Online Multicalibration and No-Regret Learning

Georgy Noarov

March 17, 2022

Penn
University of Pennsylvania

Agenda

At a Glance

- Online Multivalid Learning: Means, Moments and Prediction Intervals (joint with Varun Gupta, Chris Jung, Mallesh Pai and Aaron Roth) (`https://arxiv.org/abs/2101.01739`)

Our Results $\in$ Uncertainty Quantification $\cap$ Fair ML

- We show how to obtain uncertainty guarantees in the contextual online adversarial setting...
- With respect to arbitrary collections of “population groups” (= subsets of context space)

$$\text{Our Technique} = \overbrace{\text{Simple}}^{\text{Minimax theorem is the main ingredient}} \;\&\; \underbrace{\text{General}}_{\text{Yields many flavors of uncertainty estimates}} \text{ Game-Theoretic Framework}$$

On Accuracy and Uncertainty

- Traditionally: "Model gets it right most of the time"
- Recent focus: "Model knows when it does not know"

On Uncertainties and Subpopulations

Given your features x, our model predicts your expected disease severity in two days time is $f(x)$.

How sure are you?

I have a 95% prediction interval that your severity will be in $[\ell(x), u(x)]$.

Hmmm...

On Uncertainties and Subpopulations

$[\ell(x), u(x)]$ is a 95% marginal prediction interval.

But I'm part of a demographic group representing less than 5% of the population...

On Uncertainties and Subpopulations

What about for people like me?

For African Americans under the age of 50,
the 95% prediction interval is [a,b]

For women with a family history of diabetes,
the 95% prediction interval is [c,d]

What does this mean for me?

For people with egg allergies and no history of smoking,
the 95% prediction interval is [e, f]

How Strong Can Uncertainty Guarantees Be?

- Ideally: Conditional guarantees
  - $f(x) = \mathbb{E}[y|x]$
  - $g(x) = \mathbb{E}[(y - \mathbb{E}[y|x])^2|x]$
  - $\Pr_y[y \in [\ell(x), u(x)]|x] = 0.95$
- Hardly possible in rich feature spaces: any given x probably seen at most once (and probably, never)
  - A “statistical” way out: make a strong parametric assumption, such as $\mathbb{E}[y|x] = \langle \theta, x \rangle$, and estimate uncertainty of θ
- More realistically: Marginal guarantees
  - Calibration: $f(x) = \mathbb{E}[y|f(x)]$
  - Marginal Moment: $\mathbb{E}_{(x,y)}[(y - f(x))^2]$
  - Marginal Coverage: $\Pr_{(x,y)}[y \in [\ell(x), u(x)]] = 0.95$

Distributional Assumptions

- Standard assumption in Conformal Prediction theory:
  - Exchangeability — future looks like past
  - Only slightly weaker than iid.
- Such an assumption is often unrealistic:
  - Time series data — disease severity can change over time, future depends on past
  - Covariate shift — as disease moves through population, demographics can change
  - Label shift — better medicine can change outcome distribution conditional on patient's features
  - Strategic effects — lending, hiring, admissions classifiers may need to be frequently retrained, as candidates manipulate their features to exploit current deployed version

Our Work

- We give stronger-than-marginal guarantees:
  - Marginal guarantees (e.g. prediction interval coverage) that hold *for every group G* (= subset of feature space) in chosen, potentially large/complex, collection $\mathcal{G}$
- We assume nothing about the data:
  - Data arrives online, potentially adversarially selected
  - Allows for correlates, covariate shift, strategic effects, ...

Setting

- Space of contexts $\mathcal{X}$
- A collection of groups $\mathcal{G} \subseteq 2^{\mathcal{X}}$
  - Can be large and overlapping

In rounds $t = 1 \dots T$:

1. Adversary picks a joint distribution over context-label pairs $(x_t, y_t) \in \mathcal{X} \times [0, 1]$
2. Learner observes realized context x_t
3. Learner makes a prediction regarding $y_t | x_t$:
   - Mean $\bar{\mu}_t$ (Our guarantee: Multicalibration)
   - Mean & k^{th} moment: $(\bar{\mu}_t, \bar{m}_t^k)$ (Mean-Moment Multicalibration)
   - Prediction interval: (ℓ_t, u_t) (Prediction Interval Multivalidity)
4. Learner observes realized label y_t

Our Plan Now...

1. Define online multicalibration
2. Cast Learner-Adversary interaction as repeated game
3. Show multicalibration bounds Learner gets by playing well
4. Explicitly show how to *efficiently* solve repeated game

Multicalibration

- At each round t, Learner predicts mean $\bar{\mu}_t$
- Partition $[0,1] = B(1) \cup \ldots \cup B(n)$, where $B(i) = [\frac{i-1}{n}, \frac{i}{n})$
  - Roughly, $\bar{\mu}_t \in B(i) \iff \bar{\mu}_t \approx \frac{2i-1}{2n}$
- "Regular" online calibration: For $i \in [n]$, over rounds where Learner predicted $\approx \frac{i}{n}$, the average true label $\approx \frac{i}{n}$
- Multicalibration [Hebert-Johnson, Kim, Reingold, Rothblum]: For all groups $G \in \mathcal{G}$, be calibrated over rounds where $x_t \in G$

Definition (Online Multicalibration)

- For $G \in \mathcal{G}, i \in [n], s \in [T]$, let

$$V_s^{G,i} = \sum_{t=1}^{s} 1[x_t \in G, \bar{\mu}_t \in B(i)] \cdot (y_t - \bar{\mu}_t)$$

- Learner is (α, n)-*multicalibrated* if

$$\frac{1}{T} \max_{G \in \mathcal{G}, i \in [n]} |V_T^{G,i}| \leq \alpha$$

- What is the structure of this definition?
  - Asks to satisfy $2|G|n$ constraints of the form: $\pm V_T^{G,i} \leq \alpha T$
  - Each constraint is a sum of terms linear in $y_1, \ldots, y_T$
- To make this tractable, there are 3 hurdles to overcome:
  1. Overall objective $\max |V_T^{G,i}|$ analytically inconvenient: non-smooth!
  2. All constraints are sums *over all rounds* $1 \ldots T$ — but Learner needs to make decisions affecting these constraints *at every round*
  3. At every round, Learner has to predict label mean *without any knowledge* of Adversary's conditional distribution $y_t|x_t$ — how is this possible?

Making Multicalibration Tractable

Pick a Surrogate Loss...

- We are unhappy that overall objective
  1. is a *max* of many constraints — hence nonsmooth
  2. takes into account all of the rounds — while Learner's decisions are local to each round
- Let's solve both these issues!
- First, instead of $\max_{G \in \mathcal{G}, i \in [n]} |V_T^{G,i}|$, switch to *surrogate* loss

$$L_T = \sum_{G \in \mathcal{G}, i \in [n]} \exp(\eta V_T^{G,i}) + \exp(-\eta V_T^{G,i})$$

  - Known as *softmax* — smoothly approximates maximum
- Now that our loss is smooth, we can bound its increase at any round t via Taylor ($e^x \leq 1 + x + x^2$ for $|x| \leq \frac{1}{2}$):

$$\Delta_t(y_t, \bar{\mu}_t) = L_t - L_{t-1} \leq \eta(y_t - \bar{\mu}_t) C_{t-1}^{\text{bucket of } \bar{\mu}_t}(x_t) + 2\eta^2 L_{t-1},$$

where for $i \in [n]$, let $C_{t-1}^i = \sum_{G \ni x_t} \exp(\eta V_{t-1}^{G,i}) - \exp(-\eta V_{t-1}^{G,i})$

Making Multicalibration Tractable

Be Greedy...

- Loss increments bounded as:

$$\Delta_t(y_t, \bar{\mu}_t) \leq A \cdot (y_t - \bar{\mu}_t) C_{t-1}^{\text{bucket of } \bar{\mu}_t}(x_t) + B$$

- If at all rounds t, Learner gets prediction $\bar{\mu}_t$ close to true label y_t, then all Δ_t will be small
- If all Δ_t small, can bound L_T by "telescoping"
- Make Learner greedy — at round t, chooses $\bar{\mu}_t$ so as to make Δ_t small *without regard* to future rounds

Let us focus on any round t...

- Imagine Learner-Adversary interaction as 2-player game:
  - Contested quantity is $u(y_t, \bar{\mu}_t) = (y_t - \bar{\mu}_t) C_{t-1}^{\text{bucket of } \bar{\mu}_t}(x_t)$
  - Learner wants it to be small, goes first and plays (distribution over) $\bar{\mu}_t$
  - Adversary wants it to be large, goes second and plays y_t
- Learner seems handicapped here... False impression!

Making Multicalibration Tractable

...And Win the Game!

- **Minimax Theorem of Zero-Sum Games**: For any 2 player zero sum game with finite action sets A_1, A_2 and utility u, order of play does not matter:

$$\min_{q_1 \in \Delta A_1} \max_{a_2 \in A_2} u(q_1, a_2) = \max_{q_2 \in \Delta A_2} \min_{a_1 \in A_1} u(a_1, q_2) = \text{value of game}$$

- To use this result, restrict Learner's action set to be finite: $\{0, \frac{1}{rn}, \ldots, 1\}$ for some r
- Thus despite going first, Learner has a *minimax optimal* strategy, which obtains the best possible bound on Δ_t:
  - As good as if Learner first got to see Adversary's label y_t
- But if Learner knew true label, would make $|y_t - \bar{\mu}_t| \leq \frac{1}{2rn}$
  - Minimax strategy achieves this even though Learner actually goes first!
- Now unwrap: Plug $\frac{1}{2rn}$ back into bound on $\Delta_t \Rightarrow$ bound L_T $\Rightarrow$ bound $\max |V_T^{G,i}| \Rightarrow$ get multicalibration guarantee

Existential Guarantees

Theorem

Given: collection of groups $\mathcal{G}$, n buckets, $\epsilon > 0$.
Can make Learner (α, n)-multicalibrated, where:

- (High-Probability Bound) $\forall \lambda \in (0,1)$,

$$\alpha \leq (4+\epsilon)\sqrt{\frac{2}{T}\ln\left(\frac{2|\mathcal{G}|n}{\lambda}\right)} \text{ with prob. } 1-\lambda;$$

- (In-Expectation Bound)

$$\mathbb{E}[\alpha] \leq (2+\epsilon)\sqrt{\frac{2}{T}\ln(2|\mathcal{G}|n)}.$$

- If every x_t is in $\leq d$ groups, can replace $|\mathcal{G}|$ with d

Algorithm: A Multicalibrated Learner

At each round $t = 1, \ldots, T$:

- For all $i \in [n]$, compute "historical quantities"

$$C_{t-1}^{i}(x_t) = \sum_{G \ni x_t} \exp\left(\eta V_{t-1}^{G,i}\right) - \exp\left(-\eta V_{t-1}^{G,i}\right)$$

- There are 3 cases:
  - $C_{t-1}^{i}(x_t) > 0$ for all $i \in [n]$: Predict $\bar{\mu}_t = 1$ with prob. 1
  - $C_{t-1}^{i}(x_t) < 0$ for all $i \in [n]$: Predict $\bar{\mu}_t = 0$ with prob. 1
  - $C_{t-1}^{j}(x_t) \cdot C_{t-1}^{j+1}(x_t) \leq 0$ for some $j \in [n]$:

$$\text{Predict } \bar{\mu}_t = \begin{cases} \frac{j}{n} - \frac{1}{rn} & \text{with prob. } q_t = \frac{|C_{t-1}^{j+1}(x_t)|}{|C_{t-1}^{j}(x_t)| + |C_{t-1}^{j+1}(x_t)|}, \\ \frac{j}{n} & \text{with remaining prob.} \end{cases}$$

Our Plan Now...

1. Define prediction interval multivalidity
2. Derive existential guarantees — same steps as before
3. Display efficient algorithm that achieves these guarantees

Prediction Interval Multivalidity

- At each round, Learner predicts interval (ℓ_t, u_t)
  - Bucketing: $(\ell, u) \in B(i,j) \iff \ell \approx \frac{i}{n}$ & $u \approx \frac{j}{n}$
- Goal: For $\delta \in (0,1)$, predict (ℓ, u) s.t. $\Pr[y \in (\ell, u)] \approx 1 - \delta$
  - And this should hold conditional on $x_t \in G$ for all $G \in \mathcal{G}$
- For all $G \in \mathcal{G}$, $(i,j) \in [n] \times [n]$, $s \in [T]$, let

$$V_s^{G,(i,j)} = \sum_{t=1}^{s} 1[x_t \in G, (\ell_t, u_t) \in B(i,j)] \cdot \big(\overbrace{1[y_t \in (\ell_t, u_t)]}^{\text{coverage} \approx 1-\delta} - (1-\delta)\big)$$

Definition (Prediction Interval Multivalidity)

Learner's prediction intervals are (α, n)-*multivalid* if

$$\frac{1}{T} \max_{G \in \mathcal{G}, (i,j) \in [n] \times [n]} |V_T^{G,(i,j)}| \leq \alpha$$

- Also assume Adversary ρ-smooth: $\Pr[y_t \in [a,b]$ of len $\leq \frac{1}{rn}] \leq \rho$
  - E.g. if label always 1, any (ℓ, u) has coverage 0/1, not $1-\delta$

Deriving Prediction Interval Guarantees

We know the drill:

- Switch to softmax loss

$$L_T = \sum_{G,(i,j)} \exp(\eta V_T^{G,(i,j)}) + \exp(-\eta V_T^{G,(i,j)})$$

- Bound one-step differences $\Delta_t((\ell_t, u_t), y_t) = L_t - L_{t-1}$ using $e^x \leq 1 + x + x^2$ for $|x| \leq \frac{1}{2}$
- Consider zero-sum game with payoff = (bound on Δ_t), where Learner is the min player
- Assuming Learner plays minimax optimally, get existential bounds via telescoping
  - Consider: If Learner knew true label distribution, easy to build good prediction interval!
- Give efficient alg for finding minimax optimal strategy

Existential Guarantees

Theorem

Given: n buckets, coverage param δ, adversary ρ-smooth.
Can get: (α, n)-multivalid $(1-\delta)\%$ prediction intervals, where:

- (High-Probability Bound) $\forall \lambda \in (0,1)$,

$$\alpha \leq \rho + 4\sqrt{\frac{2}{T} \ln \left(\frac{2|\mathcal{G}|n^2}{\lambda} \right)} \text{ with prob. } 1-\lambda;$$

- (In-Expectation Bound)

$$\mathbb{E}[\alpha] \leq \rho + 2\sqrt{\frac{2\ln(2|\mathcal{G}|n^2)}{T}}.$$

Algorithm: Multivalid Prediction Intervals

At each round $t = 1, \dots, T$:

- For all $(i, j) \in [n] \times [n]$, compute "historical quantities"

$$C_{t-1}^{\ell,u}(x_t) = \sum_{G \ni x_t} \exp\left(\eta V_{t-1}^{G,(i,j)}\right) - \exp\left(-\eta V_{t-1}^{G,(i,j)}\right)$$

- Solve the following LP (using Ellipsoid with sep. oracle) — "Find distr. Q_t^L over intervals that works against any adversarial distr. Q^A over labels":

$$\min_{Q_t^L \in \mathcal{P}_{\text{int}}^{rn}} \gamma \text{ s.t.}$$

$$\forall Q^A \in \hat{\mathcal{Q}}_{\rho,rn} : \sum_{y \in \mathcal{P}^{rn}} Q^A(y) \sum_{(\ell,u) \in \mathcal{P}_{\text{int}}^{rn}} Q_t^L(\ell, u)\, C_{t-1}^{\ell,u}(x_t)\, (1[y \in (\ell, u)] - (1 - \delta)) \leq \gamma,$$

$$\sum_{(\ell,u) \in \mathcal{P}_{\text{int}}^{rn}} Q_t^L(\ell, u) = 1,$$

$$\forall (\ell, u) \in \mathcal{P}_{\text{int}}^{rn} : Q_t^L(\ell, u) \geq 0.$$

Linear Program to compute Learner's round-t minimax strategy.

- Predict interval (ℓ_t, u_t) sampled from Q_t^L.

Conclusion

- We have presented a general technique which obtains *stronger-than-marginal* uncertainty guarantees in an *adversarial* online setting
- Both *in-expectation* and *high-probability* guarantees are available
- Besides providing *existential guarantees*, our algorithms derived in this framework are *efficiently implementable*

Future Directions

- We believe our results may be extended along a few directions:
  - From binary to multilabel settings
  - From prediction intervals to *prediction sets*
- Can the prediction intervals algorithm be implemented even *more* efficiently to make it fully practical?
  - Currently uses Ellipsoid method (efficient but slow!)
- Experimental evaluation of our algorithms
  - Do they empirically perform better on distributional data?

Subsequent Work

- Our game-theoretic framework can be naturally extended to more general online environments with convex-concave vector losses
- See `https://arxiv.org/abs/2108.03837`
- In the same manner as for multicalibration, it lets one easily design efficient algorithms for:
  - Very general no-regret settings (external, internal, sleeping experts, adaptive, multigroup, ...)
  - Blackwell approachability on polytopes
  - ...

No-Regret Learning, Classical Setting

Learner has a finite set of pure actions ("experts") $\mathcal{A}$.
In rounds $t = 1 \ldots T$:

1. Learner picks a distribution x_t over actions $\mathcal{A}$
2. Adversary picks bounded vector of losses for each action: $r^t \in [0, 1]^{\mathcal{A}}$
3. Learner samples action for this round: $a_t \sim x_t$
4. Learner experiences loss $r^t_{a_t}$ for this round, and gets to observe the entire vector of losses r^t.
   - If Learner could only observe $r^t_{a_t}$ (loss of taken action), that would be called bandit feedback

Learner's goal is to ensure that external regret is sublinear in T:

$$\underbrace{\sum_{t=1}^{T} r^t_{a_t}}_{\text{total incurred loss}} - \underbrace{\min_{j \in \mathcal{A}} \sum_{t=1}^{T} r^t_j}_{\text{benchmark}} = o(T) \quad \text{in expectation.}$$

Exponential Weights Update

- Want algorithm giving Learner sublinear external regret:

$$\underbrace{\sum_{t=1}^{T} r^t_{a_t}}_{\text{total incurred loss}} - \underbrace{\min_{j \in \mathcal{A}} \sum_{t=1}^{T} r^t_j}_{\text{benchmark}} = o(T) \quad \text{in expectation}$$

- Take "history" into account
- The more losses action $a \in \mathcal{A}$ accumulates before round t, the less Learner wants to pick a in round t
- Exponential Weights Update with rate $\eta \in (0, 1/2)$:
  At round t, play action $a \in \mathcal{A}$ with prob. proportional to

$$\exp\left(-\eta \sum_{s=1}^{t-1} r^s_a\right)$$

- Theorem: With $\eta \approx \frac{1}{\sqrt{T}}$, EWU has external regret $O(\sqrt{T})$

Other Famous No-Regret Benchmarks

- **Internal Regret:** Family of "substitution" maps $\mathcal{M} \subset \mathcal{A}^{\mathcal{A}}$: identity map $\mu_{\text{id}}(a) = a$, and for each pair of actions $i \neq j$, map $\mu_{i \to j}$ s.t. $\mu_{i \to j}(i) = j$ and $\mu_{i \to j}(a) = a$ for $a \neq i$.

$$\max_{\mu \in \mathcal{M}} \sum_{t=1}^{T} r^t_{a_t} - r^t_{\mu(a_t)}$$

- **Adaptive Regret:** Do (almost) as well as locally best action on each time interval.

$$\max_{1 \leq t_1 \leq t_2 \leq T} \max_{j \in \mathcal{A}} \sum_{t=t_1}^{t_2} r^t_{a_t} - r^t_j$$

- **Sleeping Experts Regret:** At each round t, only a subset $\mathcal{A}_t \subseteq \mathcal{A}$ of actions are "awake" (= available to Learner).

$$\max_{j \in \mathcal{A}} \sum_{t : j \in \mathcal{A}_t} r^t_{a_t} - r^t_j.$$

Deriving Sublinear-Regret for All These Benchmarks and More: General Setting

In rounds $t = 1 \dots T$, Learner accumulates a d-dimensional coordinate-wise bounded ($||\cdot||_\infty \leq C$) loss vector.

1. Before round t, Adversary selects an *environment*:
   1.1 Learner's and Adversary's convex compact action sets $\mathcal{X}_t$, $\mathcal{Y}_t$ embedded into a finite-dimensional Euclidean space;
   1.2 Continuous vector loss $\ell^t(\cdot,\cdot) : \mathcal{X}_t \times \mathcal{Y}_t \to [-C, C]^d$, with convex-concave coordinates $\ell^t_j(\cdot,\cdot) : \mathcal{X}_t \times \mathcal{Y}_t \to [-C, C]$.
2. Learner selects some $x_t \in \mathcal{X}^t$.
3. Adversary observes x_t, and responds with some $y_t \in \mathcal{Y}^t$.
4. Learner suffers (and observes) loss vector $\ell^t(x_t, y_t)$.

Adversary-Moves-First Regret:

$$\max_{j\in[d]} \sum_{t=1}^{T} \ell^t_j(x_t, y_t) - \sum_{t=1}^{T} w^t,$$

where $w^t = \sup_{y_t \in \mathcal{Y}_t} \min_{x_t \in \mathcal{X}_t} \max_{j\in[d]} \ell^t_j(x_t, y_t)$.

Adversary-Moves-First Regret

$$\max_{j\in[d]} \sum_{t=1}^{T} \ell_j^t(x_t, y_t) - \sum_{t=1}^{T} w^t,$$

$$\text{where } w^t = \sup_{y_t\in\mathcal{Y}_t} \min_{x_t\in\mathcal{X}_t} \max_{j\in[d]} \ell_j^t(x_t, y_t).$$

- Encodes that Learner cares about minimizing max coordinate in accumulated (= summed over all rounds) loss vector
- $\sum_t w^t$ is the benchmark: w^t is the best Learner could do in round t if Adversary told his strategy in advance (Adversary-Moves-First value)

Deriving No-AMF-Regret

- AMF regret is equivalently $\max_{j\in[d]} R_j^T$, where $R_j^t = \sum_{s=1}^t \ell_j^s(x_s, y_s) - \sum_{s=1}^t w^s$
- This max is nonsmooth, so instead track softmax loss:

$$L^t = \sum_{j\in[d]} \exp\left(\eta R_j^t\right)$$

- Can show a Taylor bound:

$$L^t \leq \left(4\eta^2 C^2 + 1\right) L^{t-1} + \eta \underbrace{\sum_{j\in[d]} \exp\left(\eta R_j^{t-1}\right) \cdot \left(\ell_j^t\left(x_t, y_t\right) - w^t\right)}_{:=u^t(x_t, y_t)}$$

- Learner should play $x_t \in \arg\min_{x\in\mathcal{X}_t} \max_{y\in\mathcal{Y}_t} u^t(x, y)$
- Via minimax theorem + defn of w^t, turns out: $\min_{x\in\mathcal{X}_t} \max_{y\in\mathcal{Y}_t} u^t(x, y) = 0$
- Hence with optimal play, $L^T \leq \left(4\eta^2 C^2 + 1\right)^T d$ for all t, and can obtain

$$\text{AMF Regret } \leq 4C\sqrt{T \ln d}.$$

Reducing External to AMF Regret

- External Regret: $\sum_{t=1}^T r_{a_t}^t - \min_{j\in\mathcal{A}} \sum_{t=1}^T r_j^t$.
- Equivalently:
$$\max_{j\in\mathcal{A}} \sum_{t=1}^{T} r_{a_t}^t - r_j^t$$
- So define: Learner's space $\mathcal{X}_t =$ distributions over $\mathcal{A}$, Adversary's space $\mathcal{Y}_t =$ vectors $r^t \in [0,1]^{\mathcal{A}}$, loss vector ℓ^t with $d = |\mathcal{A}|$ dims, with coordinate j equal to:
$$\ell_j^t(a, r^t) = r_a^t - r_j^t.$$
- Each Adversary-Moves-First value is:
$w^t = \sup_{r^t\in\mathcal{Y}_t} \min_{a_t\in\mathcal{A}} \max_{j\in[d]} r_a^t - r_j^t = 0$
  - If Learner knew action losses r^t for round t, would just pick $a = \arg\min_j r_j^t$ and get regret 0 in that round

$$\text{AMF Regret } = \max_{j\in\mathcal{A}} \sum_{t=1}^{T} \underbrace{\ell_j^t(x_t, r_t)}_{\mathbb{E}_{a_t\sim x_t}[r_{a_t}^t - r_j^t]} - \underbrace{\sum_{t=1}^{T} w^t}_{=0} \implies \mathbb{E}[\text{External Regret}].$$

What Is This Algorithm?

- We use the AMF minimization algorithmic procedure, on the "external-regret" loss coordinates $\ell_j^t(a, r^t) = r_{a_t}^t - r_j^t$
- Recall that Learner solves for
  $x_t \in \arg\min_{x \in \mathcal{X}_t} \max_{y \in \mathcal{Y}_t} u^t(x, y)$, where
  $u^t(x_t, y_t) = \sum_{j \in [d]} \exp\left(\eta R_j^{t-1}\right) \cdot \left(\ell_j^t(x_t, y_t) - w^t\right)$.
- Recall that
  $R_j^{t-1} = \sum_{s=1}^{t-1} \ell_j^s(x_s, y_s) - \sum_{s=1}^{t} w^s = \sum_{s=1}^{t-1} (r_{a_s}^s - r_j^s)$

$$x_t \in \arg\min_{x \in \Delta\mathcal{A}} \max_{r^t \in [0,1]^{|\mathcal{A}|}} \sum_{j \in \mathcal{A}} \frac{\exp\left(\eta \sum_{s=1}^{t-1} (r_{a_s}^s - r_j^s)\right)}{\sum_{i \in \mathcal{A}} \exp\left(\eta \sum_{s=1}^{t-1} (r_{a^s}^s - r_i^s)\right)} \mathop{\mathbb{E}}_{a \sim x} [r_a^t - r_j^t],$$

$$= \arg\min_{x \in \Delta\mathcal{A}} \max_{r^t \in [0,1]^{|\mathcal{A}|}} \sum_{j \in \mathcal{A}} \frac{\exp\left(-\eta \sum_{s=1}^{t-1} r_j^s\right)}{\sum_{i \in \mathcal{A}} \exp\left(-\eta \sum_{s=1}^{t-1} r_i^s\right)} \mathop{\mathbb{E}}_{a \sim x} [r_a^t - r_j^t],$$

$$= \arg\min_{x \in \Delta\mathcal{A}} \max_{r^t \in [0,1]^{|\mathcal{A}|}} \mathop{\mathbb{E}}_{a \sim x, j \sim \mathrm{EW}_\eta(\pi^{t-1})} [r_a^t - r_j^t],$$

- The unique distribution x_t that makes this minimax objective 0 is the EW distribution!

Summary

- The AMF framework is a very general online learning tool for getting low regret relative to vector objectives
- We derived online multicalibration + prediction intervals
- Next, we discovered how this framework results in another motivation for the Exponential Weights algorithm
  - So now you know that EW is just the result of a greedy Learner playing to minimize the short-term increase in the softmax surrogate loss!
- And in fact we can similarly derive efficient No-X-Regret algorithms for every benchmark X that we know of

Thank you!

Penn
UNIVERSITY of PENNSYLVANIA