Statistics and Samples in Distributional Reinforcement Learning

Mark Rowland, Robert Dadashi, Saurabh Kumar, Rémi Munos, Marc G. Bellemare, Will Dabney

ICML 2019

DeepMind

Google Research
Brain team

Distributional Reinforcement Learning

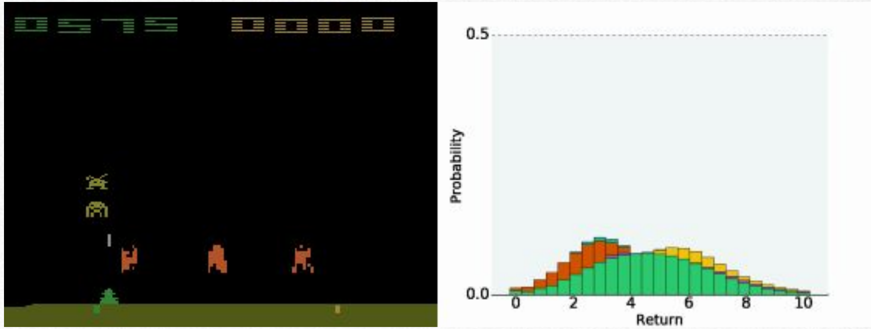
Distributional RL aims to learn
full return distributions.

Return distribution:

$$Z^\pi(x,a) = \sum_{t=0}^{\infty} \gamma^t R_t \,\Big|\, X_0 = x, A_0 = a$$

Distributional Bellman equation:

$$Z^\pi \stackrel{\mathcal{D}}{=} T^\pi Z^\pi$$

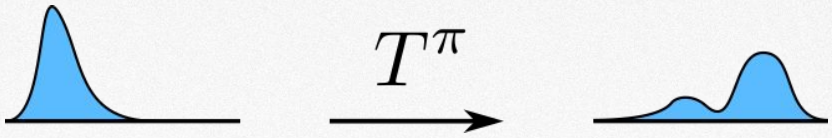


[Bellemare et al., 2017]

Distributional Reinforcement Learning

In practice, we often work with **parametric approximate distributions**.

Non-parametric

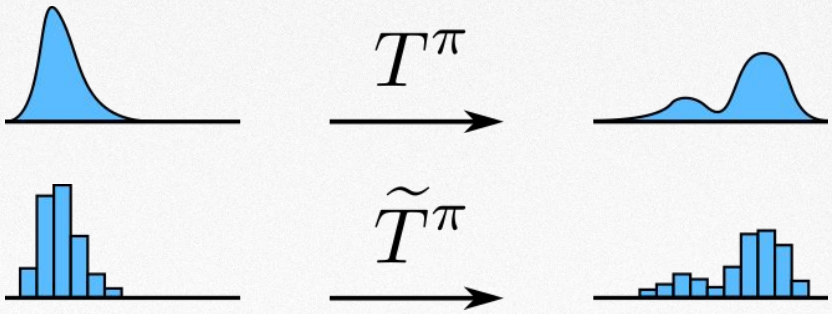


Distributional Reinforcement Learning

In practice, we often work with **parametric approximate distributions**.

Non-parametric

Categorical [Bellemare et al., 2017]



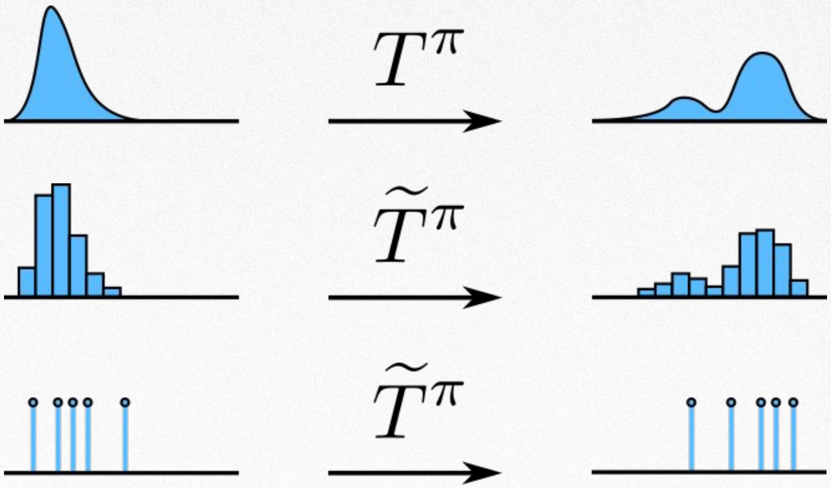
Distributional Reinforcement Learning

In practice, we often work with **parametric approximate distributions**.

Non-parametric

Categorical [Bellemare et al., 2017]

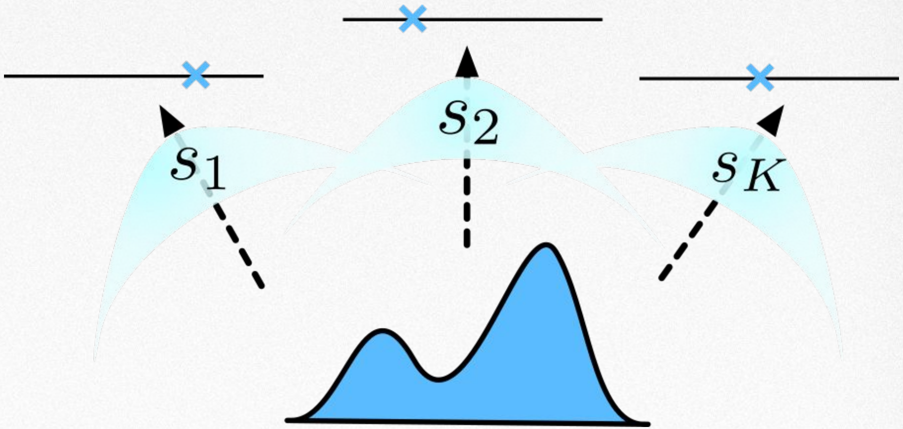
Dirac deltas [Dabney et al., 2018]



Main Contribution: An Alternative Perspective

Distributional RL algorithms learn **statistical functionals** of the return distribution.

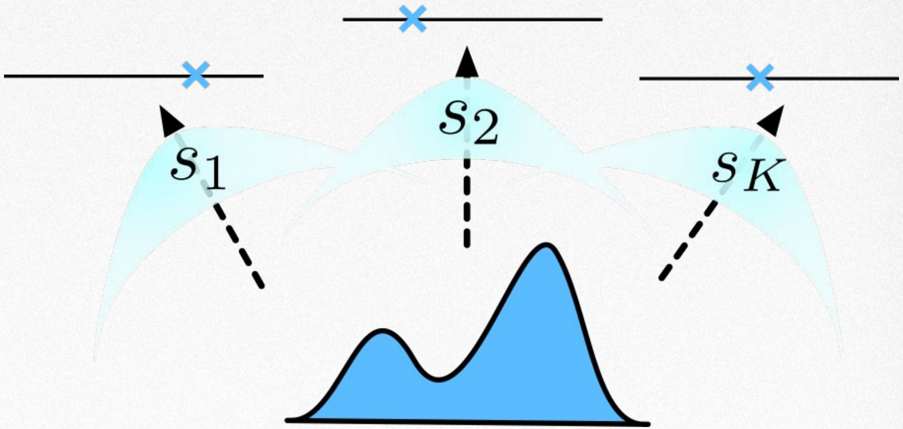
- Moments, tail probabilities, expectations, etc.



Main Contribution: An Alternative Perspective

Distributional RL algorithms learn **statistical functionals** of the return distribution.

- Moments, tail probabilities, expectations, etc.



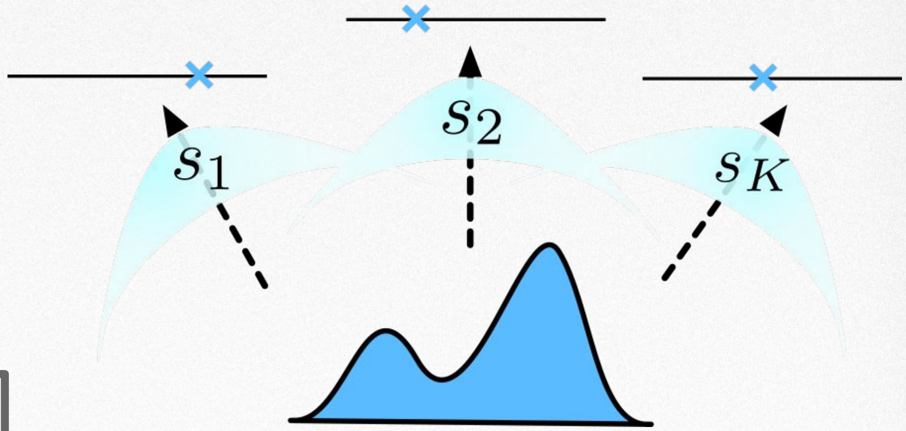
Theory: What properties of return distributions can be learnt through dynamic programming?

Algorithmic: A general framework for approximate learning of statistics.

Main Contribution: An Alternative Perspective

Distributional RL algorithms learn **statistical functionals** of the return distribution.

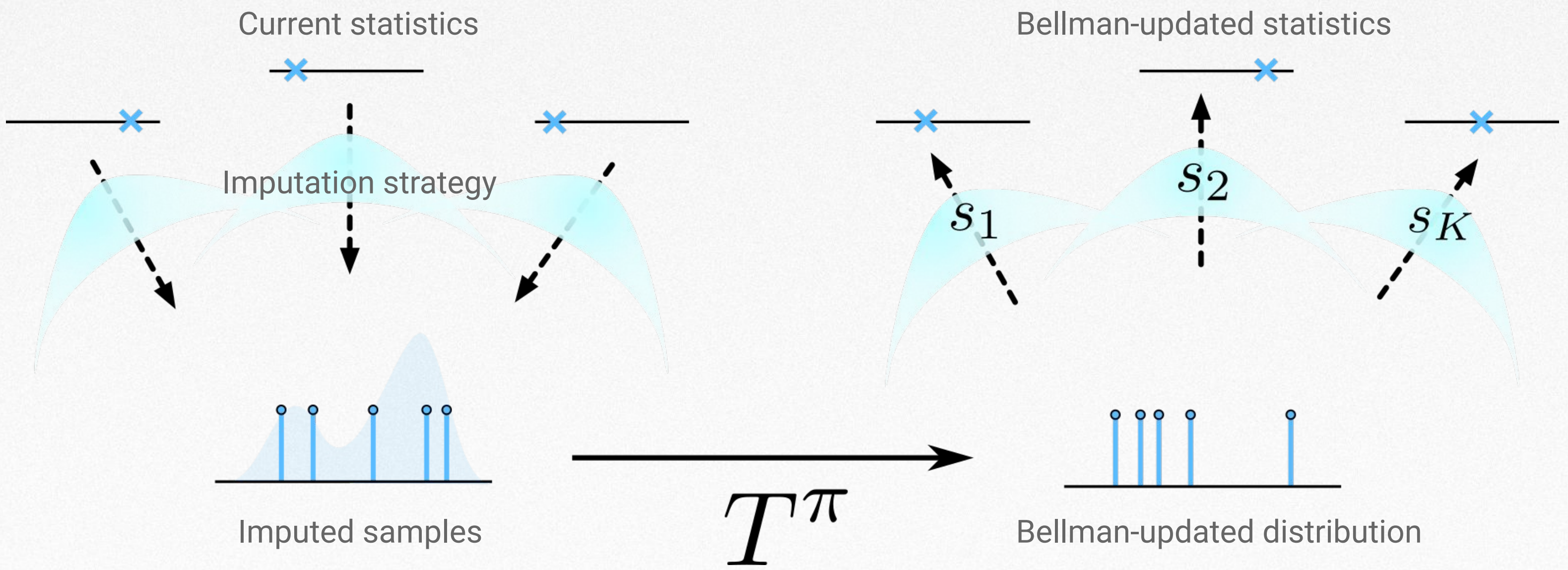
- Moments, tail probabilities, expectations, etc.



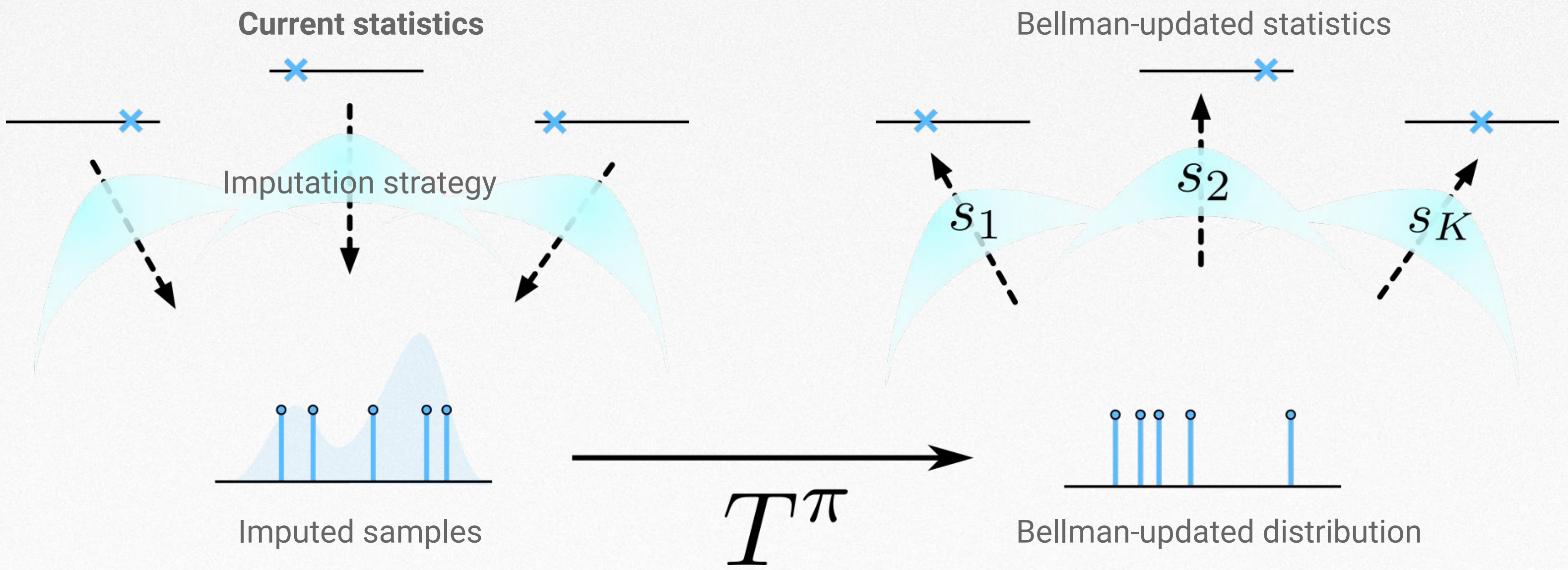
Theory: What properties of return distributions can be learnt through dynamic programming?

Algorithmic: A general framework for approximate learning of statistics.

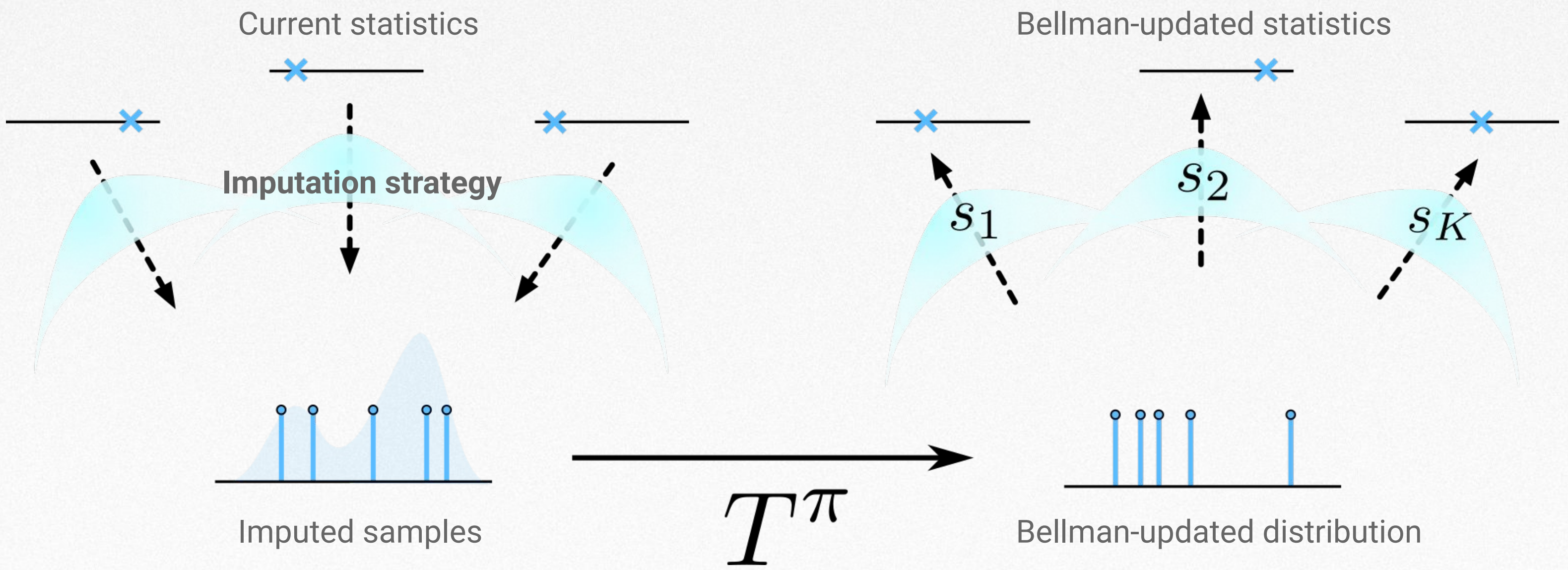
A General Framework for Distributional RL Algorithms



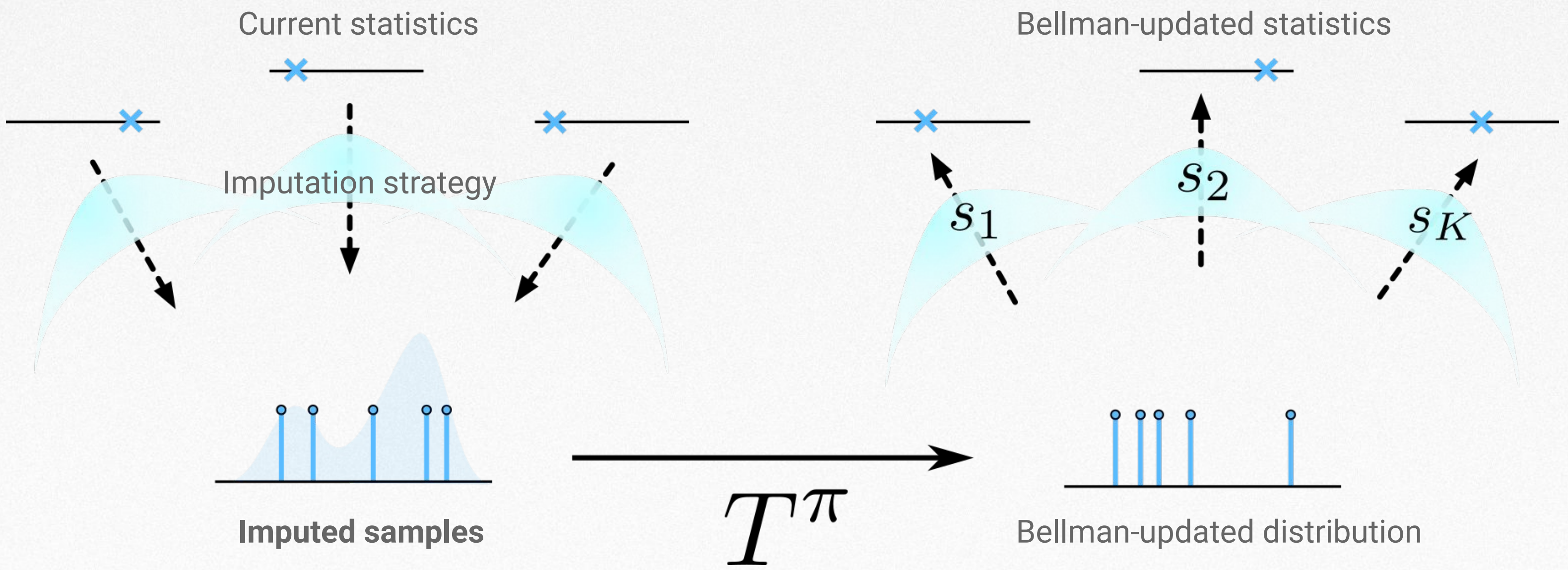
A General Framework for Distributional RL Algorithms



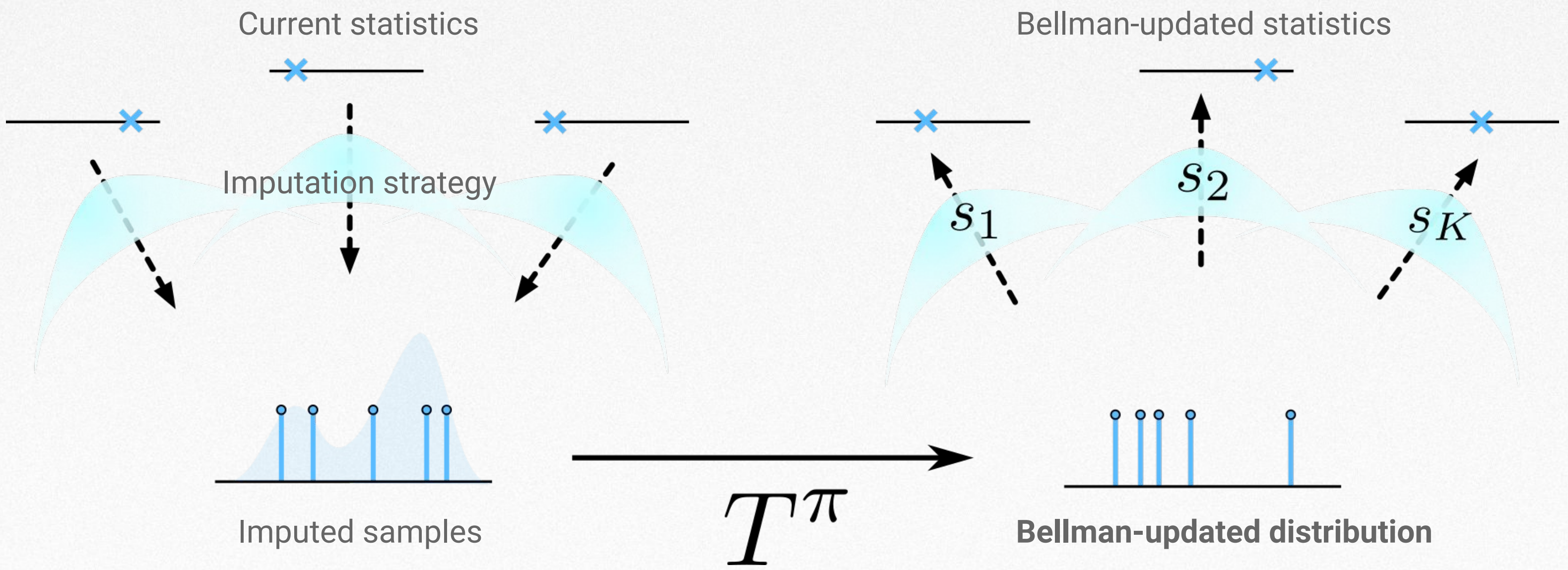
A General Framework for Distributional RL Algorithms



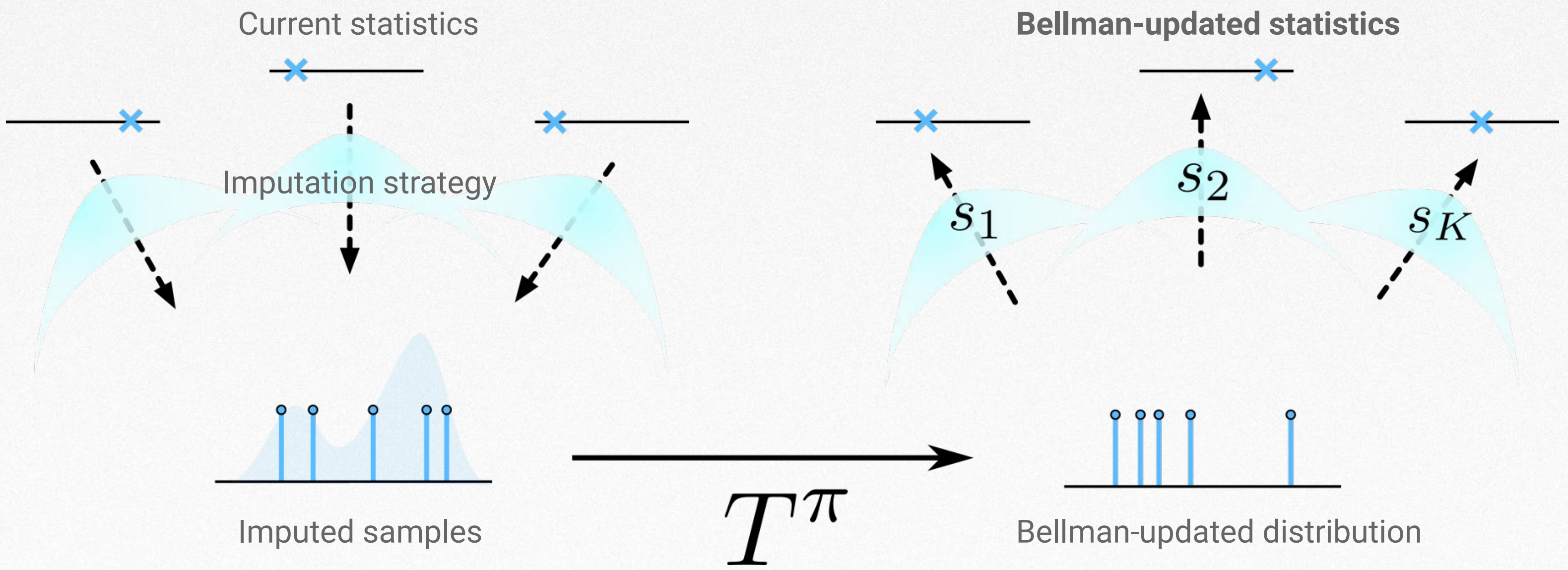
A General Framework for Distributional RL Algorithms



A General Framework for Distributional RL Algorithms



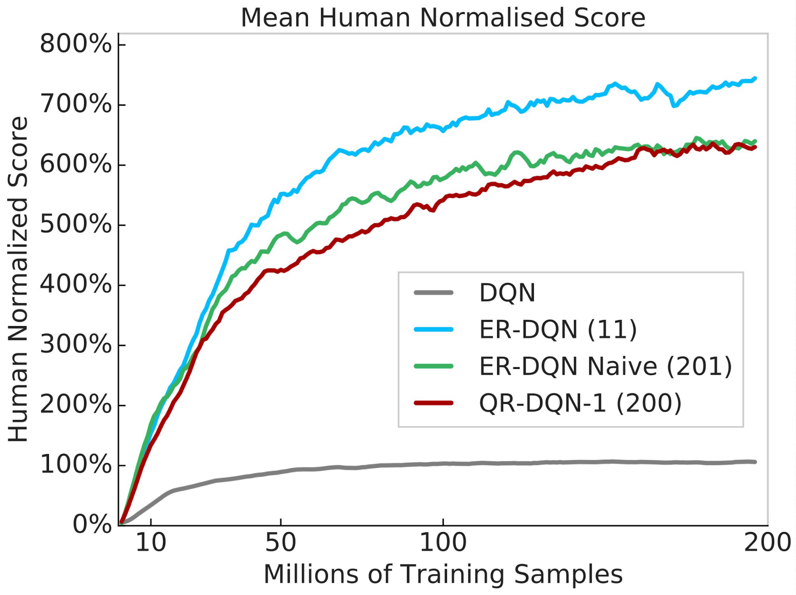
A General Framework for Distributional RL Algorithms



Application: Expectiles

We apply this framework to learn **expectiles** of return distributions.

New deep RL agent: **Expectile Regression DQN (ER-DQN)**, with **improved mean performance** on Atari-57 relative to QR-DQN.



Summary

A new perspective on distributional RL

Theoretical progress on what it is possible to learn

A general framework for distributional RL algorithms

THANK YOU

Poster #113