Purely greedy exploitation

- **Algorithm:** Toss the coin having the maximum empirical reward
- **Consider two coins:**
  - **Coin 1** with probability of head **0.4**
  - **Coin 2** with probability of head **0.6**
- **With probability 0.24, Coin 1** gives head and **Coin 2** gives tail in the first toss
- **The algorithm then sticks to the sub-optimal Coin 1**
- **Overall regret is atleast** $0.048 \times T$ **which is linear !!**

Having sub-linear regret

Can we get sub-linear regret i.e. $\lim_{T \to \infty} \frac{\text{Regret}_T}{T} = 0$?

Recall that:

$$\frac{1}{T} \times \text{Expected total reward till time } T = \text{Mean reward of best arm } (\mu^\star) - \frac{\text{Regret}}{T}$$

Thus having $\frac{\text{Reg}_T}{T} \to 0$ implies:

$$\frac{1}{T}\text{Expected total reward} \to \text{Mean of best arm } (\mu^\star)$$