state 0, 1, or 2, and replacing it when it is in state 3. The minimum expected total discounted cost of the system starting in state i, $i = 0, 1, 2, 3$, and evolving for one period is given by 0, 1,000, 3,000, and 6,000, respectively. The optimal solution to the two-period machine-maintenance model is

Period 1 Leave machine alone when it is in state 0 or 1.
Overhaul machine when it is in state 2.
Replace machine when it is in state 3.

Period 2 Leave machine alone when it is in state 0, 1, or 2.
Replace machine when it is in state 3.

The minimum expected total discounted cost of the system starting in state i, $i = 0, 1, 2, 3$, and evolving for two periods is given by 1,294, 2,688, 4,900, and 6,000, respectively. Finally, the optimal solution to the three-period model is

Period 1 *and* *Period 2* { Leave machine alone when it is in state 0 or 1.
Overhaul machine when it is in state 2.
Replace machine when it is in state 3.

Period 3 Leave machine alone when it is in state 0, 1, or 2.
Replace machine when it is in state 3.

The minimum expected total discounted costs over three periods, if the system starts in state i, $i = 0, 1, 2, 3$, are given by 2,730, 4,041, 6,419, and 7,165, respectively.

20.6 A Water-Resource Model

A multipurpose dam is used for generating electric power as well as for flood control. The capacity of the dam is 3 units. The probability distribution of the quantity of water, W_t, that flows into the dam during month t (for $t = 0, 1, \ldots$) is given by $P_W(m)$, where

$$P_W(0) = P\{W = 0\} = \tfrac{1}{6}$$
$$P_W(1) = P\{W = 1\} = \tfrac{1}{3}$$
$$P_W(2) = P\{W = 2\} = \tfrac{1}{3}$$
$$P_W(3) = P\{W = 3\} = \tfrac{1}{6}.$$

For the purpose of generating electric power, 1 unit of water is required. At the beginning of each month, water is released from the dam. The first unit is used to generate electric power and then used for irrigation purposes, the latter function being worth \$100,000. If additional units are released, they can also be used for irrigation purposes, and each unit is worth \$100,000. If the dam contains less than 1 unit at the beginning of a month, additional power must be purchased at a cost of \$300,000. If at any time the water in the dam exceeds the capacity of 3 units, the excess water is released through the spillways at no cost or gain.

A release policy is sought. Policies are to be compared on the basis of expected discounted cost, with discount factor $\alpha = 0.99$. The *policy improvement algorithm* will be used.

Let X_t denote the amount of water in the dam at time t. Then $X_t = 0, 1, 2, 3$. The natural laws of motion for this system (no water released) are given by the transition matrix:

State	0	1	2	3
0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
1	0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$
2	0	0	$\frac{1}{6}$	$\frac{5}{6}$
3	0	0	0	1

For example, the element in the second row and fourth column, p_{13}, is obtained as follows: If the dam contains 1 unit of water now, then for it to contain 3 units of water a month later, 2 or 3 units of water must flow into the dam during the month (recall that dam capacity is 3 units, so that a flow of 3 units will result in 1 unit being released through the spillways). This occurs with probability $\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$.

There are three possible decisions that can be made at the beginning of each month:

Decision	Action
1	Release 1 unit
2	Release 2 units
3	Release 3 units

It is clear that releasing no units is not a sensible action, because 1 unit is needed for electric power generation anyway. Thus a policy calls for determining how many units to release as a function of the quantity of water found in the dam. A typical policy R_1 might call for releasing all the water in the dam if it contains 0, 1, or 2 units, and releasing 2 units if it contains 3 units. The resultant transition matrix is given by

State	0	1	2	3
0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
1	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
3	0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

Of course, a policy that calls for releasing 3 units when there is only 1 unit in the dam is to be interpreted as calling for releasing all the available water. Necessary cost information can be obtained from the following data:

State	Decision	Cost (In Hundred Thousands)
0	1	3
	2	3
	3	3
1	1	−1
	2	−1
	3	−1
2	1	−1
	2	−2
	3	−2
3	1	−1
	2	−2
	3	−3

The policy R_1 will be used in the value-determination step (step 1) of the policy improvement algorithm. Using the cost information just given, the values of C_{ik_1} are

$$C_{0k_1} = 3$$
$$C_{1k_1} = -1$$
$$C_{2k_1} = -2$$
$$C_{3k_1} = -2.$$

The following four equations must be solved:

$$V_0(R_1) = 3 + 0.99[\tfrac{1}{6}V_0(R_1) + \tfrac{1}{3}V_1(R_1) + \tfrac{1}{3}V_2(R_1) + \tfrac{1}{6}V_3(R_1)]$$
$$V_1(R_1) = -1 + 0.99[\tfrac{1}{6}V_0(R_1) + \tfrac{1}{3}V_1(R_1) + \tfrac{1}{3}V_2(R_1) + \tfrac{1}{6}V_3(R_1)]$$
$$V_2(R_1) = -2 + 0.99[\tfrac{1}{6}V_0(R_1) + \tfrac{1}{3}V_1(R_1) + \tfrac{1}{3}V_2(R_1) + \tfrac{1}{6}V_3(R_1)]$$
$$V_3(R_1) = -2 + 0.99[\qquad \tfrac{1}{6}V_1(R_1) + \tfrac{1}{3}V_2(R_1) + \tfrac{1}{2}V_3(R_1)].$$

The simultaneous solution of these equations results in the values

$$V_0(R_1) = -103.881,\ V_1(R_1) = -107.881,$$

and

$$V_2(R_1) = -108.881,\ V_3(R_1) = -110.358.$$

Step 2 can now be applied. We want to find an improved policy R_2 that has the property that $d_0(R_2) = k_2^0$, $d_1(R_2) = k_2^1$, $d_2(R_2) = k_2^2$, and $d_3(R_2) = k_2^3$ minimizes the following expressions:

(0) $$C_{0k_2^0} + 0.99[-103.881p_{00}(k_2^0) - 107.881p_{01}(k_2^0) - 108.881p_{02}(k_2^0) - 110.358p_{03}(k_2^0)]$$

(1) $$C_{1k_2^1} + 0.99[-103.881p_{10}(k_2^1) - 107.881p_{11}(k_2^1) - 108.881p_{12}(k_2^1) - 110.358p_{13}(k_2^1)]$$

$$(2) \quad C_{2k_2^2} + 0.99[-103.881p_{20}(k_2^2) - 107.881p_{21}(k_2^2) - 108.881p_{22}(k_2^2) - 110.358p_{23}(k_2^2)]$$

$$(3) \quad C_{3k_2^3} + 0.99[-103.881p_{30}(k_2^3) - 107.881p_{31}(k_2^3) - 108.881p_{32}(k_2^3) - 110.358p_{33}(k_2^3)].$$

To find k_2^0, the best decision when the system is in state 0, it is necessary to evaluate the first expression for all possible decisions. It is clear that when the system is in state 0 (dam empty), there is no choice among the decisions because they all are equivalent. The data for the necessary calculations for evaluating expression (0) follow:

State 0

Decision	$p_{00}(k_2)$	$p_{01}(k_2)$	$p_{02}(k_2)$	$p_{03}(k_2)$	C_{0k_2}	Total Value of Expression 0
1, 2, 3	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$	3	−103.881

Similarly, for state 1 there is no choice among the decisions because they are all equivalent. The data for the necessary calculations for evaluating follow:

State 1

Decision	$p_{10}(k_2)$	$p_{11}(k_2)$	$p_{12}(k_2)$	$p_{13}(k_2)$	C_{1k_2}	Total Value of Expression 1
1, 2, 3	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$	−1	−107.881

For the remaining two states, the appropriate transition probabilities and costs generally depend upon the decisions made. The data for the necessary calculations for finding the best decisions, given the dam is in state 2 or 3, follow:

State 2

Decision	$p_{20}(k_2)$	$p_{21}(k_2)$	$p_{22}(k_2)$	$p_{23}(k)_2$	C_{2k_2}	Total Value of Expression 2
1	0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$	−1	−109.358
2, 3	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$	−2	−108.881

State 3

Decision	$p_{30}(k_2)$	$p_{31}(k_2)$	$p_{32}(k_2)$	$p_{33}(k_2)$	C_{3k_2}	Total Value of Expression 3
1	0	0	$\frac{1}{6}$	$\frac{5}{6}$	−1	−110.011
2	0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$	−2	−110.358
3	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$	−3	−109.881

Thus $d_0(R_2) = k_2^0 = d_1(R_2) = k_2^1 = 1, 2,$ or 3; $d_2(R_2) = k_2^2 = 1$; and $d_3(R_2) = k_2^3 = 2$. Hence policy R_2 calls for releasing all the water when there is 1 unit in the dam, 1 unit of water when there are 2 units available in the dam, and 2 units when there are 3 units available in the dam. This policy differs from R_1, so that another iteration is required. For the value-determination step, the equations that must now be solved are

$$V_0(R_2) = 3 + 0.99[\tfrac{1}{6}V_0(R_2) + \tfrac{1}{3}V_1(R_2) + \tfrac{1}{3}V_2(R_2) + \tfrac{1}{6}V_3(R_2)]$$

$$V_1(R_2) = -1 + 0.99[\tfrac{1}{6}V_0(R_2) + \tfrac{1}{3}V_1(R_2) + \tfrac{1}{3}V_2(R_2) + \tfrac{1}{6}V_3(R_2)]$$

$$V_2(R_2) = -1 + 0.99[\tfrac{1}{6}V_1(R_2) + \tfrac{1}{3}V_2(R_2) + \tfrac{1}{2}V_3(R_2)]$$

$$V_3(R_2) = -2 + 0.99[\tfrac{1}{6}V_1(R_2) + \tfrac{1}{3}V_2(R_2) + \tfrac{1}{2}V_3(R_2)].$$

The simultaneous solution of these equations results in the values $V_0(R_2) = -119.642$, $V_1(R_2) = -123.642$, $V_2(R_2) = -125.119$, and $V_3(R_2) = -126.119$.

Step 2 can now be applied. We want to find an improved policy R_3 that has the property that $d_0(R_3) = k_3^0$, $d_1(R_3) = k_3^1$, $d_2(R_3) = k_3^2$, and $d_3(R_3) = k_3^3$ minimizes the following expressions:

(0) $C_{0k_3^0} + 0.99[-119.642p_{00}(k_3^0) - 123.642p_{01}(k_3^0) - 125.119p_{02}(k_3^0) - 126.119p_{03}(k_3^0)]$

(1) $C_{1k_3^1} + 0.99[-119.642p_{10}(k_3^1) - 123.642p_{11}(k_3^1) - 125.119p_{12}(k_3^1) - 126.119p_{13}(k_3^1)]$

(2) $C_{2k_3^2} + 0.99[-119.642p_{20}(k_3^2) - 123.642p_{21}(k_3^2) - 125.119p_{22}(k_3^2) - 126.119p_{23}(k_3^2)]$

(3) $C_{3k_3^3} + 0.99[-119.642p_{30}(k_3^3) - 123.642p_{31}(k_3^3) - 125.119p_{32}(k_3^3) - 126.119p_{33}(k_3^3)]$.

The data on the transition matrices and the costs from the previous iteration can again be used; the resulting values of the expression are

Decision	Value of Expression 0	Value of Expression 1	Value of Expression 2	Value of Expression 3
1	−119.642	−123.642	−125.119	−125.693
2	−119.642	−123.642	−124.642	−126.119
3	−119.642	−123.642	−124.642	−125.642

Thus $d_0(R_3) = k_3^0 = d_1(R_3) = k_3^1 = 1, 2,$ or 3; $d_2(R_3) = k_3^2 = 1$; and $d_3(R_3) = k_3^3 = 2$. Hence policy R_3 and policy R_2 are identical, and the optimal release policy calls for releasing all the water when there is 1 unit in the dam, 1 unit of water when there are 2 units available in the dam, and 2 units when there are 3 units available in the dam.

Of course, direct enumeration would have been just as simple a technique to use in this situation, but the policy improvement algorithm was used for illustrative purposes.