por Isadora de Oliveira

* Análise Probabilística e Algoritmos Aleatorizados

* Problema da Contratação

Entrevistar tem um custo baixo, digamos c_i enquanto contratar é caro e custa c_h. Se m for o número de pessoas contratadas, o custo total associado a esse algoritmo é $O(c_i n + c_h m)$. Independente de quantas pessoas contratamos n canditados e portanto sempre incorreremos no custo $c_i n$ associado a entrevistar. Assim nos concentraremos na análise de $c_h m$ o custo de contratação. Essa quantidade vária com cada execução do algoritmo.

- Análise do pior caso

No pior caso contratamos cada candidato que entrevistamos, isso ocorre se os canditados vierem crescente de qualidade e nesse caso contratamos n vezes para um custo total de contratação $O(c_h n)$.

* Análise Probabilística

Utilização da probabilidade na análise de problemas. Usamos análise probabilística para analisar o tempo de execução de um algoritmo. Para fazer isso temos de conhecer ou supor a distribuição das entradas, em seguida analisar o algoritmo calculando o tempo de execução para o caso médio e tomar a média sobre a distribuição das entradas possível.

Supomos que podemos comparar dois candidatos quaisquer decidir qual é o mais bem qualificado, isto é existe uma ordem total para os candidatos. Assim classificamos cada candidato com um número exclusivo de 1...n usando ordenação (i) para denotar o posto do candidato i. A lista ordenada <ordenação (1), ordenação (2), ..., ordenação (n)> é uma permutação da lista <1, 2,... n>.

Dizer que os candidatos chegam em ordem aleatória equivale a dizer que essa lista de classificação chegam em ordem aleatória das n! permutações de 1...n.

Alternativamente dizemos que as classificações formam uma permutação aleatória uniforme isto é cada uma das n! permutações possíveis aparece com igual probabilidade.

* Algoritmos Aleatorizados

Podemos usar probabilidade e aleatoriedade como ferramentas para projeto e análise de algoritmos, aleatorizando o comportamento de parte do algoritmo.

Para desenvolver um algoritmo aleatorizado para o problema da contratação devemos ter maior controle sobre a ordem em que entrevistamos os candidatos. A agência de empregos tem n candidatos e envia uma lista dos candidatos com antecedência, em vez de confiar em uma suposição de que os candidatos virão em ordem aleatória, rodamos um algoritmo para que a lista de fato esteja aleatória.

Dizemos que um algoritmo é aleatorizado se seu comportamento for determinado não apenas por sua entrada mas também por valores produzidos por um gerador de números aleatórios.

Uma chamada RANDOM(a,b) retorna um inteiro entre a, b inclusive sendo cada inteiro igualmente provavel. Por exemplo, RANDOM(0,1) produz 0 com probabilidade 1/2 e produz 1 com probabilidade 1/2. Uma chamada RANDOM(3,7) retorna 3, 4, 5, 6 e 7 com probabilidade 1/5. Cada inteiro retornado por RANDOM é independente dos inteiros retornados em chamadas anteriores.
(Na prática a maioria dos ambientes de programação oferece um gerador de números pseudoaleatórios: um algoritmo determinístico que retorna números que "parecem" aleatórios estritamente.
Distinguimos esse algoritmo daqueles, cuja a entrada é aleatória referindo ao tempo de execução do algoritmo aleatorizado como um tempo de execução esperado.
Tempo de execução do caso médio: refere-se a entrada do algoritmo.
Tempo de execução esperado: próprio algoritmo faz escolhas aleatórias.

* Variáveis aleatórias indicadoras

Nos dão um método conveniente para converter probabilidade em expectativas. Suponha que temos um espaço amostral S e um evento A. Então, a variável aleatória indicadora I{A} associada ao evento A é definida como:

$$I\{A\} = \begin{cases} 1 & \text{se } A \text{ ocorrer} \\ 0 & \text{se } A \text{ não ocorrer} \end{cases}$$

Como exemplo simples, vamos determinar o número esperado de caras que obtemos quando lançamos uma moeda não viciada.

Nosso espaço amostral é $S = \{H, T\}$ com $Pr\{H\} = Pr\{T\} = \frac{1}{2}$

Então podemos definir uma variável aleatória indicadora X_H, associado ao resultado "cara" do lançamento da moeda que é o evento H. Essa variável conta o número de caras obtidas nesse lancamento.

$$X_H = I\{H\} = \begin{cases} 1 & \text{se } H \text{ ocorrer} \\ 0 & \text{se } T \text{ ocorrer} \end{cases}$$

O número esperado de caras obtidas em um lançamento da moeda é simplesmente o valor esperado de nossa variável indicadora X_H.

$$\begin{aligned} E[X_H] &= E[I\{H\}] \\ &= 1\,Pr\{H\} + 0\,Pr\{T\} \\ &= 1 \cdot (1/2) + 0 \cdot (1/2) \\ &= 1/2 \end{aligned}$$

Desse modo o número esperado de caras obtidas por um lançamento de uma moeda não viciada é $1/2$.

* Lema 5.1

Dado um espaço amostral S e um evento A no espaço amostral S, seja $X_A = I\{A\}$. Então, $E\{X_A\} = Pr\{A\}$.

Temos: $E[X_A] = E[I\{A\}]$
$= 1 \cdot Pr\{A\} + 0 \cdot Pr\{\bar{A}\}$
$= Pr\{A\}$
onde $\bar{A}$ denota $S-A$, o complemento de A.

Fazemos X_i a variável aleatória indicadora associada ao evento no qual o i-ésimo lançamento dá cara $X_i = I\{$ o i-ésimo lancamento resulta no evento $H\}$. Seja X a variável aleatória que denota o nº total de caras no n lançamentos da moeda, de modo que:

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right]$$

Essa equação dá a esperança da soma de n variáveis aleatorias indicadoras.

Agora é fácil calcular o número esperado de caras.

$$E[X] = E\left[\sum_{i=0}^{n} X_i\right]$$
$$= \sum_{i=1}^{n} E[X_i]$$
$$= \sum_{i=0}^{n} 1/2$$
$$= n/2$$

* Apêndice C

* Permutação

Uma permutação de um conjunto finito S é uma sequência ordenada de todos os elementos de S, sendo que cada elemento aparece exatamente uma vez. Por exemplo, se $S = \{a, b, c\}$ então S tem seis permutações

$$abc, acb, bac, bca, cab, cba$$

Há $n!$ permutações de um conjunto de n elementos, visto que o primeiro elemento da sequência pode ser escolhido de n modos, o segundo $n-1$ modos, o terceiro de $n-2$ modos, e assim por diante.

O número de K-permutações (K é o nº de elementos) é:

$$n(n-1)(n-2)\ldots(n-K+1) = \frac{n!}{(n-K)!}$$