ORIE 3310/5310 — 12 March 2013

Lecture 15:

Today:
- Stochastic DP.

- So far, we have been looking at deterministic Dynamic programming problems. This means that
  given the current state and the decision, the state at the next stage is completely determined

  Ex: If the initial inventory at the beginning of period 1 is 0, and we produce $X_1 = 5$ units, then the initial inventory at the beginning of period 2 is completely determined: It is $I + X_1 - d_1$, where d_1 = demand in period 1, known.

- In Stochastic dynamic programming problems, however, given the current state and a decision, the state at the next stage is not completely determined. Instead, there is a probability distribution that specifies a set of possible states at the next stage.

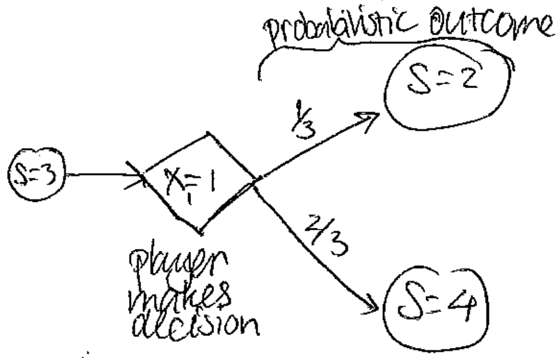
EX A betting game.

- Player starts with three chips.
- The game has three plays
- At each play, the player can bet any number of her chips
  - with probability $\frac{1}{3}$, the player loses # chips that are bet
  - with probability $\frac{2}{3}$, the player wins # chips that are bet.
- Goal: to finish the game with (at least) 5 chips.

- First, let's do a small example to see how this game is played.

<u>Ex</u>: Start with $s=3$ chips. → can bet 0, 1, 2, or 3 chips.
<u>Play #1</u>: bet $x_1 = 1$ chip.

Then, will end up with 2 chips with prob $= 1/3$, and
will end up with 4 chips with prob $= 2/3$.



<u>Note</u>: Must <u>wait</u> to see the outcome before moving on to Play #2.

Suppose: $s = 2$. → can bet 0, 1, or 2 chips.
<u>Play #2</u>: bet $x_2 = 1$ chip.
Then will end up w/ 1 chip w/ prob $= 1/3$
3 chips w/ prob $= 2/3$.

Then, suppose $s = 3$
Play #3: bet $x_3 = 2$ chips.
Then will end up with 1 chip w/ prob $= 1/3$
3 chips w/ prob $= 2/3$.

→ Can "think of this as a 2-player game: player vs. "nature"
where "nature" is a probabilistic player.

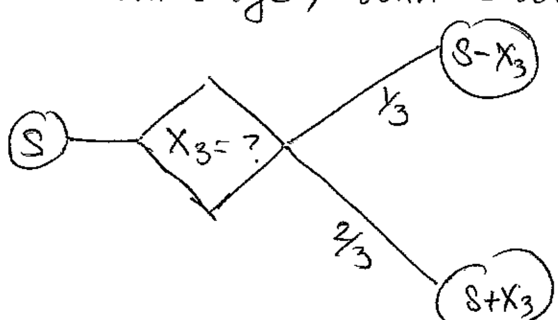
- A Dynamic programming approach :
  - Stage $k \leftrightarrow$ play #k.
  - State $\leftrightarrow$ remaining #coins at beginning of the stage.
    $\rightarrow S_1 = \{3\}$, $S_2 = \{0, 1, \ldots, \}$, $S_3 = \{0, \ldots, 5\}$.
  - At stage k, state s,
    decide : x_k = # coins to bet at stage k.
    $x_k \in \{0, 1, \ldots, \min\{s, 5-s\}\}$.
    <u>Note</u> : x_k can be any # from $0, 1, \ldots, s$ s.t. $s - x_k \geq 0$ and $s + x_k \leq 5$. | $\therefore$ $x_k \leq s$ and $x_k \leq 5 - s$.
- Optimization function :

  $f_k^*(s)$ = max probability of having at least 5 coins at the end, given that s coins remain at the beginning of stage k.

  $\downarrow$

  Note: At stage k, we consider plays $k, k+1, \ldots,$ end.

- Suppose $k = 3$ = last stage, with s coins remaining:



- Make decision $x_3 \in \{0, 1, \ldots, \min\{s, 5-s\}\}$.
- Since this is the last stage, then the probability of winning the game is simple to compute :
  - Check if $s - x_3 \geq 5$ or $s + x_3 \geq 5$.
  - Then, prob of winning the game is $= \begin{cases} 1 & \text{if } s - x_3 \geq 5, s + x_3 \geq 5 \\ 2/3 & \text{if } s - x_3 < 5, s + x_3 \geq 5 \\ 0 & \text{if } s - x_3 < 5, s + x_3 < 5 \end{cases}$

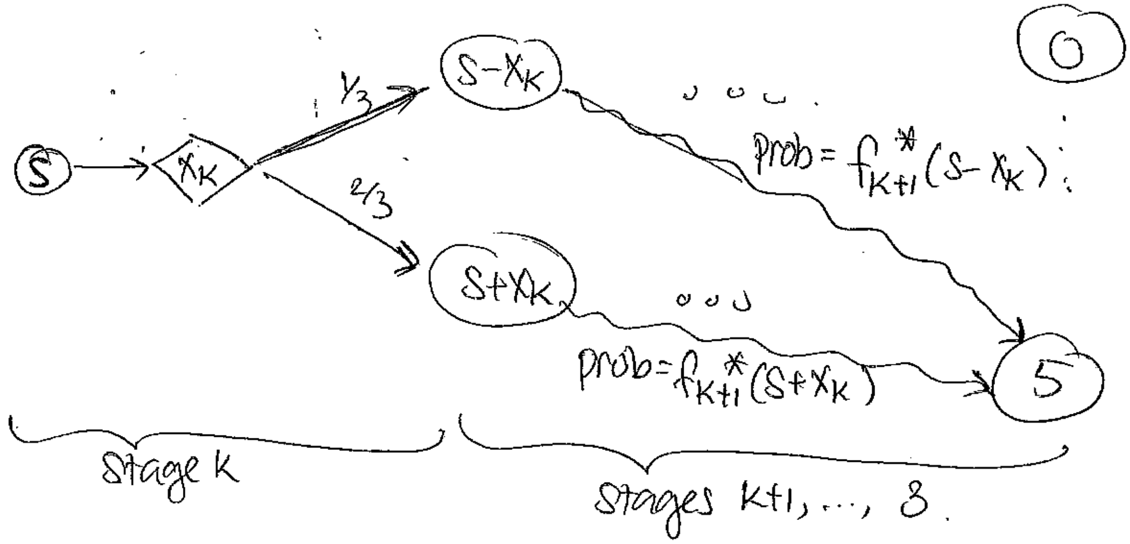
→ Choose x_3 that maximizes the probability of winning.

$$
\begin{aligned}
f_3^*(0) &= 0 & \quad x_3^* &= 0 \\
f_3^*(1) &= 0 & \quad x_3^* &= 0, 1 \\
f_3^*(2) &= 0 & \quad x_3^* &= 0, 1, 2 \\
f_3^*(3) &= 2/3 & \quad x_3^* &= 2 \\
f_3^*(4) &= 2/3 & \quad x_3^* &= 1 \\
f_3^*(5) &= 1 & \quad x_3^* &= 0
\end{aligned}
$$

These are the boundary conditions.

- Recurrence relation:

$$f_k^*(s) = \max_{\substack{x_k = 0, 1, \ldots, \\ \min\{5, 5-s\}}} \left\{ \begin{array}{l} \text{prob that we'll have 5 coins at end of 3rd play} \\ \ldots \text{ remain at beginning} \\ \text{of stage } k \text{, and we } \ldots\ x_k \text{ coins.} \end{array} \right\}$$



∴ Starting w/ s coins at stage k, if we bet x_k coins, the prob. of having 5 coins at the end is:

$$\frac{1}{3} \cdot f_{k+1}^*(s-x_k) + \frac{2}{3} f_{k+1}^*(s+x_k). \qquad (1)$$

So, choose x_k that maximizes (1):

$$f_k^*(s) = \max_{x_k \in \{0, \ldots, \min\{5, 5-s\}\}} \left\{ \frac{1}{3} f_{k+1}^*(s-x_k) + \frac{2}{3} f_{k+1}^*(s+x_k) \right\}.$$

- Computation:

<u>Stage 3</u> → bdry conditions.

<u>Stage 2</u>:

$$f_2^*(0) = \tfrac{1}{3} f_3^*(0) + \tfrac{2}{3} f_3^*(0) = 0 \qquad x_2^* = 0.$$

$$f_2^*(1) = \max_{x_2 \in \{0,1\}} \left\{ \tfrac{1}{3} f_3^*(1-x_K) + \tfrac{2}{3} f_3^*(1+x_K) \right\}$$

$$= \max \left\{ \underbrace{\tfrac{1}{3} f_3^*(1) + \tfrac{2}{3} f_3^*(1)}_{x_2=0},\ \underbrace{\tfrac{1}{3} f_3^*(0) + \tfrac{2}{3} f_3^*(2)}_{x_2=1} \right\}$$

$$= 0 \qquad x_2^* = 0, 1.$$

$$f_2^*(2) = \max \left\{ \tfrac{1}{3} f_3^*(2) + \tfrac{2}{3} f_3^*(2),\ \tfrac{1}{3} f_3^*(1) + \tfrac{2}{3} f_3^*(3),\ \tfrac{1}{3} f_3^*(0) + \tfrac{2}{3} f_3^*(4) \right\}$$

$$= 4/9 \qquad x_2^* = 1 \text{ or } 2.$$

$$f_2^*(3) = \max \left\{ \tfrac{1}{3} f_3^*(3) + \tfrac{2}{3} f_3^*(3),\ \tfrac{1}{3} f_3^*(2) + \tfrac{2}{3} f_3^*(4),\ \tfrac{1}{3} f_3^*(1) + \tfrac{2}{3} f_3^*(5) \right\} = \tfrac{2}{3} \qquad x_2^* = 2.$$

$$f_2^*(4) = \max \left\{ \tfrac{1}{3} f_3^*(4) + \tfrac{2}{3} f_3^*(4),\ \tfrac{1}{3} f_3^*(3) + \tfrac{2}{3} f_3^*(5) \right\}$$

$$= \tfrac{1}{3} \cdot \tfrac{2}{3} + \tfrac{2}{3} = \tfrac{8}{9} \qquad x_2^* = 1.$$

$$f_2^*(5) = \cancel{\max}\ \tfrac{1}{3} f_3^*(5) + \tfrac{2}{3} f_3^*(5) = 1 \qquad x_2^* = 0.$$

Stage 1:

$f_1^*(0) = 0 \qquad x_1^* = 0.$

$f_1^*(1) = \max\{ f_2^*(1),\ \frac{1}{3}f_2^*(0) + \frac{2}{3}f_2^*(1) \} = 0 \qquad x_1^* = 0.$

$f_1^*(2) = \max\{ f_2^*(2), \frac{1}{3}f_2^*(1) + \frac{2}{3}f_2^*(3),\ \underbrace{\frac{1}{3}f_2^*(0) + \frac{2}{3}f_2^*(4)} \}$
$= \frac{16}{27} \quad x_1^* = 2$

$f_1^*(3) = \max\{ f_2^*(3), \frac{1}{3}f_2^*(2) + \frac{2}{3}f_2^*(4), \frac{1}{3}f_2^*(1) + \frac{2}{3}f_2^*(5) \}.$
$= \max\{ \frac{2}{3},\ \underbrace{\frac{1}{3}\cdot\frac{4}{9} + \frac{2}{3}\cdot\frac{8}{9}},\ \frac{2}{3} \}$
$= \frac{20}{27},\ x_1^* = 1.$

$f_1^*(4) = \max\{ f_2^*(4), \frac{1}{3}f_2^*(3) + \frac{2}{3}f_2^*(5) \}.$
$= \max\{ \frac{8}{9},\ \frac{1}{3}\cdot\frac{2}{3} + \frac{2}{3}\cdot 1 \} = \frac{8}{9} \qquad x_1^* = 0, 1.$

$f_1^*(5) = 1 \quad x_1^* = 0.$

DP Table: $f_K^*(s)$.

	Stage 1		Stage 2		Stage 3	
$s = 0$	0	$x_1^* = 0$	0,	$x_2^* = 0$	0.	$x_3^* = 0$.
1	0	$x_1^* = 0$	0,	$x_2^* = 0, 1$	0	$x_3^* = 0, 1$
2	16/27	$x_1^* = 2$	4/9,	$x_2^* = 1, 2$	0	$x_3^* = 0, 1, 2$
3	20/27	$x_1^* = 1$	2/3,	$x_2^* = 2$	2/3	$x_3^* = 2$
4	8/9	$x_1^* = 0, 1$	8/9,	$x_2^* = 1$	2/3	$x_3^* = 1$
5	1	$x_1^* = 0$	1,	$x_2^* = 0$	1	$x_3^* = 0$

In Game Theory "Lingo": this table gives you a "complete strategy" for maximizing your probability of winning the game. That is, at each play, you only need to see how many coins remain and play x_K^* according to the above table.