Joao Felipe Guedes da Silva
guedes.joaofelipe@poli.ufrj.br

CPS767 - Algoritmo de Monte Carlo
e Cadeias de Markov
Lista 1

2020/03/26

1 *Filhos e Filhas*: Considere um casal que tem dois descendentes e que as chances de cada um deles ser filho ou filha são iguais. Responda às perguntas abaixo:

1.1 Calcule a probabilidade dos descendentes formarem um casal (ou seja, um filho e uma filha)

.
Seja A o elemento "descendente ser filha" e O o elemento "descendente ser filho. Assim, como as chances de cada um deles ser filho ou filha são iguais:

$$P_A = P_O = 1/2$$

Considerando que o casal terá dois descendentes, o espaço amostral de possíveis pares de descendentes é:

$$S = [\{A, O\}, \{A, A\}, \{O, O\}, \{O, A\}]$$

Os eventos de interesse, portanto, são: $\{A, O\}$ ou $\{O, A\}$.
A probabilidade de cada um desses eventos é:

$$P[\{A, O\}] = P[A \wedge O] = 1/2 \cdot 1/2 = 1/4$$
$$P[\{O, A\}] = P[O \wedge A] = 1/2 \cdot 1/2 = 1/4$$

Logo, a probabilidade de se ter um casal é:

$$P_{casal} = P[\{A, O\} \vee \{O, A\}] = 1/4 + 1/4 = 1/2$$

1.2 Calcule a probabilidade de ao menos um dos descendentes ser filho.

Neste caso, os eventos de interesse são: $\{A, O\}, \{O, O\}, \{O, A\}]$.
Portanto,

$$P_{1\ filho} = P[\{A, O\} \vee \{O, O\} \vee \{O, A\}] = 1/4 + 1/4 + 1/4 = 3/4$$

1.3 Calcule a probabilidade das duas serem filhas dado que uma é filha (cuidado!)

Dado que 1 é filha, o espaço amostral é reduzido para o sub-conjunto:

$$S_A = [\{A, O\}, \{A, A\}, \{O, A\}]$$

.

$$\begin{aligned} P[\{A, A\}|S_A] &= \frac{P[\{A, A\} \cap S_A]}{P[S_A]} \\ &= \frac{P[\{A, A\}]}{P[S_A]} \\ &= \frac{1/4}{3/4} \\ &= 1/3 \end{aligned} \quad (1)$$

1.4 Calcule a probabilidade dos descendentes nascerem no mesmo dia (assuma que a chance de nascer em um determinado dia é igual a qualquer outro).

Neste caso, podemos ignorar o sexo do descendente e analisar apenas o dia em que ele/ela nasceu.

Considerando que o ano possui 365 dias e a probabilidade do descendente nascer no dia i é $P_i = 1/365, \forall i$, então:

$$P[D_1 = i \wedge D_2 = i]$$

Onde D_1 e D_2 são as datas de nascimento do primeiro e segundo descendente, respectivamente. Considerando que o nascimento de ambos sejam eventos independentes:

$$P[D_1 = i \wedge D_2 = i] = P[D_1 = i] \cdot P[D_2 = i] = \Big(1/365\Big)^2$$

2 Dado

Considere um icosaedro (um sólido Platônico de 20 faces) tal que a chance de sair a face $i = 1, \ldots, 20$ seja linearmente proporcional a i. Ou seja, $P[X = i] = ci$ para alguma constante c, onde X é uma variável aleatória que denota a face do dado. Responda às perguntas abaixo:

2.1 Determine o valor de c.

$$f_X(i) = P[X = i] = ci$$

Sabemos que uma variável aleatória possui a seguinte restrição:

$$\sum_{i=i}^{20} f_X(i) = 1$$

Consequentemente,

$$\begin{aligned}
\sum_{i=i}^{20} f_X(i) &= \sum_{i=1}^{20} ci \\
&= c \cdot (1 + 2 + \ldots + 20) \\
&= c \cdot \Big[\frac{20 \cdot (1 + 20)}{2}\Big] \\
&= c \cdot 210 \\
&= 1
\end{aligned}$$

Logo, $c = 1/210$.

2.2 Calcule o valor esperado de X (obtenha também o valor numérico).

$$\begin{aligned}
E[X] &= \sum_{i=1}^{20} i \cdot f_X(i) \\
&= \frac{1}{210} \sum_{i=i}^{20} i \cdot i \\
&= \frac{1}{210} (1 + 2^2 + \ldots + 20^2) \\
&= 13.67
\end{aligned}$$

2.3 Calcule a probabilidade de X ser maior do que seu valor esperado.

os eventos de interesse que fazem com que $X > E[X]$ são: $\{14, 15, \ldots, 20\}$. Assim, a probabilidade de que isso aconteça é:

$$P[X > E[X]] = P[\{14, 15, \ldots, 20\}] = \sum_{i=14}^{20} \frac{i}{210} = 57\%$$

2.4 Calcule a variância de X (obtenha também o valor numérico).

Pela definição,

$$\begin{aligned}\sigma_X^2 &= E[(X - E[X])^2] \\ &= E[X^2 - 2X \cdot E[X] + E^2[X]] \\ &= E[X^2] - E^2[X]\end{aligned}$$

Considerando $g(X) = X^2$, temos:

$$\begin{aligned}E[X^2] &= E[g(X)] \\ &= \sum_{i=1}^{20} g(i) \cdot f_X(i) \\ &= \frac{1}{210} \sum_{i=1}^{20} i^2 \cdot i \\ &= \frac{1}{210}(1 + 2^3 + \ldots + 20^3) \\ &= 210\end{aligned}$$

Portanto,

$$\begin{aligned}\sigma_X^2 &= E[X^2] - E^2[X] \\ &= 210 - 13.67^2 \\ &= 23.22\end{aligned}$$

2.5 Repita os últimos três itens para o caso do dado ser uniforme, ou seja, $P[X = i] = 1/20, i = 1, \ldots, 20$. Qual dado possui maior variância? Compare os resultados.

2.5.1 Determine o valor de c.

Neste caso,

$$f_X(i) = P[X = i] = c \cdot i = 1/20$$

. Portanto, c = $i/20$.

2.5.2 Calcule o valor esperado de X (obtenha também o valor numérico).

Aplicando a definição, temos:

$$\begin{aligned}E[X] &= \sum_{i=1}^{20} i \cdot f_X(i) \\ &= \sum_{i=1}^{20} i \cdot \frac{1}{20} \\ &= 10.5\end{aligned}$$

2.5.3 Calcule a probabilidade de X ser maior do que seu valor esperado.

Os eventos de interesse que fazer com que $X > E[X]$ são: $\{11, 12, \ldots, 20\}$. Desta forma, a probabilidade de X ser maior do que seu valor esperado é:

$$P[X > E[X]] = P[\{11, 12, \ldots, 20\}] = \sum_{i=11}^{20} \frac{1}{20} = \frac{9}{20} = 45\%$$

2.5.4 Calcule a variância de X (obtenha também o valor numérico).

Calculando $E[X^2]$:

$$\begin{aligned}
E[X^2] &= \sum_{i=1}^{20} i^2 \cdot f_X(i) \\
&= \sum_{i=1}^{20} i^2 \cdot \frac{1}{20} \\
&= \frac{1}{20}(1 + 2^2 + \ldots + 20^2) \\
&= 143.5
\end{aligned}$$

Portanto, o valor da variância é:

$$\begin{aligned}
\sigma_X^2 &= E[X^2] - E^2[X] \\
&= 143.5 - 10.5^2 \\
&= 33.25
\end{aligned}$$

Como podemos ver, o dado com fazer de probabilidade uniforme possui uma variância **maior**. Isso ocorre pois, no primeiro caso, onde a probabilidade da face é proporcional ao seu valor, as chances de sairem faces com valor alto são maiores. Desta forma, a maioria das jogadas terá faces com valores altos e, consequentemente, a variância será menor e no entorno de uma média mais alta.

3 Dado em ação

Considere a versão uniforme do dado acima, ou seja, $P[X = i] = 1 = 20, i = 1, \ldots, 20$. Seja Y uma variável aleatória indicadora da primalidade da face do dado. Ou seja, $Y = 1$ quando o X é um número primo, e $Y = 0$ caso contrário. Responda às perguntas abaixo

3.1 Determine $Y[P = 1]$.

$Y[P = 1]$ ocorre quando qualquer um dos elementos abaixo ocorre:

$$\{2, 3, 5, 7, 11, 13, 17, 19\}$$

Como todos os elementos possuem igual probabilidade, $Y[P = 1] = 8/20 = 2/5$.

3.2 Considere que o dado será jogado n vezes. Seja Y_i a indicadora da primalidade da i-ésima rodada, para $i = 1, \ldots, n$, e defina $Z = \sum_{i=1}^{n} Y_1$. Repare que Z é uma variável aleatória que denota o número de vezes que o resultado é primo. Determine a distribuição de Z, ou seja, $P[Z = k]$, para $k = 0, \ldots, n$. Que distribuição é esta?

Considerando que a variável aleatória Y pode ser modelada como uma Bernoulli com probabilidade $Y[P = 1] = 2/5$, então a variável Z será uma soma de várias Bernoulli:

$$Z = \sum_{i=1}^{n} Bernoulli(2/5)$$

Consequentemente, a variável Z terá uma distribuição Binomial com parâmetros n e $p = 2/5$.

$$Z = Binom(n, 2/5)$$

3.3 Considere que o dado será jogado até que um número primo seja obtido. Seja Y_i a indicadora da primalidade da i-ésima rodada, para $i = 1, \ldots$, e defina $Z = min_{\{i|Y_i=1\}}$. Repare que Z denota o número de vezes que o dado é jogado até que o resultado seja um número primo. Determine a distribuição de Z, ou seja $P[Z = k]$, para $k = 1, \ldots$. Que distribuição é esta?

Neste caso, a variável Y continua sendo modelada como uma Bernoulli de parâmetro $2/5$:

$$Y = Bernoulli(2/5)$$

No entanto, a variável Z terá uma distribuição geométrica com parâmetro $2/5$:

$$Z = Geom(2/5)$$

4 Cobra

Considere três imagens tiradas em uma floresta, I_1, I_2, e I_3. Em apenas uma das imagens existe uma pequena cobra. Um algoritmo de detecão de cobras em imagens detecta a cobra na imagem i com probabilidade α_i. Suponha que o algoritmo não encontrou a cobra na imagem I_1. Defina o espaço amostral e os eventos apropriados e use regra de Bayes para determinar:

4.1 A probabilidade da cobra estar na imagem I_1.

Modelaremos este problema da seguinte forma: os elementos do espaço amostral são imagens que, quando obtidas em sequência, formam eventos. Esses eventos, portanto, são tríados no espaço amostral.

Uma variável aleatória E neste espaço amostral mapeia cada evento no conjunto de números inteiros $\{1, 2, 3\}$. Assim, denotaremos E_i como o algoritmo detectar a cobra no elemento I_i e mapear em i.

Uma segunda variável aleatória diz respeito à existência (ou não) da propriedade "a imagem possui uma cobra". Assim, quando existe uma cobra na imagem i, denotamos $C_i = 1$. Caso contrário, C_i = 0.

Portanto, segue as definições:

- C_i: cobra está na imagem I_i
- E_i: algoritmo detecta a cobra na imagem I_i

onde $i \in \{1, 2, 3\}$. Do enunciado, temos as seguintes probabilidades condicionais:

$$P[E_i|C_i] = \alpha_i, \forall i \in \{1, 2, 3\}$$

Consequentemente,

$$P[\neg E_i|C_i] = (1 - \alpha_i), \forall i \in \{1, 2, 3\}$$

Consideraremos que a probabilidade da cobra estar em qualquer imagem é uniforme, ou seja:

$$P[C_i] = 1/3, \forall i \in \{1, 2, 3\}$$

E, além disso, consideraremos que não há falsos positivos, ou seja:

$$P[C_i|E_i] = 1, \forall i \in \{1, 2, 3\}$$

Queremos descobrir $P[C_1|\neg E_1]$, ou seja, a probabilidade de termos um falso negativo. Pela regra de Bayes, temos que:

$$P[C_1|\neg E_1] = \frac{P[\neg E_1|C_1]P[C_1]}{P[E_1]} = \frac{(1 - \alpha_1) \cdot 1/3}{P[\neg E_1]} = \frac{(1 - \alpha_1) \cdot 1/3}{1 - P[E_1]}$$

Ainda pela regra de Bayes:

$$P[E_1] = \frac{P[E_1|C_1] \cdot P[C_1]}{P[C_1|E_1]} = \frac{\alpha_1 \cdot 1/3}{1}$$

Portanto:

$$P[C_1|\neg E_1] = \frac{1-\alpha_1}{3-\alpha_1}$$

4.2 A probabilidade da cobra estar na imagem I_2.

Neste caso, queremos descobrir $P[C_2|\neg E_1]$:

$$P[C_2|\neg E_1] = \frac{P[\neg E_1|C_2]P[C_1]}{P[E_1]} = \frac{\Big(1 - P[E_1|C_2]\Big)P[C_1]}{P[E_1]}$$

Pela regra da probabilidade total, $P[E_1] = \sum_{j=1}^{3} P[E_1|C_j]P[C_j] = 1/3 \cdot \sum_{j=1}^{3} P[E_1|C_j]$

Pela regra da probabilidade total, $P[C_i] = \sum_{j=1}^{3} P[C_i|E_j]P[E_j] = 1/3$

5 Sem memória

Seja $X \approx Geo(p)$ uma variável aleatória Geométrica com parâmetro p. Mostre que a distribuição geométrica não tem memória. Ou seja, dado que $X > k$, o número de rodadas adicionais até que o evento de interesse ocorra possui a mesma distribuição (dica: formalize esta afirmação).

Para este problema, seria útil termos a CCDF da distribuição exponencial, que nos dá a probabilidade $P[X > k]$. Assim, iniciaremos calculando a CDF desta distribuição:

$$\begin{aligned}
P[X < k] &= \sum_{i=1}^{k} P[X = i] \\
&= \sum_{i=1}^{k} p(1-p)^{i-1} \\
&= p\sum_{i=1}^{k} (1-p)^{i-1} \\
&= p\frac{1 \cdot [(1-p)^{k-1} - 1]}{1 - p - 1} \\
&= 1 - (1-p)^{k-1}
\end{aligned}$$

Consequentemente, a CCDF desta distribuição é dada por:

$$P[X > k] = 1 - P[X \leq k] = (1-p)^{k-1}$$

Por outro lado, queremos analisar o número de rodadas s adicionais dado que um evento ocorreu na rodada k. Assim, utilizando a definição da probabilidade condicional, temos:

$$P[X > s + k|X > k] = \frac{P[X > s + k \cap X > k]}{P[X > k]}$$

Intuitivamente, vemos que a interseção $P[X > s + k \cap X > k]$ para s ¿ 0 é $P[X > s + k]$. Logo,

$$\begin{aligned}
P[X > s + k|X > k] &= \frac{P[X > s + k]}{P[X > k]} \\
&= \frac{(1-p)^{s+k-1}}{(1-p)^{k-1}} \\
&= (1-p)^{s-1} \\
&= P[X > s]
\end{aligned}$$

Portanto, vemos que a distribuição do número de rodadas s adicionais segue a mesma distribuição.

6 Considere que o processo de chegada do ônibus 485 no ponto do CT seja bem representado por um processo de Poisson. Ou seja, $X \approx Poi(\lambda; t)$ denota o número (aleatório) de ônibus que chegam ao ponto em um intervalo de tempo t com taxa média de chegada igual à λ. Assuma que $\lambda = 10$ ônibus por hora.

6.1 Determine a probabilidade de não chegar nenhum ônibus em um intervalo de 30 minutos.

A pdf de uma variável aleatória que segue uma distribuição de Poisson possui a seguinte formulação:

$$f(k, \lambda) = P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$$

Assim, ela nos indica a probabilidade de k eventos ocorrerem em um intervalo t dado que os eventos ocorrem a uma taxa λ (medido em t).

Como o processo de chegada do 485 acontece a uma taxa $\lambda = 10$ ônibus por hora, podemos considerar que em 30 minutos ele ocorre a uma taxa de $\lambda_1 = 5$ ônibus a cada meia hora. Desta forma, para determinar a probabilidade de não chegar nenhum ônibus em um intervalo de 30 minutos, fazemos:

$$P[X = 0 \; para \; \lambda_1] = \frac{\lambda_1^0 e^{-5}}{0!} = e^{-5}$$

6.2 Determine a probabilidade de 3 ônibus chegarem dentro de um intervalo de 5 minutos.

Analogamente, temos que $\lambda_2 = 10 \cdot 5/60 = 5/6$ ônibus a cada 5 minutos. A probabilidade de que 3 ônibus cheguem neste intervalo é:

$$P[X = 3 \; para \; \lambda_2] = \frac{\lambda_2^3 e^{-\lambda_2}}{3!} = 4.19\%$$

6.3 Determine a probabilidade da média ocorrer, ou seja, de chegarem exatamente 10 ônibus em uma hora.

$$P[X = 10 \; para \; \lambda] = \frac{\lambda^{10} e^{-\lambda}}{10!} = 12.5\%$$

7 Propriedades

Seja X e Y duas variáveis aleatórias discretas. Mostre as seguintes equivalências usando as definições:

7.1 $E[X] = E[E[X|Y]]$, conhecida como regra da torre da esperança

Seja O_X e O_Y os espaços de valoress que as variáveis X e Y podem ter, respectivamente. Assim, a probabilidade condicional pode ser definida da seguinte forma:

$$P[X|Y] = \frac{P[X = x; Y = y]}{P[Y = y]}$$

Além disso, pela definição temos a seguinte formulação de valor esperado:

$$E[X|Y = y] = \sum x \cdot P[X|Y = y], \forall y \in O_Y$$

Aqui, considera-se que Y é uma variável aleatória que faz com que o valor esperado $E[X|Y = y]$ também seja uma variável aleatória. Desta forma, pode-se aplicar o valor esperado desta expressão:

$$
\begin{aligned}
E[E[X|Y=y]] &= \sum_{x\in O_X} E[X|Y=y]P[Y=y] \\
&= \sum_{x\in O_X}\sum_{y\in O_Y} x\cdot P[X|Y=y]P[Y=y] \\
&= \sum_{x\in O_X}\sum_{y\in O_Y} x\cdot \frac{P[X=x;Y=y]}{P[Y=y]}P[Y=y] \\
&= \sum_{x\in O_X}\sum_{y\in O_Y} x\cdot P[X=x;Y=y] \\
&= \sum_{x\in O_X} x\cdot P[X=x] \\
&= E[X]
\end{aligned}
$$

7.2 $Var[X] = E[X^2] - E[X]^2$

$$
\begin{aligned}
Var[X] &= E[(X-E[X])^2] \\
&= E[X^2 - 2X\cdot E[X] + E[X]^2] \\
&= E[X^2] - 2E[X]E[X] + E[X]^2 \\
&= E[X^2] - E[X]^2
\end{aligned}
$$

8 Caras em sequência *

Considere uma moeda enviesada, tal que o probabilidade do resultado ser cara é p (e coroa $1-p$). Considere o número de vezes que a moeda precisa ser jogada para obtermos k caras consecutivas. Por exemplo, na sequência $COCOCCOOCOCCC$ a moeda teve que ser jogada 13 vezes até o aparecimento de $k=3$ caras consecutivas, onde $C = cara$ e $O = coroa$. Seja N_k a variável aleatória que denota esta quantidade. Qual é o número médio de vezes que a moeda precisa ser jogada para obtermos k caras consecutivas, ou seja, qual é o valor esperado de N_k? Dica: monte uma recursão e use a regra da torre da esperança.

Considere Y uma variável indicadora auxiliar que indica a 1ª ocorrência de uma coroa. Neste caso, Y segue uma distribuição geométrica com parâmetro $1-p$ que pode assumir valores de 0 a ∞.

Caso $Y \geq k+1$, então k caras seguidas ocorreram. Caso contrário, ao menos 1 coroa aconteceu no índice $Y=y$ e a contagem de caras consecutivas é reiniciada.

Isto pode ser formulado matematicamente da seguinte forma:

$$
E[N_k|Y] = \begin{cases} y + E[N_k], y \leq k \\ k, y \geq k+1 \end{cases}
$$

O valor de $E[N_k]$, portanto, pode ser obtido aplicando-se a propriedade da torre da esperança:

$$
\begin{aligned}
E[N_k] &= E[E[N_k|Y]] \\
&= \sum_{y=1}^{\infty} E[N_k|Y=y]P[Y=y] \\
&= \sum_{y=1}^{k}(y+E[N_k])P[Y=y] + \sum_{y=k+1}^{\infty} kP[Y=y]
\end{aligned}
$$

Uma vez que Y segue uma distribuição geométrica: $P[Y=y] = (1-(1-p))^{y-1} = p^{y-1}$
Assim,

$$
\begin{aligned}
E[N_k] &= \sum_{y=1}^{k}(y+E[N_k])p^{y-1} + \sum_{y=k+1}^{\infty} kp^{y-1} \\
&= \sum_{y=1}^{k} yp^{y-1} + E[N_k]\sum_{y=1}^{k} p^{y-1} + k\sum_{y=k+1}^{\infty} p^{y-1}
\end{aligned}
$$

O terceiro termo da soma é uma PG infinita que começa em p^k com razão p. Assim,

$$k\sum_{y=k+1}^{\infty} p^{y-1} = k\frac{p^k}{1-p}$$

O segundo termo da soma é uma PG finita com razão p que vai de 1 a $k-1$:

$$E[N_k]\sum_{y=1}^{k} p^{y-1} = E[N_k]\Big(\frac{1\cdot(p^k-1)}{p-1}\Big)$$

O primeito termo da soma é uma PAG finita com razão da PG p e razão da PA 1. Chamaremos este termo de PAG.
Portanto,

$$E[N_k] = \frac{PAG + k\frac{p^k}{1-p}}{1-\frac{p^k-1}{p-1}}$$