A Sequential Stochastic Assignment Problem with Random Number of Jobs

Yaakov Malinovsky

University of Maryland, Baltimore County, USA

joint work with

Alexander Goldenshluger, University of Haifa, Israel

Assaf Zeevi, Columbia University, USA

36th Annual Quality and Productivity Research Conference

June 10-13, 2019

American University

Washington, DC

Outline

- Sequential Stochastic Assignment Problem with Fixed Number of Jobs
- Sequential Stochastic Assignment Problem with Random Number of Jobs
- Selecting one of the k Best Values with Random Number of Alternatives

Part 1: Sequential Stochastic Assignment Problem with Fixed Number of Jobs

Classical Formulation [a]

- Suppose that n jobs arrive sequentially in time.
- The tth job, $1 \leqslant t \leqslant n$, is identified with a random variable Y_t which is observed.
- The jobs must be assigned to n persons which have known "values" $p_1, \cdots, p_n$.
- If the tth job is assigned to the jth person then a reward of $p_j Y_t$ is obtained and the person j becomes unavailable.
- The goal: to maximize expected total reward

$$S_n(\pi) := E \sum_{t=1}^{n} p_{\pi_t} Y_t.$$

[a] Derman, Lieberman & Ross (1972). A sequential stochastic assignment problem. *Management Science*, **1**8, 349–355.

Notations

- Assume that $Y_1, \ldots, Y_n$ are integrable independent random variables defined on probability space $(\Omega, \mathscr{F}, \mathrm{P})$.
- let F_t [a] be the distribution function of Y_t, $t = 1, \ldots, n$.
- Let $\mathscr{Y}_t$ denote the σ–field generated by $(Y_1, \ldots, Y_t)$:
  $\mathscr{Y}_t = \sigma(Y_1, \ldots, Y_t)$, $1 \leqslant t \leqslant n$.
- $\pi = (\pi_1, \ldots, \pi_n)$ is a permutation of $\{1, \ldots, n\}$ defined on $(\Omega, \mathscr{F})$.
- We say that π *is an assignment policy* if $\{\pi_t = j\} \in \mathscr{Y}_t$ for every $1 \leqslant j \leqslant n$ and $1 \leqslant t \leqslant n$:
  π is a policy if it is non–anticipating relative to the filtration $\mathscr{Y} = \{\mathscr{Y}_t, 1 \leqslant t \leqslant n\}$ so that tth job is assigned on the basis of information in $\mathscr{Y}_t$.

---

[a] Albright, S. C., Jr. (1972). Stochastic sequential assignment problems. *Technical report*, **147**, Stanford University.

Formal Statement: Problem (AP1)

- Given a vector $p = (p_1, \dots, p_n)$, with $p_1 \leqslant p_2 \leqslant \cdots \leqslant p_n$,
- we want to maximize *the total expected reward* $S_n(\pi) := \mathrm{E}\sum_{t=1}^{n} p_{\pi_t} Y_t$ with respect to $\pi \in \Pi(\mathscr{Y})$.
- The policy π^* is called *optimal* if $S_n(\pi^*) = \sup_{\pi \in \Pi(\mathscr{Y})} S_n(\pi)$.

Useful representation:

$$\sum_{t=1}^{n} p_{\pi_t} Y_t = \sum_{t=1}^{n} \sum_{j=1}^{n} p_j Y_t \mathbf{1}\{\pi_t = j\} = \sum_{j=1}^{n} p_j Y_{\nu_j};$$

- ν_j denotes the index of the job to which the jth person is assigned: $\{\nu_j = t\} = \{\pi_t = j\}$, $1 \leqslant t \leqslant n$, $1 \leqslant j \leqslant n$.

Backward Induction Solution

- Theorem (DLR, 1972; Albright, 1972):
  - There exist real numbers $-\infty \equiv a_{0,n} \leqslant a_{1,n} \leqslant \cdots \leqslant a_{n-1,n} \leqslant a_{n,n} \equiv \infty$ such that on the first step, when $Y_1 \sim F_1$ is observed, the optimal policy is

$$\pi_1^* = \sum_{j=1}^{n} j\mathbf{1}\{Y_1 \in (a_{j-1,n}, a_{j,n}]\}.$$

  - $\{a_{j,n}\}_{j=1}^{n}$ do not depend on $p_1, \ldots, p_n$ and are determined by

$$a_{j,n+1} = \int_{a_{j-1,n}}^{a_{j,n}} z\mathrm{d}F_1(z) + a_{j-1,n}F_1(a_{j-1,n}) + a_{j,n}[1 - F_1(a_{j,n})],$$

$j = 1, \ldots, n$, where $-\infty \cdot 0 \equiv 0 \equiv \infty \cdot 0$.

Backward Induction Solution (con't)

- At the end of the first stage the assigned p is removed from the feasible set and the process repeats with the next observation, where the above calculation is then performed relative to the distribution F_2 and real numbers $-\infty \equiv a_{0,n-1} \leqslant a_{1,n-1} \leqslant \cdots \leqslant a_{n-2,n-1} \leqslant a_{n-1,n-1} \equiv \infty$ are determined, and so on.
- Moreover,

$$a_{j,n+1} = \mathrm{E}Y_{\nu_j}, \ \forall 1 \leqslant j \leqslant n,$$

  i.e., $a_{j,n+1}$ is the expected value of the job which is assigned to the jth person.

Remark and Example

- By backward induction we determine a triangular array, where we use F_{n-t+2} to determine $\{a_{\cdot,t}\}$:

  $a_{1,2}$

  $a_{1,3},\ a_{2,3}$

  $a_{1,4},\ a_{2,4},\ a_{3,4}$

  $\vdots$

  $a_{1,n},\ a_{2,n},\ldots,a_{n-1,n}$

  $a_{1,n+1},\ a_{2,n+1},\ldots,a_{n,n+1}\ \Rightarrow S_n(\pi^*) = p_1 \cdot a_{1,n+1} + \cdots + p_n \cdot a_{n,n+1}$

- Example: $X_1 \sim X_2 \sim X_3 \sim Uniform[0,1]$

  $a_{1,2} = 1/2$

  $a_{1,3} = 3/8,\ a_{2,3} = 5/8$

  $a_{1,4} = 39/128,\ a_{2,4} = 39/128,\ a_{3,4} = 89/128 \Rightarrow$

$$S_3 = p_1 \cdot 39/128 + p_2 \cdot 1/2 + p_3 \cdot 89/128.$$

Part 2: Sequential Stochastic Assignment Problem with Random Number of Jobs

Problem (AP2)

- Let N be a positive integer-valued random variable with known distribution $\gamma = \{\gamma_k\}$, $\gamma_k = \mathrm{P}(N = k)$, $k = 1, \ldots, N_{\max}$, where $N_{\max}$ can be infinite.
- Let $Y_1, Y_2, \ldots$ be an infinite sequence of integrable independent random variables with distributions $F_1, F_2, \ldots$, independent of N.
- Given real numbers $p_1 \leqslant \ldots \leqslant p_{N_{\max}}$ the objective is to maximize the expected total reward

$$S_\gamma(\pi) = \mathrm{E} \sum_{t=1}^{N} p_{\pi_t} Y_t$$

over all policies $\pi \in \Pi(\mathscr{Y})$.

Random Number of Jobs

- Theorem: [a]
- In Problem (AP2) assume that $N_{\max} < \infty$ and let

$$\tilde{Y}_t := Y_t \sum_{k=t}^{N_{\max}} \gamma_k, \quad t = 1, \ldots, N_{\max}$$

.

- For any $\pi \in \Pi(\mathscr{Y})$ one has

$$S_\gamma(\pi) = \mathrm{E} \sum_{t=1}^{N_{\max}} p_{\pi_t} \tilde{Y}_t,$$

and the optimal policy in Problem (AP2) coincides with the optimal policy in Problem (AP1) associated with fixed horizon $n = N_{\max}$ and job sizes $\tilde{Y}_1, \ldots, \tilde{Y}_{N_{\max}}$.

[a] Goldenshluger, A., Malinovsky, Y., Zeevi, A. (2019). A Unified Approach for Solving Sequential Selection Problems. *arXiv:1901.04183*.

Proof

- For any $\pi \in \Pi(\mathscr{Y})$ we have
  $S_\gamma(\pi) = \mathrm{E}\sum_{t=1}^{N} p_{\pi_t} Y_t = \sum_{t=1}^{N_{\max}} \mathrm{E}[p_{\pi_t} Y_t \mathbf{1}(N \geqslant t)]$,
- and

$$\mathrm{E}[p_{\pi_t} Y_t \mathbf{1}(N \geqslant t)] = \mathrm{E}\sum_{k=t}^{N_{\max}} \mathrm{E}\Big\{[p_{\pi_t} Y_t \mathbf{1}(N = k)] \,|\, \mathscr{Y}_t\Big\} = \mathrm{E}\Big\{p_{\pi_t} Y_t \sum_{k=t}^{N_{\max}} \gamma_k\Big\}$$

$$= \mathrm{E}\Big\{p_{\pi_t} \tilde{Y}_t\Big\},$$

where we have used the fact that π_t is $\mathscr{Y}_t$–measurable, and N is independent of $\mathscr{Y}_t$.

- Therefore $\mathrm{E}\sum_{t=1}^{N} p_{\pi_t} Y_t = \mathrm{E}\sum_{t=1}^{N_{\max}} p_{\pi_t} \tilde{Y}_t$.
- Note that $\tilde{Y}_t$ are independent random variables, and σ-fields $\tilde{\mathscr{Y}}_t$ and $\mathscr{Y}_t$ are identical. This implies the stated result.

Part 3: Selecting one of the k Best Values with Random Number of Alternatives

Sequential Selection Problems

- Let $X_1, X_2, \ldots$ be an infinite sequence of independent identically distributed continuous random variables defined on a probability space $(\Omega, \mathscr{F}, \mathrm{P})$.
- $$R_t := \sum_{j=1}^{t} \mathbf{1}(X_t \leqslant X_j), \;\; A_{t,n} := \sum_{j=1}^{n} \mathbf{1}(X_t \leqslant X_j), \quad t = 1, \ldots, n.$$
- Let $\mathscr{R}_t := \sigma(R_1, \ldots, R_t)$ and $\mathscr{X}_t := \sigma(X_1, \ldots, X_t)$ denote the σ–fields generated by $R_1, \ldots, R_t$ and $X_1, \ldots, X_t$
- $\mathscr{R} = (\mathscr{R}_t, 1 \leqslant t \leqslant n)$ and $\mathscr{X} = (\mathscr{X}_t, 1 \leqslant t \leqslant n)$ are the corresponding filtrations.
- The class of all stopping times of a filtration $\mathscr{Y} = (\mathscr{Y}_t, 1 \leqslant t \leqslant n)$ will be denoted $\mathscr{T}(\mathscr{Y})$; i.e., $\tau \in \mathscr{T}(\mathscr{Y})$ if $\{\tau = t\} \in \mathscr{Y}_t$ for all $1 \leqslant t \leqslant n$.

Average Reward

- Fixed n: Problem (A1): Let n be a fixed positive integer, and let $q : \{1, 2, \dots, n\} \to \mathbb{R}$ be a reward function. The average reward of a stopping rule $\tau \in \mathscr{T}(\mathscr{R})$ is $V_n(q; \tau) := \mathrm{E}q\big(A_{\tau,n}\big)$, and we want to find the rule $\tau_* \in \mathscr{T}(\mathscr{R})$ such that

$$V_n^*(q) := \max_{\tau \in \mathscr{T}(\mathscr{R})} V_n(q; \tau) = \mathrm{E}q\big(A_{\tau_*,n}\big).$$

- Random N: Problem (A2): $\gamma_k = \mathrm{P}(N = k)$, $k = 1, 2, \dots, N_{\max}$, $N \perp \{X_t, t \geqslant 1\}$. Let $q : \{1, 2, \dots, N_{\max}\} \to \mathbb{R}$.

$$V_\gamma(q; \tau) := \mathrm{E}\big[q(A_{\tau,N})\mathbf{1}(\tau \leqslant N)\big].$$

We want to find the stopping rule $\tau_* \in \mathscr{T}(\mathscr{R})$ such that

$$V_\gamma^*(q) := \max_{\tau \in \mathscr{T}(\mathscr{R})} V_\gamma(q; \tau) = V_\gamma(q; \tau_*).$$

Fixed n: Gusein-Zade Stopping Problem [a]

- Selecting One of the k Best Values: $q(a) = q_{\mathrm{gz}}^{(k)}(a) := \mathbf{1}\{a \leqslant k\}$, and the problem is to maximize $\mathrm{P}\{A_{\tau,n} \leqslant k\}$ with respect to $\tau \in \mathscr{T}(\mathscr{R})$.
- The optimal policy: is determined by k natural numbers

$$1 \leqslant \pi_1 \leqslant \pi_2 \leqslant \cdots \leqslant \pi_k$$

  and proceeds as follows: pass the first $\pi_1 - 1$ observations and among the subsequent $\pi_1, \pi_1 + 1, \ldots, \pi_2 - 1$ choose the first best observation; if it does not exists then among the set of observations $\pi_2, \pi_2 + 1, \ldots, \pi_3 - 1$ choose one of the two best, etc.
- Example (n=30, k=3): $\pi_1 = 11$, $\pi_2 = 18$, $\pi_3 = 24$ and

[a] Gusein-Zade, S. M. (1966). The problem of choice and the optimal stopping rule for a sequence of independent trials. *Theory Probab. Appl.*, **1**1, 472–476.

$\max_{\tau \in \mathscr{T}(\mathscr{R})} \mathrm{P}\{A_{\tau,30} \leqslant 3\} = 0.73492.$

An Auxiliary Optimal Stopping Problem: Problem (B)

- Let $Y_1, \dots, Y_n$ be a sequence of integrable independent real-valued random variables with corresponding distributions $F_1, \dots, F_n$.

- For a stopping rule $\tau \in \mathscr{T}(\mathscr{Y})$ define $W_n(\tau) := \mathrm{E}Y_\tau$. The objective is to find the stopping rule $\tau_* \in \mathscr{T}(\mathscr{Y})$ such that

$$W_n^* := \max_{\tau \in \mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau = W_n(\tau_*) = \mathrm{E}Y_{\tau_*}.$$

DLR (1972) Solution of Problem (B)

- Consider Problem (AP1) with $p_1 = 0, p_2 = 0, \ldots, p_n = 1$ and by Theorem (DLR, 1972), at step t the optimal policy assign p_n to the job Y_t only if $Y_t > a_{n-t,n-(t-1)}$ and $\cdots$
- Let $\{b_t,\ t \geqslant 1\}$ be the sequence of real numbers defined recursively by
  - $b_1 = -\infty,\ \ b_2 = \mathrm{E}Y_n,$
  - $b_{t+1} = \int_{b_t}^{\infty} z \mathrm{d}F_{n-t+1}(z) + b_t F_{n-t+1}(b_t),\ \ t = 2, \ldots, n.$
- Let

$$\tau_* = \min\{1 \leqslant t \leqslant n : Y_t > b_{n-t+1}\};$$

then

$$W_n^* = \mathrm{E}Y_{\tau_*} = \max_{\tau \in \mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau = b_{n+1}.$$

Reduction: Problems (A1) $\Rightarrow$ Problem (B)

- Fixed Horizon n

  Let

$$I_{t,n}(r) := \sum_{a=r}^{n-t+r} q(a) \frac{\binom{a-1}{r-1}\binom{n-a}{t-r}}{\binom{n}{t}} = \mathrm{E}\{q(A_{t,n}) \,|\, R_t = r\}, \quad r = 1, \ldots, t. \quad (1)$$

$$Y_t := I_{t,n}(R_t), \quad t = 1, \ldots, n. \quad (2)$$

- Theorem: the optimal stopping rule τ_* solving Problem (B) with random variables $\{Y_t\}$ given in (1)–(2) also solves Problem (A1):

$$V_n(q; \tau_*) = \max_{\tau \in \mathscr{T}(\mathscr{R})} \mathrm{E}q(A_{\tau,n}) = \max_{\tau \in \mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau.$$

Proof

* First we note that for any stopping rule $\tau \in \mathscr{T}(\mathscr{R})$ one has $\mathrm{E}q(A_{\tau,n}) = \mathrm{E}Y_\tau$, where $Y_t := \mathrm{E}[q(A_{t,n})|\mathscr{R}_t]$.

*

$$\begin{aligned}\mathrm{E}q(A_{\tau,n}) &= \sum_{k=1}^{n} \mathrm{E}q(A_{\tau,n})\mathbf{1}\{\tau = k\} = \sum_{k=1}^{n} \mathrm{E}q(A_{k,n})\mathbf{1}\{\tau = k\} \\ &= \sum_{k=1}^{n} \mathrm{E}\Big[\mathbf{1}\{\tau = k\}\mathrm{E}\{q(A_{k,n})|\mathscr{R}_k\}\Big] = \sum_{k=1}^{n} \mathrm{E}[\mathbf{1}\{\tau = k\}Y_k] = \mathrm{E}Y_\tau,\end{aligned}$$

where we have used the fact that $\{\tau = k\} \in \mathscr{R}_k$. This implies that $\max_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}q(A_{\tau,n}) = \max_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}Y_\tau$.

* To prove the theorem it suffices to show only that

$$\max_{\tau\in\mathscr{T}(\mathscr{R})} \mathrm{E}Y_\tau = \max_{\tau\in\mathscr{T}(\mathscr{Y})} \mathrm{E}Y_\tau. \tag{3}$$

Proof (Con't)

* Clearly,

$$\mathscr{Y}_t \subset \mathscr{R}_t, \quad \forall 1 \leqslant t \leqslant n. \tag{4}$$

* Because $R_1, \ldots, R_n$ are independent random variables, and $Y_t = I_{t,n}(R_t)$, $\forall t$ we have that for any $s, t \in \{1, \ldots, n\}$ with $s < t$

$$\mathrm{P}\{G_t \,|\, \mathscr{Y}_s\} = \mathrm{P}\{G_t \,|\, \mathscr{R}_s\}, \quad \forall G_t \in \mathscr{Y}_t. \tag{5}$$

* The statement (3) follows from (4), (5) and Theorem 5.3. [a] This concludes the proof.

---

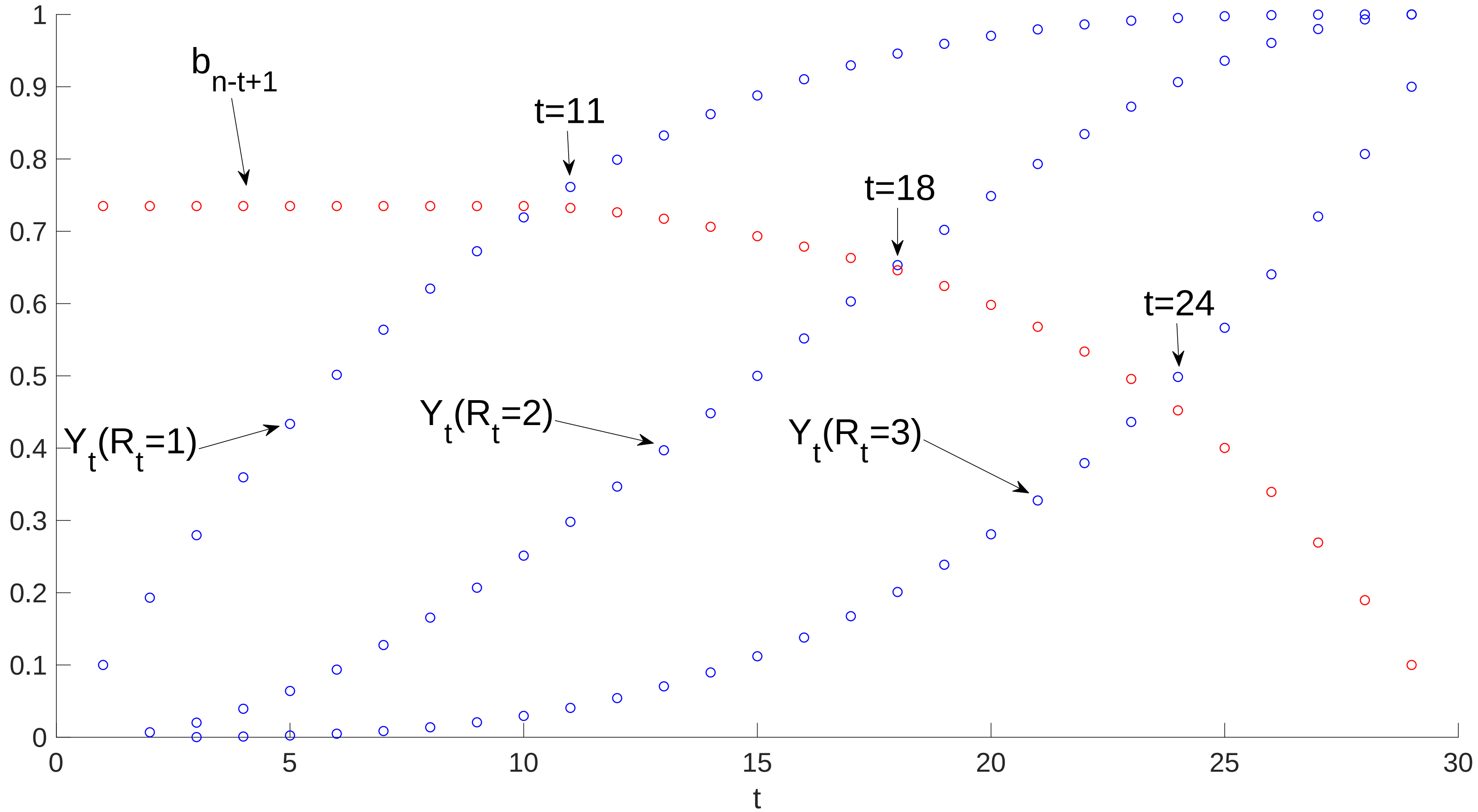
[a] Chow, Y. S., Robbins, H. and Siegmung, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton Mifflin Company, Boston.

Numerical Values

n	k	$P(n,k)$	$E(n,k)/n$	n	k	$P(n,k)$	$E(n,k)/n$
100	2	0.57956	0.68645	1,000	2	0.57417	0.68966
	5	0.86917	0.60871		5	0.86123	0.60988
	10	0.98140	0.54236		10	0.97703	0.54434
	15	0.99755	0.50428		15	0.99609	0.50893
10,000	2	0.57363	0.68927	50,000	2	0.57358	0.68923
	5	0.86043	0.61014		5	0.86036	0.61018
	10	0.97658	0.54496		10	0.97654	0.54500
	15	0.99592	0.50947		15	0.99591	0.50950

Table 1: Optimal probabilities $P(n,k)$ and the normalized expected time elapsed until stopping $E(n,k)/n$ for selecting one of the k best values.

Optimal Strategy for $n = 30, \; k = 3$



$V_n^*(q) = 0.73492$

Reduction: Problems (A2) $\Rightarrow$ Problem (B)

- Random Horizon

  Let

$$J_t(r) := \sum_{k=t}^{N_{\max}} \gamma_k I_{t,k}(r), \quad r = 1, \ldots, t. \tag{6}$$

  Define

$$Y_t := J_t(R_t) = \sum_{k=t}^{N_{\max}} \gamma_k I_{t,k}(R_t), \quad t = 1, \ldots, N_{\max}. \tag{7}$$

- Theorem: let $N_{\max} < \infty$; then the optimal stopping rule τ_* solving Problem (B) with fixed horizon $N_{\max}$ and random variables $\{Y_t\}$ given in (6)–(7) provides the optimal solution to Problem (A2):

$$V_\gamma^*(q) = \max_{\tau \in \mathscr{T}(\mathscr{R})} V_\gamma(q; \tau) = \max_{\tau \in \mathscr{T}(\mathscr{Y})} \mathrm{E} Y_\tau = W_{N_{\max}}(\tau^*).$$

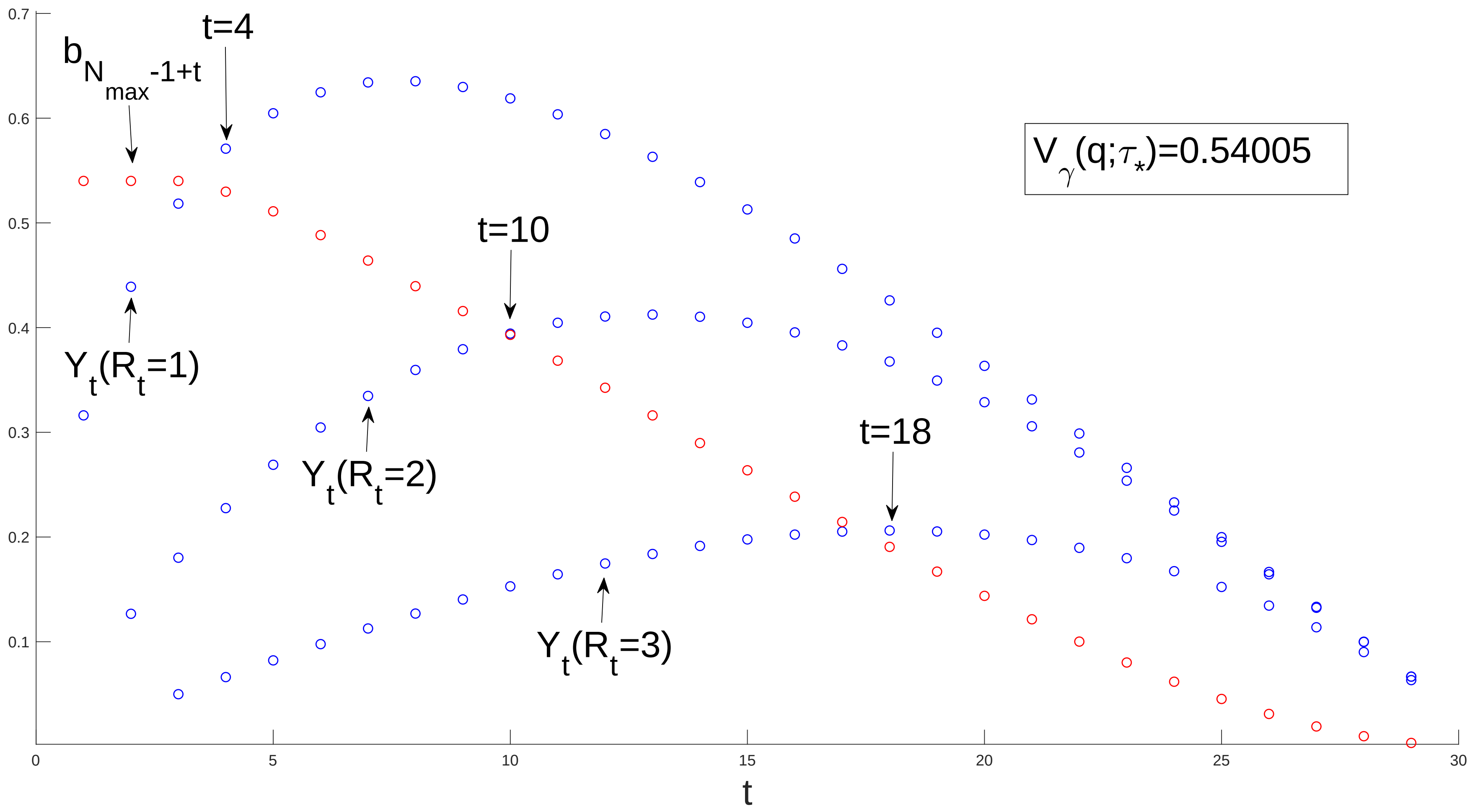
Proof

* In Problem (A2) the reward for stopping at time t is $\tilde{q}(A_{t,N}) = q(A_{t,N})\mathbf{1}\{N \geqslant t\}$.

$$
\begin{aligned}
&\mathrm{E}\{q(A_{t,N})\mathbf{1}\{N \geqslant t\} \,|\, R_1 = r_1, \dots, R_{t-1} = r_{t-1}, R_t = r\} \\
&= \sum_{k=t}^{N_{\max}} \mathrm{E}\Big\{q(A_{t,N})\mathbf{1}\{N = k\} \,|\, R_1 = r_1, \dots, R_{t-1} = r_{t-1}, R_t = r\Big\} \\
&= \sum_{k=t}^{N_{\max}} \mathrm{E}\Big\{\mathbf{1}\{N = k\}\mathrm{E}\big[q(A_{t,k}) \,|\, N = k, R_t = r\big]\Big\} \\
&= \sum_{k=t}^{N_{\max}} \gamma_k \sum_{a=r}^{k-t+r} q(a) \frac{\binom{a-1}{r-1}\binom{k-a}{t-r}}{\binom{k}{t}} = \sum_{k=t}^{N_{\max}} \gamma_k I_{t,k}(r) =: J_t(r). \qquad (8)
\end{aligned}
$$

* Together with (7) this implies that $\mathrm{E}\tilde{q}(A_{\tau,N}) = \mathrm{E}J_\tau(R_\tau) = \mathrm{E}Y_\tau$ for any $\tau \in \mathscr{T}(\mathscr{R})$. The remainder of the proof proceeds along the lines of the proof of Theorem for fixed horizon n.

Optimal Strategy for $N \sim Uniform\{1, 2, \ldots, 30\},\ k = 3$



Concluding Remarks

- The proposed framework is applicable to sequential selection problems that can be reduced to settings with independent observations and additive reward function. In particular:
  - selection problems with no-information, rank-dependent rewards and fixed or random horizon,
  - selection problems with full information when the random variables $\{X_t\}$ are observable, and the reward for stopping at time t is a function of the current observation X_t only,
  - multiple choice problems with random horizon and additive reward.

Concluding Remarks (con't)

- The proposed framework is not applicable to the following sequential selection problems:
  - for instance, settings with rank-dependent reward and full information as in Gnedin (2007) [a] cannot be reduced to optimal stopping of a sequence of independent random variables
  - multiple choice problem with zero-one reward, where the problem of maximizing the probability of selecting k best alternatives; see, e.g., Rose (1982) [b] where the problem of maximizing the probability of selecting k best alternatives was considered.

[a] Gnedin, A. V. (2007). Optimal stopping with rank-dependent loss. *J. Appl. Probab.*, **4**4, 996–1011.

[b] Rose, J. S. (1972). A problem of optimal choice and assignment. *Oper. Res.*, **3**0, 172–181.

Thank You !