Lesson 15.

Stochastic dynamic programming

1 The problem

The Hit-and-Miss Manufacturing Company has received an order to supply one item of a particular type. However, manufacturing this item is difficult, and the customer has specified such stringent quality requirements that the company may have to produce more than one item to obtain an item that is acceptable.

The company estimates that each item of this type will be acceptable with probability 1/2 and defective with probability 1/2. Each item costs $100 to produce, and excess items are worthless. In addition, a setup cost of $300 must be incurred whenever the production process is setup for this item. The company has time to make no more than 3 production runs, and at most 5 items can be produced in each run. If an acceptable item has not been obtained by the end of the third production run, the manufacturer is in breach of contract and must pay a penalty of $1600.

The objective is to determine how many items to produce in each production run in order to minimize the total expected cost.

2 Warm up

Example 1. Suppose the manufacturer produces x items in a single production run.

- What is the probability that at least one of these items is acceptable?

- What is the expected number of acceptable items?

3 Modeling the problem

- This problem involves **stochastic** – that is, random – data

⇒ We can't model the problem using the dynamic programming frameworks from previous lessons, which only deal with **deterministic** data

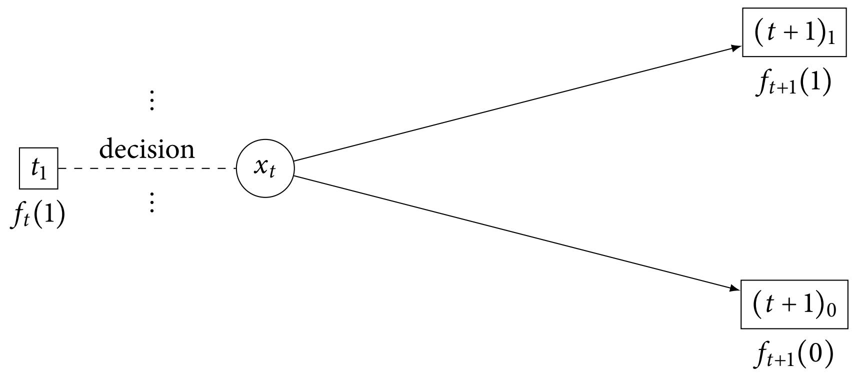
  - In particular, we can't model this problem as a shortest path problem, because the edges and edge lengths would be random

- However, we can still write a recursive representation of this problem
- Stages:

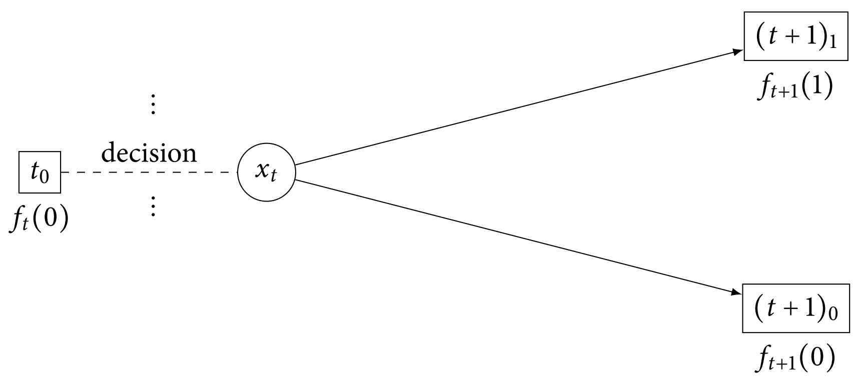
- States:

- Allowable decisions x_t at stage t and state n:

- Sketch of basic structure:
  - When the state $n = 1$:



  - When the state $n = 0$:



- Note that the state in the next stage depends on (i) the decision and (ii) randomness

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

- Cost-to-go recursion:

- Boundary conditions:

- Desired cost-to-go function value:

- 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 cost-to-go function values $f_t(n)$ for $t = 1, 2, 3$ and $n = 0, 1$, as well as the decision x_t^* that attained each value:

t	n	$f_t(n)$	x_t^*
1	0	0	0
1	1	675	2
2	0	0	0
2	1	70	2
3	0	0	0
3	1	80	3

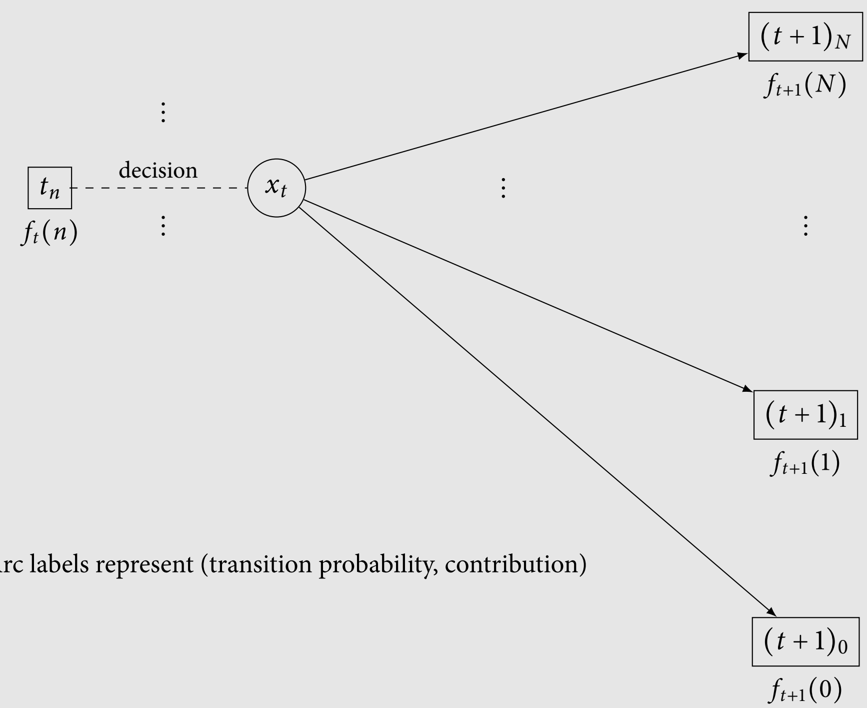
- Based on this, what should the company's policy be?

- What is the company's total expected cost?

4 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



- **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

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

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