4 Writing down the model

- What we really want: a production policy with minimum expected cost
- Let's write down the recursive representation of our model/diagram
- We can then solve this recursion by working backwards and determine the minimum expected cost and associated optimal policy
- Stages:

Stage t represents month t $(t=1,2)$ or the end of the decision-making process $(t=3)$

- States:

State n represents having n batches in inventory at the beginning of the month $(n=0,1,2)$

- Transition probability $p(m \mid n, t, x_t)$ of moving from state n to state m in stage t under decision x_t:

$$p(m \mid n, t, x_t) = \begin{cases} \frac{1}{4} & \text{if } m = \min\{n + x_t, 2\} \quad (d_t = 0 \text{ w.p. } \tfrac{1}{4}) \\ \frac{3}{4} & \text{if } m = \max\{n + x_t - 2, 0\} \quad (d_t = 2 \text{ w.p. } \tfrac{3}{4}) \\ 0 & \text{o/w} \end{cases}$$

$n + x_t - 2$

- Contribution $c(m \mid n, t, x_t)$ of moving from state n to state m in stage t under decision x_t:

$$c(m \mid n, t, x_t) = \begin{cases} 3x_t + 1\min\{n + x_t, 2\} & \text{if } m = \min\{n + x_t, 2\} \\ 3x_t + 1\max\{n + x_t - 2, 0\} + \underbrace{5\max\{2 - (n + x_t), 0\}}_{*} & \text{if } m = \max\{n + x_t - 2, 0\} \\ \text{not defined.} & \text{otherwise} \end{cases}$$

If $(n+x_t) \geq 2$, then demand is met and $* = 0$.
If $(n+x_t) < 2$, then $2-(n+x_t)$ batches of demand are not met, and $* = 2-(n+x_t)$

- Allowable decisions x_t at stage t and state n:

Let x_t = number of batches to produce in month t

x_t must satisfy: $\quad x_t \in \{0, 1\}$

- In words, the value-to-go $f_t(n)$ at stage t and state n is:

$f_t(n)$ = minimum total expected production and holding cost for months $t, t+1, \ldots$ w/ n batches of inventory at the start of month t

for $t = 1, 2, 3$ and $n = 0, 1, 2$

- Boundary conditions:

$$f_3(n) = 0 \qquad \text{for } n = 0, 1, 2$$

- Value-to-go recursion:

$$f_t(n) = \min_{x_t \in \{0,1\}} \left\{ \begin{array}{l} \frac{1}{4}\left[3x_t + 1 \min\{n + x_t, 2\} + f_{t+1}\left(\min\{n + x_t, 2\}\right)\right] \\ + \frac{3}{4}\left[3x_t + 1 \max\{n + x_t - 2, 0\} + 5 \max\{2 - (n + x_t), 0\} \right. \\ \qquad \left. + f_{t+1}\left(\max\{n + x_t - 2, 0\}\right)\right] \end{array} \right\}$$

for $t = 1, 2$ and $n = 0, 1, 2$

- Desired value-to-go function value: $f_1(1)$

5 Interpreting the value-to-go function

- We can solve this recursion just like with a deterministic DP: start at the boundary conditions and work backwards
- For this problem, we get the following value-to-go function values $f_t(n)$ for $t = 1, 2$ and $n = 0, 1, 2$, as well as the decision x_t^* that attained each value:

t	n	$f_t(n)$	x_t^*
1	0	13.125	1
1	1	8.875	1
1	2	5.875	0
2	0	7	1
2	1	3.5	1
2	2	0.5	0

- Based on this, what should the company's policy be?

Month 1: $n = 1 \Rightarrow \left.\begin{array}{l} f_1(1) = 8.875 \\ x_1^* = 1 \end{array}\right\} \Rightarrow$ produce 1 batch.

Month 2: If $n = 2$, $\left.\begin{array}{l} f_2(2) = 0.5 \\ x_2^* = 0 \end{array}\right\} \Rightarrow$ produce 0 batches

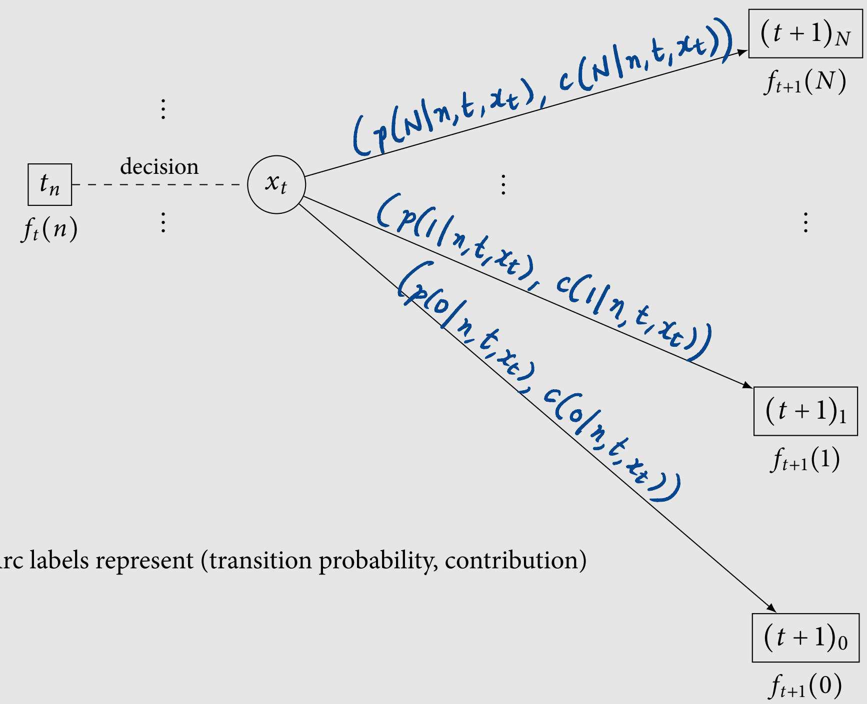
If $n = 0$, $\left.\begin{array}{l} f_2(0) = 7 \\ x_2^* = 1 \end{array}\right\} \Rightarrow$ produce 1 batch.

- What is the company's total expected cost? $f_1(1) = 8.875$

6 Stochastic dynamic programs, more generally

Stochastic dynamic program

- **Stages** $t = 1, 2, \ldots, T$ and **states** $n = 0, 1, 2, \ldots, N$
- Allowable **decisions** x_t at each stage t and state n
- **Transition probability** $p(m \mid n, t, x_t)$ of moving from state n to state m in stage t under decision x_t
- **Contribution** $c(m \mid n, t, x_t)$ for moving from state n to state m in stage t under decision x_t



Arc labels represent (transition probability, contribution)

- **Value-to-go** function $f_t(n)$ at each stage t and state n
- **Boundary conditions** on $f_T(n)$ for each state n
- **Recursion** on $f_t(n)$ at stage t and state n

$$f_t(n) = \overset{\text{(or max)}}{\min_{x_t \text{ allowable}}} \left\{ \sum_{m=0}^{N} p(m \mid n, t, x_t) \Big[c(m \mid n, t, x_t) + f_{t+1}(m) \Big] \right\}$$

for $t = 1, 2, \ldots, T-1$ and $n = 0, 1, \ldots, N$

- **Desired value-to-go**, usually $f_1(m)$ for some state m