CHAPTER 24

Probabilistic Dynamic Programming

Chapter Guide. This chapter assumes familiarity with deterministic dynamic programming (DP) in Chapter 10. The main elements of a probabilistic DP model are the same as in the deterministic case—namely, the probabilistic DP model also decomposes the problem into stages with states and alternatives associated with each stage. The difference is the probabilistic nature of the states and the returns, which necessitates basing the optimization criterion on the expected value or some other probabilistic measure. Probabilistic DP arises particularly in the treatment of stochastic inventory models and in Markovian decision processes. These two topics are treated separately in Chapters 16 and 25. The examples presented here illustrate the stochastic nature of DP.

This chapter includes 3 solved examples, 9 end-of-section problems, and 1 case.

24.1 A GAME OF CHANCE

A variation of the Russian roulette game calls for spinning a wheel marked along the perimeter with n consecutive numbers: 1 to n. The probability that the wheel will stop at number i after one spin is p_i. A player pays \$$x$ for the privilege of spinning the wheel a maximum of m times. The payoff to the player is twice the number produced in the *last* spin. Assuming that the game (of up to m spins each) is repeated a reasonably large number of times, devise an optimal strategy for the player.

We can construct the problem as a DP model using the following definitions:

1. *Stage* i is represented by spin i, $i = 1, 2, \ldots, m$.
2. The *alternatives* at each stage include either spinning the wheel once more or ending the game.
3. The *state* j of the system at stage i is one of the numbers 1 to n obtained in the *last* spin.

Let

$f_i(j)$ = Maximum expected return given that the game is at stage (spin) i and that j is the outcome of the *last* spin

Thus,

$$\begin{pmatrix}\text{Expected payoff at stage } i\\ \text{given last spin's outcome } j\end{pmatrix} = \begin{cases} 2j, & \text{if game ends}\\ \sum_{k=1}^{n} p_k f_{i+1}(k), & \text{if game continues}\end{cases}$$

The recursive equation may then be written as

$$f_{m+1}(j) = 2j$$

$$f_i(j) = \max\left\{\begin{matrix}\text{End:} & 2j,\\ \text{Spin:} & \sum_{k=1}^{n} p_k f_{i+1}(k)\end{matrix}\right\} i = 2, 3, \ldots, m$$

$$f_1(0) = \sum_{k=1}^{n} p_k f_2(k)$$

The rationale for the recursive equation is that at the first spin ($i = 1$), the state of the system is $j = 0$, because the game has just started. Hence, $f_1(0) = p_1 f_2(1) + p_2 f_2(2) + \cdots + p_n f_2(n)$. After the last spin ($i = m$), the game must end regardless of the outcome j of the mth spin. Thus, $f_{m+1}(j) = 2j$.

The recursive calculations start with f_{m+1} and terminate with $f_1(0)$, thus producing $m + 1$ computational stages. Because $f_1(0)$ is the expected return from all m spins, and given that the game costs \$$x$, the net return is $f_1(0) - x$.

Example 24.1-1

Suppose that the perimeter of the Russian roulette wheel is marked with the numbers 1 to 5. The probability p_i of stopping at the number i is given by $p_1 = .3$, $p_2 = .25$, $p_3 = .2$, $p_4 = .15$, and $p_5 = .1$. The player pays \$5 for a maximum of four spins. Determine the optimal strategy for each of the four spins and the expected net return.

Stage 5

$$f_5(j) = 2j, j = 1, 2, 3, 4, \text{ or } 5$$

Decision: End if $j = 1, 2, 3, 4$, or 5, $f_5(j) = 2j$

Stage 4

$$\begin{aligned} f_4(j) &= \max\{2j, p_1 f_5(1) + p_2 f_5(2) + p_3 f_5(3) + p_4 f_5(4) + p_5 f_5(5)\}\\ &= \max\{2j, .3 \times 2 + .25 \times 4 + .2 \times 6 + .15 \times 8 + .1 \times 10\}\\ &= \max\{2j, 5\}\\ &= \begin{cases} 5, \text{ if } j = 1 \text{ or } 2\\ 2j, \text{ if } j = 3, 4, \text{ or } 5\end{cases}\end{aligned}$$

Decision: Spin if $j = 1$ or 2, $f_4(j) = 5$. End if $j = 3, 4$, or 5, $f_4(j) = 2j$.

Stage 3

$$
\begin{aligned}
f_3(j) &= \max\{2j, p_1f_4(1) + p_2f_4(2) + p_3f_4(3) + p_4f_4(4) + p_5f_4(5)\} \\
&= \max\{2j, .3 \times 5 + .25 \times 5 + .2 \times 6 + .15 \times 8 + .1 \times 10\} \\
&= \max\{2j, 6.15\} \\
&= \begin{cases} 6.15, \text{ if } j = 1, 2, \text{ or } 3 \\ 2j, \text{ if } j = 4 \text{ or } 5 \end{cases}
\end{aligned}
$$

Decision: Spin if $j = 1, 2,$ or $3, f_3(j) = 6.15$. End if $j = 4$ or $5, f_3(j) = 2j$.

Stage 2

$$
\begin{aligned}
f_2(j) &= \max\{2j, p_1f_3(1) + p_2f_3(2) + p_3f_3(3) + p_4f_3(4) + p_5f_3(5)\} \\
&= \max\{2j, .3 \times 6.15 + .25 \times 6.15 + .2 \times 6.15 + .15 \times 8 + .1 \times 10\} \\
&= \max\{2j, 6.8125\} \\
&= \begin{cases} 6.8125, \text{ if } j = 1, 2, \text{ or } 3 \\ 2j, \text{ if } j = 4 \text{ or } 5 \end{cases}
\end{aligned}
$$

Decision: Spin if $j = 1, 2,$ or $3, f_3(j) = 6.8125$. End if $j = 4$ or $5, f_3(j) = 2j$.

Stage 1

$$
\begin{aligned}
f_1(0) &= p_1f_2(1) + p_2f_2(2) + p_3f_2(3) + p_4f_2(4) + p_5f_2(5) \\
&= .3 \times 6.8125 + .25 \times 6.8125 + .2 \times 6.8125 + .15 \times 8 + .1 \times 10 \\
&= 7.31
\end{aligned}
$$

Decision: Spin, $f_1(0) = \$7.31$.

From the preceding calculations, the optimal solution is

Spin no.	Optimal strategy
1	Game starts, spin
2	Continue if spin 1 produces 1, 2, or 3; else, end game
3	Continue if spin 2 produces 1, 2, or 3; else, end game
4	Continue if spin 3 produces 1 or 2; else, end game
	Expected net return = \$7.31 − \$5.00 = \$2.31

PROBLEM SET 24.1A

1. In Example 24.1–1, suppose that the wheel is marked with the numbers 1 to 8 and will stop at any of these numbers with equal probabilities. Assuming that each game includes a total of five spins, develop an optimal strategy for the game.

*__2.__ I would like to sell my used car to the highest bidder. Studying the market, I have concluded that I am likely to receive three types of offers with equal probabilities: low at about \$1050, medium at about \$1900, and high at about \$2500. I advertise the car for up to 3 consecutive days. At the end of each day, I will decide whether or not to accept the best offer made that day. What should be my optimal strategy regarding the acceptance of an offer?

24.2 INVESTMENT PROBLEM

An individual wishes to invest up to $\$C$ thousand in the stock market over the next n years. The investment plan calls for buying the stock at the start of the year and selling it at the end of the same year. Accumulated money may then be reinvested (in whole or in part) at the start of the following year. The degree of risk is represented by expressing the return probabilistically. A study of the market shows that the return on investment is affected by m market conditions and that condition k yields a return r_k (positive, zero, or negative) with probability p_k, $k = 1, 2, \ldots, m$. How should the amount C be invested to realize the highest accumulation at the end of n years?

Define

$$x_i = \text{Available funds at the start of year } i,\ i = 1, 2, \ldots, n\ (x_1 = C)$$
$$y_i = \text{Invested funds at the start of year } i\ (y_i \leq x_i)$$

The elements of the DP model can be described as

1. *Stage* i is represented by year i.
2. The *alternatives* at stage i are given by y_i.
3. The *state* at stage i is given by x_i.

Let

$$f_i(x_i) = \text{Maximum expected funds for years } i,\ i+1, \ldots, \text{ and } n, \text{ given } x_i \text{ at the start of year } i$$

For market condition k, we have

$$x_{i+1} = (1 + r_k)y_i + (x_i - y_i) = x_i + r_k y_i,\ k = 1, 2, \ldots, m$$

Given that market condition k occurs with probability p_k, the DP recursive equation is written as

$$f_i(x_i) = \max_{0 \leq y_i \leq x_i} \left\{ \sum_{k=1}^{m} p_k f_{i+1}(x_i + r_k y_i) \right\},\ i = 1, 2, \ldots, n$$

where $f_{n+1}(x_{n+1}) = x_{n+1}$ because no investment occurs after year n.

For year n, we have

$$f_n(x_n) = \max_{0 \leq y_n \leq x_n} \left\{ \sum_{k=1}^{m} p_k (x_n + r_k y_n) \right\}$$
$$= x_n + \max_{0 \leq y_n \leq x_n} \left\{ \left(\sum_{k=1}^{m} p_k r_k \right) y_n \right\}$$

Letting

$$\bar{r} = \sum_{k=1}^{m} p_k r_k$$

we get

$$y_n = \begin{cases} 0, & \text{if } \bar{r} \le 0 \\ x_n, & \text{if } \bar{r} > 0 \end{cases}$$

$$f_n(x_n) = \begin{cases} x_n, & \text{if } \bar{r} \le 0 \\ (1 + \bar{r})x_n, & \text{if } \bar{r} > 0 \end{cases}$$

Example 24.2-1

In the investment model, suppose that you want to invest \$10,000 over the next 4 years. There is a 50% chance that you will double your money, a 20% chance that you will break even, and a 30% chance that you will lose the invested amount. Devise an optimal investment strategy.

Using the notation of the model, we have

$$C = \$10{,}000, n = 4, m = 3$$
$$p_1 = .4, p_2 = .2, p_3 = .4$$
$$r_1 = 1, r_2 = 0, r_3 = -1$$

Stage 4

$$\bar{r} = .5 \times 1 + .2 \times 0 + .3 \times -1 = .2$$

Thus,

$$f_4(x_4) = 1.2x_4$$

The optimum solution is summarized as

	Optimum solution	
State	$f_4(x_4)$	y_4^*
x_4	$1.2x_4$	x_4

Stage 3

$$f_3(x_3) = \max_{0 \le y_3 \le x_3} \{p_1 f_4(x_3 + r_1 y_3) + p_2 f_4(x_3 + r_2 y_3) + p_3 f_4(x_3 + r_3 y_3)\}$$
$$= \max_{0 \le y_3 \le x_3} \{.5 \times 1.2(x_3 + y_3) + .2 \times 1.2(x_3 + 0y_3) + .3 \times 1.2[x_3 + (-1)y_3]\}$$
$$= \max_{0 \le y_3 \le x_3} \{1.2x_3 + .24y_3\}$$
$$= 1.44x_3$$

Thus, we get

	Optimum solution	
State	$f_3(x_3)$	y_2^*
x_3	$1.44x_3$	x_3

Stage 2

$$f_2(x_2) = \max_{0 \le y_2 \le x_2} \{p_1 f_3(x_2 + r_1 y_2) + p_2 f_3(x_2 + r_2 y_2) + p_3 f_3(x_2 + r_3 y_2)\}$$

$$= \max_{0 \le y_2 \le x_2} \{.5 \times 1.44(x_2 + y_2) + .2 \times 1.44(x_2 + 0y_2) + .3 \times 1.44[x_2 + (-1)y_2]\}$$

$$= \max_{0 \le y_2 \le x_2} \{1.44x_2 + .288y_2\}$$

$$= 1.728x_2$$

Thus, we get

	Optimum solution	
State	$f_2(x_2)$	y_2^*
x_2	$1.728x_2$	x_2

Stage 1

$$f_1(x_1) = \max_{0 \le y_1 \le x_1} \{p_1 f_2(x_1 + r_1 y_1) + p_2 f_2(x_1 + r_2 y_1) + p_3 f_2(x_1 + r_3 y_1)\}$$

$$= \max_{0 \le y_1 \le x_1} \{.5 \times 1.728(x_1 + y_1) + .2 \times 1.728(x_1 + 0y_1) + .3 \times 1.728[x_1 + (-1)y_1]\}$$

$$= \max_{0 \le y_1 \le x_1} \{1.728x_1 + .3456y_1\}$$

$$= 2.0736x_1$$

Thus, we get

	Optimum solution	
State	$f_1(x_1)$	y_1^*
x_1	$2.0736x_1$	x_1

The optimal investment policy can thus be summarized as follows: Because $y_i^* = x_i$ for $i = 1$ to 4, the optimal solution calls for investing all available funds at the start of each year. The accumulated funds at the end of 4 years are $2.0736x_1 = 2.0736(\$10{,}000) = \$20{,}376$.

Actually, it can be shown by induction that the problem has the following general solution at stage i, $i = 1, 2, \ldots, n$.

$$f_i(x_i) = \begin{cases} x_i, & \text{if } \bar{r} \le 0 \\ (1 + \bar{r})^{n-i+1}, & \text{if } \bar{r} > 0 \end{cases}$$

$$y_i = \begin{cases} 0, & \text{if } \bar{r} \le 0 \\ x_i, & \text{if } \bar{r} > 0 \end{cases}$$

PROBLEM SET 24.2A

*__1.__ In Example 24.2–1, find the optimal investment strategy, assuming that the probabilities p_k and return r_k vary for the 4 years according to the following data:

Year	r_1	r_2	r_3	p_1	p_2	p_3
1	2	1	.5	.1	.4	.5
2	1	0	−1	.4	.4	.2
3	4	−1	−1	.2	.4	.4
4	.8	.4	.2	.6	.2	.2

2. A 10-m^3 compartment is available for storing three items. The volumes needed to store 1 unit of items 1, 2, and 3 are 2, 1, and 3 m^3, respectively. The probabilistic demand for the items is described as follows:

	Probability of demand		
No. of units	*Item 1*	*Item 2*	*Item 3*
1	.5	.3	.3
2	.5	.4	.2
3	.0	.2	.5
4	.0	.1	.0

The shortage costs per unit for items 1, 2, and 3 are \$8, \$10, and \$15, respectively. How many units of each item should be held in the compartment?

3. HiTec has just started to produce supercomputers for a limited period of 4 years. The annual demand, D, for the new computer is described by the following distribution:

$$p(D = 1) = .5, p(D = 2) = .3, p(D = 3) = .2$$

The production capacity of the plant is three computers annually at the cost of \$5 million each. The actual number of computers produced per year may not equal the demand exactly. An unsold computer at the end of a year incurs \$1 million in storage and maintenance costs. A loss of \$2 million occurs if the delivery of a computer is delayed by 1 year. HiTec will not accept new orders beyond year 4 but will continue production in year 5 to satisfy any unfilled demand at the end of year 4. Determine the optimal annual production schedule for HiTec.

*4. The PackRat Outdoors Company owns three sports centers in downtown Little Rock. On Easter Day, bicycle riding is a desirable outdoors activity. The company owns a total of eight rental bikes to be allocated to the three centers with the objective of maximizing revenues. The demand for the bikes and the hourly rental cost to customers vary by location and are described by the following distributions:

	Probability of demand		
No. of bikes	*Center 1*	*Center 2*	*Center 3*
0	.1	.02	0
1	.2	.03	.15
2	.3	.10	.25
3	.2	.25	.30
4	.1	.30	.15
5	.1	.15	.10
6	0	.05	.025
7	0	.05	.025
8	0	.05	0
Rental cost/hr ($)	6	7	5

How should PackRat allocate the eight bikes to the three centers?

24.3 MAXIMIZATION OF THE EVENT OF ACHIEVING A GOAL

Section 24.2 deals with maximizing the optimal expected return. Another useful objective is the maximization of the probability of achieving a certain level of return. We use the investment situation in Section 24.2 to illustrate the application of the new criterion.

As in Section 24.2, the definitions of the *stage, i, alternative,* y_i, and *state,* x_i, remain the same. The new criterion maximizes the probability of realizing a sum of money, S, at the end of n years. Define

$f_i(x_i)$ = Probability of realizing the amount S given x_i is the amount of funds available at the start of year i and that an optimal policy is implemented for years $i,\ i + 1,\ \ldots,$ and n

The DP recursive equation is thus given as

$$f_n(x_n) = \max_{0 \le y_n \le x_n} \left\{ \sum_{k=1}^{m} p_k P\{x_n + r_k y_n \ge S\} \right\}$$

$$f_i(x_i) = \max_{0 \le y_i \le x_i} \left\{ \sum_{k=1}^{m} p_k f_{i+1}(x_i + r_k y_i) \right\},\ i = 1, 2, \ldots, n - 1$$

The recursive formula is based on the conditional probability law

$$P\{A\} = \sum_{k=1}^{m} P\{A|B_k\} P\{B_k\}$$

In this case, $f_{i+1}(x_i + r_k y_i)$ plays the role of $P\{A|B_k\}$.

Example 24.3-1

An individual wants to invest \$2000. Available options include doubling the amount invested with probability .3 or losing all of it with probability .7. Investments are sold at the end of the year, and reinvestment, in whole or part, starts again at the beginning of the following year. The process is repeated for three consecutive years. The objective is to maximize the probability of realizing \$4000 at the end of the third year. For simplicity, assume that all investments are in multiples of \$1000.

Using the notation of the model, we say that $r_1 = 1$ with probability .3, and $r_2 = -1$ with probability .7.

Stage 3. At stage 3, the state x_3 can be as small as \$0 and as large as \$8000. The minimum value is realized when the entire investment is lost, and the maximum value occurs when the investment is doubled at the end of each of the first 2 years. The recursive equation for stage 3 is thus written as

$$f_3(x_3) = \max_{y_3=0,1,\ldots,x_3} \{.3P\{x_3 + y_3 \geq 4\} + .7P\{x_3 - y_3 \geq 4\}\}$$

where $x_3 = 0, 1, \ldots, 8$.

Table 24.1 details the computations for stage 3. All the shaded entries are infeasible because they do not satisfy the condition $y_3 \leq x_3$. Also, in carrying out the computations, we notice that

$$P\{x_3 + y_3 \geq 4\} = \begin{cases} 0, \text{ if } x_3 + y_3 < 4 \\ 1, \text{ if otherwise} \end{cases}$$

$$P\{x_3 - y_3 \geq 4\} = \begin{cases} 0, \text{ if } x_3 - y_3 < 4 \\ 1, \text{ if otherwise} \end{cases}$$

Although Table 24.1 shows that alternative optima exist for $x_3 = 1, 3, 4, 5, 6, 7$, and 8, the optimum (last) column provides only the smallest optimum y_3. The assumption here is that the investor is not going to invest more than what is absolutely necessary to achieve the desired goal.

Stage 2

$$f_2(x_2) = \max_{y_2=0,1,\ldots,x_2} \{.3f_3(x_2 + y_2) + .7f_3(x_2 - y_2)\}$$

The associated computations are given in Table 24.2.

Stage 1

$$f_1(x_1) = \max_{y_1=0,1,2} \{.3f_2(x_1 + y_1) + .7f_2(x_1 - y_1)\}$$

Table 24.3 provides the associated computations.

The optimum strategy is determined in the following manner: Given the initial investment $x_1 = \$2000$, stage 1 (Table 24.3) yields $y_1 = 0$, which means that no investment should be made in year 1. The decision not to invest in year 1 leaves the investor with \$2000 at the start of year 2. From stage 2 (Table 24.2), $x_2 = 2$ yields $y_2 = 0$, indicating once again that no investment should occur in year 2. Next, using $x_3 = 2$, stage 3 (Table 24.1) shows that $y_3 = 2$, which calls for investing the entire amount in year 3. The associated maximum probability for realizing the goal $S = 4$ is $f_1(2) = .3$.

TABLE 24.1

	$.3P(x_3 + y_3 \geq 4) + .7P(x_3 - y_3 \geq 4)$									*Optimum*	
x_3	$y_3 = 0$	1	2	3	4	5	6	7	8	f_3	y_3
0	$.3 \times 0 + .7 \times 0 = 0$									0	0
1	$.3 \times 0 + .7 \times 0 = 0$	$.3 \times 0 + .7 \times 0 = 0$								0	0
2	$.3 \times 0 + .7 \times 0 = 0$	$.3 \times 0 + .7 \times 0 = 0$	$.3 \times 1 + .7 \times 0 = .3$							.3	2
3	$.3 \times 0 + .7 \times 0 = 0$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$						.3	1
4	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$					1	0
5	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$				1	0
6	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.31 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$			1	0
7	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$		1	0
8	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 1 = 1$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	$.3 \times 1 + .7 \times 0 = .3$	1	0

TABLE 24.2

	$.3f_3(x_2+y_2)+.7f_3(x_2-y_2)$					Optimum	
x_2	$y_2 = 0$	1	2	3	4	f_2	y_2
0	$.3 \times 0 +$ $.7 \times 0 = 0$					0	0
1	$.3 \times 0 +$ $.7 \times 0 = 0$	$.3 \times .3 +$ $.7 \times 0 = .9$				.09	1
2	$.3 \times .3 +$ $.7 \times .3 = .3$	$.3 \times .3 +$ $.7 \times 0 = .09$	$.3 \times 1 +$ $.7 \times 0 = .3$			.30	0
3	$.3 \times .3 +$ $.7 \times .3 = .3$	$.3 \times 1 +$ $.7 \times .3 = .51$	$.3 \times 1 +$ $.7 \times 0 = .3$	$.3 \times 1 +$ $.7 \times 0 = .3$		.51	1
4	$.3 \times 1 +$ $.7 \times 1 = 1$	$.3 \times 1 +$ $.7 \times .3 = .51$	$.3 \times 1 +$ $.7 \times .3 = .51$	$.3 \times 1 +$ $.7 \times 0 = .3$	$.3 \times 1 +$ $.7 \times 0 = .3$	1	0

TABLE 24.3

	$.3f_2(x_1+y_1)+.7f_2(x_1-y_1)$			Optimum	
x_1	$y_1 = 0$	1	2	f_1	y_1
2	$.3 \times .3 + .7 \times .3 = .3$	$.3 \times .51 + .7 \times .09 = .216$	$.3 \times 1 + .7 \times 0 = .3$	.3	0

PROBLEM SET 24.3A

1. In Example 24.3-1, stage 1 indicates the alternative optima $y_1 = 0$ and $y_1 = 2$. Show that $y_1 = 2$ (i.e., invest all in year 1) will still lead to the same probability of achieving the desired goal.
2. Solve Example 24.3-1 using the following data: The investor's goal is to maximize the probability of realizing at least \$6000 at the end of year 3. The amount available to the investor is \$1000 and the probability of doubling the money in any year is .6.

*3. You play a game in a casino where you bet a certain amount of money and flip a fair coin twice. For each \$1 you bet, the casino pays you \$3 (that gives you a net gain of \$2) if the outcome is *HH*. Otherwise, you lose the amount you bet. Assuming that you have a total of \$1, determine the game strategy, given that the objective is to maximize the probability of ending up with \$4 after three games.

REFERENCES

Bertsekas, D., *Dynamic Programming: Deterministic and Stochastic Models*, Prentice Hall, Upper Saddle River, NJ, 1987.

Cooper, L., and M. Cooper, *Introduction to Dynamic Programming*, Pergamon Press, New York, 1981.

Smith, D., *Dynamic Programming: A Practical Introduction*, Ellis Horwood, London, 1991.