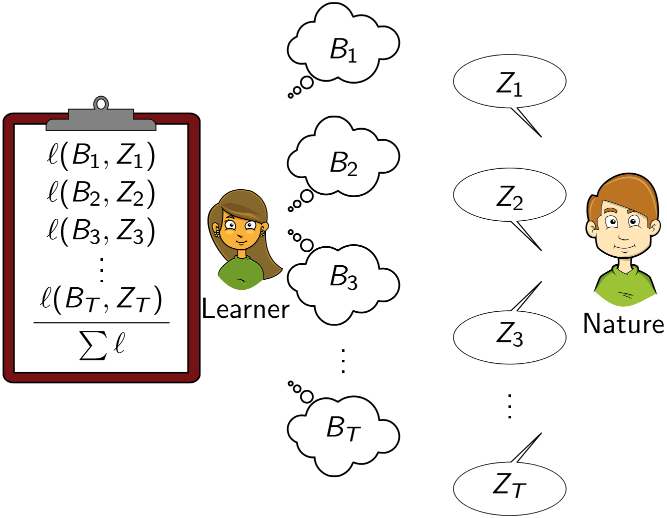
Online Prediction of Stochastic Sequences with High Probability Regret Bounds

Matthias Frey, Jonathan H. Manton, and Jingge Zhu

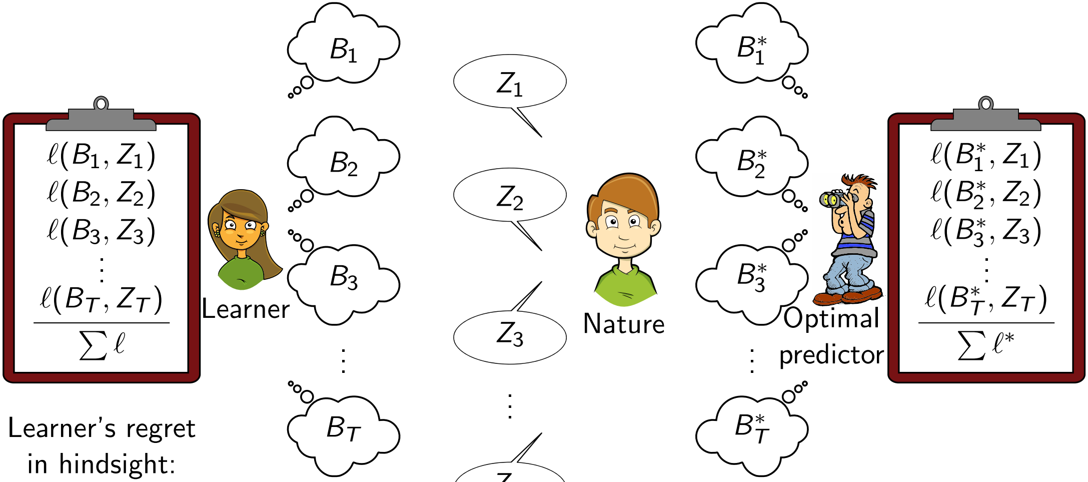
Department of Electrical and Electronic Engineering, The University of Melbourne

The Fourteenth International Conference on Learning Representations (ICLR 2026)

Problem Setup



Problem Setup



Learner's regret in hindsight:
$\Delta_T := \sum \ell - \sum \ell^*$

Applications

- Prediction of wireless channel realizations
- Trajectory predictions in autonomous driving
- Investment decisions
- Recommendation systems

Classical Results (1)

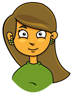
Optimal predictor:
$b_t^*(z_1, \ldots, z_{t-1}) \in \arg\min_{b \in \mathcal{B}} \mathbb{E}_P\left(\ell(b, Z_t) | Z_1 = z_1, \ldots, Z_{t-1} = z_{t-1}\right)$

Learner's prediction policy:
$b_t(z_1, \ldots, z_{t-1}) \in \arg\min_{b \in \mathcal{B}} \mathbb{E}_Q\left(\ell(b, Z_t) | Z_1 = z_1, \ldots, Z_{t-1} = z_{t-1}\right)$

Performance measure: $\Delta := \frac{1}{T} \sum_{t=1}^{T} \left(\ell\left(b(Z_1, \ldots, Z_{t-1}), Z_t\right) - \ell\left(b^*(Z_1, \ldots, Z_{t-1}), Z_t\right) \right)$

Classical Results (1)

Optimal predictor:
$b_t^*(z_1,\dots,z_{t-1}) \in \arg\min_{b\in\mathcal{B}} \mathbb{E}_P\left(\ell(b,Z_t)|Z_1=z_1,\dots,Z_{t-1}=z_{t-1}\right)$



Learner's prediction policy:
$b_t(z_1,\dots,z_{t-1}) \in \arg\min_{b\in\mathcal{B}} \mathbb{E}_Q\left(\ell(b,Z_t)|Z_1=z_1,\dots,Z_{t-1}=z_{t-1}\right)$

Performance measure: $\Delta := \frac{1}{T}\sum_{t=1}^{T}\Big(\ell\big(b(Z_1,\dots,Z_{t-1}),Z_t\big) - \ell\big(b^*(Z_1,\dots,Z_{t-1}),Z_t\big)\Big)$

Merhav-Feder 1998

If ℓ bounded in $[0,L]$: $\mathbb{E}_P\Delta \le L\sqrt{\frac{D(P||Q)}{2T}}$.

Classical Results (2)

Merhav-Feder 1998

If ℓ bounded in $[0, L]$: $\mathbb{E}_P \Delta \leq L\sqrt{\frac{D(P||Q)}{2T}}$.

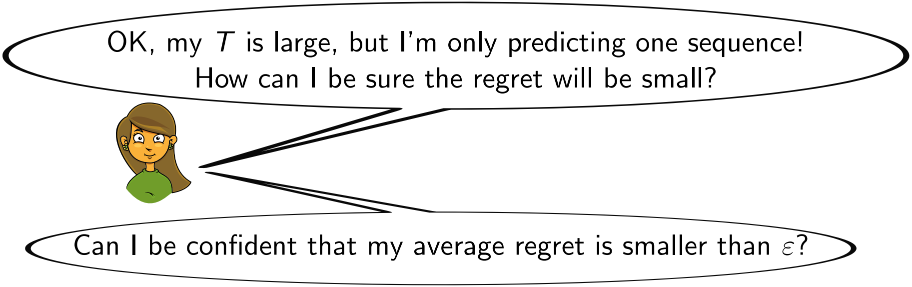
Bound for divergence term

Assume $P = P_{\theta_0}$ where $\theta_0 \in \Theta$ with parametrized family $(P_\theta)_{\theta\in\Theta}$; $Q := \int P_\theta w(d\theta)$.

- If $\mathcal{Z}$ countable and $w(\theta) > 0$ for all $\theta \in \Theta$: $D\left(P||Q\right) \leq -\log w(\theta_0)$ (Hutter 2003)
- If θ_0 in support of w; P i.i.d.: $D\left(P||Q\right) = \mathcal{O}(\log T)$ (Clarke & Barron 1990)
- If θ_0 in supp. of w; P Markov w/ memory: $D\left(P||Q\right) = \mathcal{O}(\log T)$ (Atteson 1999)

What About High Probability?

So we know: $\mathbb{E}_P \Delta \to 0$ as $T \to \infty$.



Δ can be negative, so a bound in expectation does not guarantee that the average regret will vanish during a single run!

Can we get bounds of the form $P(\Delta \geq \varepsilon) \leq \delta$?

Theoretical Results

High-probability regret bound

$$P\left(\Delta \geq 2L\sqrt{\frac{D(P||Q)}{T}} \cdot \frac{1}{\sqrt{\delta}} + \frac{2\sqrt{2}L}{\sqrt{T}}\sqrt{\log \frac{2}{\delta}}\right) < \delta$$

Theoretical Results

High-probability regret bound

$$P\left(\Delta \geq 2L\sqrt{\frac{D(P||Q)}{T}} \cdot \frac{1}{\sqrt{\delta}} + \frac{2\sqrt{2}L}{\sqrt{T}}\sqrt{\log\frac{2}{\delta}}\right) < \delta$$

Impossibility result

Let $C \in (0,\infty)$, $\beta \in [0,1/2)$, $\varepsilon : \mathbb{N} \times (0,1] \to [0,\infty)$. If $\forall \delta \in (0,1]$: $\varepsilon(T,\delta) \xrightarrow{T\to\infty} 0$, then there is a prediction problem, a Q, and a $\delta \in (0,1)$ with

$$P\left(\frac{\Delta_T}{T} \geq C \cdot \sqrt{\frac{D(P||Q)}{T}} \cdot \frac{1}{\delta^{\beta}} + \varepsilon(T,\delta)\right) \geq \delta.$$

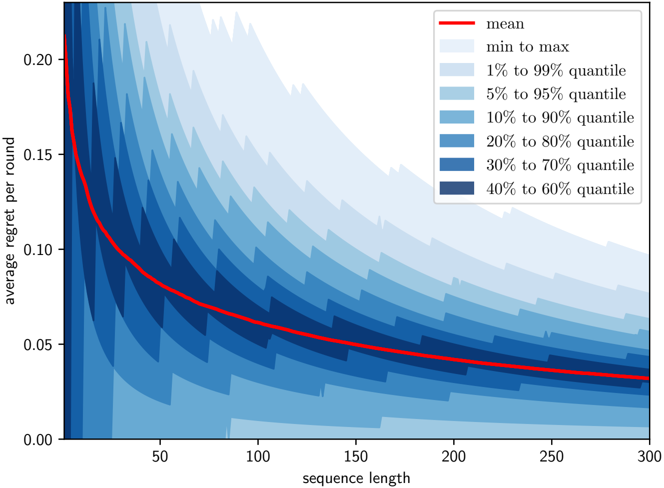
Convergence Rates

Assume $P = P_{\theta_0}$ where $\theta_0 \in \Theta$ with parametrized family $(P_\theta)_{\theta \in \Theta}$; $Q := \int P_\theta w(d\theta)$.

	In expectation (Merhav & Feder 1998)	With probability $\geq 1 - \delta$ (our work)
$\mathcal{Z}$ countable (Hutter 2003)	$\mathcal{O}(T^{-1/2})$	$\mathcal{O}(T)^{-1/2}\delta^{-1/2})$
$P_{\mathbb{N}}$ i.i.d. (Clarke & Barron 1990)	$\mathcal{O}((T/\log T)^{-1/2})$	$\mathcal{O}((T/\log T)^{-1/2}\delta^{-1/2})$
$P_{\mathbb{N}}$ Markov with memory (Atteson 1998)	$\mathcal{O}((T/\log T)^{-1/2})$	$\mathcal{O}(T/\log T)^{-1/2}\delta^{-1/2})$

Numerical Results

ℓ: Classification Loss, $Z_1, \ldots, Z_T$: Markov chain with 2 states and memory 3



Conclusion and Future Directions

- High-probability question conclusively answered for "standard" set of assumptions
- Lemma seems to suggest that under suitable additional assumptions, it may be possible to get $\sqrt{\log \frac{1}{\delta}}$ dependence?
- Practical feasibility/tractability of analyzed algorithm