}{\sqrt{\frac{c-1}{c}}} e^{\frac{-2\left[\lambda - \frac{T}{2}\right]^2}{T(c-1)}} - 1 \right| + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

Temos então uma maneira (surpreendentemente) mais fácil de descrever a $d(N)$.

$$d(cT) = \frac{1}{2}\sum_{\lambda=0}^{T}\frac{1}{2^T}\binom{T}{\lambda}\left|\frac{1}{\sqrt{\frac{c-1}{c}}}e^{\frac{-2\left[\lambda-\frac{T}{2}\right]^2}{T(c-1)}} - 1\right| + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

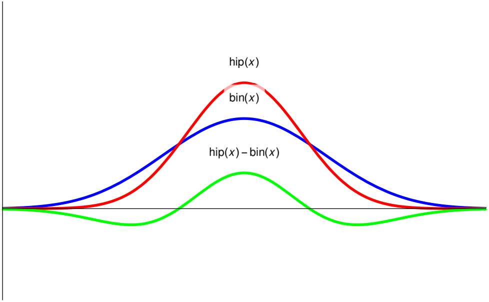
Temos então uma maneira (surpreendentemente) mais fácil de descrever a $d(N)$.

$$d(cT) = \frac{1}{2}\sum_{\lambda=0}^{T}\frac{1}{2^T}\binom{T}{\lambda}\left|\frac{1}{\sqrt{\frac{c-1}{c}}}e^{\frac{-2\left[\lambda-\frac{T}{2}\right]^2}{T(c-1)}} - 1\right| + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

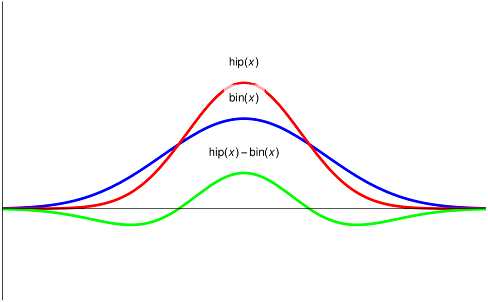
Para auxiliar nas futuras passagens, vamos escrever

$$\frac{1}{\sqrt{\frac{c-1}{c}}}e^{\frac{-2\left[\lambda-\frac{T}{2}\right]^2}{T(c-1)}} = f(\lambda, T)$$

Os pontos de mudança de sinal ainda são difíceis de encontrar, entretanto sabemos que existem pelos gráficos.



Os pontos de mudança de sinal ainda são difíceis de encontrar, entretanto sabemos que existem pelos gráficos.



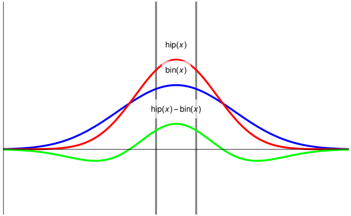
E ainda, μ_H e μ_B são simétricas, logo esses pontos estão equidistantes de $T/2$.

Podemos então definir s_T como pontos em que:

$$\left|\frac{T}{2} - \lambda\right| < s_T \implies f(\lambda, T) - 1 \geq 0$$

e assim

$$d(N) = \sum_{\lambda=T/2-s_T}^{T/2+s_T} \frac{1}{2^T}\binom{T}{\lambda}\left[\frac{1}{\sqrt{\frac{c-1}{c}}} e^{\frac{-2\left[\lambda-\frac{T}{2}\right]^2}{T(c-1)}} - 1\right] + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$



Teorema

Para λ na forma

$$\lambda = \frac{T}{2} - \alpha\frac{\sqrt{T}}{2}$$

então o comportamento assintótico de $\binom{T}{\lambda}$ é:

$$\binom{T}{\lambda} = 2^T\sqrt{\frac{2}{\pi T}} \cdot e^{\frac{-\alpha^2}{2}} + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

Teorema

Para λ na forma

$$\lambda = \frac{T}{2} - \alpha\frac{\sqrt{T}}{2}$$

então o comportamento assintótico de $\binom{T}{\lambda}$ é:

$$\binom{T}{\lambda} = 2^T\sqrt{\frac{2}{\pi T}} \cdot e^{\frac{-\alpha^2}{2}} + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

note que $\alpha = 2\left[\frac{\lambda - T/2}{\sqrt{T}}\right]$

Dessa forma, escrevendo $s_T = \alpha_T \frac{\sqrt{T}}{2}$ temos que $d(N)$ é:

$$d(cT) = \sum_{\lambda=\frac{T}{2}-\frac{\alpha_T\sqrt{T}}{2}}^{\frac{T}{2}+\frac{\alpha_T\sqrt{T}}{2}} \frac{1}{\sqrt{T}} \frac{2}{\sqrt{2\pi}} e^{\frac{-2\left(\frac{T}{2}-\lambda\right)^2}{T}} \left[\frac{1}{\sqrt{\frac{c-1}{c}}} e^{\frac{-2\left(\frac{T}{2}-\lambda\right)^2}{T(c-1)}} - 1 \right] + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

Dessa forma, escrevendo $s_T = \alpha_T \frac{\sqrt{T}}{2}$ temos que $d(N)$ é:

$$d(cT) = \sum_{\lambda = \frac{T}{2} - \frac{\alpha_T \sqrt{T}}{2}}^{\frac{T}{2} + \frac{\alpha_T \sqrt{T}}{2}} \frac{1}{\sqrt{T}} \frac{2}{\sqrt{2\pi}} e^{\frac{-2\left(\frac{T}{2} - \lambda\right)^2}{T}} \left[\frac{1}{\sqrt{\frac{c-1}{c}}} e^{\frac{-2\left(\frac{T}{2} - \lambda\right)^2}{T(c-1)}} - 1 \right] + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

Após algumas manipulações temos que:

$$d(cT) = \sum_{\lambda = -\frac{\alpha_T \sqrt{T}}{2}}^{\frac{\alpha_T \sqrt{T}}{2}} \frac{1}{\sqrt{T}} \frac{1}{\sqrt{2\pi}} \left[\frac{1}{\sqrt{\frac{1}{4} \frac{c-1}{c}}} e^{\frac{-\left(\lambda/\sqrt{T}\right)^2}{2(1/4)\frac{c-1}{c}}} - \frac{1}{\sqrt{\frac{1}{4}}} e^{\frac{-\left(\lambda/\sqrt{T}\right)^2}{2(1/4)}} \right] + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

$$d(cT) = \sum_{\lambda=-\frac{\alpha_T\sqrt{T}}{2}}^{\frac{\alpha_T\sqrt{T}}{2}} \frac{1}{\sqrt{T}}\frac{1}{\sqrt{2\pi}}\left[\frac{1}{\sqrt{\frac{1}{4}\frac{c-1}{c}}}e^{\frac{-\left(\lambda/\sqrt{T}\right)^2}{2(1/4)\frac{c-1}{c}}} - \frac{1}{\sqrt{\frac{1}{4}}}e^{\frac{-\left(\lambda/\sqrt{T}\right)^2}{2(1/4)}}\right] + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

Podemos aplicar somas de Riemann no intervalo $[-\frac{\alpha_T}{2}, \frac{\alpha_T}{2}]$ particionado em segmentos de tamanho $\sqrt{T}$. Resultando em:

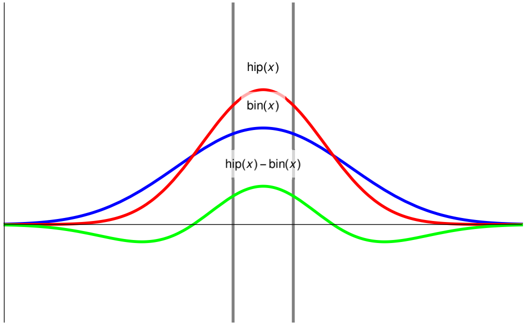
$$d(cT) = \int_{-\frac{\alpha_T}{2}}^{\frac{\alpha_T}{2}} \frac{1}{\sqrt{T}}\frac{1}{\sqrt{2\pi}}\left[\frac{1}{\sqrt{\frac{1}{4}\frac{c-1}{c}}}e^{\frac{-x^2}{2(1/4)\frac{c-1}{c}}} - \frac{1}{\sqrt{\frac{1}{4}}}e^{\frac{-x^2}{2(1/4)}}\right]dx + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

$$d(cT) = \int_{-\frac{\alpha_T}{2}}^{\frac{\alpha_T}{2}} \frac{1}{\sqrt{T}} \frac{1}{\sqrt{2\pi}} \left[\frac{1}{\sqrt{\frac{1}{4}\frac{c-1}{c}}} e^{\frac{-x^2}{2(1/4)\frac{c-1}{c}}} - \frac{1}{\sqrt{\frac{1}{4}}} e^{\frac{-x^2}{2(1/4)}} \right] dx + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

Notando as gaussianas acima, podemos escrever a integral como a função erro:

$$d(cT) = \operatorname{erf}\left(\frac{\alpha_T}{2\sqrt{\frac{c-1}{2c}}}\right) - \operatorname{erf}\left(\frac{\alpha_T}{\sqrt{2}}\right) + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

Tendo agora uma expressão para a acumulada de uma vizinhança de $T/2$



Pode-se enxergar que esses α_T convergem para algum ponto, denotado por σ_c, quanto T cresce.

Após encontrar a expressão:

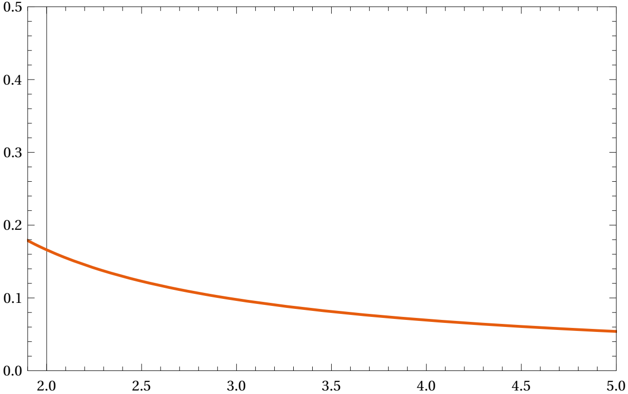
$$\sigma_c = \sqrt{(c-1)\log\left[\frac{c}{1-c}\right]}$$

Após encontrar a expressão:

$$\sigma_c = \sqrt{(c-1)\log\left[\frac{c}{1-c}\right]}$$

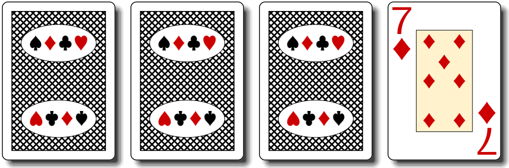
Obtemos que:

$$d(cT) = \mathrm{erf}\left(\frac{\sigma_c}{2\sqrt{\frac{c-1}{2c}}}\right) - \mathrm{erf}\left(\frac{\sigma_c}{\sqrt{2}}\right) + \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$



Com c no eixo horizontal, associado à $d(cT)$ no eixo vertical

O que ainda pode ser explorado



Cor	Movimento atribuído
♣♠	↑
♥♦	↓

Naipe	Movimento atribuído
♣	↑↑
♥	↑↓
♠	↓↑
♦	↓↓

Considere ↑ (1) e ↓ (0).

Carta	Movimento atribuído
A♠	$\cdots 0000000001$
A♣	$\cdots 0000000010$
A♥	$\cdots 0000000011$
A♦	$\cdots 0000000100$
2♠	$\cdots 0000000101$
2♣	$\cdots 0000000110$
$\vdots$	$\vdots$

Suponha $N = 16 = 2^4$. Podemos escolher 4 passos por carta.

Carta	Movimento atribuído
A♠	0001
A♣	0010
A♥	0011
A♦	0100
$\vdots$	$\vdots$

Suponha $N = 16 = 2^4$. Podemos escolher 4 passos por carta.

Carta	Movimento atribuído
A♠	0001
A♣	0010
A♥	0011
A♦	0100
⋮	⋮

Devemos nos preocupar, contudo, com caminhos do tipo
A♥A♠A♠$\cdots$

Naipe	Movimento atribuído
♣	$\uparrow$
♥	$\downarrow$
♠	$\rightarrow$
♦	$\leftarrow$

Divisão inicial em duas cores:

A♣ 2♣ 3♣ 4♣ 5♣ 6♣ 7♣ 8♣ 9♣ 10♣ J♣ Q♣ K♣

A♠ 2♠ 3♠ 4♠ 5♠ 6♠ 7♠ 8♠ 9♠ 10♠ J♠ Q♠ K♠

A♥ 2♥ 3♥ 4♥ 5♥ 6♥ 7♥ 8♥ 9♥ 10♥ J♥ Q♥ K♥

A♦ 2♦ 3♦ 4♦ 5♦ 6♦ 7♦ 8♦ 9♦ 10♦ J♦ Q♦ K♦

Após a primeira retirada (7♣):

A♣ 2♣ 3♣ 4♣ 5♣ 6♣ 8♣ 9♣ 10♣ J♣ Q♣ K♣

A♠ 2♠ 3♠ 4♠ 5♠ 6♠ 7♠ 8♠ 9♠ 10♠ J♠ Q♠ K♠

A♥ 2♥ 3♥ 4♥ 5♥ 6♥ 7♥ 8♥ 9♥ 10♥ J♥ Q♥ K♥

A♦ 2♦ 3♦ 4♦ 5♦ 6♦ 7♦ 8♦ 9♦ 10♦ J♦ Q♦ K♦

Após a primeira retirada (7♣):

A♣ 2♣ 3♣ 4♣ 5♣ 6♣ 8♣ 9♣ 10♣ J♣ Q♣ K♣

A♠ 2♠ 3♠ 4♠ 5♠ 6♠ 7♠ 8♠ 9♠ 10♠ J♠ Q♠ K♠

A♥ 2♥ 3♥ 4♥ 5♥ 6♥ 7♥ 8♥ 9♥ 10♥ J♥ Q♥ K♥

A♦ 2♦ 3♦ 4♦ 5♦ 6♦ 7♦ 8♦ 9♦ 10♦ J♦ Q♦ K♦

Carta com valor nulo: A♥

Após a segunda retirada (3♣):

	A♥	A♣	2♣	4♣	5♣	6♣	8♣	9♣	10♣	J♣	Q♣	K♣
A♠	2♠	3♠	4♠	5♠	6♠	7♠	8♠	9♠	10♠	J♠	Q♠	K♠

	2♥	3♥	4♥	5♥	6♥	7♥	8♥	9♥	10♥	J♥	Q♥	K♥
A♦	2♦	3♦	4♦	5♦	6♦	7♦	8♦	9♦	10♦	J♦	Q♦	K♦

- A simulação produz caminhos (de tamanho aleatório) perfeitos!

- A simulação produz caminhos (de tamanho aleatório) perfeitos!

$$T = N - \mathbb{P}_r$$

em que $\mathbb{P}_r$ é Poisson-Binomial distribuída

- A simulação produz caminhos (de tamanho aleatório) perfeitos!

$$T = N - \mathbb{P}_r$$

em que $\mathbb{P}_r$ é Poisson-Binomial distribuída

- $\mu = \sum_i^n p_i = \sum_i^n \frac{1}{2(K-i)} = H_{2K} - \frac{1}{2}H_K \sim O(\log(K))$

- A simulação produz caminhos (de tamanho aleatório) perfeitos!

$$T = N - \mathbb{P}_r$$

em que $\mathbb{P}_r$ é Poisson-Binomial distribuída

- $\mu = \sum_i^n p_i = \sum_i^n \frac{1}{2(K-i)} = H_{2K} - \frac{1}{2}H_K \sim O(\log(K))$
- $\sigma^2 = \sum_i^n (1-p_i)p_i = \sum_i^n \left(1 - \frac{1}{2(K-i)}\right)\frac{1}{2(K-i)} \sim O(\log(K))$

[1] D. Bayer, P. Diaconis. *Trailing the dovetail shuffle to its lair*. The Annals of Applied Probability, 1992.

[2] T. Franco. *Princípios de Combinatória e Probabilidade*. Coleção Matemática Universitária, 2020.

[3] J. Spencer. *Asymptopia*. Student Mathematical Library. American Mathematical Society, 2014.