CE064 - MODELOS MARKOVIANOS
(TÓPICOS ESPECIAIS EM ESTATÍSTICA IV)
Departamento de Estatística
Universidade Federal do Paraná

Lista No.1

Henrique Aparecido Laureano
GRR: 20115307

Março de 2016

Sumário

Exercício 1 [7, pg 53]

Uma urna contém inicialmente cinco bolas pretas e cinco bolas brancas. A seguinte experiência é repetida indefinidamente: Uma bola é retirada da urna; se a bola for branca, ela é colocada de volta na urna, caso contrário ela é deixada de fora. Considere X_n o número de bolas restantes na urna após a n-ésima retirada da urna.

a) É $\{X_n\}$ um processo de Markov? Se assim for, encontrar as probabilidades de transição adequadas e faça o grafo correspondente.

$\{X_n\}$ é um processo de Markov? Sim. Por quê?

De maneira informal podemos dizer que um processo estocástico é uma sequência de variáveis aleatórias que evoluem de alguma maneira no tempo.

Um processo estocástico é um processo de Markov quando ele satisfaz a propriedade de Markov.

A propriedade de Markov é definida pela exigência de que

$$P(X_{n+1} = x_{n+1}|X_0 = x_0, \cdots, X_n = x_n) = P(X_{n+1} = x_{n+1}|X_n = x_n),$$

ou seja, dado o estado atual, os estados passados não tem influência sobre o futuro.

Tá, mas por que $\{X_n\}$ satisfaz a propriedade de Markov?

Vamos elaborar um cenário:

Primeiro algumas definições:

A retirada $n+1$ é a 4a retirada e $x_{n+1} = 8$.

Na primeira retirada temos uma bola branca, logo, ela volta pra urna. Na segunda retirada temos uma bola preta, que saí da urna, e na terceira outra bola branca, que volta pra urna.

O que temos pode ser representado por $P(X_4 = 8|X_0 = 10, X_1 = 10, X_2 = 9, X_3 = 9)$.

Então, se

$$P(X_4 = 8|X_0 = 10, X_1 = 10, X_2 = 9, X_3 = 9) = P(X_4 = 8|X_3 = 9),$$

$\{X_n\}$ é um processo de Markov.

Dada as definições da experiência, se a urna tem 9 bolas (4 pretas e 5 brancas), a probabilidade dela ficar com 8 bolas é de 4/9.

Logo,

$$P(X_4 = 8|X_0 = 10, X_1 = 10, X_2 = 9, X_3 = 9) = \frac{4}{9} = P(X_4 = 8|X_3 = 9)$$

e $\{X_n\}$ é um processo de Markov.

Quais as probabilidades de transição?

Para facilitar a visualização, é apresentada a matriz de probalidades de transição P

$$P = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & \cdots & d \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ d \end{array} & \begin{pmatrix} p_{0,0} & p_{0,1} & p_{0,2} & \cdots & p_{0,d} \\ p_{1,0} & p_{1,1} & p_{1,2} & \cdots & p_{1,d} \\ p_{2,0} & p_{2,1} & p_{2,2} & \cdots & p_{2,d} \\ \vdots & \vdots & \vdots & & \vdots \\ p_{d,0} & p_{d,1} & p_{d,2} & \cdots & p_{d,d} \end{pmatrix} \end{array}$$

em que $S = \{0, 1, 2, \cdots, d\}$ é o espaço de estados, e $p_{0,0}$, por exemplo, é a probabilidade de ir do estado 0 pro estado 0 em um passo, i.e., num instante de tempo, nesse caso a retirada de uma bola da urna.

Na experiência aqui estudada o espaço de estados é dado por $S = \{10, 9, 8, 7, 6, 5\}$, i.e., os possíveis números de bolas restantes na urna. Como as bolas brancas são sempre recolocadas, o mínimo possível é 5.

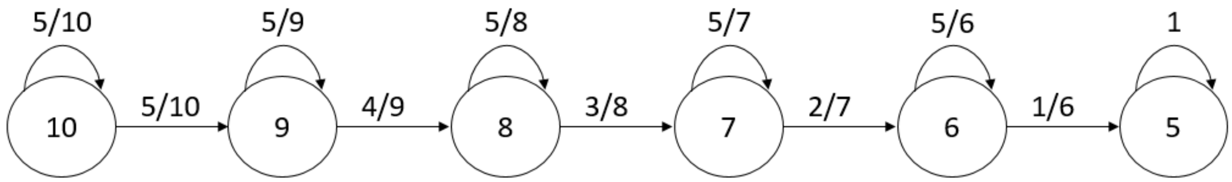
A matriz de probabilidades de transição é dada por

$$P = \begin{array}{c c} & \begin{array}{cccccc} 10 & 9 & 8 & 7 & 6 & 5 \end{array} \\ \begin{array}{c} 10 \\ 9 \\ 8 \\ 7 \\ 6 \\ 5 \end{array} & \begin{pmatrix} 5/10 & 5/10 & 0 & 0 & 0 & 0 \\ 0 & 5/9 & 4/9 & 0 & 0 & 0 \\ 0 & 0 & 5/8 & 3/8 & 0 & 0 \\ 0 & 0 & 0 & 5/7 & 2/7 & 0 \\ 0 & 0 & 0 & 0 & 5/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

Começamos com 10 bolas na urna e temos probabilidade 1/2 de permanecer com 10, já que temos 5 bolas brancas. A probabilidade de ficar com 9 também é 1/2, já que as outras cinco bolas são pretas.

Essa mesma ideia é aplicada para os demais estados, i.e., para os demais números de bolas na urna.

O grafo correspodente é apresentado abaixo.



Nas flechas temos as probabilidades de transição.

b) Será que as probabilidades de transição dependem de n?

As probabilidades já mostradas são em um passo, i.e., em uma retirada da urna.

Quando o número de retiradas da urna, n, aumenta, as probabilidades de transição podem ser alteradas.

Um dado **Teorema** em processos de Markov diz que para cada $n = 1, 2, \cdots$, a n-ésima potência P^n da matriz de probabilidades de transição P tem na linha i e coluna j a probabilidade $p_{i,j}^{(n)}$, i.e., a probabilidade do processo passar do estado i para o estado j em n passos.

Exemplo em R:

No código abaixo é mostrado que em 2 passos, $n = 2$, as probabilidades de transição mudam, ou seja, elas dependem de n.

```
library(markovchain)
# mpt: matriz de probabilidades de transição
mpt <- new("markovchain", name = "Experiência da urna"
           , states = as.character(10:5)
           , transitionMatrix = matrix(c(5/10, 5/10, 0, 0, 0, 0
                                         , 0, 5/9, 4/9, 0, 0, 0
                                         , 0, 0, 5/8, 3/8, 0, 0
                                         , 0, 0, 0, 5/7, 2/7, 0
                                         , 0, 0, 0, 0, 5/6, 1/6
                                         , 0, 0, 0, 0, 0, 1)
                                       , byrow = TRUE, ncol = 6))
mpt**2
```

```
Experiência da urna^2
 A  6 - dimensional discrete Markov Chain with following states
 10 9 8 7 6 5
 The transition matrix   (by rows)  is defined as follows
     10         9         8         7         6          5
10 0.25 0.5277778 0.2222222 0.0000000 0.0000000 0.00000000
```

```
9  0.00 0.3086420 0.5246914 0.1666667 0.0000000 0.00000000
8  0.00 0.0000000 0.3906250 0.5022321 0.1071429 0.00000000
7  0.00 0.0000000 0.0000000 0.5102041 0.4421769 0.04761905
6  0.00 0.0000000 0.0000000 0.0000000 0.6944444 0.30555556
5  0.00 0.0000000 0.0000000 0.0000000 0.0000000 1.00000000
```

Exercício 2 [18, pg 54]

Seja $\{X_n : n \geq 0\}$ uma Cadeia de Markov. Mostre que

$$P(X_0 = x_0 | X_1 = x_1, \cdots, X_n = x_n) = P(X_0 = x_0 | X_1 = x_1).$$

Temos aqui uma probabilidade condicional, $P(X_0 = x_0 | X_1 = x_1, \cdots, X_n = x_n)$, logo

$$P(X_0 = x_0 | X_1 = x_1, \cdots, X_n = x_n) = \frac{P(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n)}{P(X_1 = x_1, \cdots, X_n = x_n)}.$$

Já um dado **Teorema** nos diz que para uma Cadeia de Markov $\{X_n : n \geq 0\}$,

$$\begin{aligned} &P(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n) = \\ &\qquad P(X_n = x_n)P(X_{n-1} = x_{n-1} | X_n = x_n) \cdots P(X_1 = x_1 | X_2 = x_2)P(X_0 = x_0 | X_1 = x_1). \end{aligned}$$

Dessa forma

$$\begin{aligned} &P(X_0 = x_0 | X_1 = x_1, \cdots, X_n = x_n) = \\ &\qquad \frac{P(X_n = x_n)P(X_{n-1} = x_{n-1} | X_n = x_n) \cdots P(X_1 = x_1 | X_2 = x_2)P(X_0 = x_0 | X_1 = x_1)}{P(X_n = x_n)P(X_{n-1} = x_{n-1} | X_n = x_n) \cdots P(X_1 = x_1 | X_2 = x_2)}. \end{aligned}$$

Ou seja,

$$P(X_0 = x_0 | X_1 = x_1, \cdots, X_n = x_n) = P(X_0 = x_0 | X_1 = x_1).$$

Exercício 3 [4, pg 79]

Uma Cadeia de Markov a três estados tem a seguinte como matriz de probabilidades de transição:

$$P = \begin{pmatrix} 0.25 & 0.5 & 0.25 \\ 0.4 & 0.6 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

a) Qual é o valor aproximado de $p_{1,3}^{(100)}$? Que interpretação você dá a esse resultado?

$P^{(100)}$:

```
mpt <- new("markovchain", name = "Cadeia a três estados"
           , states = as.character(1:3)
           , transitionMatrix = matrix(c(.25, .5, .25
                                        , .4, .6, 0
                                        , 1, 0, 0)
                                      , byrow = TRUE, ncol = 3))
mpt**100

Cadeia a três estados^100
 A  3 - dimensional discrete Markov Chain with following states
 1 2 3
 The transition matrix   (by rows)  is defined as follows
    1   2   3
1 0.4 0.5 0.1
2 0.4 0.5 0.1
3 0.4 0.5 0.1
```

O valor aproximado de $p_{1,3}^{(100)}$ é igual a 0.1.

Isso significa que 0.1 é a probabilidade da cadeia passar do estado 1 para o estado 3 em 100 passos.

b) Qual é a probabilidade de que após o terceiro passo a cadeia esteja no estado 3 se o vetor de probabilidades inicial é $(1/3, 1/3, 1/3)$?

A probabilidade desejada pode ser obtida com a aplicação da seguinte equação

$$P(X_{n+1} = y) = \sum_{x \in S} P(X_n = x) p_{x,y}. \tag{1}$$

Conhecemos a distribuição de X_0, $(1/3, 1/3, 1/3)$, logo podemos usar o resultado em (1) para encontrar a distribuição de X_1. Em seguida, sabendo a distribuição de X_1, utilizamos (1) novamente para encontrar a distribuição de X_2. Da mesma forma, encontramos a distribuição de X_3 aplicando mais uma vez a relação encontrada em (1).

Com o código em R abaixo temos a distribuição de X_3

```
round(
  c(1/3, 1/3, 1/3)*(mpt**3)
  , 3)

         1     2     3
[1,] 0.427 0.481 0.092
```

A probabilidade de que após o terceiro passo a cadeia esteja no estado 3 é de 0.092 (9%, arredondando).

Exercício 4 [15 pg 81]

Resposta Imunológica **Um estudo de resposta imunológica em coelhos classificou os coelhos em quatro grupos de acordo com a intensidade da reposta imune[1]. De uma semana para a seguinte, os coelhos alteram a classificação de um grupo para o outro, de acordo com a seguinte matriz de transição:**

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 5/7 & 2/7 & 0 & 0 \\ 0 & 1/2 & 1/3 & 1/6 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/4 & 3/4 \end{pmatrix} \end{array}$$

a) Qual a proporção dos coelhos no grupo 1 que ainda estavam no grupo 1 cinco semanas mais tarde?

Em uma semana, um passo, 5/7 (71.43%) dos coelhos no grupo 1 ainda estavam no grupo 1.

```
mpt <- new("markovchain", name = "Resposta imunológica"
           , states = as.character(1:4)
           , transitionMatrix = matrix(c(5/7, 2/7, 0, 0
                                        , 0, 1/2, 1/3, 1/6
                                        , 0, 0, 1/2, 1/2
                                        , 0, 0, 1/4, 3/4)
                                      , byrow = TRUE, ncol = 4))
mpt**5
```

[1] McGilchrist, C.A., C.W. Aisbett, and S. Cooper. (1989.) "A Markov Transition Model in the Analysis of the Immune Response". Journal of Theoretical Biology, Vol. 138, pp. 17-21.

```
Resposta imunológica^5
 A  4 - dimensional discrete Markov Chain with following states
 1 2 3 4
 The transition matrix   (by rows)  is defined as follows
          1         2         3         4
1 0.1859344 0.2062459 0.2642237 0.3435960
2 0.0000000 0.0312500 0.3430990 0.6256510
3 0.0000000 0.0000000 0.3339844 0.6660156
4 0.0000000 0.0000000 0.3330078 0.6669922
```

Em cinco semanas, 18.59% dos coelhos no grupo 1 ainda estavam no grupo 1.

b) Na primeira semana, havia nove coelhos no primeiro grupo, 4 no segundo e nenhum no terceiro e quarto grupo. Quantos coelhos seria de esperar em cada grupo após 4 semanas?

Podemos imaginar o número de coelhos em cada grupo na primeira semana como uma espécie de distribuição inicial.

```
round(
  c(9, 4, 0, 0)*(mpt**4)
  , 0)

     1 2 3 4
[1,] 2 3 3 5
```

Após 4 semanas é de se esperar 2 coelhos no primeiro grupo, 3 no segundo e terceiro grupo, e cinco no quarto grupo.

c) Ao investigar a matriz de transição elevada a potências cada vez maiores, faça uma suposição razoável para a probabilidade a longo tempo que um coelho no grupo 1 ou 2 ainda esteja no grupo 1 ou 2 depois de um tempo arbitrariamente longo. Explique por que esta resposta é razoável.

No código R fornecido a seguir a matriz de probabilidades de transição foi elevada a três diferentes potências, 100, 1250 e 2500.

Na primeira potência, 100, que já é alta, as probabilidades dos coelhos permanecerem nos grupos 1 e 2 é bem baixa, 2 e 3e-15. Conforme aumentamos a potências essa probabilidades vão diminuindo, chegando a ser igual a zero com a potência 2500.

Ou seja, conforme a potência aumenta as probabilidades diminuem, chegando a zero.

Mas o que isso significa e por que isso é razoável de se considerar?

Cada potência é um passo, i.e., uma semana. Cada estado é um grupo definido de acordo com a intensidade da resposta imunológica dos ratos, logo, conforme o tempo passa é de se esperar que a resposta imunológica dos ratos diminua, já que eles vão ficando mais velhos, chegando ao ponto de ser impossivel (probabilidade zero) permanecer nos grupos 1 e 2.

```
mpt**100
```

```
Resposta imunológica^100
 A  4 - dimensional discrete Markov Chain with following states
 1 2 3 4
 The transition matrix   (by rows)  is defined as follows
             1            2         3         4
1 2.438914e-15 3.251885e-15 0.3333333 0.6666667
2 0.000000e+00 7.888609e-31 0.3333333 0.6666667
3 0.000000e+00 0.000000e+00 0.3333333 0.6666667
4 0.000000e+00 0.000000e+00 0.3333333 0.6666667
```

```
mpt**1250
```

```
Resposta imunológica^1250
 A  4 - dimensional discrete Markov Chain with following states
 1 2 3 4
 The transition matrix   (by rows)  is defined as follows
              1             2         3         4
1 2.187537e-183 2.916716e-183 0.3333333 0.6666667
2  0.000000e+00  0.000000e+00 0.3333333 0.6666667
3  0.000000e+00  0.000000e+00 0.3333333 0.6666667
4  0.000000e+00  0.000000e+00 0.3333333 0.6666667
```

```
mpt**2500
```

```
Resposta imunológica^2500
 A  4 - dimensional discrete Markov Chain with following states
 1 2 3 4
 The transition matrix   (by rows)  is defined as follows
  1 2         3         4
1 0 0 0.3333333 0.6666667
2 0 0 0.3333333 0.6666667
3 0 0 0.3333333 0.6666667
4 0 0 0.3333333 0.6666667
```

Exercício 5 [16, pg 82]

Criminologia **Num estudo com homens criminosos em Filadélfia descobriram que a probabilidade de que um tipo de ataque seja seguido por um outro tipo pode ser descrito pela seguinte matriz de transição[2].**

$$P = \begin{array}{c} \\ \text{Outro} \\ \text{Injúria} \\ \text{Roubo} \\ \text{Dano} \\ \text{Misto} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{Outro} & \text{Injúria} & \text{Roubo} & \text{Dano} & \text{Misto} \end{array} \\ \begin{pmatrix} 0.645 & 0.099 & 0.152 & 0.033 & 0.071 \\ 0.611 & 0.138 & 0.128 & 0.033 & 0.090 \\ 0.514 & 0.067 & 0.271 & 0.030 & 0.118 \\ 0.609 & 0.107 & 0.178 & 0.064 & 0.042 \\ 0.523 & 0.093 & 0.183 & 0.022 & 0.179 \end{pmatrix} \end{array}$$

a) Para um criminoso que comete roubo, qual é a probabilidade de que o seu próximo crime também seja um roubo?

0.271.

A probabilidade de que um criminoso no estado 'Roubo' permaneça em um passo no estado 'Roubo' é de 27%.

b) Para um criminoso que comete roubo, qual é a probabilidade de que seu segundo crime depois do atual também seja um roubo?

Estamos interessados na probabilidade de que um criminoso que se encontra no estado 'Roubo' esteja no estado 'Roubo' no segundo passo.

No código abaixo as probabilidades de transição em dois passos são calculadas.

```r
mpt <- new("markovchain", name = "Criminologia"
           , states = c("Outro", "Injúria", "Roubo", "Dano", "Misto")
           , transitionMatrix = matrix(c(0.645, 0.099, 0.152, 0.033, 0.071
                                        , 0.611, 0.138, 0.128, 0.033, 0.090
                                        , 0.514, 0.067, 0.271, 0.030, 0.118
                                        , 0.609, 0.107, 0.178, 0.064, 0.042
                                        , 0.523, 0.093, 0.183, 0.022, 0.179)
                                      , byrow = TRUE, ncol = 5))
mpt**2
```

```
Criminologia^2
```

[2]Stander, Julian, et al. (1989). "Markov Chain Analysis and Specialization in Criminal Careers". The British Journal of Criminology, Vol. 29, No.4, pp. 319-335.

```
 A  5 - dimensional discrete Markov Chain with following states
 Outro Injúria Roubo Dano Misto
 The transition matrix   (by rows)  is defined as follows
           Outro  Injúria    Roubo     Dano    Misto
Outro   0.611872 0.097835 0.170771 0.032786 0.086736
Injúria 0.611372 0.100010 0.167568 0.032649 0.088401
Roubo   0.591745 0.092473 0.187079 0.031819 0.096884
Dano    0.610616 0.097737 0.173580 0.033988 0.084079
Misto   0.595235 0.095873 0.177666 0.031164 0.100062
```

A probabilidade de interesse é de 18.7%.

c) Se essas tendências continuarem, quais são as probabilidades de longo prazo para cada tipo de crime?

No código abaixo a matriz de probabilidades de transição é elevada a três diferentes potências, 5, 10 e 50. Independente da potência as probabilidades são iguais, com pequenas diferenças a partir da quarta, quinta casa decimal. Conforme a potência aumenta as diferenças vão diminuindo.

```
mpt**5

Criminologia^5
 A  5 - dimensional discrete Markov Chain with following states
 Outro Injúria Roubo Dano Misto
 The transition matrix   (by rows)  is defined as follows
            Outro    Injúria     Roubo       Dano      Misto
Outro   0.6068045 0.09693739 0.1739969 0.03249919 0.08976204
Injúria 0.6068110 0.09693969 0.1739905 0.03249940 0.08975945
Roubo   0.6067279 0.09691951 0.1740497 0.03249389 0.08980899
Dano    0.6068019 0.09693608 0.1740006 0.03249930 0.08976212
Misto   0.6067506 0.09692644 0.1740299 0.03249491 0.08979816

mpt**10

Criminologia^10
 A  5 - dimensional discrete Markov Chain with following states
 Outro Injúria Roubo Dano Misto
 The transition matrix   (by rows)  is defined as follows
            Outro    Injúria     Roubo      Dano      Misto
Outro   0.6067869 0.09693348 0.1740085 0.0324979 0.08977320
Injúria 0.6067869 0.09693348 0.1740085 0.0324979 0.08977320
Roubo   0.6067869 0.09693348 0.1740085 0.0324979 0.08977321
Dano    0.6067869 0.09693348 0.1740085 0.0324979 0.08977320
Misto   0.6067869 0.09693348 0.1740085 0.0324979 0.08977321
```

```
mpt**50
```

```
Criminologia^50
 A  5 - dimensional discrete Markov Chain with following states
 Outro Injúria Roubo Dano Misto
 The transition matrix   (by rows)  is defined as follows
            Outro    Injúria     Roubo      Dano     Misto
Outro   0.6067869 0.09693348 0.1740085 0.0324979 0.0897732
Injúria 0.6067869 0.09693348 0.1740085 0.0324979 0.0897732
Roubo   0.6067869 0.09693348 0.1740085 0.0324979 0.0897732
Dano    0.6067869 0.09693348 0.1740085 0.0324979 0.0897732
Misto   0.6067869 0.09693348 0.1740085 0.0324979 0.0897732
```

Todas as probabilidades de transição são idênticas, independente do estado de origem.

A longo prazo a maior probabilidade é de que um criminoso pratique um outro tipo de crime, com 60.68%. A segunda maior probabilidade é de que um criminoso cometa um roubo, com 17.4%.

Com menor probabilidade, 3.25%, se espera que um criminoso pratique um dano.