Finite time-horizon MDP
Asset Selling Problem

The decision maker has an asset to sell. She sequentially receives N offers $w_1, \ldots, w_N$ i.i.d. with density $f(w)$ (w.r.t. the Lebesgue measure on $\mathbb{R}_+$). At time i, after receiving the offer w_i, the decision maker has to decide whether to accept the offer or to reject it. If the offer is accepted, the reward is $(1+r)^{N-i} w_i$, where $r > 0$ denotes the interest rate. Once she accepted an offer, the subsequent offers do not matter. The problem is to define a strategy maximizing the expected reward. We use the formalism of finite time-horizon to identify such strategy.

- Time horizon: N;
- Possible states: 0 if the decision maker has not accepted an offer yet, 1 otherwise;
- Actions: A (accept) or R (reject);
- Reward at time k:
  - If the state is 1, the reward is 0,
  - If the state is 0, the reward is defined as follows:

$$r_k(0, A, w_k) = (1+r)^{N-k} w_k,$$

$$r_k(1, D, w_k) = 0,$$

    for $D = A$ or R, for all w_k.
- Transitions: if the selected action is A, the state moves to 1, if it is R, it stays the same.

We denote by $V_k(w_k)$ the expected reward from time k to N, when the offer at time k is w_k. Bellman's equations provide the following recursion:

$$V_k(w_k) = 0, \quad \text{if the state is } 1,$$

$$V_k(w_k) = \max((1+r)^{N-k} w_k, \mathbb{E}[V_{k+1}(w)]), \quad \text{if the state is } 0.$$

We deduce that the optimal strategy is thershold-based. If the asset has not been sold at time k, the optimal decision is to accept the offer w_k if and only if:

$$w_k \geq \alpha_k = (1+r)^{k-N} \mathbb{E}[V_{k+1}(w)].$$

Now the thresholds α_k can be computed using backward induction.
First note that $\alpha_N = 0$ and $V_N(w) = w$. Assume that the threshold α_{k+1} is known. We have:

$$\begin{aligned}
\alpha_k &= (1+r)^{k-N}\mathbb{E}[V_{k+1}(w)] \\
&= (1+r)^{-1}\mathbb{E}[\max(w, \alpha_{k+1})].
\end{aligned}$$

Hence:

$$\alpha_k = (1+r)^{-1}\left[\alpha_{k+1}\int_0^{\alpha_{k+1}} f(w)dw + \int_{\alpha_{k+1}}^{\infty} wf(w)dw\right].$$