Translation from MC-style to our usual notation

$$P(X_{n+1} = i_{n+1} \,|\, X_n = i_n, \ldots, X_0 = i_0)$$
$$= f_{X_{n+1} \,|\, X_n, \ldots, X_0}(i_{n+1} \,|\, i_n, \ldots, i_0)$$

$$p_{ij} = P(X_{n+1} = j \,|\, X_n = i) = f_{X_{n+1} \,|\, X_n}(j \,|\, i)$$

$$a_i = P(X_0 = i) = f_{X_0}(i)$$

The Joint Mass Function of a MC

$$f_{X_0,X_1,X_2,X_3,X_4}(x_0,x_1,x_2,x_3,x_4)$$
$$= f_{X_0}(x_0) f_{X_1 \,|\, X_0}(x_1 \,|\, x_0) f_{X_2 \,|\, X_0,X_1}(x_2 \,|\, x_0,x_1)$$
$$\cdot f_{X_3 \,|\, X_0,X_1,X_2}(x_3 \,|\, x_0,x_1,x_2)$$
$$\cdot f_{X_4 \,|\, X_0,X_1,X_2,X_3}(x_4 \,|\, x_0,x_1,x_2,x_3)$$
$$= f_{X_0}(x_0) f_{X_1 \,|\, X_0}(x_1 \,|\, x_0) f_{X_2 \,|\, X_1}(x_2 \,|\, x_1)$$
$$\cdot f_{X_3 \,|\, X_2}(x_3 \,|\, x_2) f_{X_4 \,|\, X_3}(x_4 \,|\, x_3)$$

(if Markov)

$$= f_{X_0}(x_0) f(x_1 \,|\, x_0) f(x_2 \,|\, x_1) f(x_3 \,|\, x_2) f(x_4 \,|\, x_3)$$

(if time homogeneous)

$$= a_{x_0}\, p_{x_0 x_1}\, p_{x_1 x_2}\, p_{x_2 x_3}\, p_{x_3 x_4}$$

(in MC-style notation)

Finding Marginal Distributions

Marginal for X_4

$$f_{X_4}(x_4) = P(X_4 = x_4)$$

$$= \sum_{x_3=1}^{m} \sum_{x_2=1}^{m} \sum_{x_1=1}^{m} \sum_{x_0=1}^{m} a_{x_0} p_{x_0 x_1} p_{x_1 x_2} p_{x_2 x_3} p_{x_3 x_4}$$

$$= \sum_{x_3=1}^{m} \sum_{x_2=1}^{m} \sum_{x_1=1}^{m} \left(\sum_{x_0=1}^{m} a_{x_0} p_{x_0 x_1} \right) p_{x_1 x_2} p_{x_2 x_3} p_{x_3 x_4}$$

$$= \sum_{x_3=1}^{m} \sum_{x_2=1}^{m} \sum_{x_1=1}^{m} b_{x_1} p_{x_1 x_2} p_{x_2 x_3} p_{x_3 x_4}$$

where $\mathbf{b} = \mathbf{a}\mathbf{P}$

($\mathbf{a}$,$\mathbf{b}$ are row vectors; $\mathbf{P}$ is a square matrix)

$$= \sum_{x_3=1}^{m} \sum_{x_2=1}^{m} \left(\sum_{x_1=1}^{m} b_{x_1} p_{x_1 x_2} \right) p_{x_2 x_3} p_{x_3 x_4}$$

$$= \sum_{x_3=1}^{m} \sum_{x_2=1}^{m} c_{x_2} p_{x_2 x_3} p_{x_3 x_4}$$

where $\mathbf{c} = \mathbf{b}\mathbf{P} = (\mathbf{a}\mathbf{P})\mathbf{P} = \mathbf{a}\mathbf{P}^2$

$$= \sum_{x_3=1}^{m} \left(\sum_{x_2=1}^{m} c_{x_2}\, p_{x_2 x_3} \right) p_{x_3 x_4}$$

$$= \sum_{x_3=1}^{m} d_{x_3}\, p_{x_3 x_4}$$

where $d = cP = (aP^2)P = aP^3$

$$= e_{x_4}$$

where $e = dP = (aP)^3 P = aP^4$.

Requires computation $O(4m^2)$
(versus $O(m^5)$ if no structure).

In general, computing the marginal f_{X_n} requires computation $O(nm^2)$ (versus $O(m^{n+1})$ if no structure).

Result: $P(X_n = i) = (aP^n)_i =$ the i-th entry of the row vector aP^n.