Randomized Algorithms

Ioannis Caragiannis (this time) and Kasper Green Larsen

Today

- Prophet inequality
- Secretary problem
- Martingales

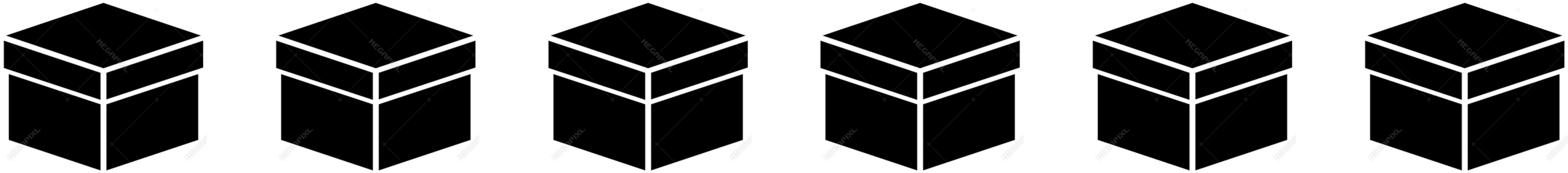
Optimal stopping

Problem setting

- There are n **independent random variables** $X_1, X_2, \ldots, X_n$
- **We know their distributions** upfront, but not their realizations
- Realizations are revealed **one-by-one**
- Goal: Find a **stopping rule**, i.e., when seeing X_i, decide either to stop and get reward X_i or to move on to the next items
- Objective: **maximize the expected reward**

An example

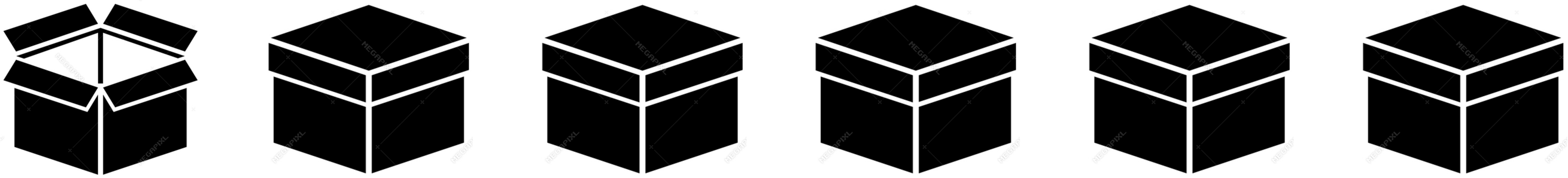
- $n = 6$



$U[0,10]$ $U[1,7]$ $U[5,8]$ $U[3,6]$ $U[2,5]$ $U[0,3]$

An example

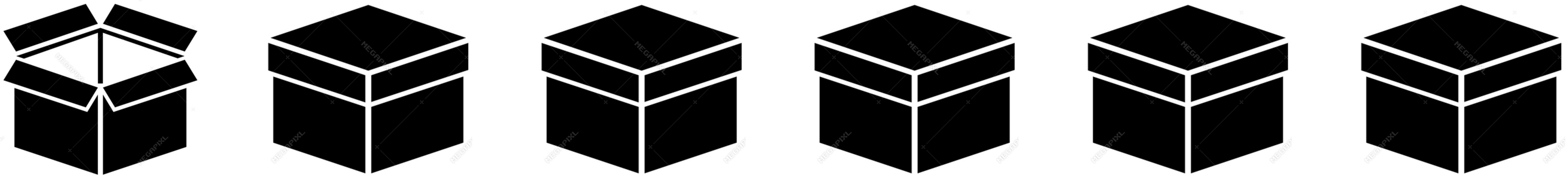
- $n = 6$



2 $\quad U[1,7] \quad U[5,8] \quad U[3,6] \quad U[2,5] \quad U[0,3]$

An example

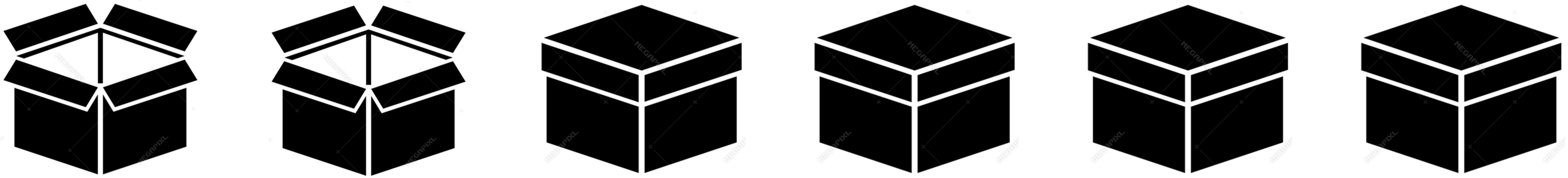
- $n = 6$



$U[1,7]$ $U[5,8]$ $U[3,6]$ $U[2,5]$ $U[0,3]$

An example

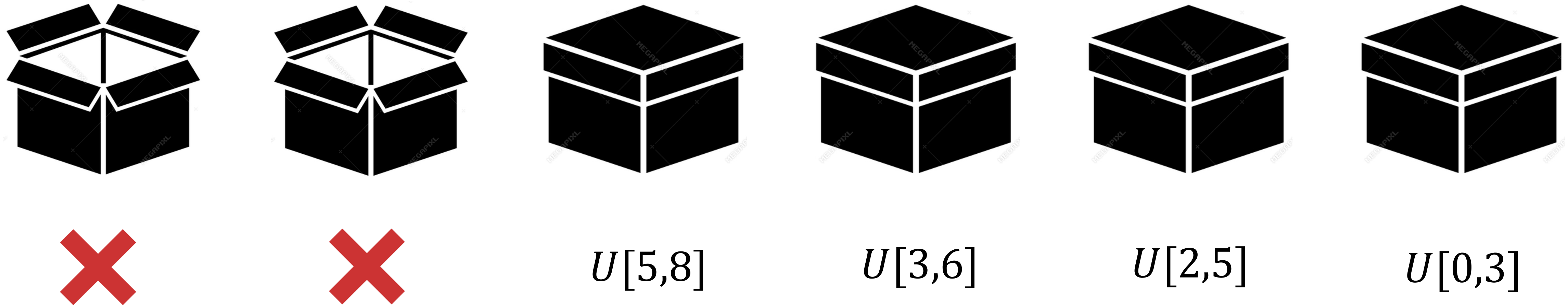
- $n = 6$



❌ **3** $U[5,8]$ $U[3,6]$ $U[2,5]$ $U[0,3]$

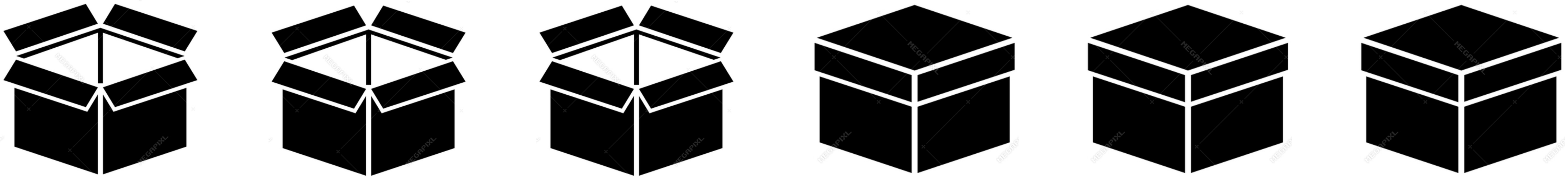
An example

- $n = 6$



An example

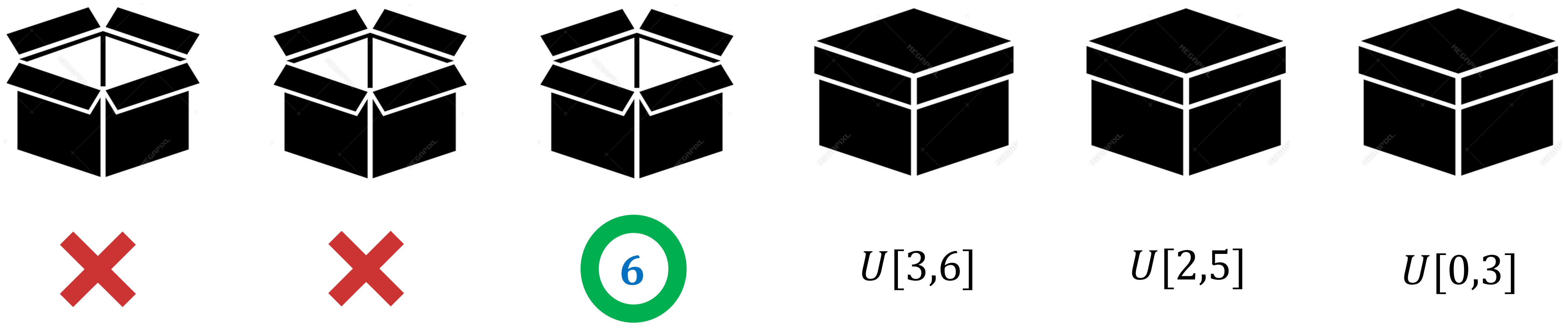
- $n = 6$



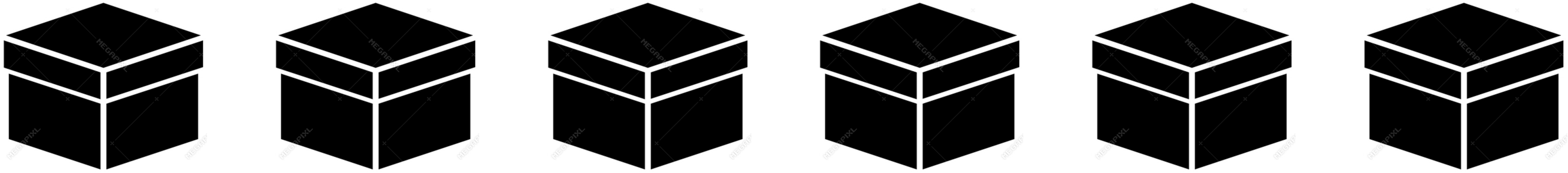
6 $U[3,6]$ $U[2,5]$ $U[0,3]$

An example

- $n = 6$

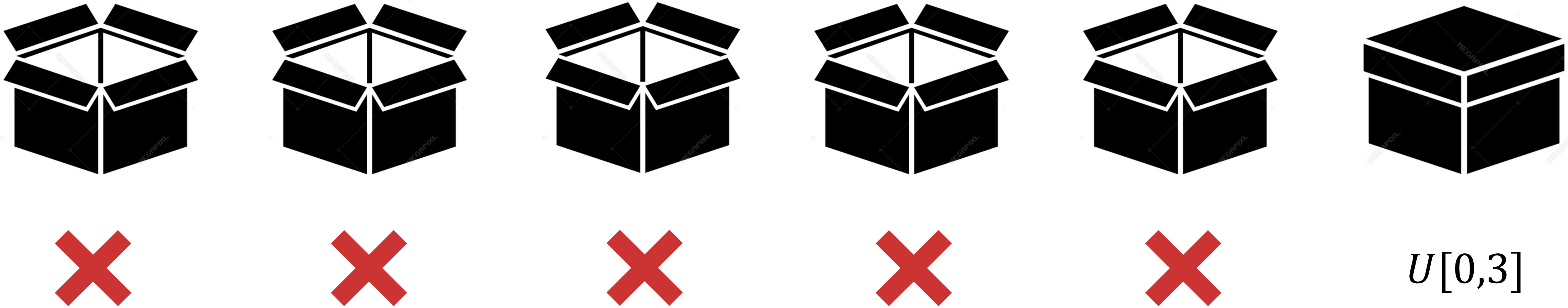


The optimal strategy: backward induction



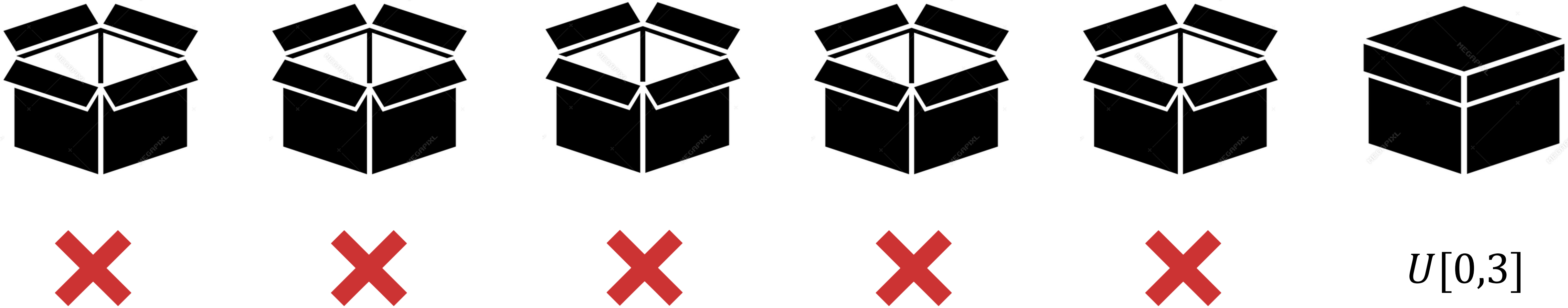
$U[0,10]$ $U[1,7]$ $U[5,8]$ $U[3,6]$ $U[2,5]$ $U[0,3]$

The optimal strategy: backward induction



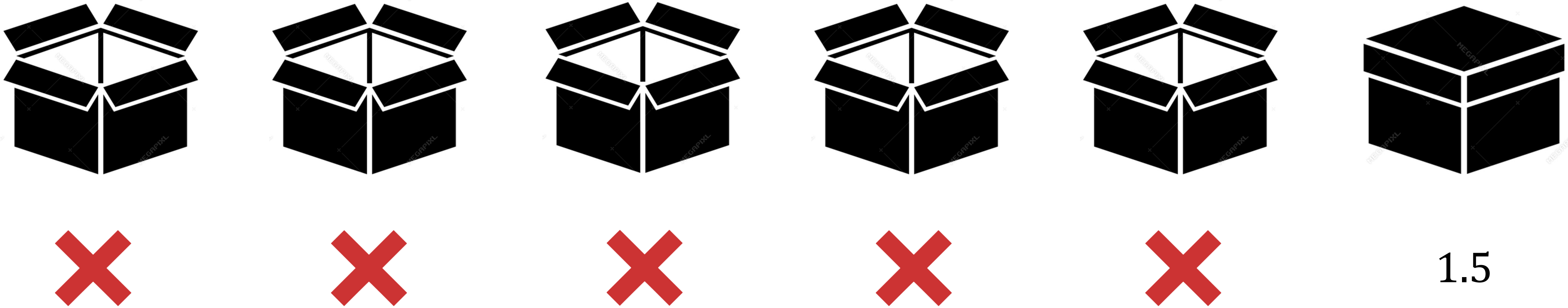
$U[0,3]$

The optimal strategy: backward induction



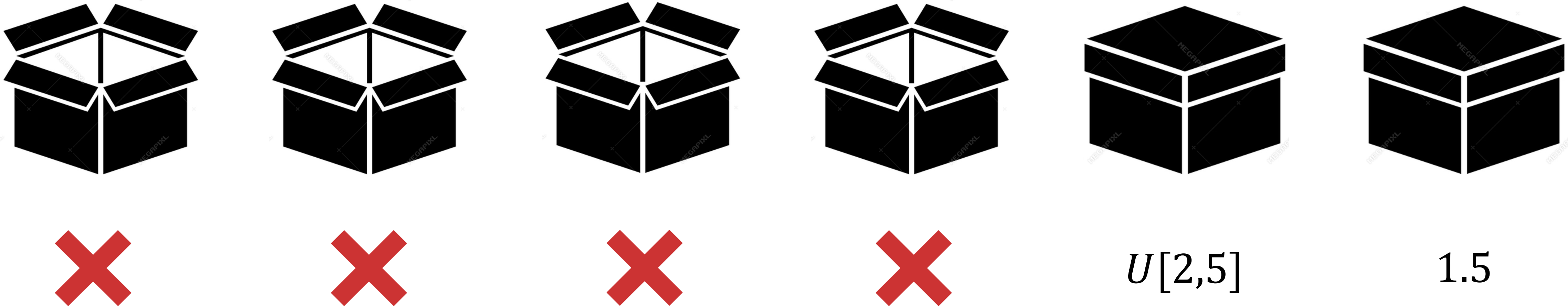
- At step 6, accept any reward

The optimal strategy: backward induction



- At step 6, accept any reward

The optimal strategy: backward induction

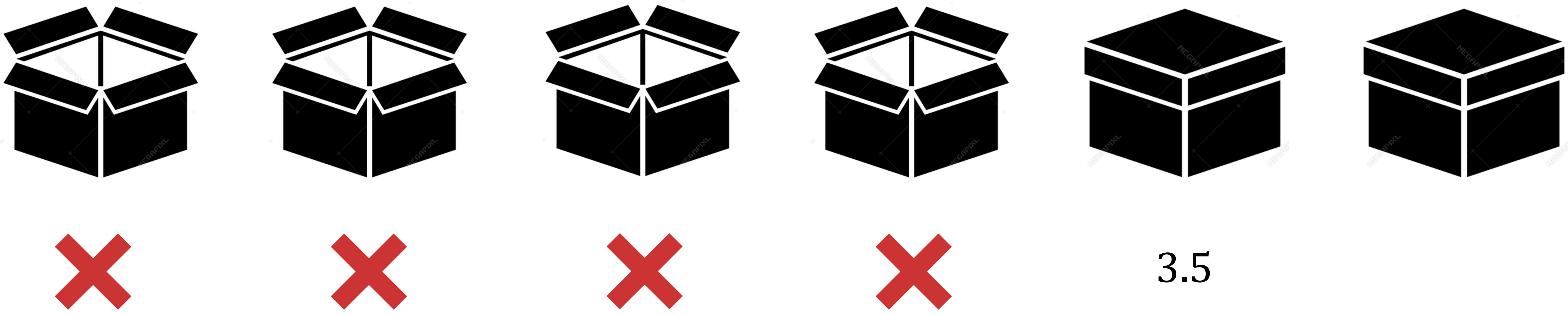


- At step 6, accept any reward

The optimal strategy: backward induction

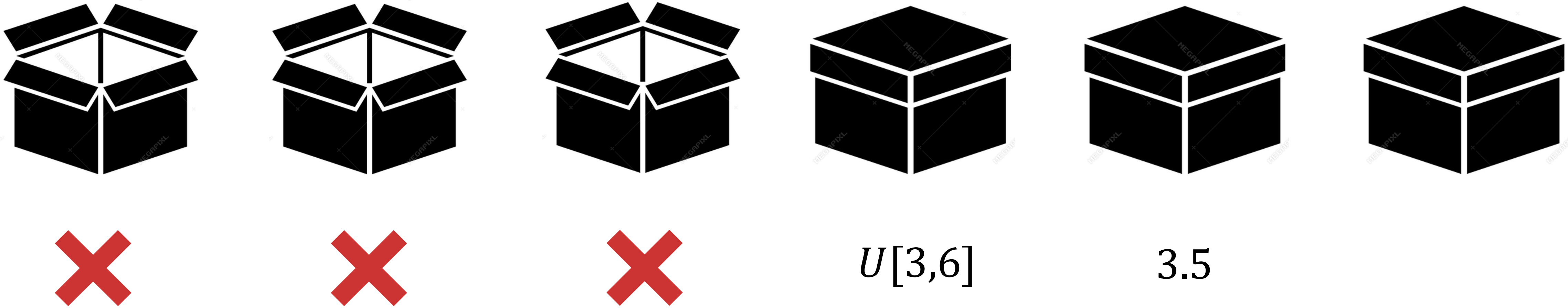
- At step 5, accept any reward

The optimal strategy: backward induction



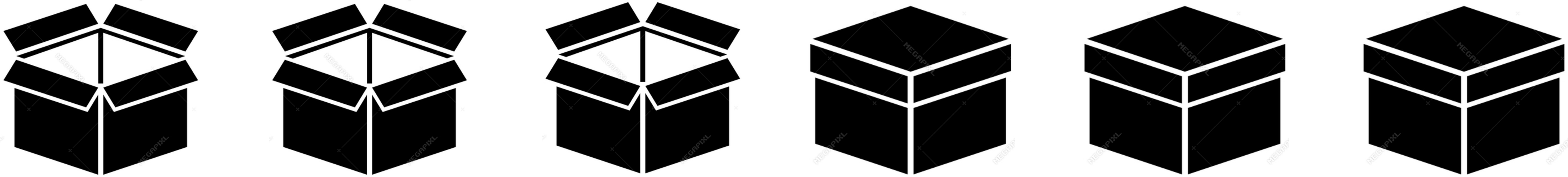
- At step 5, accept any reward

The optimal strategy: backward induction



- At step 5, accept any reward

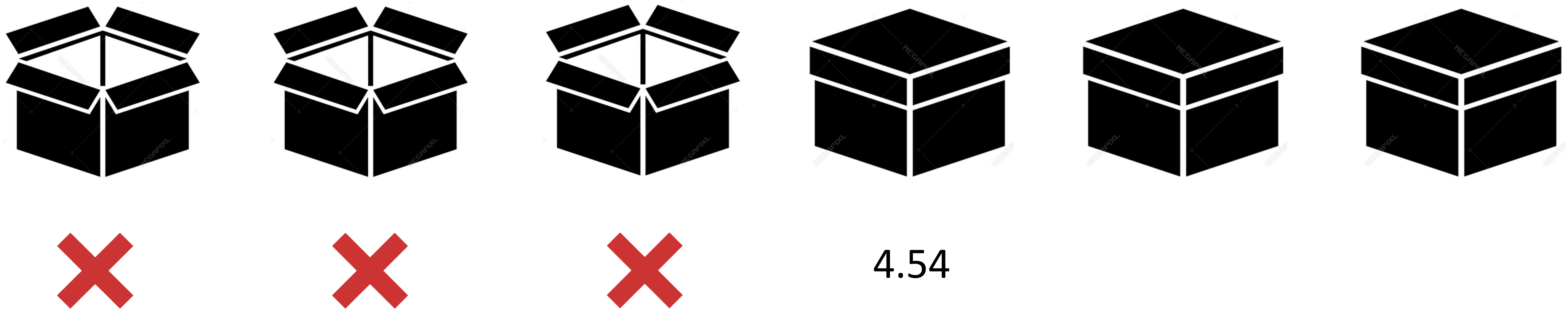
The optimal strategy: backward induction



$U[3,6]$ 3.5

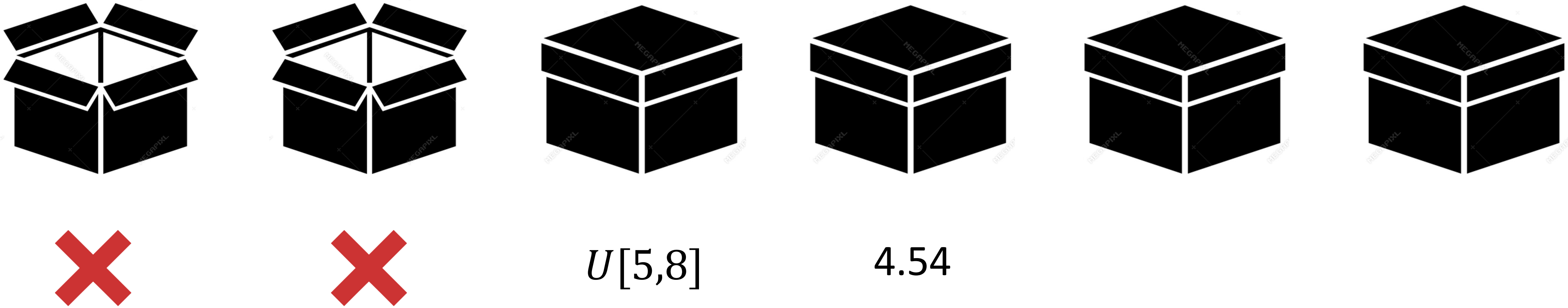
- At step 5, accept any reward
- At step 4, accept a reward if it is higher than 3.5

The optimal strategy: backward induction



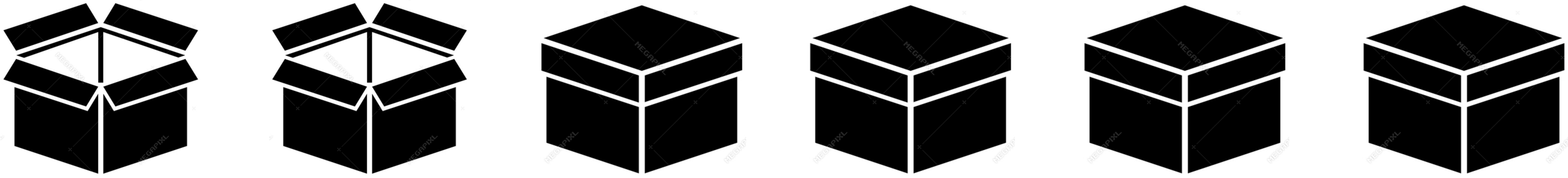
- At step 5, accept any reward
- At step 4, accept a reward if it is higher than 3.5

The optimal strategy: backward induction



- At step 5, accept any reward
- At step 4, accept a reward if it is higher than 3.5

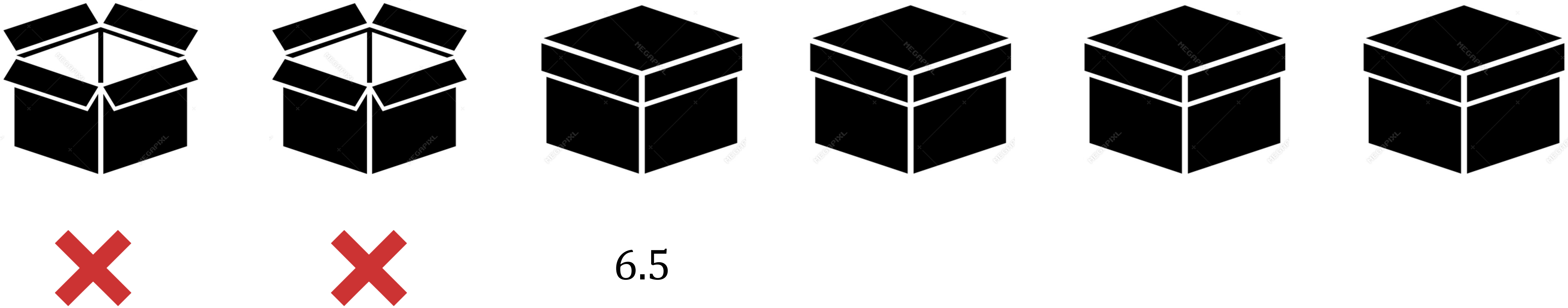
The optimal strategy: backward induction



$U[5,8]$ 4.54

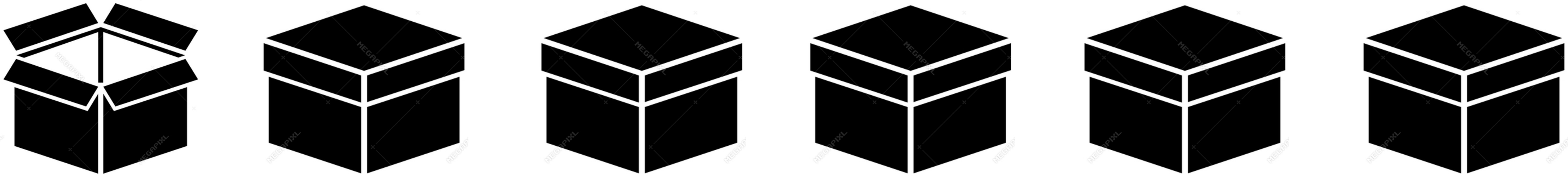
- At step 3, accept any reward

The optimal strategy: backward induction



- At step 3, accept any reward

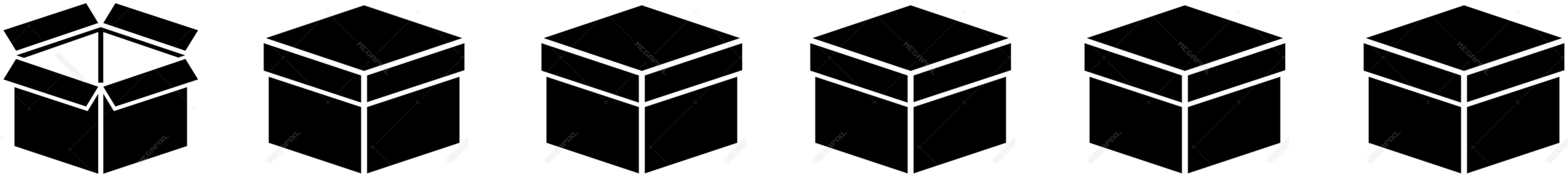
The optimal strategy: backward induction



$U[1,7]$ $\quad$ 6.5

- At step 3, accept any reward

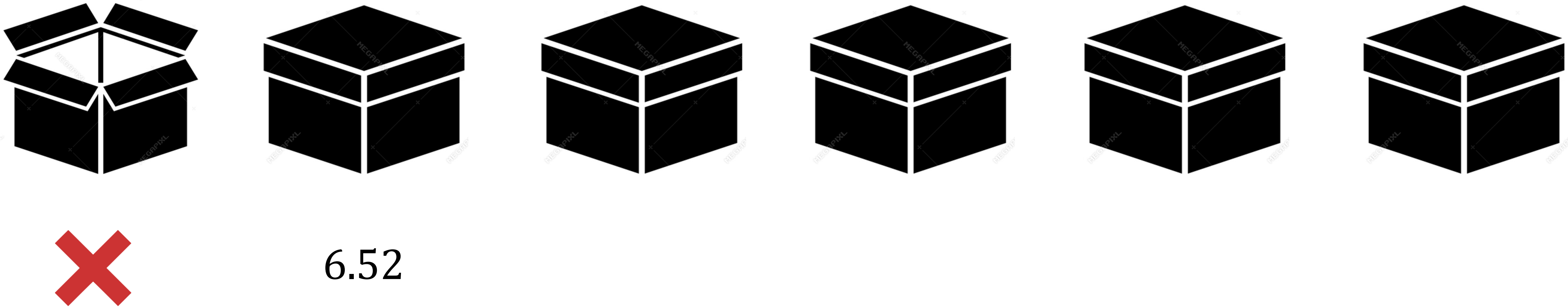
The optimal strategy: backward induction



$U[1,7]$ $\quad$ 6.5

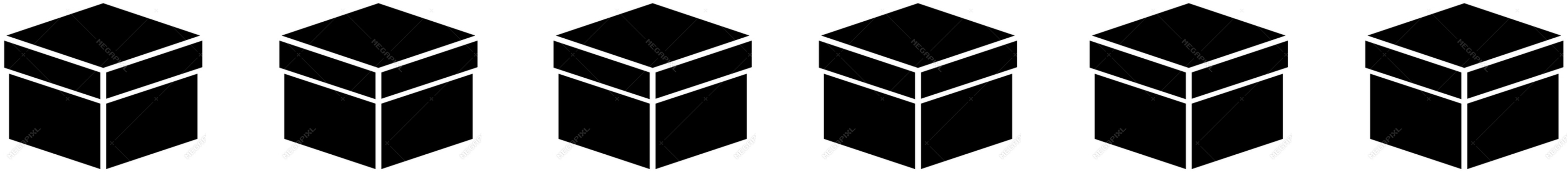
- At step 3, accept any reward
- At step 2, accept a reward if it is higher than 6.5

The optimal strategy: backward induction



- At step 3, accept any reward
- At step 2, accept a reward if it is higher than 6.5

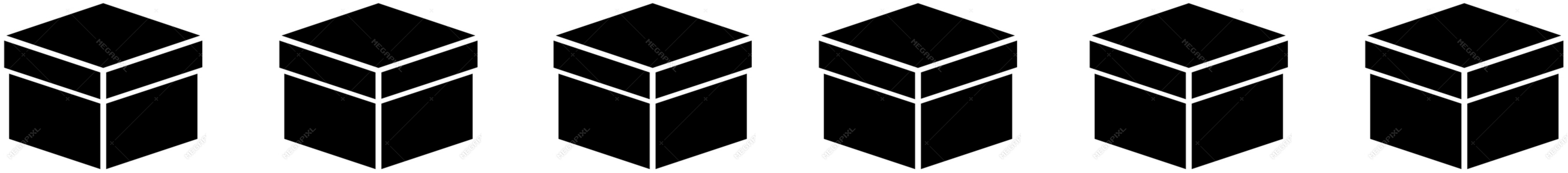
The optimal strategy: backward induction



$U[0,10]$ 6.52

- At step 3, accept any reward
- At step 2, accept a reward if it is higher than 6.5

The optimal strategy: backward induction



$U[0,10]$ 6.52

- At step 3, accept any reward
- At step 2, accept a reward if it is higher than 6.5
- At step 1, accept a reward if it is higher than 6.52

The prophet inequality

Alternatives

- The optimal strategy can be too complicated to implement
- So, consider **simpler strategies**
- Comparing to the optimal strategy can be too challenging
- So, **compare to $\mathbb{E}\left[\max_i X_i\right]$** instead

Prophet inequality

- **Gambler**: Use a single threshold τ and accept the first value that is above this threshold
- **Prophet**: Knows $\max_i X_i$

- Compare the expected reward of the **best strategy** for the gambler to the **expected reward of a prophet**

Prophet inequality

Theorem: There exists a threshold strategy so that the expected reward of the gambler is **at least half** that of the prophet

This bound is **best possible**:

- $X_1 = 1$ with certainty
- $X_2 = 1/\varepsilon$ with probability ε, $X_2 = 0$ with probability $1 - \varepsilon$

Analysis of a threshold strategy

- Let τ be the **median** of the distribution of $\max_i X_i$
- Assume that there is no point mass at τ
- Define $x^+ = \max\{0, x\}$

Analysis of a threshold strategy

Analysis of a threshold strategy

$$\mathbb{E}\left[\max_i X_i\right]$$

Analysis of a threshold strategy

$$\mathbb{E}\left[\max_i X_i\right] = \mathbb{E}\left[\tau + \max_i\{X_i - \tau\}\right]$$

obvious

Analysis of a threshold strategy

$$\mathbb{E}\left[\max_i X_i\right] = \mathbb{E}\left[\tau + \max_i\{X_i - \tau\}\right] \le \tau + \mathbb{E}\left[\max_i\{(X_i - \tau)^+\}\right]$$

obvious

$x \le x^+$

Analysis of a threshold strategy

$$\mathbb{E}\left[\max_i X_i\right] = \mathbb{E}\left[\tau + \max_i \{X_i - \tau\}\right] \leq \tau + \mathbb{E}\left[\max_i \{(X_i - \tau)^+\}\right] \leq \tau + \mathbb{E}\left[\sum_{i=1}^{n} (X_i - \tau)^+\right]$$

obvious

$x \leq x^+$

max of non-negative values is at most their sum

Analysis of a threshold strategy

ALG

Analysis of a threshold strategy

$$ALG = \sum_{i=1}^{n} \mathbb{E}\left[X_i \middle| X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

Analysis of a threshold strategy

$$ALG = \sum_{i=1}^{n} \mathbb{E}\left[X_i | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \sum_{i=1}^{n} \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right] + \sum_{i=1}^{n} \mathbb{E}\left[X_i - \tau | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

Analysis of a threshold strategy

$$ALG = \sum_{i=1}^{n} \mathbb{E}\left[X_i | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \sum_{i=1}^{n} \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right] + \sum_{i=1}^{n} \mathbb{E}\left[X_i - \tau | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau] \cdot \Pr\left[X_j < \tau, \forall j < i\right]$$

Analysis of a threshold strategy

$$ALG = \sum_{i=1}^{n} \mathbb{E}[X_i | X_i \geq \tau, X_j < \tau, \forall j < i] \cdot \Pr[X_i \geq \tau, X_j < \tau, \forall j < i]$$

$$= \tau \cdot \sum_{i=1}^{n} \Pr[X_i \geq \tau, X_j < \tau, \forall j < i] + \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau, X_j < \tau, \forall j < i] \cdot \Pr[X_i \geq \tau, X_j < \tau, \forall j < i]$$

$$= \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau] \cdot \Pr[X_j < \tau, \forall j < i]$$

definition of $\Pr\left[\max_i X_i \geq \tau\right]$

independence of the X_i's

Analysis of a threshold strategy

$$ALG = \sum_{i=1}^{n} \mathbb{E}\left[X_i | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \sum_{i=1}^{n} \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right] + \sum_{i=1}^{n} \mathbb{E}\left[X_i - \tau | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau] \cdot \Pr\left[X_j < \tau, \forall j < i\right]$$

$$\geq \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \Pr\left[\max_i X_i < \tau\right] \cdot \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau]$$

$$\Pr\left[X_j < \tau, \forall j < i\right] \geq \Pr\left[\max_i X_i < \tau\right]$$

Analysis of a threshold strategy

$$ALG = \sum_{i=1}^{n} \mathbb{E}\left[X_i | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \sum_{i=1}^{n} \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right] + \sum_{i=1}^{n} \mathbb{E}\left[X_i - \tau | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau] \cdot \Pr\left[X_j < \tau, \forall j < i\right]$$

$$\geq \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \Pr\left[\max_i X_i < \tau\right] \cdot \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau]$$

$$= \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \Pr\left[\max_i X_i < \tau\right] \cdot \sum_{i=1}^{n} \mathbb{E}[(X_i - \tau)^+]$$

Analysis of a threshold strategy

$$ALG = \sum_{i=1}^{n} \mathbb{E}\left[X_i | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \sum_{i=1}^{n} \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right] + \sum_{i=1}^{n} \mathbb{E}\left[X_i - \tau | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau] \cdot \Pr\left[X_j < \tau, \forall j < i\right]$$

$$\geq \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \Pr\left[\max_i X_i < \tau\right] \cdot \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau]$$

$$= \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \Pr\left[\max_i X_i < \tau\right] \cdot \sum_{i=1}^{n} \mathbb{E}[(X_i - \tau)^+] = \frac{1}{2} \cdot \left(\tau + \mathbb{E}\left[\sum_{i=1}^{n} (X_i - \tau)^+\right]\right)$$

Analysis of a threshold strategy

$$ALG = \sum_{i=1}^{n} \mathbb{E}\left[X_i | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \sum_{i=1}^{n} \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right] + \sum_{i=1}^{n} \mathbb{E}\left[X_i - \tau | X_i \geq \tau, X_j < \tau, \forall j < i\right] \cdot \Pr\left[X_i \geq \tau, X_j < \tau, \forall j < i\right]$$

$$= \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau] \cdot \Pr\left[X_j < \tau, \forall j < i\right]$$

$$\geq \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \Pr\left[\max_i X_i < \tau\right] \cdot \sum_{i=1}^{n} \mathbb{E}[X_i - \tau | X_i \geq \tau] \cdot \Pr[X_i \geq \tau]$$

$$= \tau \cdot \Pr\left[\max_i X_i \geq \tau\right] + \Pr\left[\max_i X_i < \tau\right] \cdot \sum_{i=1}^{n} \mathbb{E}[(X_i - \tau)^+] = \frac{1}{2} \cdot \left(\tau + \mathbb{E}\left[\sum_{i=1}^{n} (X_i - \tau)^+\right]\right) \geq \frac{1}{2} \cdot \mathbb{E}\left[\max_i X_i\right]$$

An alternative proof of the prophet inequality

An alternative strategy

- Let $\boldsymbol{p_i}$ be the probability that X_i is the maximum among $X_1, X_2, \ldots, X_n$
- I.e., $\sum_{i=1}^{n} p_i = 1$
- Let $\boldsymbol{\tau_i}$ be the p_i-th percentile of X_i, i.e., p_i is such that $\Pr[X_i \geq \tau_i] = p_i$
- Define $v_i(p_i) = \mathbb{E}[X_i | X_i \geq \tau_i]$, the expected value of X_i, conditioned on it lying in the p_i-th percentile
- Then, $\mathbb{E}\left[\max_i X_i\right] \leq \sum_{i=1}^{n} v_i(p_i) \cdot p_i$

An alternative strategy: a first attempt

- Algorithm: At step i, if $X_i \geq \tau_i$, accept with probability $q_i = 1/2$

An alternative strategy: a first attempt

- Algorithm: At step i, if $X_i \geq \tau_i$, accept with probability $q_i = 1/2$
- Let r_i be the probability that we reach step i

$$ALG = \sum_{i=1}^{n} r_i \cdot q_i \cdot \Pr[X_i \geq \tau_i] \cdot \mathbb{E}[X_i | X_i \geq \tau_i] = \sum_{i=1}^{n} r_i \cdot q_i \cdot p_i \cdot v_i(p_i)$$

An alternative strategy: a first attempt

- Algorithm: At step i, if $X_i \geq \tau_i$, accept with probability $q_i = 1/2$
- Let r_i be the probability that we reach step i

$$ALG = \sum_{i=1}^{n} r_i \cdot q_i \cdot \Pr[X_i \geq \tau_i] \cdot \mathbb{E}[X_i | X_i \geq \tau_i] = \sum_{i=1}^{n} r_i \cdot q_i \cdot p_i \cdot v_i(p_i)$$

- At each step j, the algorithm accepts with probability $p_j/2$
- The prob. that it has accepted before step i is at most $\sum_{j=1}^{i-1} p_j/2 \leq 1/2$
- Hence, $r_i \geq 1/2$
- $ALG \geq \frac{1}{4} \cdot \sum_{i=1}^{n} p_i \cdot v_i(p_i) \geq \frac{1}{4} \cdot \mathbb{E}\left[\max_i X_i\right]$

An alternative strategy: a better attempt

- At step i, if $X_i \geq \tau_i$, accept with probability $q_i = \left(2 - \sum_{j<i} p_i\right)^{-1}$

$$ALG = \sum_{i=1}^{n} r_i \cdot q_i \cdot \Pr[X_i \geq \tau_i] \cdot \mathbb{E}[X_i | X_i \geq \tau_i] = \sum_{i=1}^{n} r_i \cdot q_i \cdot p_i \cdot v_i(p_i)$$

- Then, $r_i = 1 - \frac{1}{2} \cdot \sum_{j<i} p_i$ and $r_i \cdot q_i = 1/2$
- Why? Observe that $r_1 = 1$ and $r_{i+1} = r_i(1 - p_i \cdot q_i)$
- Hence, $ALG \geq \frac{1}{2} \cdot \sum_{i=1}^{n} p_i \cdot v_i(p_i) \geq \frac{1}{2} \cdot \mathbb{E}\left[\max_i X_i\right]$

Accepting up to k items

Accepting up to k items

- Let $\boldsymbol{p_i}$ be the probability that X_i is among the top k values among $X_1, X_2, \dots, X_n$
- I.e., $\sum_{i=1}^{n} p_i = k$
- Let $\boldsymbol{\tau_i}$ be the p_i-th percentile of X_i, i.e., p_i is such that $\Pr[X_i \geq \tau_i] = p_i$
- Define $v_i(p_i) = \mathbb{E}[X_i | X_i \geq \tau_i]$
- Then, $\mathbb{E}[\text{sum of top } k \text{ values of } X_i' \text{s}] \leq \sum_{i=1}^{n} v_i(p_i) \cdot p_i$
- Algorithm: At step i, accept with probability $\mathbf{1 - \delta}$, until k items have been accepted in total
- How small can δ be?

Accepting up to k items: analysis

- Let r be the probability that all steps are executed

$$ALG \geq \sum_{i=1}^{n} r \cdot (1-\delta) \cdot \Pr[X_i \geq \tau_i] \cdot \mathbb{E}[X_i | X_i \geq \tau_i] = r \cdot (1-\delta) \cdot \sum_{i=1}^{n} p_i \cdot v_i(p_i)$$

Accepting up to k items: analysis

- Let r be the probability that all steps are executed

$$ALG \geq \sum_{i=1}^{n} r \cdot (1-\delta) \cdot \Pr[X_i \geq \tau_i] \cdot \mathbb{E}[X_i | X_i \geq \tau_i] = r \cdot (1-\delta) \cdot \sum_{i=1}^{n} p_i \cdot v_i(p_i)$$

- Let Y_i denote whether the algorithm accepts at step i

Accepting up to k items: analysis

- Let r be the probability that all steps are executed

$$ALG \geq \sum_{i=1}^{n} r \cdot (1-\delta) \cdot \Pr[X_i \geq \tau_i] \cdot \mathbb{E}[X_i | X_i \geq \tau_i] = r \cdot (1-\delta) \cdot \sum_{i=1}^{n} p_i \cdot v_i(p_i)$$

- Let Y_i denote whether the algorithm accepts at step i
- $\sum_{i=1}^{n} Y_i$ is a sum of independent binary r.v.s with $\Pr[Y_i = 1] = (1 - \delta) p_i$

Accepting up to k items: analysis

- Let r be the probability that all steps are executed

$$ALG \geq \sum_{i=1}^{n} r \cdot (1-\delta) \cdot \Pr[X_i \geq \tau_i] \cdot \mathbb{E}[X_i | X_i \geq \tau_i] = r \cdot (1-\delta) \cdot \sum_{i=1}^{n} p_i \cdot v_i(p_i)$$

- Let Y_i denote whether the algorithm accepts at step i
- $\sum_{i=1}^{n} Y_i$ is a sum of independent binary r.v.s with $\Pr[Y_i = 1] = (1-\delta)p_i$
- Hence, $\mathbb{E}[\sum_{i=1}^{n} Y_i] = (1-\delta)k$

Accepting up to k items: analysis

- Let r be the probability that all steps are executed
$$ALG \geq \sum_{i=1}^{n} r \cdot (1-\delta) \cdot \Pr[X_i \geq \tau_i] \cdot \mathbb{E}[X_i | X_i \geq \tau_i] = r \cdot (1-\delta) \cdot \sum_{i=1}^{n} p_i \cdot v_i(p_i)$$
- Let Y_i denote whether the algorithm accepts at step i
- $\sum_{i=1}^{n} Y_i$ is a sum of independent binary r.v.s with $\Pr[Y_i = 1] = (1-\delta)p_i$
- Hence, $\mathbb{E}[\sum_{i=1}^{n} Y_i] = (1-\delta)k$
- Using a Chernoff bound, $\Pr[\sum_{i=1}^{n} Y_i \geq k] \leq \exp\left(-\frac{\delta^2 k}{3}\right)$

Accepting up to k items: analysis

- Let r be the probability that all steps are executed

$$ALG \geq \sum_{i=1}^{n} r \cdot (1-\delta) \cdot \Pr[X_i \geq \tau_i] \cdot \mathbb{E}[X_i | X_i \geq \tau_i] = r \cdot (1-\delta) \cdot \sum_{i=1}^{n} p_i \cdot v_i(p_i)$$

- Let Y_i denote whether the algorithm accepts at step i
- $\sum_{i=1}^{n} Y_i$ is a sum of independent binary r.v.s with $\Pr[Y_i = 1] = (1-\delta) p_i$
- Hence, $\mathbb{E}[\sum_{i=1}^{n} Y_i] = (1-\delta)k$
- Using a Chernoff bound, $\Pr[\sum_{i=1}^{n} Y_i \geq k] \leq \exp\left(-\frac{\delta^2 k}{3}\right)$
- Hence, $r \geq 1 - \exp\left(-\frac{\delta^2 k}{3}\right)$

Accepting up to k items: analysis

- Setting $\delta = \sqrt{\frac{3 \ln k}{k}}$, we have

$$r \cdot (1-\delta) \geq \left(1 - \frac{1}{k}\right) \cdot \left(1 - \sqrt{\frac{3 \ln k}{k}}\right) \geq 1 - 2\sqrt{\frac{3 \ln k}{k}}$$

Accepting up to k items: analysis

- Setting $\delta = \sqrt{\frac{3 \ln k}{k}}$, we have

$$r \cdot (1-\delta) \geq \left(1 - \frac{1}{k}\right) \cdot \left(1 - \sqrt{\frac{3 \ln k}{k}}\right) \geq 1 - 2\sqrt{\frac{3 \ln k}{k}}$$

- Hence,

$$ALG \geq r \cdot (1-\delta) \cdot \sum_{i=1}^{n} p_i \cdot v_i(p_i) \geq r \cdot (1-\delta) \cdot \mathbb{E}[\text{sum of top } k \text{ values of } X_i'\text{s}]$$

$$\geq \left(1 - 2\sqrt{\frac{3 \ln k}{k}}\right) \cdot \mathbb{E}[\text{sum of top } k \text{ values of } X_i'\text{s}]$$

The secretary problem

Problem setting

- There are n **items**, with (distinct) unknown non-negative values
- Items are presented **one-by-one**, in a **uniformly random order**
- Upon seeing an item, we can **either pick it and stop, or continue** with the next
- Goal: maximize the probability of picking the item with the **largest value**

- Theorem: There is a strategy that picks the best item with probability at least $\mathbf{1/e}$

A simple strategy

- Algorithm: **Ignore the first $n/2$ items** and then pick the first item that is larger than all of them
- The algorithm succeeds if **the best item appears in the last $n/2$ items** and **the second best item appears among the first $n/2$ items**
- The probability that the best item appears in the last $n/2$ items is 1/2
- Conditioned on this event, the probability that the second best item appears among the first $n/2$ items is at least ½
- Overall, the algorithm picks the best item with probability **1/4**

An improved analysis

- Algorithm: **Ignore the first τ items** and then pick the first item that is larger than all of them
- Y_i: the event indicating that the best item is revealed in the i-th step
- Z_i: the event indicating that the best among the $i-1$ first items actually appears in some of the first τ steps

$$\Pr[\text{success}] = \sum_{i=\tau+1}^{n} \Pr[Y_i \wedge Z_i] = \sum_{i=\tau+1}^{n} \Pr[Y_i] \cdot \Pr[Z_i] = \sum_{i=\tau+1}^{n} \frac{1}{n} \cdot \frac{\tau}{i-1}$$

$$= \frac{\tau}{n} \cdot \sum_{i=\tau}^{n-1} \frac{1}{i} \geq \frac{\tau}{n} \cdot \ln \frac{n}{\tau}$$

- Selecting $\tau \approx n/e$, we get $\Pr[\text{success}] \geq 1/e$

Martingales

Definitions

A martingale is a sequence of r.v.'s X_0, X_1, ..., of bounded expectation such that for every $i \geq 0$,

$$\mathbb{E}[X_{i+1} | X_0, X_1, \ldots, X_i] = X_i$$

More generally, a sequence of r.v.'s Z_0, Z_1, ..., is a martingale with respect to a sequence X_0, X_1, ..., if for every $i \geq 0$, the following conditions hold:

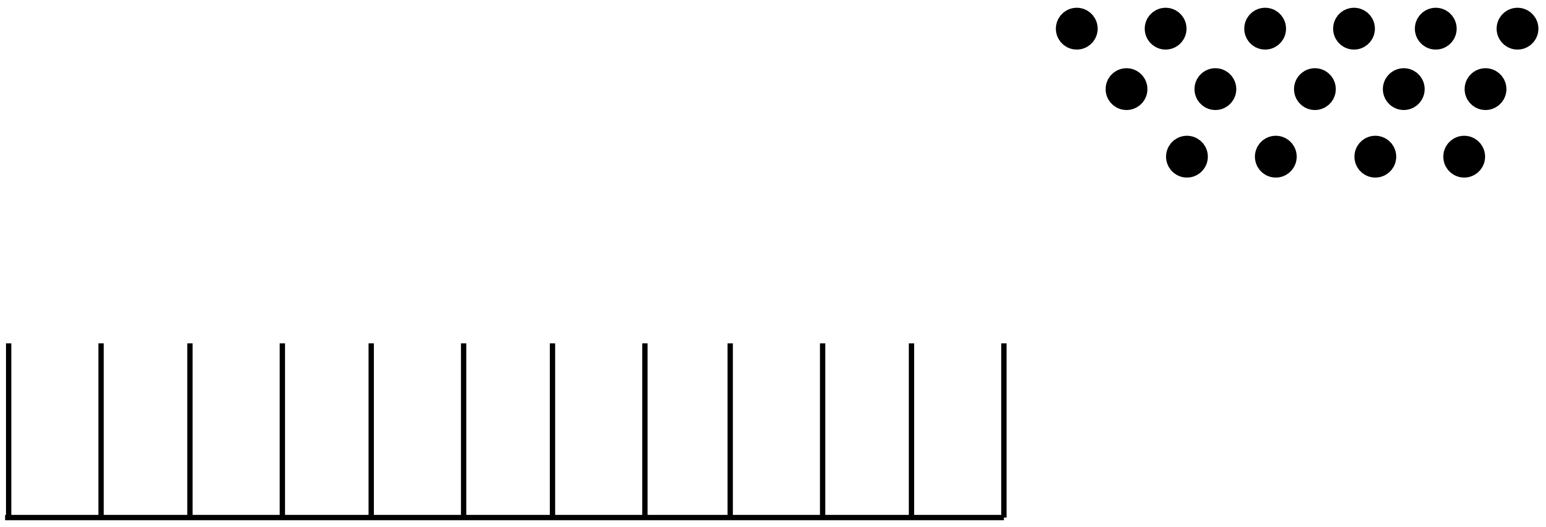
- Z_i is a function of $X_0, X_1, \ldots, X_i$
- $\mathbb{E}[|Z_i|] < \infty$
- $\mathbb{E}[Z_{i+1} | X_0, X_1, \ldots, X_i] = Z_i$

An example: gambler's fortune

- A **gambler** plays a sequence of **fair games**
- Let X_i denote the outcome of each game; fairness implies that $\mathbb{E}[X_i] = 0$
- Let Z_i denote the profits/loses up to step i
- We have that $\mathbb{E}[Z_{i+1}|X_0, X_1, \dots, X_i] = Z_i + \mathbb{E}[X_{i+1}] = Z_i$, i.e., the sequence Z_0, Z_1, …, is a **martingale**

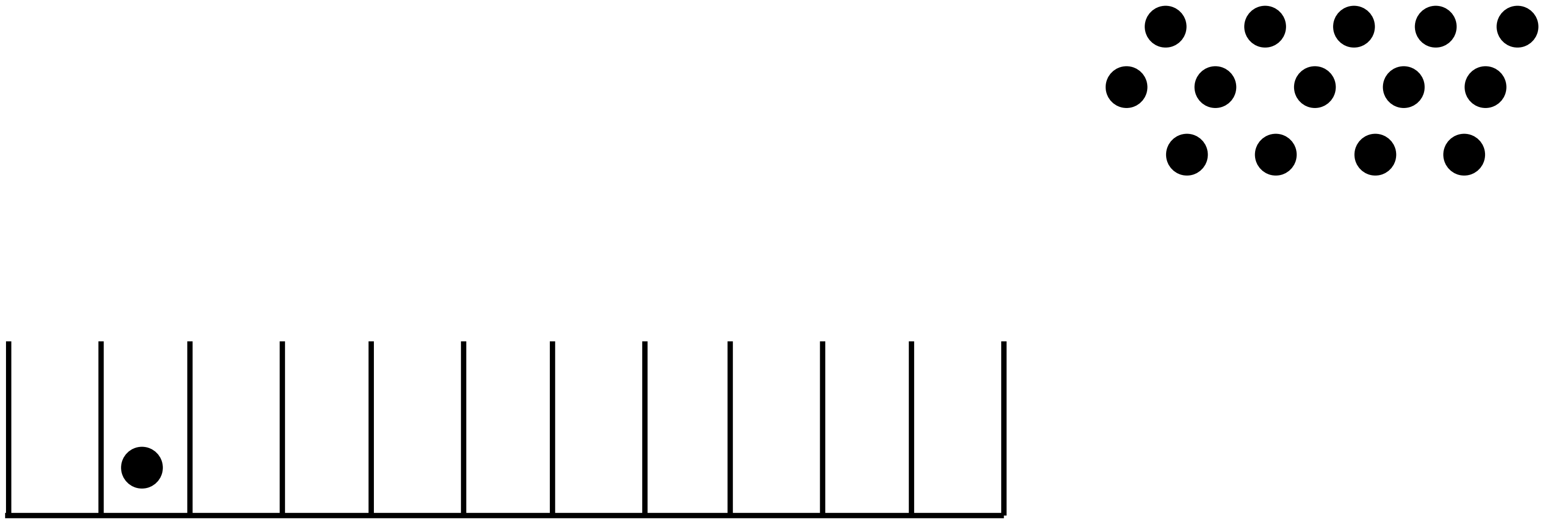
Another example: balls-to-bins

- Suppose we **throw m balls into n bins** ind/ly and uniformly at random



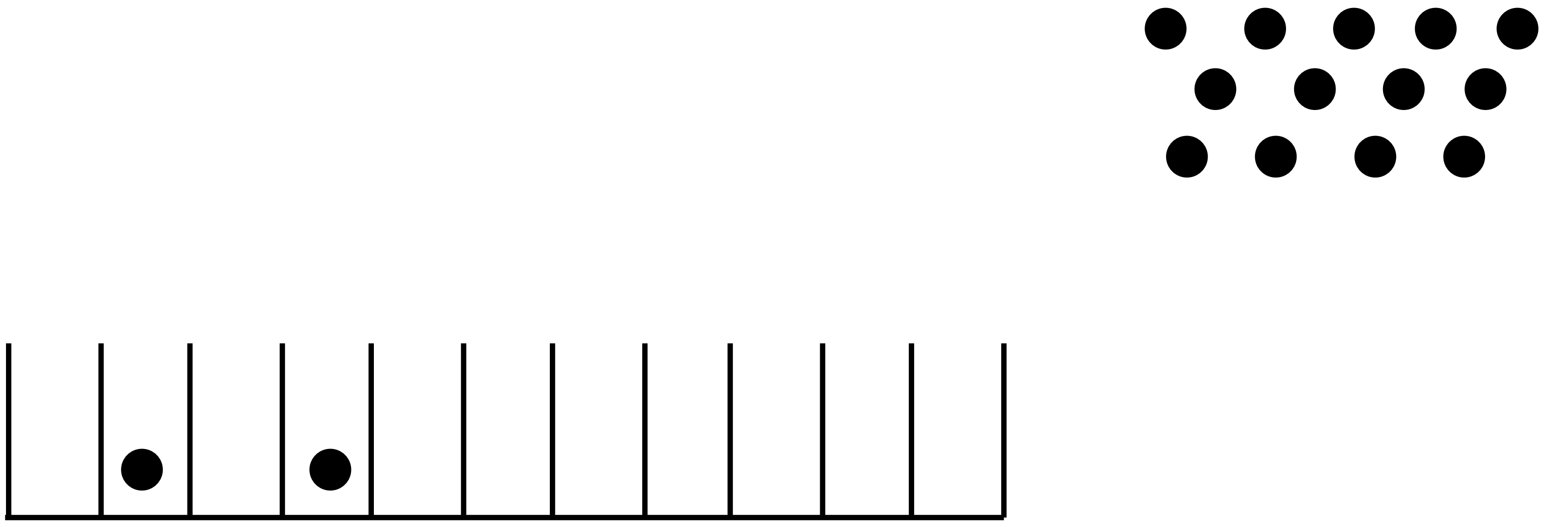
Another example: balls-to-bins

- Suppose we **throw $\boldsymbol{m}$ balls into $\boldsymbol{n}$ bins** ind/ly and uniformly at random



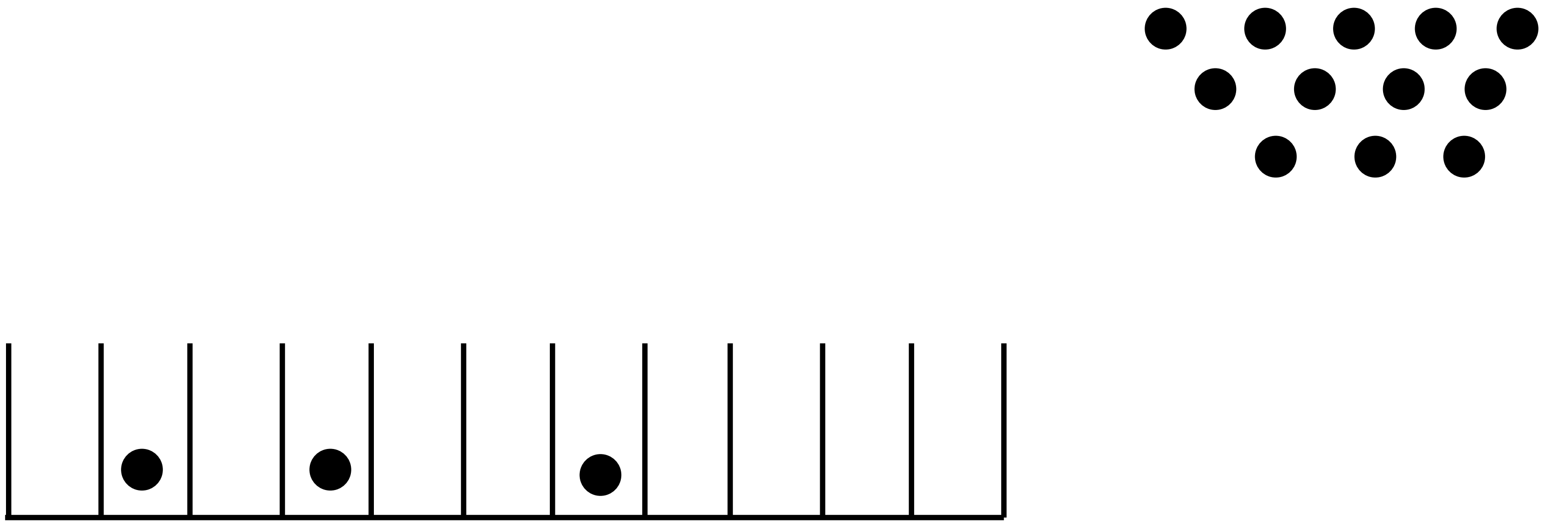
Another example: balls-to-bins

- Suppose we **throw m balls into n bins** ind/ly and uniformly at random



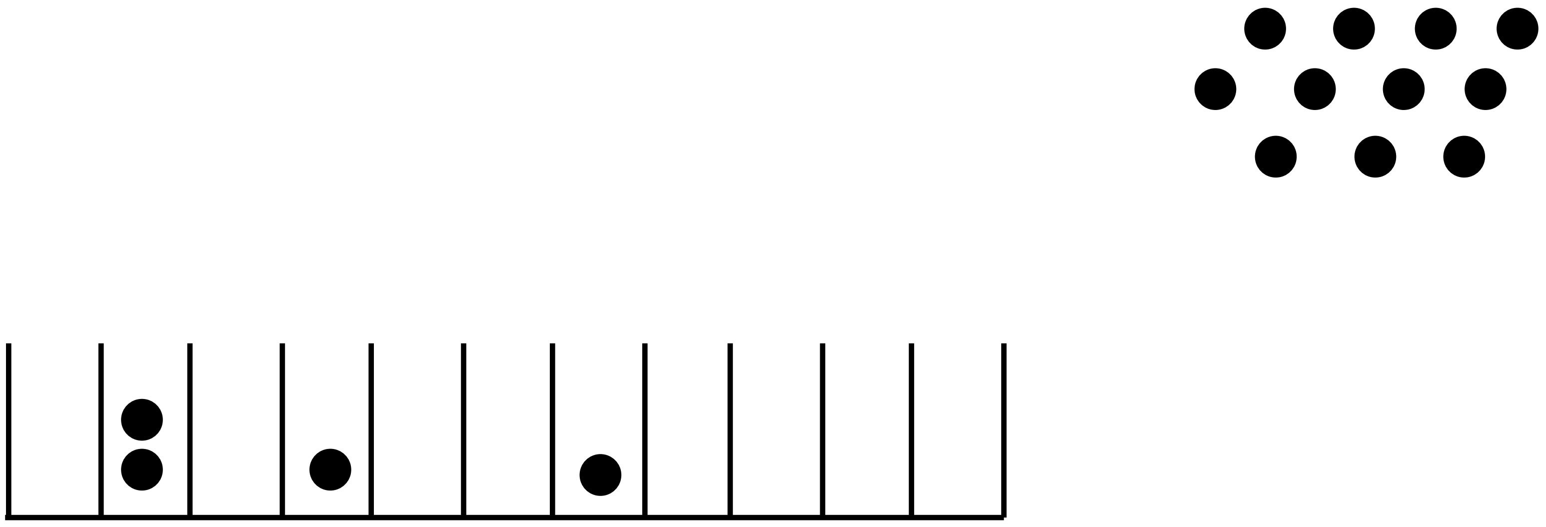
Another example: balls-to-bins

- Suppose we **throw m balls into n bins** ind/ly and uniformly at random



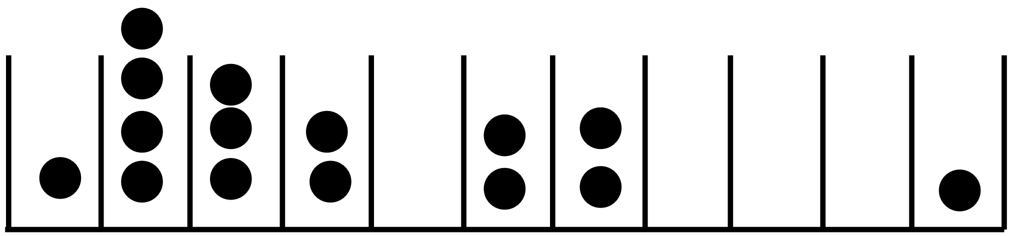
Another example: balls-to-bins

- Suppose we **throw m balls into n bins** ind/ly and uniformly at random



Another example: balls-to-bins

- Suppose we **throw m balls into n bins** ind/ly and uniformly at random



Another example: balls-to-bins

- Suppose we **throw m balls into n bins** ind/ly and uniformly at random
- Let X_i be the r.v. representing the bin to which the i-th ball falls
- Let Y be the r.v. representing the **number of empty bins** (after all balls have been thrown)
- Then, the sequence of r.v.'s $Z_0, Z_1, \dots,$ defined as $Z_i = \mathbb{E}[Y | X_1, X_2, \dots, X_i]$ is a martingale
- Why? Clearly, Z_i is a function of $X_1, \dots, X_i$ and has bounded expectation

$$\mathbb{E}[Z_{i+1} | X_1, X_2, \dots, X_i] = \mathbb{E}[\mathbb{E}[Y | X_1, X_2, \dots, X_i, X_{i+1}] | X_1, X_2, \dots, X_i]$$
$$= \mathbb{E}[Y | X_1, X_2, \dots, X_i] = Z_i$$

Doob martingales

- The number of empty bins in the previous example defines a **Doob martingale**
- Doob martingales are processes in which we obtain a **sequence of improved estimates** of the value of a r.v. as information about it is revealed progressively
- Assume that Y is a function of the r.v.'s $X_0, X_1, \ldots$
- The sequence of the mean estimates $Z_i = \mathbb{E}[Y | X_0, X_1, \ldots, X_i]$ forms a martingale with respect to the sequence $X_0, X_1, \ldots$ (provided that the Z_i's are bounded)

Azuma-Hoeffding inequality

- Let $X_0, X_1, \dots, X_n$ be a martingale such that $|X_i - X_{i-1}| \le c_i$
- Then, for any $\lambda > 0$,

$$\Pr[X_n - X_0 \ge \lambda] \le \exp\left(-\frac{\lambda^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

and

$$\Pr[X_n - X_0 \le -\lambda] \le \exp\left(-\frac{\lambda^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

Application: Number of empty bins (with $m = n$)

- Each ball cannot change the expected number of bins by more than 1, i.e., $c_i = 1$
- Recall the definition of the Doob martingale $Z_0, Z_1, \dots,$
- $Z_0 = \mathbb{E}[Y]$ and $Z_n = Y$
- Hence, $\Pr[|Y - \mathbb{E}[Y]| \geq \varepsilon n] = \Pr[|Z_n - Z_0| \geq \varepsilon n] \leq 2\exp\left(-\frac{\varepsilon^2 n}{2}\right)$
- But what is the expected number of empty bins?

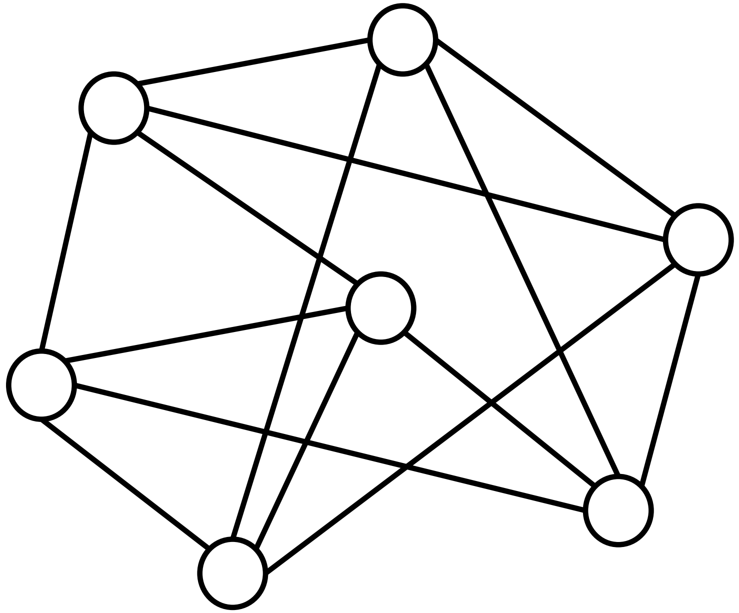
$$\mathbb{E}[Y] = n\left(1 - \frac{1}{n}\right)^n \approx n/e$$

Application: Chromatic number of random graphs

- **Random $G_{n,p}$ graph model**: n nodes, each edge exists with probability p, independently from the others

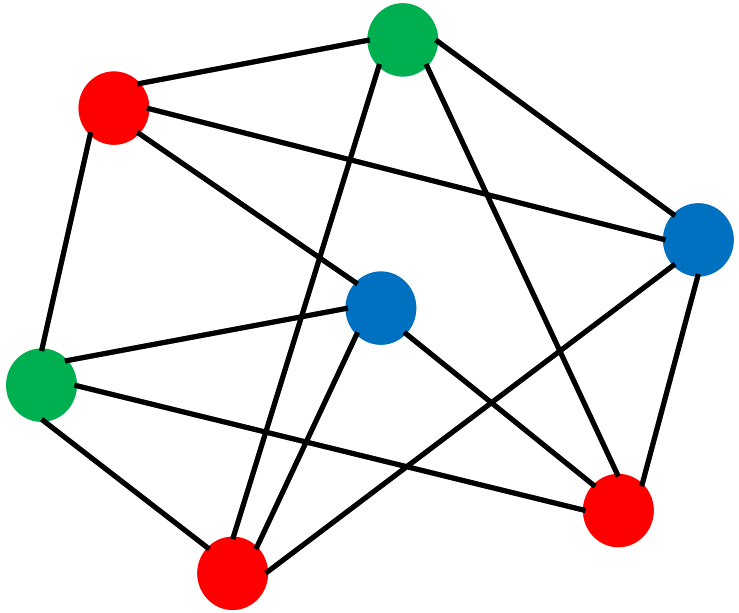
Application: Chromatic number of random graphs

- **Random $G_{n,p}$ graph model**: n nodes, each edge exists with probability p, independently from the others
- **Chromatic number** $\chi(G)$: the minimum number of colors needed to assign to the nodes of graph G so that no adjacent nodes have the same color



Application: Chromatic number of random graphs

- **Random $G_{n,p}$ graph model**: n nodes, each edge exists with probability p, independently from the others
- **Chromatic number $\chi(G)$**: the minimum number of colors needed to assign to the nodes of graph G so that no adjacent nodes have the same color



Application: Chromatic number of random graphs

- **Random $G_{n,p}$ graph model**: n nodes, each edge exists with probability p, independently from the others
- **Chromatic number** $\chi(G)$: the minimum number of colors needed to assign to the nodes of graph G so that no adjacent nodes have the same color
- Let G_i be the node-induced subgraph consisting of nodes $1, 2, \dots, i$
- Define the Doob martingale $Z_i = \mathbb{E}[\chi(G) | G_1, G_2, \dots, G_i]$, which we call the **node exposure martingale**
- Clearly, any new node does not change the expected chromatic number by more than 1, i.e., $c_i = 1$
- Hence, $\Pr[|\chi(G) - \mathbb{E}[\chi(G)]| \geq \lambda] = \Pr[|Z_n - Z_0| \geq \lambda] \leq 2\exp\left(-\frac{\lambda^2}{2n}\right)$

Last slide

- Prophet inequality
- Secretary problem
- Martingales