EE365: Example: Dynamic Pricing

Dynamic pricing

$$x_{t+1} = \begin{cases} x_t - 1 & \text{if } w_t \geq u_t \text{ and } x_t > 0 \\ x_t & \text{otherwise} \end{cases}$$

- $x_t \in \mathcal{X} = \{0, 1, \ldots, n\}$ is stock at time t
- assume one customer arrives per period (time periods are very short)
- $w_t \in \{0, 1, 2\}$ is the *reservation price*, maximum price customer willing to pay distribution is $p_i = \mathbf{Prob}(w_t = i)$

$$\begin{bmatrix} p_0 & p_1 & p_2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/3 & 1/6 \end{bmatrix}$$

- $w_t = 0$ means no sale made at any price
- $u_t \in \{1, 2\}$ is the price we ask

Rewards

- if a sale is made, we receive the price u_t

$$g_t(x,u,w) = \begin{cases} u & \text{if } w \geq u \text{ and } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

- any stock leftover after *time horizon* $t = T$ we sell for salvage price $s > 0$

$$g_T(x) = sx$$

- mean total reward

$$J = \mathbf{E}\left(\sum_{t=0}^{T-1} g_t(x_t,u_t,w_t) + g_T(x_T)\right)$$

Random rewards

with state-feedback policy $u_t = \mu_t(x_t)$ we have a reward of the form

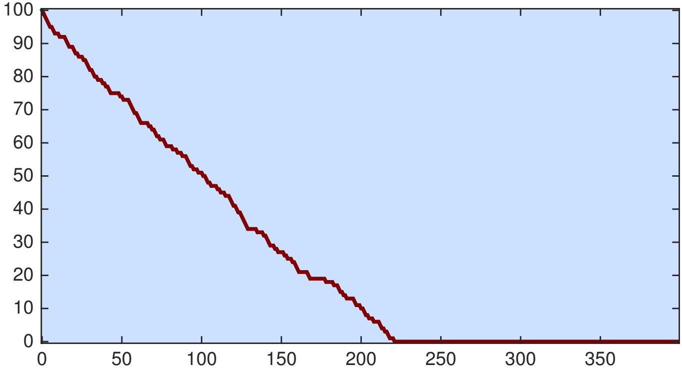
$$J = \mathbf{E}\left(\sum_{t=0}^{T-1} g_t(x_t, w_t) + g_T(x_T)\right)$$

define conditionally expected cost $h_t(x) = \mathbf{E}\big(g_t(x_t, w_t) \mid x_t = x\big)$, then

$$J = \mathbf{E}\left(\sum_{t=0}^{T-1} h_t(x_t) + g_T(x_T)\right)$$

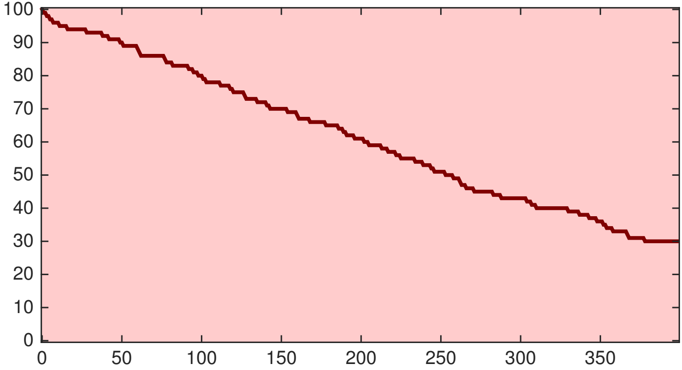
so can use value iteration for computation

Effect of prices



- $T = 400$, $n = 100$, $x_0 = 100$
- with policy $u_t = 1$, we always sell at low price (or not at all)
- almost always sell all stock ($\mathbf{Prob}(w_t \geq 1) = p_1 + p_2 = \frac{1}{2}$) so $J \approx 100$

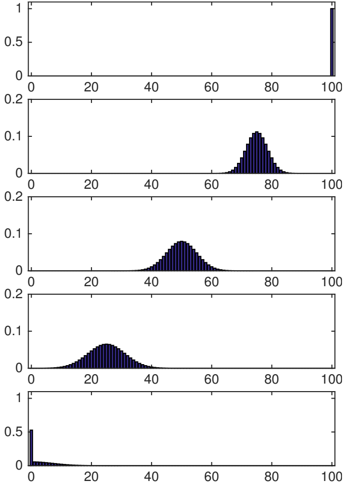
Effect of prices



- with policy $u_t = 2$, we always sell at high price (or not at all)
- on average sell 67 items $(\mathbf{Prob}(w_t \geq 2) = p_2 = \frac{1}{6})$
- so $J \approx 33 \times (0.1) + 67 \times 2 \approx 137$

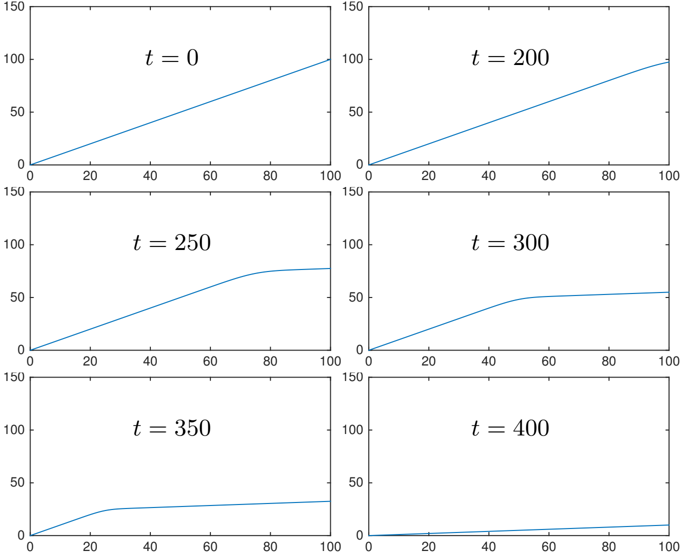
Distribution propagation

for $t = 0, 50, 100, 150, 200$, policy $u_t = 1$ on left, $u_t = 2$ on right



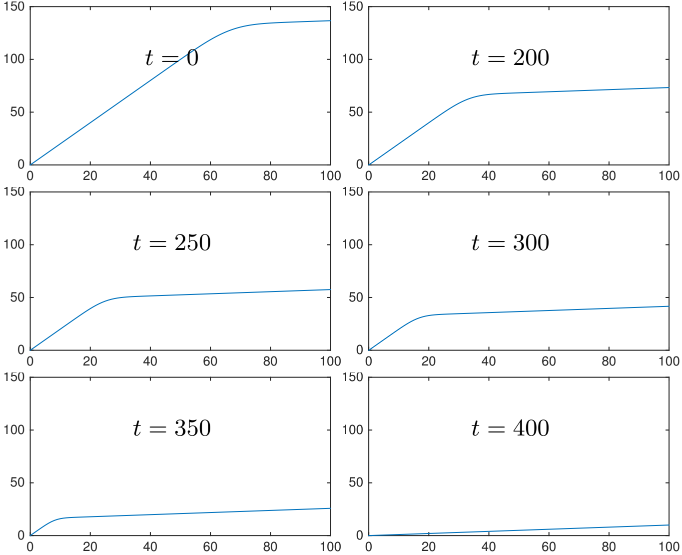
Value function

plots show $v_t(x)$ vs. x for policy $u_t = 1$



Value function

plots show $v_t(x)$ vs. x for policy $u_t = 2$



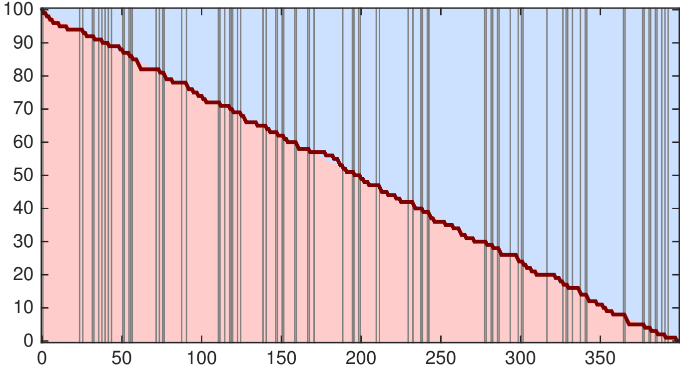
Rate control policy

- to do better, we try rate control policy $u_t = \mu_t(x_t)$ given by

$$\mu_t(x) = \begin{cases} 2 & \text{if } x/n < (T-t)/T \\ 1 & \text{otherwise} \end{cases}$$

- if the rate of sales so far is enough to make the target, sell at a high price

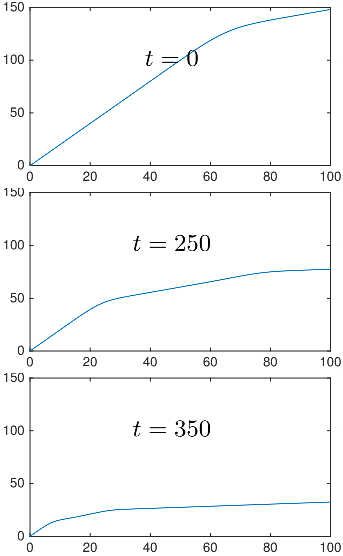
Simulation with rate control policy



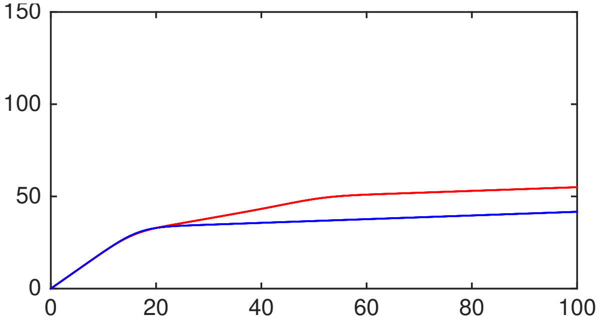
- vertical lines show times of switches between $u_t = 1$ and $u_t = 2$
- mean reward $J = 148$

Value function for rate control policy

plots show $v_t(x)$ vs. x



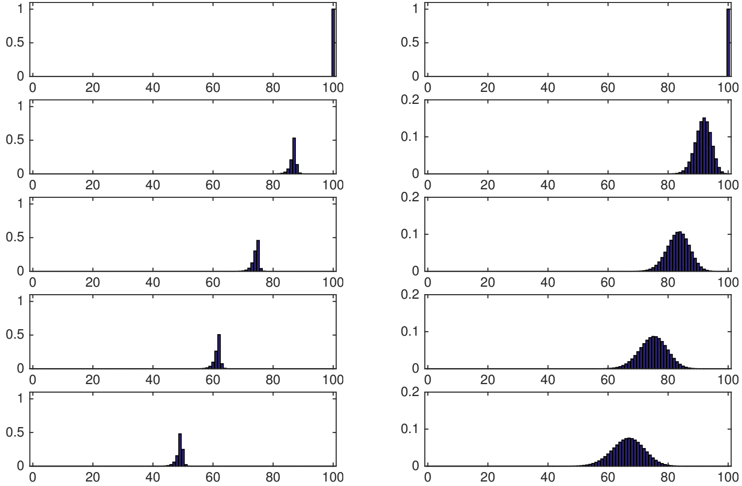
policy comparison



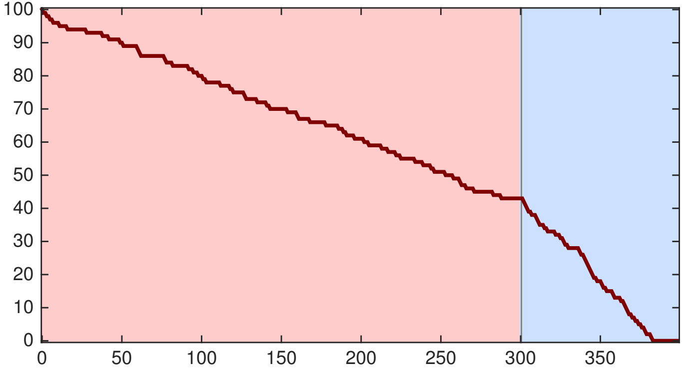
- $t = 300$, value function for rate control is red, for $u_t = 2$ in blue
- for small stock, no benefit to rate control policy
- for large stock, marginal benefit is salvage value

Distribution propagation for rate control policy

for $t = 0, 50, 100, 150, 200$, rate control policy on left, $u_t = 2$ on right

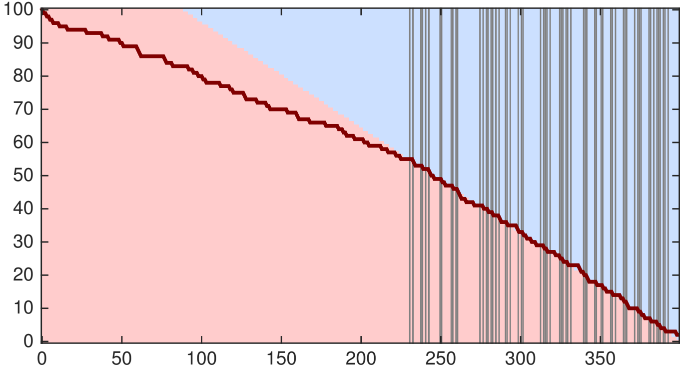


Simulation with closeout policy



- mean reward $J = 147$
- policy is $u_t = \begin{cases} 2 & \text{if } t \leq 300 \\ 1 & \text{otherwise} \end{cases}$

Simulation with optimal policy



- mean reward $J = 149.7$
- optimal policy sells at high price initially

Cost distributions via Monte Carlo