Paola de Oliveira Prado

Este trabalho apresenta um estudo de uma simulação de uma cadeia de alcance k com 100 replicações para k=0, k=1, k=2 e k=3, variando o tamanho amostral N em 100, 1.000 e 10.000. Para cada replicação da combinação k e N foi calculado o BIC (1) para k em 0, 1, 2 e 3 afim de verificar a proporção de acerto do estimador, de superestimação e subestimação para cada combinação de k e N.

$$BIC(k, X_1^N) = log \prod_{a \in A} \prod_{u \in A^k} p(a|u)^{N(ua)} \; - \frac{1}{2} \; [|A|^k \; (|A| - 1) \, log \, n] \qquad (1)$$

$$\hat{k} = \arg\max\{BIC(k, X_1^N) : 0 \; \leq k \leq \; log_{|A|}n$$

- **Caso 1**: Simulando 100 replicações de cadeia com alcance k=0

Matriz de transição: $\begin{array}{cc} 0 & 1 \\ {[0,3} & {0,7]} \end{array}$

Tabela 1: Proporção de acerto, de superestimação e subestimação do estimador.

N	k			
	0	1	2	3
100	88%	11%	1%	0
1.000	0	100%	0	0
10.000	0	100%	0	0

Simulando o processo com k=0, podemos notar que para a amostra de tamanho 100, a probabilidade de acerto do estimador foi de 88%, porém quando aumentou a amostra para 1.000 e 10.000, não houve acerto, em 100% das vezes houve superestimação do estimador para um alcance, ou seja, k=1.

- **Caso 2**: Simulando 100 replicações de cadeia com alcance k=1

Matriz de transição: $\begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 0,4 & 0,6 \\ 0,2 & 0,8 \end{bmatrix} \end{array}$

Estado inicial $\mathrm{X}_0 = 0$

Tabela 2: Proporção de acerto, de superestimação e subestimação do estimador.

N	k			
	0	1	2	3
100	51%	45%	4%	0
1.000	0	99%	1%	0
10.000	0	100%	0	0

Estado inicial $X_0 = 1$

Tabela 3: Proporção de acerto, de superestimação e subestimação do estimador.

N	k			
	0	1	2	3
100	58%	40%	2%	0
1.000	0	100%	0	0
10.000	0	100%	0	0

Para este caso, foi feita a simulação para os dois estados iniciais possíveis, em ambos os casos, para a amostra de tamanho 100, a proporção de acerto do estimador ficou por volta de 40%, enquanto a proporção de subestimação, para k=0, foi aproximadamente 50%, e a proporção de superestimação, para k=2, em torno de 3%. Quando foi aumentado o tamanho da amostra para 1.000 e 10.000, a proporção de acerto do estimador foi de aproximadamente 100%.

- **Caso 3**: Simulando 100 replicações de cadeia com alcance k=2

Matriz de transição: $\begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} & \begin{bmatrix} 0{,}5 & 0{,}5 \\ 0{,}2 & 0{,}8 \\ 0{,}3 & 0{,}7 \\ 0{,}4 & 0{,}6 \end{bmatrix} \end{array}$

Estado inicial $X_0 = c(0{,}0)$

Tabela 4: Proporção de acerto, de superestimação e subestimação do estimador.

N	k			
	0	1	2	3
100	67%	14%	19%	0
1.000	0	94%	6%	0
10.000	0	97%	3%	0

Estado inicial $X_0 = c(0,1)$

Tabela 5: Proporção de acerto, de superestimação e subestimação do estimador.

N	k			
	0	1	2	3
100	72%	8%	20%	0
1.000	0	90%	10%	0
10.000	0	96%	4%	0

Estado inicial $X_0 = c(1,0)$

Tabela 6: Proporção de acerto, de superestimação e subestimação do estimador.

N	k			
	0	1	2	3
100	72%	8%	20%	0
1.000	0	91%	9%	0
10.000	0	97%	3%	0

Estado inicial $X_0 = c(1,1)$

Tabela 7: Proporção de acerto, de superestimação e subestimação do estimador.

N	k			
	0	1	2	3
100	68%	12%	20%	0
1.000	0	91%	9%	0
10.000	0	97%	3%	0

Para este caso, foi feita a simulação para os quatros estados iniciais possíveis, em ambos os casos, para a amostra de tamanho 100, a proporção de acerto do estimador foi de 20%, enquanto a probabilidade de subestimação, para k=0, foi de aproximadamente 70%, e para k=1 foi de aproximadamente 10%. Quando aumentou a amostra para 1.000 e 10.000, a proporção de acerto do estimador foi de aproximadamente 10% e 3%, respectivamente, enquanto isso a proporção de subestimação, para k=1, variou entre 90% e 97%. Não houve casos de superestimação.

- **Caso 4**: Simulando 100 replicações de cadeia com alcance k=3

Matriz de transição: $\begin{array}{c} \\ 000 \\ 001 \\ 010 \\ 011 \\ 100 \\ 101 \\ 110 \\ 111 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 0,5 & 0,5 \\ 0,4 & 0,6 \\ 0,4 & 0,6 \\ 0,3 & 0,7 \\ 0,5 & 0,5 \\ 0,3 & 0,7 \\ 0,3 & 0,7 \\ 0,3 & 0,7 \end{bmatrix} \end{array}$

Estado inicial $X_0 = c(0,1,0)$

Tabela 8: Proporção de acerto, de superestimação e subestimação do estimador.

N	k			
	0	1	2	3
100	81%	16%	3%	0
1.000	0	99%	1%	0
10.000	0	100%	0	0

Estado inicial $X_0 = c(1,1,1)$

Tabela 9: Proporção de acerto, de superestimação e subestimação do estimador.

N	k			
	0	1	2	3
100	82%	16%	2%	0
1.000	0	99%	1%	0
10.000	0	100%	0	0

Para este caso, foi feita a simulação para apenas dois estados iniciais, em todas as simulações com os diferentes tamanhos de amostra o estimador não acertou nenhuma vez. Para amostras de tamanho 100, a proporção de subestimação do estimador foi cerca de 80%, para k=0, 16%, para k=1, e 2% para k=2.

Para amostras de tamanho 1.000 a proporção de subestimação do estimador foi cerca de 99%, para k=1, e 1%, para k=2, e de tamanho 10.000 a proporção de subestimação foi de 100% para k=1.

Considerações Finais

Pode-se notar que independente do estado inicial, as proporções de acerto, subestimação e superestimação do estimador acabam dando o resultado equivalente. O estimador foi mais eficiente para simulação de k=0, com amostras de tamanho 100, e para simulação de k=1, para amostras de tamanho acima de 1000. Na simulação de k=2 o estimador foi pouco eficiente e na simulação de k=3 não foi eficiente.