Optimal Stopping

- **Setting** - State space $S \cup \{\text{Stopped}\}$
  - Action space $A = \{0, 1\}$ stop ↗ ↖ continue
  - If $A_t = 0$, then $S_{t+i} = \text{'Stopped'}$ for all $i \geq 1$
  - Costs $C_t(x, 0) = c(x)$ (continuation cost)
    $x \neq \text{Stopped}$ $C_t(x, 1) = s(x)$ (stopping cost)
    $C_t(\text{Stopped}, \cdot) = 0$

- If we play this over a horizon T, then HJB eqn

$$V_t(x) = \min\Big[\underbrace{c(x) + \mathbb{E}[V_{t+1}(X_{t+1}) \mid X_t = x, A_t = 1]}_{a=1,\ \text{ie, continue}}, \underbrace{s(x)}_{a=0}\Big] \quad ; x \in S$$

$V_t(\text{stopped}) = 0$, $V_T(x) = 0 \ \forall x$

- 1-Step Look-Ahead Rule (1-SLA) - Stop whenever stopping now is better than stopping in next step, ie, $\forall t$ $A_t = 1$ under the event

$$T_t \triangleq \{ s(X_t) \leq c(X_t) + \mathbb{E}[s(X_{t+1}) \mid X_t = x, A_t = 1] \}$$

Thm (Chow & Robbins '61) - Suppose T_t is monotone, i.e. $T_0 \subseteq T_1 \subseteq T_2 \ldots$ a.s. Then 1-sla is optimal,

Pf - Induction on $t \in \{T-1, T-2, \ldots, 0\}$. For $t = T-1$, its clear we should stop if $s(X_t) \leq c(X_t) \Leftrightarrow X_t \in T$. Now if its true for $t \geq k$, then at time $t = k-1$, if $X_t \notin T$, then its better to continue till $t = k$; if $X_t \in T$, then $X_{t+1} \in T$ and by induction hypothesis, its better to stop.

Eg (Selling an asset) - $X_1, X_2, \dots, X_T \sim F$ are iid offers for an asset with $Var(X_t) < \infty$; Per day holding cost c; also, can recall the best past offer anytime.

- $s(X_t) = -\underbrace{\max_{t' \leq t}\{X_{t'}\}}_{= M_t}$, $c(X_t) + \mathbb{E}[s(X_{t+1})] = c - \mathbb{E}[\max(X_{t+1}, M_t) | X_1, \dots, X_t]$

$$\Rightarrow s(X_t) - c(X_t) - \mathbb{E}[s(X_{t+1})] = \underbrace{\mathbb{E}\left[(X_{t+1} - M_t)^+ | M_t\right]}_{\text{decreasing as fn of } M_t} - c$$

$\Rightarrow$ 1-sLa is optimal!

Note - Not true without recall (ie, for the 'prophet inequality')

- Corollary - If the set of 'allowed stopping points' is determined by X_t, then also we can define monotone stopping rules

Eg - Last-Success and Bruss' Odds Policy $(q_k = 1 - p_k)$

- Sequence of T independent events $X_1, X_2, \dots, X_T$, $X_k \sim Ber(p_k)$
- Objective: Maximize probability of stopping at the 'last success' i.e., $R(X_t = 1, A_t = 0) = \mathbb{P}[X_{t+k} = 0 \ \forall k \geq 1]$

- Let $\theta_i = p_i / q_i$ ('odds' for i^{th} event)
  and $O_i = \sum_{t > i} \theta_i$

Thm (Bruss 2000) - The optimal stopping policy is to stop after the last time t with $I_t = 1$ and $O_t \geq 1$

Pf - Let $T_t = \{X_t = 1$, and stopping at next success is worse$\}$

$$= \{X_t = 1 \text{ and } q_{t+1} q_{t+2}, \dots, q_T > p_{t+1} q_{t+2} \dots q_T + \dots + q_t \dots q_{T-1} p_T\}$$

$$= \{X_t = 1, \ 1 > \theta_{t+1} + \theta_{t+2} + \dots + \theta_T\}$$

Clearly T_t is monotone $\Rightarrow$ opt policy is stop after last t s.t $\sum_{t' \geq t} \theta_{t'} \geq 1$

Corollary - For secretary problem, stop after last t st. $H_t \geq 1$

Optimal Stopping in Infinite Horizon Settings

- Above we assumed T was finite; can 1-sla work if $T \to \infty$

Eg - $X_t = X_{t-1} + Z_t$, with $Z_t = \mp 1$ wp $1/2$ (symmetric RW)
$c(X_t) = 0 \ \forall t$, $s(X_t) = e^{-X_t}$
- $s(X_t) = e^{-X_t}$, $c(X_t) + \mathbb{E}[s(X_{t+1})] = e^{-X_t}(e + e^{-1})/2 = 1.54 e^{-X_t} > s(X_t)$
moreover $T_t = \mathbb{Z}$ which is monotone
- However X_t is recurrent $\Rightarrow$ eventually hits every $x \in \mathbb{Z}$
$\Rightarrow$ opt policy is to wait forever (with cost 0)...

- However, we can avoid these pathologies with additional assumptions
Lemma - If $S = \sup_x s(x) < \infty$, $C = \inf_x c(x) > 0 \Rightarrow F_T(x) \to F(x)$ as $T \nearrow \infty$
Moreover a 1-sla is optimal if and only if T_t is monotone
(see Weber section 6.3)

Eg - (Sequential Prob Ratio Test) - Suppose $X_1, X_2, \ldots \sim F$, where our prior is that $\underbrace{F = F_0}_{H_0}$ wp p_0, $\underbrace{F = F_1}_{H_1}$ wp $p_1 = 1 - p_0$.
- To get each additional sample, we pay cost r
- If we stop and declare $H_0 | H_1$ is true, we pay cost $s_0 | s_1$ if wrong

- $l_0 = \dfrac{p_1}{p_0} = \dfrac{\mathbb{P}[F = F_1]}{\mathbb{P}[F = F_0]}$, $l_n = \dfrac{\mathbb{P}[F = F_1 | X_1, \ldots, X_n]}{\mathbb{P}[F = F_0 | X_1 \ldots X_n]} = \dfrac{f_1(X_n)}{f_0(X_n)} l_{n-1}$ (likelihood ratios)
- We now want to consider a 1-sla policy
— Stopping at time t =
$$s(L_t) = \min\left(\frac{c_0 L_t}{1 + L_t}, \frac{c_1}{1 + L_t}\right) \Rightarrow (1 + L_t) s(L_t) = \min(c_0 L_t, c_1)$$

- Continuing at t and stopping in next step

$$c(L_t) + \mathbb{E}[s(L_{t+1})|L_t] = \gamma + \frac{L_t}{1+L_t}\mathbb{E}_{X_{t+1}\sim F_1}[s(L_{t+1})] + \frac{1}{1+L_t}\mathbb{E}_{X_{t+1}\sim F_0}[s(L_{t+1})]$$

$$\Rightarrow (1+L_t)(c(L_t)+\mathbb{E}[s(L_{t+1})|L_t]) = \gamma(1+L_t) + \int L_t f_1(y)\, s(L_t f_1(y)/f_0(y))\,dy + \int f_0(y)\, s(L_t f_1(y)/f_0(y))\,dy$$

$$= \gamma(1+L_t) + \int (1+L_t f_1(y)/f_0(y))\, s(L_{t+1}) f_0(y)\,dy$$

$$= \gamma(1+L_t) + \mathbb{E}_{X_{t+1}\sim F_0}[(1+L_{t+1})\, s(L_{t+1})]$$

Thus the 1-sla rule corresponds to the following:

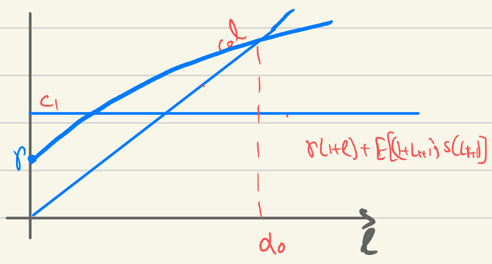
1) If $c_0 L_t \le \gamma(1+L_t) + \mathbb{E}_{F_0}[(1+L_{t+1})\, s(L_{t+1})] \Rightarrow$ Stop and declare $F=F_0$

2) If $c_1 \le \gamma(1+L_t) + \mathbb{E}_{F_0}[(1+L_{t+1})\, s(L_{t+1})] \Rightarrow$ Stop and declare $F=F_1$

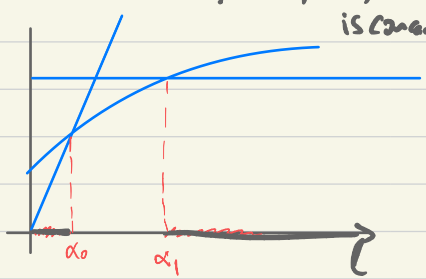
3) Else continue

- Note $(1+\ell)\, s(\ell) = \min(c_0\ell, c_1) \equiv$ concave, non-dec^t in ℓ

$\Rightarrow \mathbb{E}_{F_0}[(1+L_{t+1})\, s(L_{t+1})|L_t=\ell] = \mathbb{E}_{Y\sim F_0}[\min(c_0\ell f_1(Y)/f_0(Y), c_1)]$ is concave non-dec^t in ℓ



2 intervals



3 intervals

It can be seen from the plots that the 1-sla either has a single threshold α_0, or 2 thresholds $\alpha_0 < \alpha_1$. This is clearly monotone $\Rightarrow$ the 1-sla is optimal.