CONSIDER $R=2$. (2 action Case)

$$p_1(n+1) = (1-a)\, p_1 \qquad \text{if} \quad \alpha = \alpha_2 \quad \beta = 0$$

$$= 1 - p_2 + b\, p_2 \qquad \text{if} \quad \alpha = \alpha_2 \quad \beta = 1$$

$$= 1 - (1-a)\, p_2 \qquad \text{if} \quad \alpha = \alpha_1 \quad \beta = 0$$

$$= (1-b)\, p_1 \qquad \text{if} \quad \alpha = \alpha_1 \quad \beta = 1$$

GENERAL $L_{R,P}$ (LINEAR REWARD-PENALTY SCHEME)

$\alpha = \alpha_2,\ \beta = 1$	$\alpha = \alpha_1 \quad \beta = 0$
$p_2 \leftarrow (1-b)\, p_2$	$p_2 \leftarrow (1-a)\, p_2$
$\therefore p_1(n+1) = 1 - (1-b)\, p_2$	$\therefore p_1 \leftarrow 1 - p_2 = 1 - (1-a)\, p_2$
$= 1 - p_2 + b\, p_2.$	

SIMPLE Case. $a = b$. (Symmetric $L_{R,P}$ Scheme).

$$
\begin{aligned}
p_1(n+1) &= k p_1 && \text{if } \alpha = \alpha_1 \quad \beta = 1 \\
&= 1 - k p_2 && \text{if } \alpha = \alpha_1 \quad \beta = 0 \\
&= 1 - k p_2 && \text{if } \alpha = \alpha_2 \quad \beta = 1 \\
&= k p_1 && \text{if } \alpha = \alpha_2 \quad \beta = 0
\end{aligned}
$$

NOTE: $(\alpha_2, \beta = 0) \equiv (\alpha_1, \beta = 1)$

AND. $(\alpha_1, \beta = 0) \equiv (\alpha_2, \beta = 1)$

NOTE

p_1, p_2 are both RANDOM VARIABLES

Previously : $\underline{P}(n) = F^T \underline{P}(n-1)$ (FSSA)

NOW WE CAN'T DO THAT.

$p_1(n+1)$ can take 4 values - depending on outcome.

i.e. $p_1(n+1)$ is a r.v. with a mean, variance etc.

WHAT IS $E\left[p_1(n+1)\right]$? <u>USE ONLY P_1 in Expression</u>

$$
\begin{aligned}
p_1(n+1) &= k\, p_1 && \text{w.p. } p_1^{(n)} c_1 \\
&= 1 - k(1-p_1) && \text{w.p. } p_1^{(n)}(1-c_1) \\
&= 1 - k(1-p_1) && \text{w.p. } (1-p_1)\, c_2 \\
&= k\, p_1 && \text{w.p. } (1-p_1)(1-c_2)
\end{aligned}
$$

$\therefore$

$E\left[p_1(n+1)\right]$ is a function of p_1

$$
E\left[p_1(n+1) \mid p_1\right] = k\, p_1 \left(p_1 c_1 + (1-p_1)(1-c_2)\right) + \left(1 - k(1-p_1)\right)\left(p_1(1-c_1) + (1-p_1) c_2\right)
$$

<u>NOTE</u> : ALL p_1^2 terms luckily cancel.

$$
E\left[p_1(n+1) \mid p_1\right] = p_1 - k(c_1 + c_2)\, p_1 + k c_2
$$

$\therefore$

$$
E\left[p_1(n+1)\right] = \left[1 - k(c_1 + c_2)\right] E\left[p_1\right] + k c_2 .
$$

a linear difference equation ...

$$
y(n+1) = e\, y(n) + f .
$$

General Solution.

In the Limit. $p_1(n+1) = p_1(n)$

$$\bar{p}_1(n+1) = [1 - k(c_1+c_2)]\,\bar{p}_1(n) + k\,c_2$$

is.

$$\bar{p}_1(\infty) = [1 - k(c_1+c_2)]\,\bar{p}_1(\infty) + k\,c_2$$

$$k(c_1+c_2)\,\bar{p}_1(\infty) = k\,c_2$$

or $E[p_1(\infty)] = c_2/c_1+c_2.$

THEOREM.

1. The Symmetric $L_{R,P}$ scheme has

$$E[p_i(\infty)] = \frac{c_j}{c_i+c_j} \qquad i \neq j, \; i = 1,2$$

AND is THEREFORE EXPEDIENT.

2. THE LIMITING EXP. VALUE OF $p_i(\infty)$ is INDEPENDENT OF THE PARAMETER OF $L_{R,P}$ scheme.

IS THERE A "BETTER" WAY TO LOOK AT THIS ?

YES !

$$E[p_1(n+1) \mid p_1] = p_1 - k(C_1+C_2)p_1 + kC_2 .$$

i.e: writing it in terms of both p_1 & p_2.

$$(kC_2 = (p_1+p_2)(kC_2)$$

$$E(p_1(n+1) \mid p_1] = p_1\left[1 - k(C_1+C_2) + kC_2\right] + kC_2 p_2$$

$$= p_1(1-kC_1) + kC_2 p_2 .$$

$$\therefore E[p_1(n+1)] = (1-kC_1)E[p_1] + kC_2 E[p_2]$$

$$E[p_2(n+1)] = kC_1 E[p_1] + (1-kC_2)E[p_2]$$

$$E\begin{bmatrix} p_1(n+1) \\ p_2(n+1) \end{bmatrix} = \begin{bmatrix} 1-kC_1 & kC_1 \\ kC_2 & 1-kC_2 \end{bmatrix}^T E\begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

i.e. $E\begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$ are the "STATES OF AN ERGODIC M.C.

$$F = \begin{bmatrix} 1-kC_1 & kC_1 \\ kC_2 & 1-kC_2 \end{bmatrix}$$

markov matrix

SOLUTION is $\begin{bmatrix} \frac{C_2}{C_1+C_2} \\ \frac{C_1}{C_1+C_2} \end{bmatrix}$ just as we thought...

i.e. $E[p_1(\infty)] \to \frac{C_2}{C_1+C_2}$.

EIGENVALUES : ALWAYS ONE is UNITY.

OTHER : $(1-kC_1 + 1-kC_2) = 1 + \lambda_2$

$$\lambda_2 = 1 - k(C_1+C_2)$$

RATE OF CONVERGENCE determined by 'k'

VARIANCE

$$\mathrm{Var}\, p_1(\infty) = \frac{C_1 C_2}{(C_1+C_2)^2} \cdot \frac{(1-k)^2}{(1-k)^2 + 2k(1-k)(C_1+C_2)}$$