Simulation: Flight Market

Course: Dynamic Programming and Reinforcement Learning

Fabian Lange and Leonard Dreeßen

Mid-Term Recap: Flight Ticket Market

Problem Context:

Time Horizon: finite (take-off)

Action: Price Offer

Demand: Price and Time-Dependent

Event: Tickets sold

Reward: Sales Revenue

Simple Environment:

State Space: Discrete (0, 10) → (10, 0)

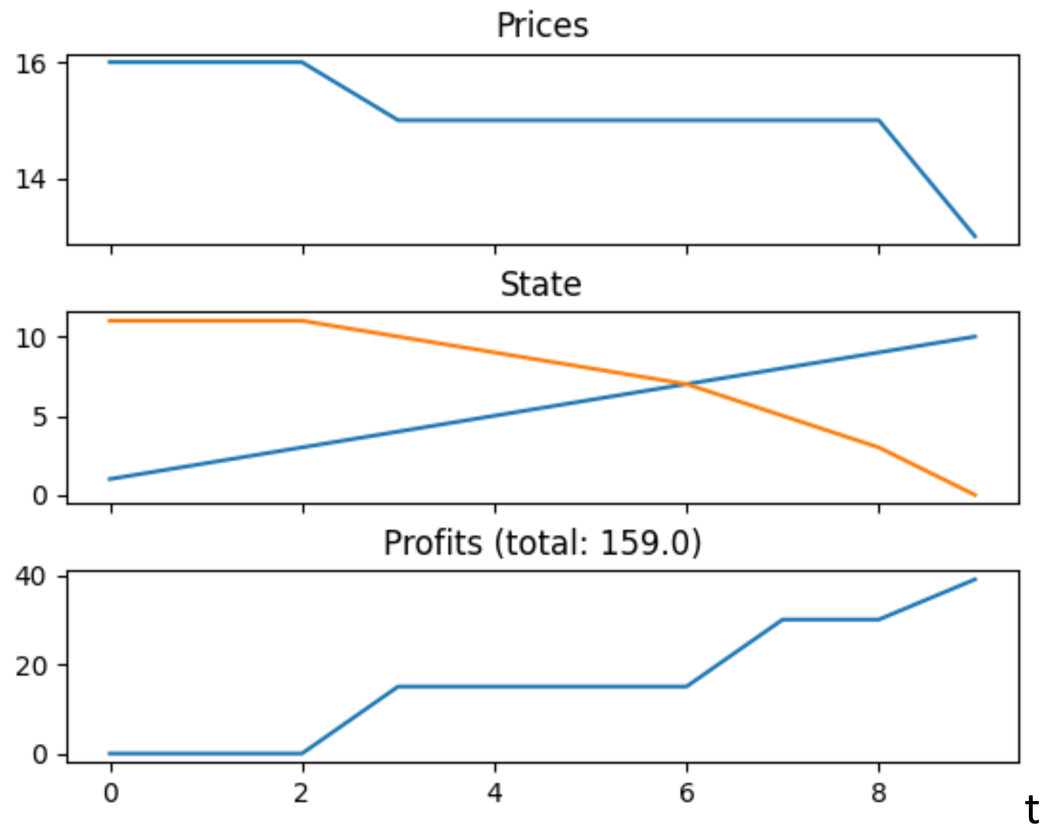
Action Space: Discrete (0, 10, 20, …, 200)

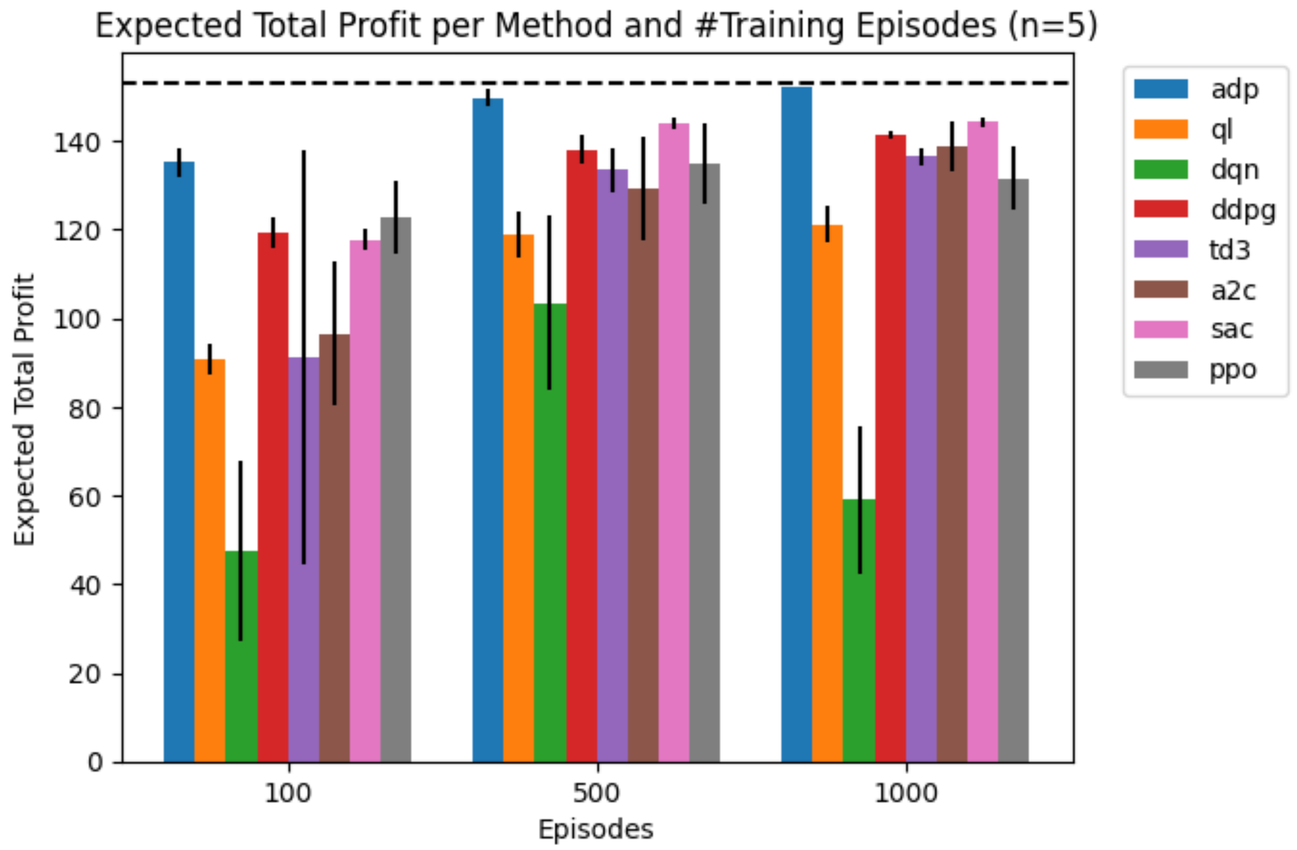
Event Space: Discrete (0, …, 10)

Reward: Price x Tickets sold

Buying Probability: $\left(1 - \frac{a}{200}\right) * (1 + \frac{t}{10})$

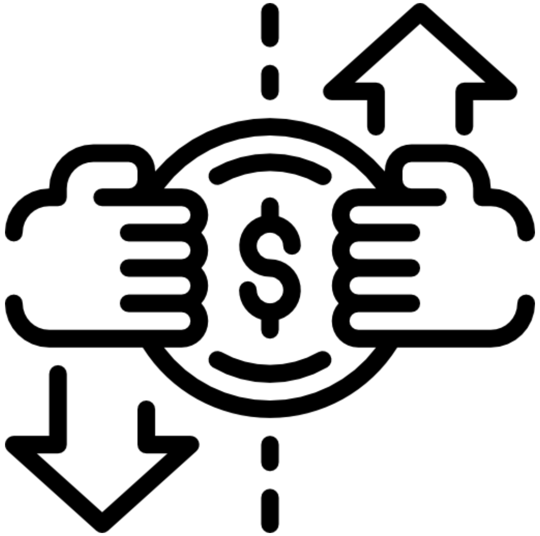
Mid-Term Recap





Research Directions

- Realistic Customer Behavior
- Estimating Demand and using Dynamic Programming methods
- Duopoly setup



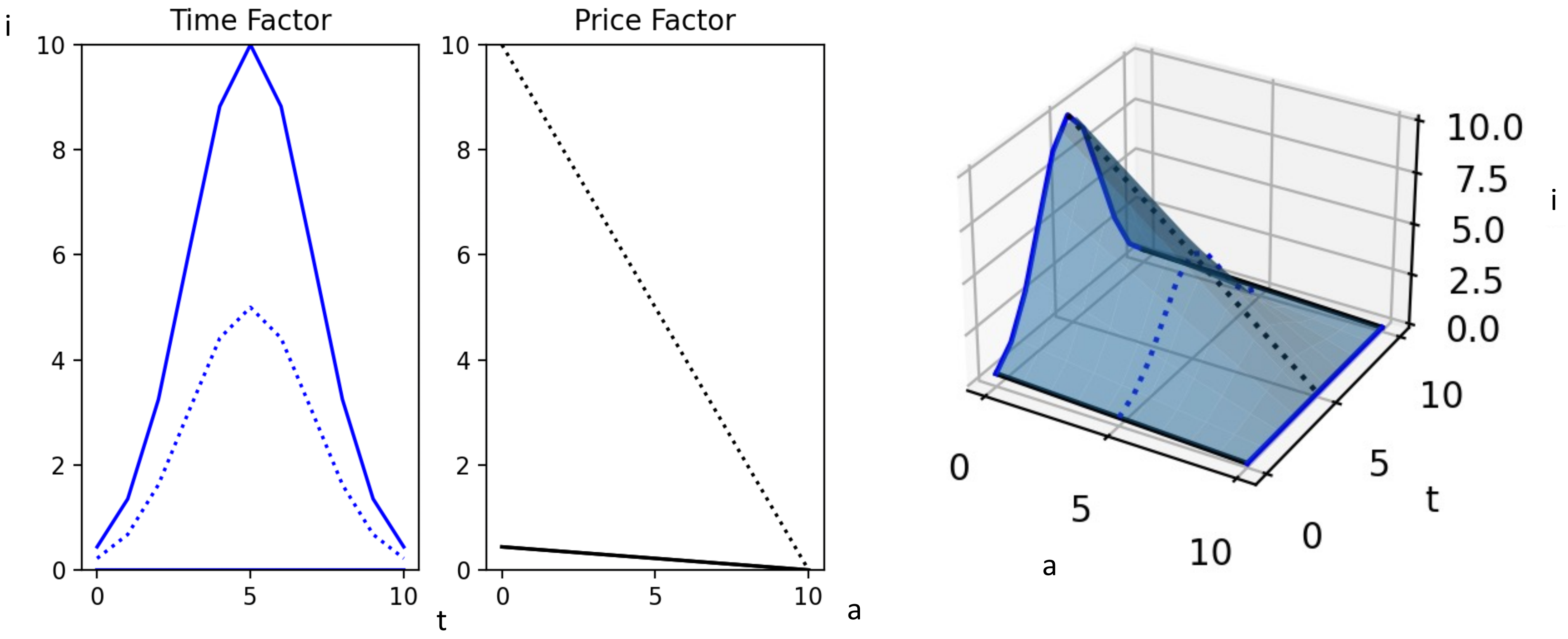
Customer Behavior

- Five types of customer behaviors implemented
  - Rational, Family, Business, Early Booking, Party
- Every combination of types possible

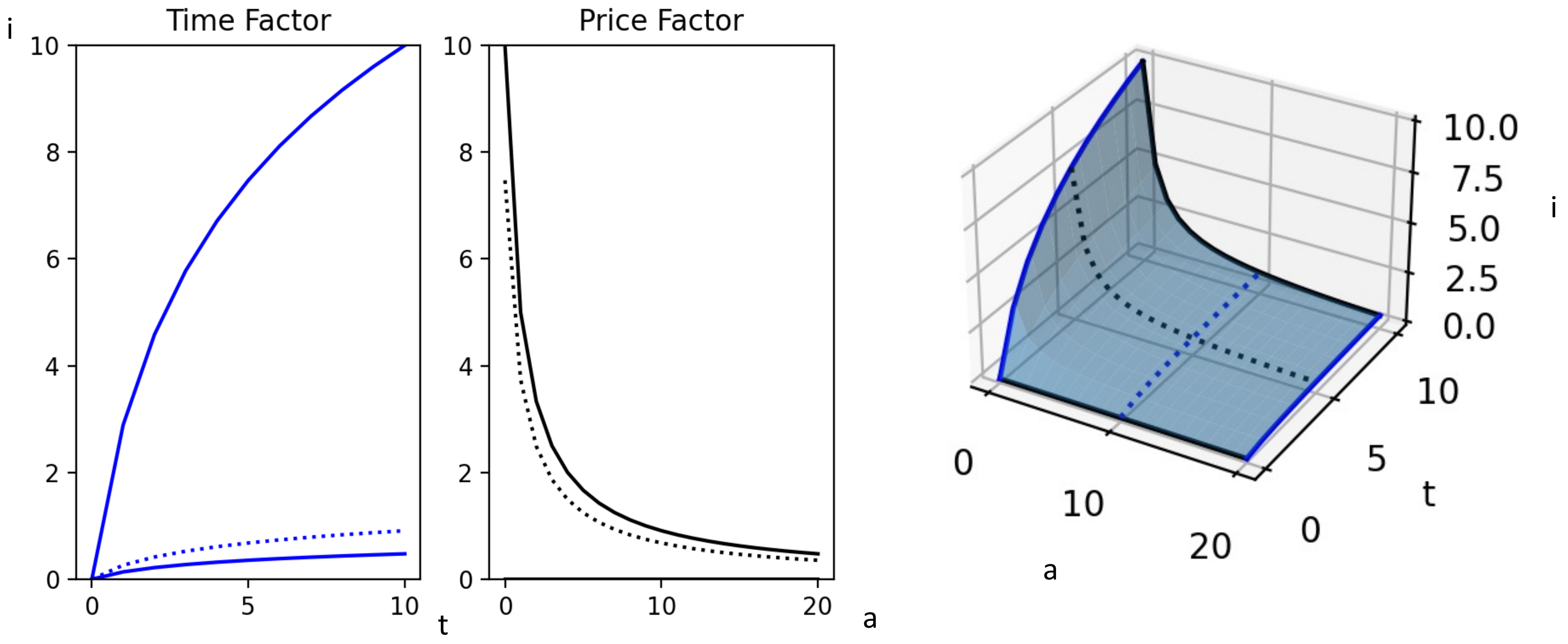
- One simulation step:

n customers arrive → split up equally into groups → calculate buying probability for each group → average all buying probabilities to get p → draw from multinomial distribution with p and n

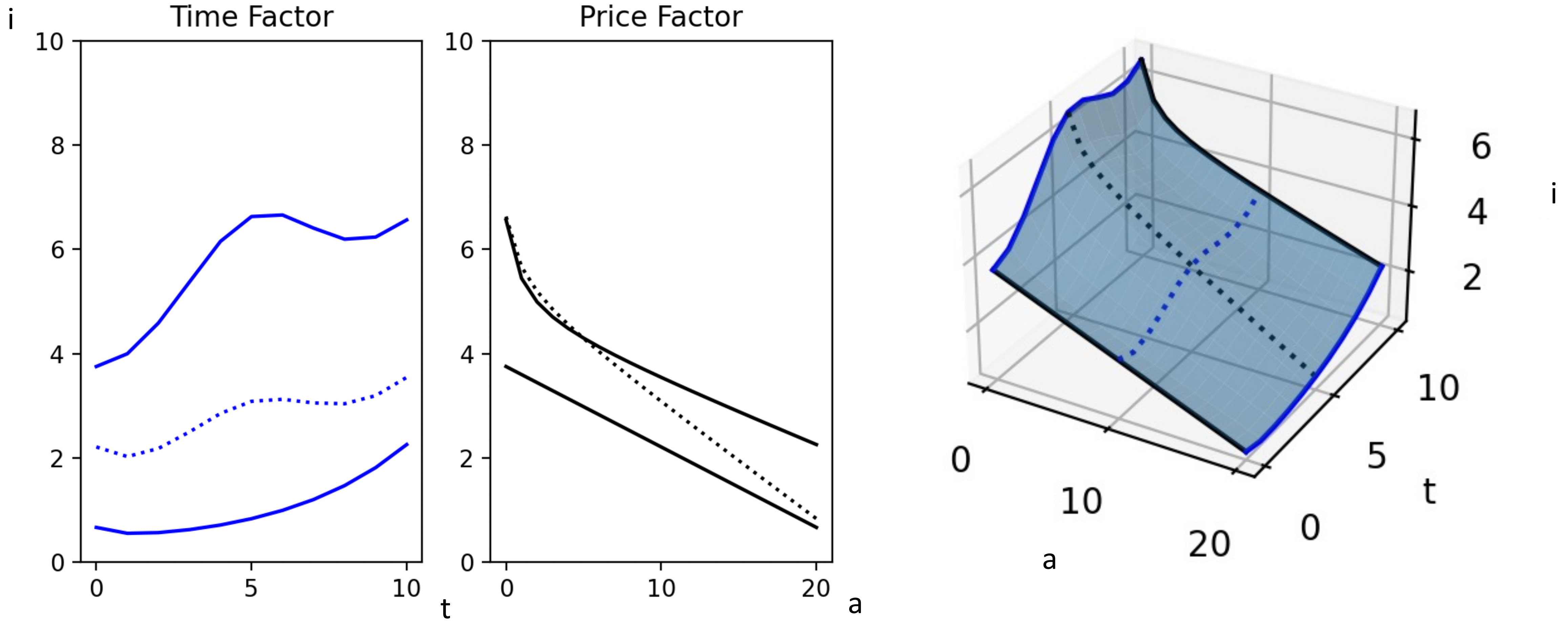
Family Customer



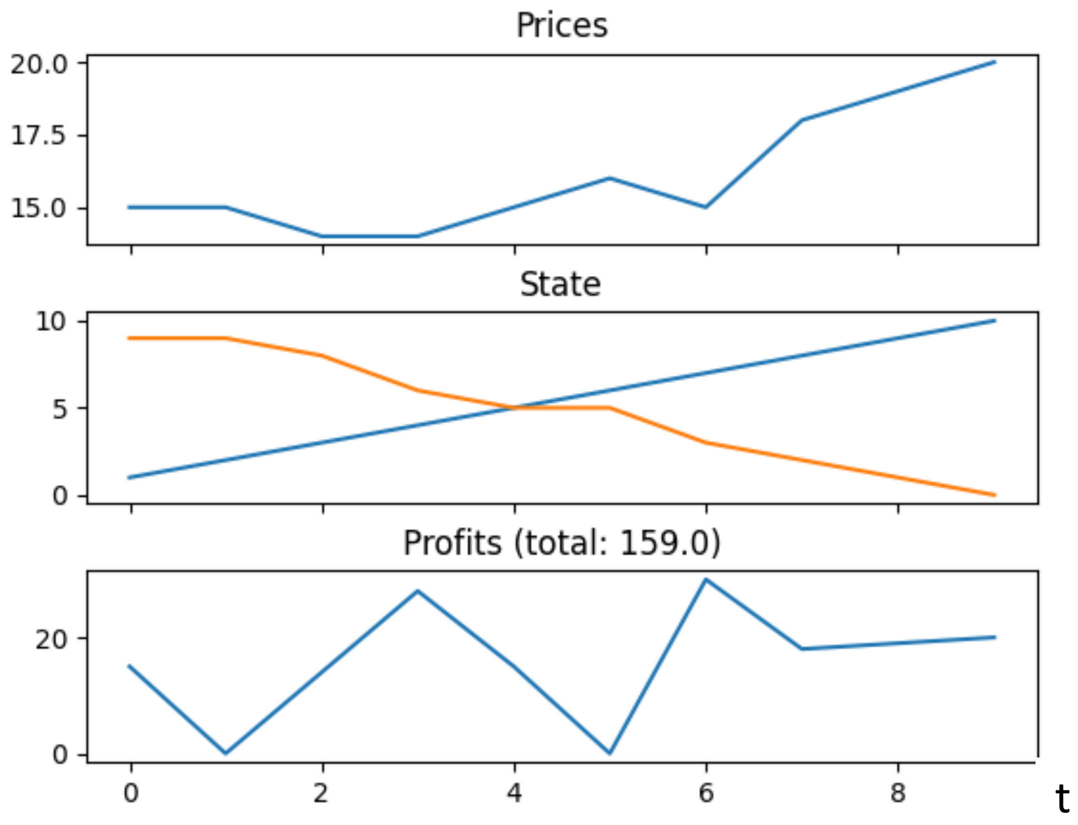
Party Customer

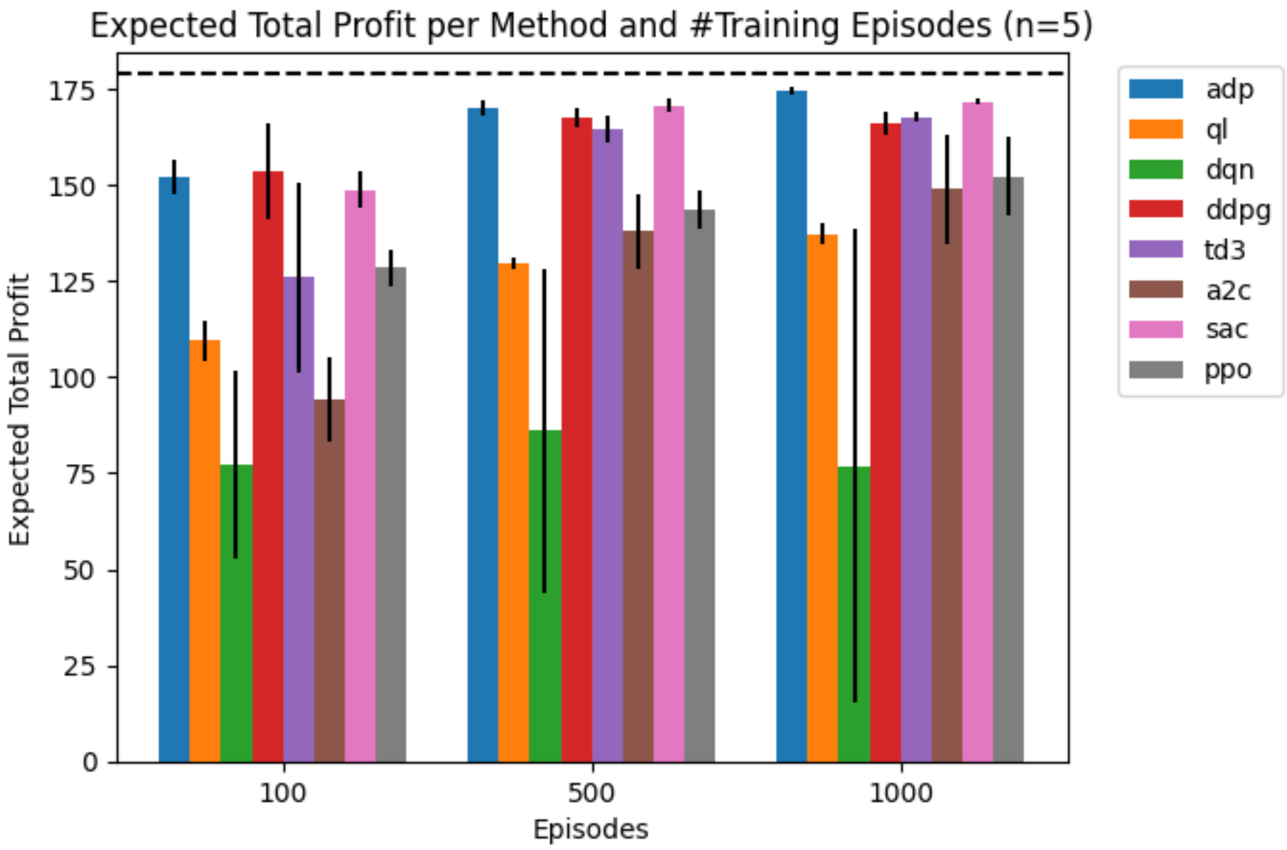


All Customers combined



New Simulation





Estimating Customer Behavior

Airlines cannot tryout prices over 500 Episodes $\sim$ 15 years

DP methods would not require playing random actions

Airlines might have historical data on customer demand

Estimating Customer Behavior

$$V_t(s) = \max_{a \in A_t(s)} \left\{ \sum_{i \in I_t} P_t(i,a,s) * \left(r_t(i,a,s) + \gamma * V_{t+1}(\Gamma_t(i,a,s))\right) \right\}$$

- Airlines know
  - state transitions, reward function, state, action space
- Airlines do not know
  - Number of arriving customers, probability distribution of customer purchases

Use remaining seats as limit

Regression Analysis

Regression Analysis

- Example dataset: play random actions over n episodes, store i, a, s
- Estimate expected sales ($\hat{\imath}$) with OLS, Ridge or LASSO Regression

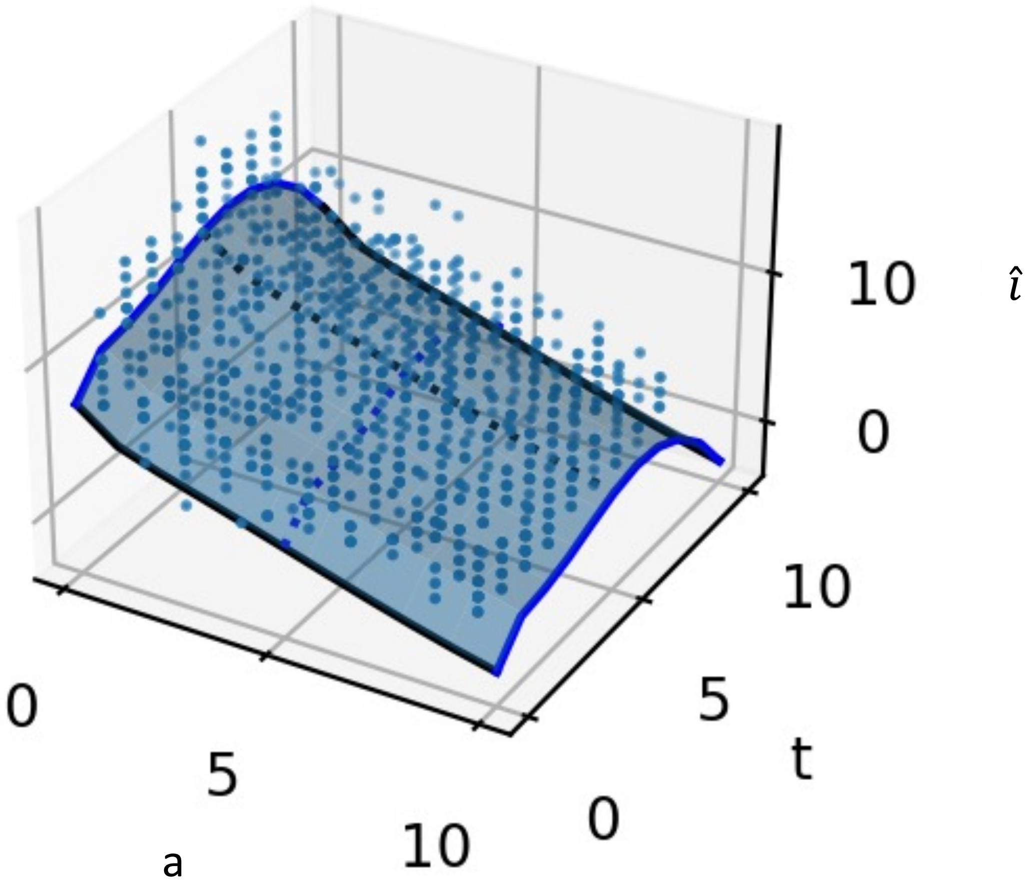
Best Performing:

- OLS & Ridge with at least 1000 datapoints
- Features: $a, t, a^2, t^2, \sqrt{a+1}, \sqrt{t+1}, \log(a+1)\,, \log(t+1)\,, a * t$

→ Without stochastic customers: $R^2 \sim 0.837$ and $\emptyset nMSE = 0.021$

→ Stochastic customers: $R^2 \sim 0.35$ and $\emptyset nMSE = 0.21$

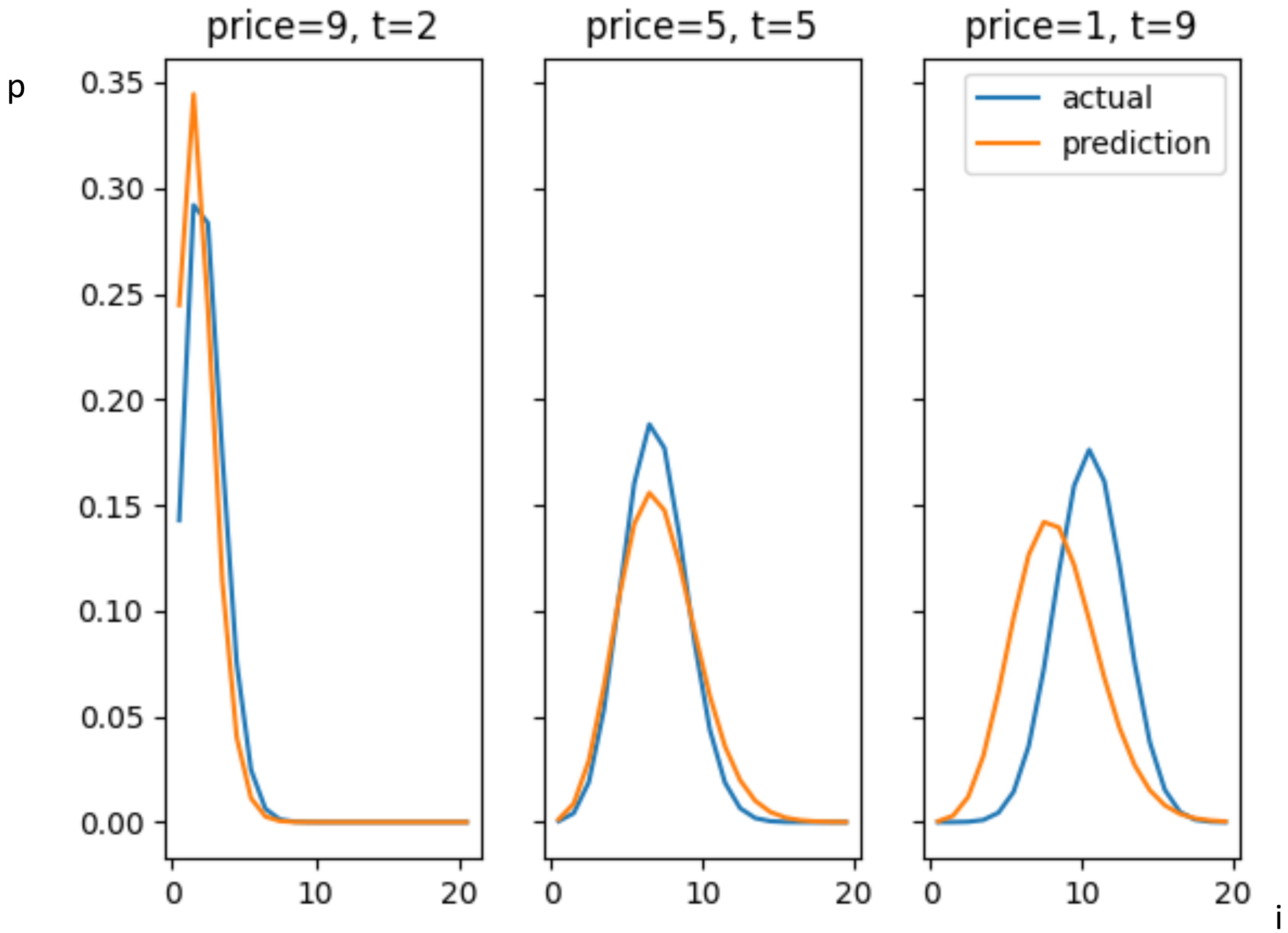
Estimation -> Probability



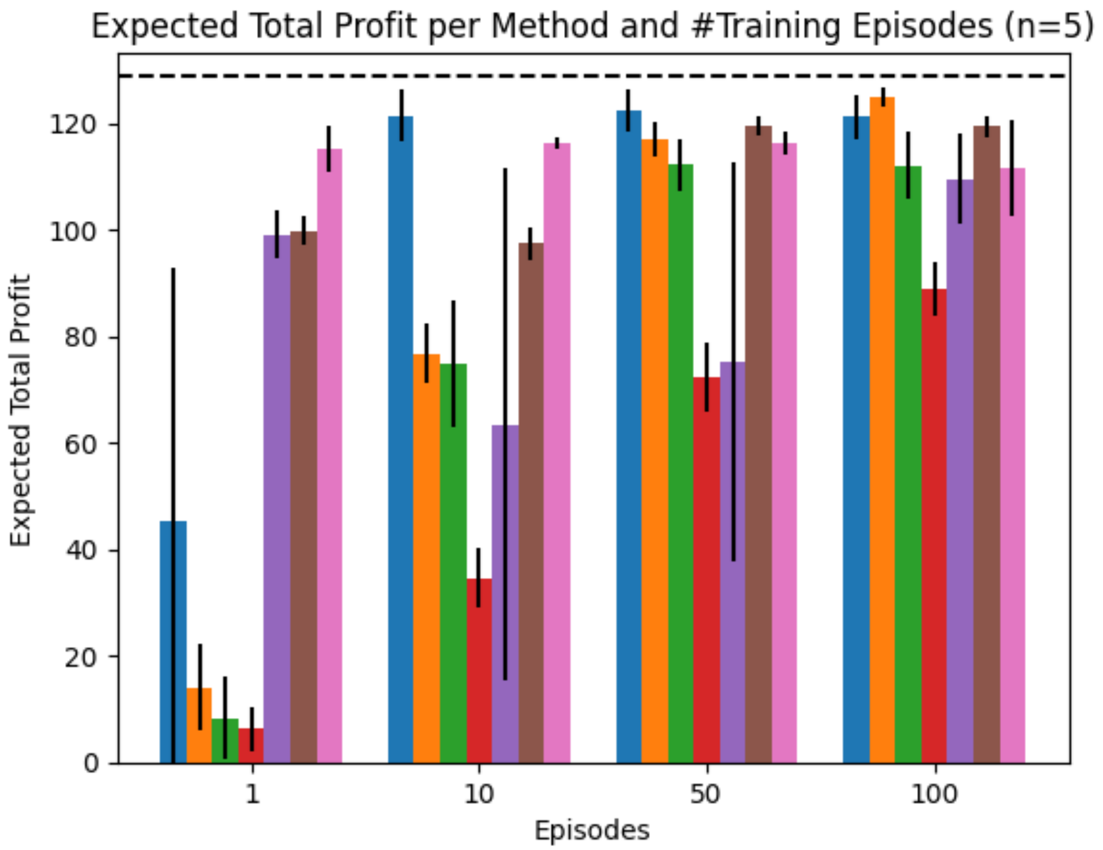
Discrete probability distribution required:

Poisson distribution with our estimation as parameter:

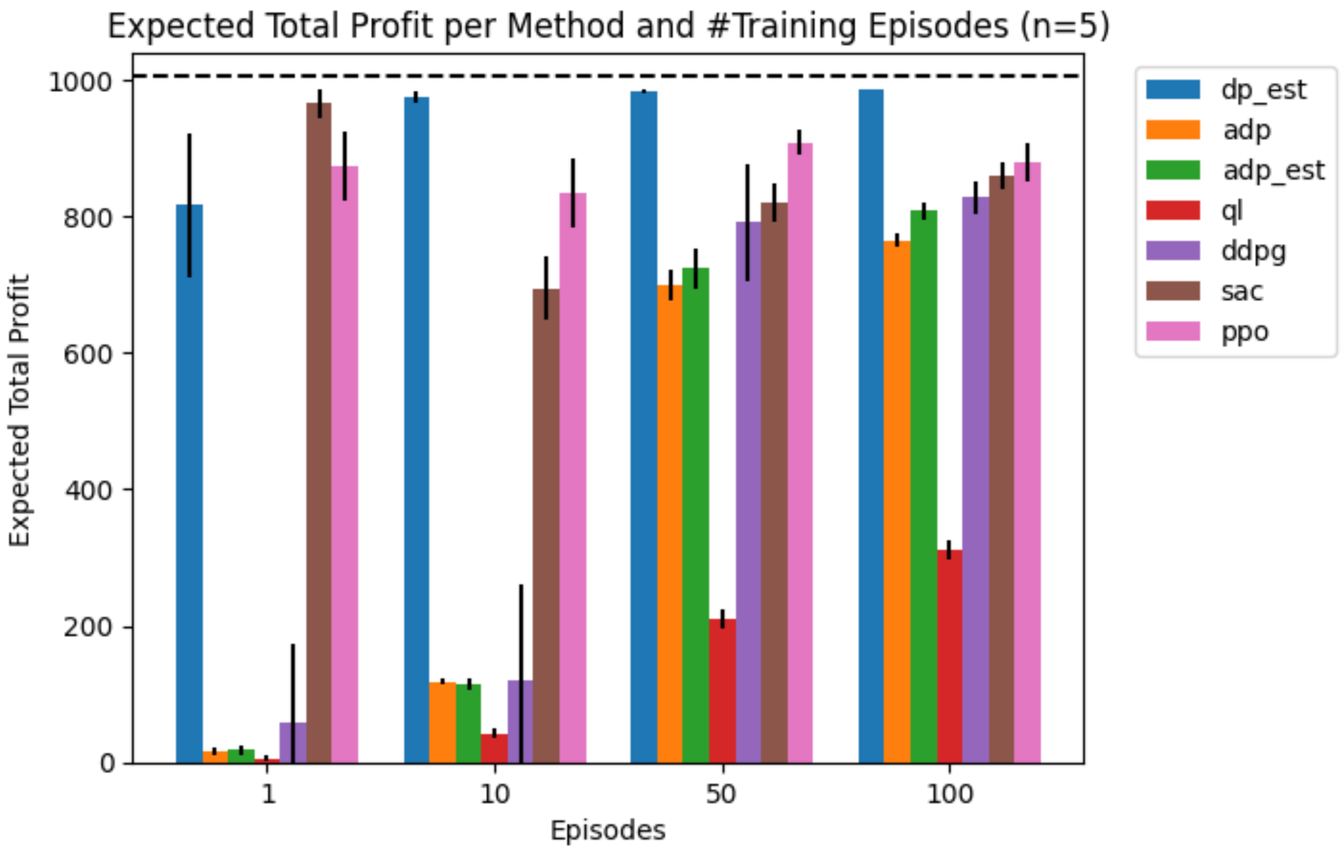
$P_\lambda(k) = \frac{\lambda^k}{k!} * e^{-\lambda}$ with $\lambda = \hat{\imath}$


p
price=9, t=2
price=5, t=5
price=1, t=9
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
actual
prediction
i

DP and ADP using Estimation



Simple Environment



Complex Environment

Duopoly - Setup

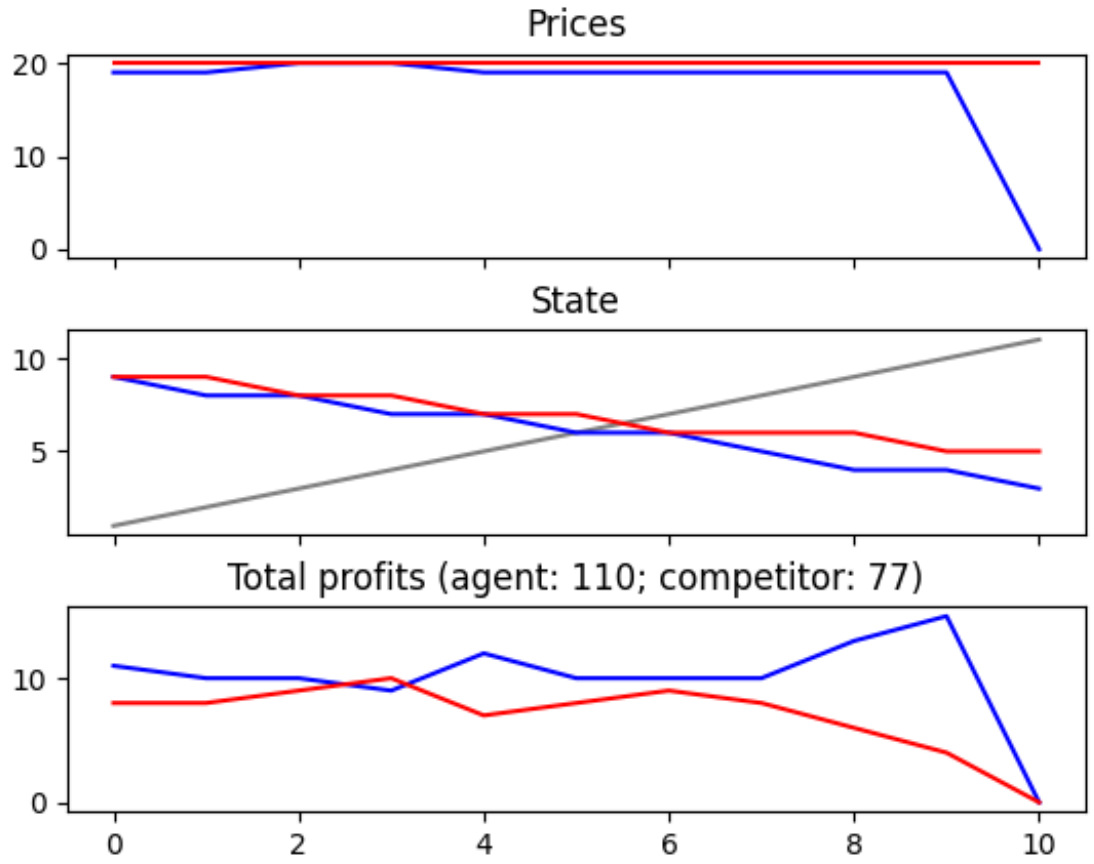
State Space: Discrete $(0, 5, 5, a_{agent}^{t\_0}, a_{comp}^{t\_0})$	**Demand**: $\frac{a_{comp}}{a_{agent}+a_{comp}}$ →
→ $(10, s_{agent}^{t\_10}, s_{comp}^{t\_10}, a_{agent}^{t\_10}, a_{comp}^{t\_10})$	$softmax(D_{agent}, D_{comp}, D_{no\_sell})$ →
Action Space: Discrete $(0, 1, 2, 3, \dots, 100)$	$P_{agent}, P_{comp}, P_{no\_sell}$ →
Event: i_{agent}, i_{comp}	$multinom()$ →
Reward: $i_{player} \cdot a_{player}$	i_{agent}, i_{comp}

Duopoly Scenario – Full Information

agent has **full knowledge** about:

- exact customer behavior
- *competitor strategy*
  - undercut agents price by 1 (but not go below 20)

$\rightarrow \min(20, a_{agent\_last} - 1)$



Duopoly Scenario – Limited Information

agent has **knowledge** about:

- own inventory
- own price and competitors price
  - assumes here a time indepedent fix price strategy

→ $30 \; \forall t$

Duopoly – Outlook

- setting duopoly fully up:
  - support all DP and RL methods we used so far
  - provide comprehensive monitoring like in monopoly
  - enable estimation of
    - consumer behavior
    - competitor strategy