## Data Generation

### Backward Recursion

Let's start from the last period, we assume that there is no scrap value in the dynamic problem, so $V_{T+1}(s_{t+1}, a_{t+1}) = 0$ for all $s \in S$ and $a \in \{0, 1\}$. We have:

$$\pi_T(s_T, a_T) = a_T[-R - \nu(1)] + (1 - a_T)[-\mu \cdot s - \nu(0)]$$

$$V_T(s_T, a_T) = \max_{a_T} [-R - \nu(1)] + (1 - a_T)[\mu \cdot s - \nu(0)] + \beta \mathbb{E}\left(\underbrace{V_{T+1}(s_{T+1}, a_{T+1}}_{=0}\right)$$

$$V_T(s_T, a_T) = \max_{a_T} \pi_T(s_T, a_T)$$

Clearly given the discrete nature of the problem, the optimal problem has the following form:

$$\begin{aligned} \sigma_T(s_T, \nu) = 1 &\iff -R - \nu_T(1) \geq -\mu \cdot s_T - \nu_T(0) \\ &\iff \mu \cdot s_T - R \geq \nu_T(1) - \nu_T(0) \end{aligned}$$

and given that $\nu(a) \sim N(0, 1)$ for $a \in \{0, 1\}$ then $\nu_T(1) - \nu_T(0) \sim N(\mu \cdot s - R, 2) \Rightarrow \Pr(a = 1|s) = 1 \Rightarrow \Phi\left(\frac{\mu \cdot s_T - R}{\sqrt{2}}\right)$.

At time $T-1$ the problem starts being dynamic, given that the action today determines the state tomorrow:

$$V_{T-1}(s_{T-1}, a_{T-1}) = \max_{a_{T-1}} a_{T-1}[-R - \nu(1)_{T-1}] + (1 - a_{T-1})[\mu \cdot s_{T-1} - \nu_{T-1}(0)] + \beta \mathbb{E}[V_T(s_T, a_T)]$$

and the transition is given by $s_T = 1$ if $a_{T-1} = 1$ and $s_T = \min\{s_{T-1} + 1, S^{max}\}$. Define $\tilde{s}_t \equiv \min\{s_{t-1} + 1, S^{max}\}$

$$\begin{aligned} v_{T-1}(s_{T-1}, 1) &= -R + \beta \mathbb{E}(V_T(1, a_T)) \\ &= -R + \beta\left(\Phi\left(\frac{\mu - R}{\sqrt{2}}\right) \cdot (-R) + \left(1 - \Phi\left(\frac{\mu - R}{\sqrt{2}}\right)\right) \cdot (-\mu)\right) \\ &= -R + \beta\left[\Phi\left(\frac{\mu - R}{\sqrt{2}}\right)[\mu - R]\right] \\ v_{T-1}(s_{T-1}, 0) &= -\mu \cdot s_{T-1} + \beta \mathbb{E}(V_T(\tilde{s}_{T-1}, a_T)) \\ &= -\mu \cdot s_{T-1} + \beta\left(\Phi\left(\frac{\mu \cdot \tilde{s}_T - R}{\sqrt{2}}\right) \cdot (-R) + \left(1 - \Phi\left(\frac{\mu \cdot \tilde{s}_T - R}{\sqrt{2}}\right)\right)(-\mu \cdot \tilde{s}_T)\right) \end{aligned}$$

Again given the unobserved shocks, the optimal decision rule at time $T - 1$ can be written as:

$$V_{T-1}(s_{T-1}, a_{T-1}) = \max_{a_{T-1}} (v_{T-1}(s_{T-1}, 1) + \nu(1), v_{T-1}(s_{T-1}, 0) + \nu(0))$$

which gives the decision rule:

$$\begin{aligned}\sigma_{T-1}(s_{T-1},\nu_{T-1})=1 &\iff v_{T-1}(s_{T-1},1)+\nu(1)\geq v_{T-1}(s_{T-1},0)+\nu(0)\\ &\iff v_{T-1}(s_{T-1},1)-v_{T-1}(s_{T-1},0)\geq \nu(0)-\nu(1)\end{aligned}$$

so that

$$\Pr\left(\sigma_{T-1}(s_{T-1},\nu)=1\right)=\Phi\left(\frac{v_{T-1}(s_{T-1},1)-v_{T-1}(s_{T-1},0)}{\sqrt{2}}\right)$$

To ease the notation we omit the state variables when referring to the optimal policy. Notice that $\Pr\left(\sigma_{T-1}=1\right)$ is of dimension $|S|\times 1$. Hence, recursively, we can define a matrix containing optimal cutoff rules. Assume we want to generate the data for $T$ periods[1], then we define $\Sigma$ as the matrix of dimension $T\times S$ containing the cutoff rules for every period:

$$\Sigma=\begin{bmatrix}\Pr\left(\sigma_T=1\right)'\\ \Pr\left(\sigma_{T-1}=1\right)'\\ \vdots\\ \Pr\left(\sigma_0=1\right)'\end{bmatrix}=\begin{bmatrix}\Phi\left(\frac{v_{T-1}(s_1,1)-v_{T-1}(s_1,0)}{\sqrt{2}}\right) & \dots & \Phi\left(\frac{v_{T-1}(s_5,1)-v_{T-1}(s_5,0)}{\sqrt{2}}\right)\\ \vdots & \dots & \vdots\\ \Phi\left(\frac{v_0(s_1,1)-v_0(s_1,0)}{\sqrt{2}}\right) & \dots & \Phi\left(\frac{v_{T0}(s_5,1)-v_0(s_5,0)}{\sqrt{2}}\right)\end{bmatrix}$$

**Forward Data Simulation**

Finally, we can simulate forward $N$ observations so that the final artificial dataset will be of dimension $N\times T$.

[1] Assume also that $S=\{1,2,3,4,5\}$ so that $S^{max}=5$.