Notas de Aula: Riffle Shuffles e Processos Estocásticos

Disciplina de Processos Estocásticos

Introdução

Como modelar o ato de embaralhar cartas? Essa pergunta aparentemente simples leva a uma bela aplicação da Teoria das Probabilidades e Processos Estocásticos, com conexões com cadeias de Markov, variação total, o paradoxo do aniversário e o problema do coletor de cupons.

1 Modelagem do Embaralhamento

Dado um baralho com n cartas numeradas de 1 a n, podemos embaralhá-lo aplicando uma permutação $\pi \in S_n$ (o grupo das permutações).

O embaralhamento ideal corresponderia a aplicar uma permutação uniformemente aleatória, mas na prática, apenas subconjuntos específicos de S_n são realizados, e com distribuições não uniformes.

2 Riffle Shuffle (Embaralhamento à la Cassino)

O **riffle shuffle** (embaralhamento por entrelaçamento) é uma modelagem realista do embaralhamento feito por crupiês:

(a) Divide-se o baralho em duas partes (com corte em t).

(b) As cartas são intercaladas uma a uma, respeitando a ordem interna de cada parte.

Formalmente, uma permutação $\pi \in S_n$ é um riffle shuffle se a sequência $(\pi(1), \ldots, \pi(n))$ consiste na intercalação de duas sequências crescentes.

Número de Riffle Shuffles

Para cada corte em t, há $\binom{n}{t}$ formas distintas de intercalar. Portanto, existem 2^n possíveis interações (mas apenas $2^n - n$ riffle shuffles distintos).

Para entender quantas permutações diferentes podem ser produzidas por um único **riffle shuffle**, vamos analisar o processo em detalhes.

Divisão do baralho

Primeiro, o baralho de n cartas é dividido em duas partes:

- A parte da **mão direita** com t cartas;
- A parte da **mão esquerda** com $n - t$ cartas.

Para cada valor possível de t entre 0 e n, há $\binom{n}{t}$ formas diferentes de intercalar as cartas das duas mãos preservando a ordem relativa dentro de cada uma. Ou seja, não embaralhamos dentro de cada metade: apenas as intercalamos.

Intercalações possíveis

Ao intercalar duas listas ordenadas de tamanhos t e $n - t$, o número de intercalamentos possíveis (também chamados de *interleavings*) é:

$$\binom{n}{t}$$

Isso porque estamos escolhendo em quais das n posições finais irão as t cartas da mão direita (as restantes serão ocupadas pelas cartas da mão esquerda).

Como t pode variar de 0 até n, o número total de possíveis intercalações que preservam a ordem interna de cada parte é:

$$\sum_{t=0}^{n} \binom{n}{t} = 2^n$$

No entanto, isso é apenas o número total de **intercalações possíveis**, e não necessariamente de **riffle shuffles distintos**.

Permutações geradas

Nem todas as intercalações geram permutações diferentes. Algumas diferentes formas de intercalar podem resultar na mesma permutação final. Além disso, o número de permutações que podem ser obtidas por um riffle shuffle é muito menor do que $n!$ (o total de permutações possíveis).

De fato, o conjunto de permutações obtidas por um riffle shuffle corresponde exatamente às permutações de n elementos que podem ser decompostas como a **intercalação de duas subsequências crescentes**. Tais permutações têm a propriedade de ter no máximo uma "queda"(ou *descent*), isto é, um índice i tal que $\pi(i) > \pi(i+1)$.

Esse conjunto especial de permutações é bastante restrito: existem exatamente $2^n - n$ permutações desse tipo, chamadas de **riffle permutations**.

Resumo

- Existem 2^n formas diferentes de intercalar duas pilhas preservando ordem interna.
- Nem todas essas formas geram permutações distintas.
- O número de permutações possíveis por um único riffle shuffle é igual a $2^n - n$.
- Isso é muito menor do que $n!$, o número total de permutações possíveis de n cartas.

Este fato destaca uma verdade importante: **um único riffle shuffle não produz uma permutação aleatória uniforme**. É preciso repetir o processo múltiplas vezes para que a distribuição se aproxime da uniforme.

3 Modelo Probabilístico de Gilbert-Shannon-Reeds

O modelo clássico e mais estudado para o riffle shuffle foi desenvolvido por **Edgar Gilbert** e **Claude Shannon**, e depois formalizado por **Jim Reeds**. Ele fornece uma descrição estocástica natural para o embaralhamento tipo riffle e nos permite analisá-lo matematicamente.

Esse modelo define uma distribuição de probabilidade Rif sobre o grupo de permutações S_n, onde n é o número de cartas.

Existem três maneiras equivalentes de descrever esse modelo:

1. Definição por permutações permitidas

O modelo define a distribuição $\mathrm{Rif}(\pi)$ como:

$$\mathrm{Rif}(\pi) = \begin{cases} \dfrac{n+1}{2^n} & \text{se } \pi = \mathrm{id} \text{ (a identidade)} \\ \dfrac{1}{2^n} & \text{se } \pi \text{ pode ser obtida por riffle shuffle (i.e., intercalação de duas sequências cres} \\ 0 & \text{caso contrário} \end{cases}$$

Essa definição mostra que: - A permutação identidade ocorre com probabilidade ligeiramente maior. - Todas as outras riffle-permutações válidas têm a mesma probabilidade. - Permutações que não podem surgir de um riffle shuffle têm probabilidade zero.

2. Modelo de corte e intercalamento probabilístico

Esse é o modelo mais intuitivo para quem já embaralhou cartas fisicamente:

1. Corta-se o baralho em dois blocos: o número t de cartas a serem colocadas na mão direita é escolhido com probabilidade

$$\mathbb{P}(t) = \frac{1}{2^n}\binom{n}{t}, \quad \text{para } 0 \leq t \leq n.$$

2. As duas metades (de tamanhos t e $n-t$) são intercaladas uma a uma: a próxima carta a ser colocada vem da mão direita com probabilidade proporcional ao número de cartas restantes nela, isto é,

$$\mathbb{P}(\text{direita}) = \frac{r}{r+\ell}, \qquad \mathbb{P}(\text{esquerda}) = \frac{\ell}{r+\ell},$$

onde r e ℓ são os números atuais de cartas nas mãos direita e esquerda, respectivamente.

Esse processo gera uma permutação $\pi \in S_n$ com distribuição $\mathrm{Rif}(\pi)$.

3. Modelo binário (shuffle inverso)

Neste modelo, cada carta é rotulada independentemente com um bit aleatório 0 ou 1, com igual probabilidade:

$$\mathbb{P}(\text{bit} = 0) = \mathbb{P}(\text{bit} = 1) = \frac{1}{2}.$$

Depois disso:

- Todas as cartas com bit 0 são colocadas no topo do baralho;
- As cartas com bit 1 são colocadas abaixo;
- Dentro de cada grupo (0s e 1s), a ordem relativa original das cartas é preservada.

Este modelo define uma **inversa** de um riffle shuffle, pois é equivalente a aplicar uma permutação cuja inversa é uma riffle shuffle válida. Por isso, ele é frequentemente usado para análise teórica, como veremos na próxima seção.

Equivalência dos modelos

As três descrições acima são probabilisticamente equivalentes, isto é, produzem exatamente a mesma distribuição Rif sobre S_n. A equivalência entre elas pode ser compreendida da seguinte forma:

- O modelo 1 é a descrição explícita da distribuição.
- O modelo 2 fornece um algoritmo estocástico para gerar permutações com essa distribuição.
- O modelo 3 transforma o problema em um contexto mais combinatório (bits aleatórios), útil para prova de teoremas.

Vantagens do modelo GSR

- Reflete bem o comportamento de embaralhamentos feitos por humanos (especialmente amadores).
- Permite análises precisas, via ferramentas de combinatória e probabilidade.

- É compatível com técnicas de stopping time e distância de variação total.

4 Regra de Parada Uniforme Forte

Uma regra de parada uniforme forte é uma variável aleatória T que representa o tempo de parada de um processo estocástico com a propriedade:

$$\mathbb{P}[X_k = \pi \mid T = k] = \frac{1}{n!}, \quad \text{para todo } \pi \in S_n.$$

Regra de Parada para Riffle Shuffles

Após k embaralhamentos inversos, cada carta adquire uma sequência binária $(b_1, b_2, \ldots, b_k)$. A regra de parada:

"Pare quando todas as cartas tiverem rótulos binários distintos."

5 Análise via Paradoxo do Aniversário

Assumindo k embaralhamentos, cada carta recebe um rótulo binário de comprimento k, totalizando 2^k possibilidades.

O problema se reduz ao paradoxo do aniversário: qual a probabilidade de n objetos aleatórios *não* colidirem em 2^k caixas?

$$\mathbb{P}[T > k] = 1 - \prod_{i=1}^{n-1} \left(1 - \frac{i}{2^k}\right).$$

Teorema 1: Após k riffle shuffles, a distância de variação total satisfaz:

$$\|\mathrm{Rif}^{*k} - U\| \leq 1 - \prod_{i=1}^{n-1} \left(1 - \frac{i}{2^k}\right).$$

6 Quantas Vezes Embaralhar?

- Para $n = 52$ cartas, temos:

k	$\|\text{Rif}^{*k} - U\|$
1	1.000
5	0.952
7	0.334
10	0.043
12	0.028

- Conclusão prática: **7 riffle shuffles** são suficientes para tornar o baralho ”quase aleatório”.

7 Conclusão

O riffle shuffle é um belo exemplo de aplicação de ferramentas probabilísticas em situações cotidianas. Modelos como o do embaralhamento inverso, regras de parada fortes e conexões com problemas clássicos (aniversário, cupons) revelam a elegância dos processos estocásticos.