1. More contextual bandits.
   - Recall: Bandits with expert advice.
   - Infinite comparison classes.
     - Examples: parameterized policies.
     - Finite approximations: $\epsilon$-covers and Exp4.
     - Constructing $\epsilon$-covers:
       (a) Lipschitz, bounded parameterization.
       (b) $\Pi$ with bounded VC-dimension.
Recall: Contextual bandits

At each round:

- See $X_t \in \mathcal{X}$.
- Choose $I_t \in \mathcal{A}$, $\mathcal{A} = \{1, \ldots, k\}$.
- Receive reward $Y_{I_t,t} \in \mathbb{R}$.

Stochastic/adversarial model for $(X, Y) \in \mathcal{X} \times \mathbb{R}^\mathcal{A}$.

Pseudo-regret:

$$\overline{R}_n = \sup_{\pi \in \Pi} \mathbb{E} \sum_{t=1}^{n} Y_{\pi(X_t),t} - \mathbb{E} \sum_{t=1}^{n} Y_{I_t,t}.$$ 

where $\Pi$ is comparison class of policies $\pi : \mathcal{X} \rightarrow \mathcal{A}$. 

Recall: Bandits with expert advice

Repeated game:

1. Adversary chooses rewards \((y_{1,t}, \ldots, y_{k,t})\).
2. Adversary presents expert advice \(\xi_t^1, \ldots, \xi_t^N \in \Delta_k\).
3. Strategy chooses the distribution of \(I_t\).
4. Strategy receives reward \(y_{I_t,t}\).
Recall: Exp4

Strategy Exp4

set \( q_1 \) uniform on \( \{1, \ldots, N\} \).

for \( t = 1, 2, \ldots, n \), observe \( \xi^1_t, \ldots, \xi^N_t \in \Delta_k \);
choose \( I_t \sim p_t \), where \( p_{i,t} = \mathbb{E}_{J \sim q_t} \xi^J_{i,t} \); observe \( \ell_{I_t,t} \).

\[
\tilde{\ell}_{i,t} = \frac{\ell_{i,t}}{p_{i,t}} 1[I_t = i], \quad \tilde{y}_{j,t} = \mathbb{E}_{I \sim \xi^J_t} \tilde{\ell}_{I,t},
\]

\[
\tilde{Y}_{j,t} = \sum_{s=1}^{t} \tilde{y}_{j,t}, \quad q_{j,t+1} = \frac{\exp(-\eta \tilde{Y}_{j,t})}{\sum_{i=1}^{N} \exp(-\eta \tilde{Y}_{i,t})}.
\]
Recall: Exp4

**Theorem:** Regret of Exp4:

\[ \eta = \sqrt{\frac{2 \log N}{nk}}, \quad \overline{R}_n \leq \sqrt{2nk \log N}. \]

\[ \eta = \sqrt{\frac{\log N}{tk}}, \quad \overline{R}_n \leq 2\sqrt{nk \log N}. \]
More interesting cases allow the comparison class $\Pi$ to be infinite. For instance, for $\mathcal{X} \subseteq \mathbb{R}^d$, we might consider linear threshold functions,

$$\pi(x) = \arg \max_{j \in \{1, \ldots, k\}} x' \theta_j,$$

where $\theta_1, \ldots, \theta_k$ are parameter vectors. Or linear threshold functions defined in terms of features of $x$ and $j \in \mathcal{A}$,

$$\pi(x) = \arg \max_{j \in \mathcal{A}} \phi(x, j)' \theta.$$

Or a probabilistic version, $\pi : \mathcal{X} \rightarrow \Delta_{\mathcal{A}}$,

$$\pi(j|x) = \frac{\exp(\phi(x, j)' \theta)}{\sum_i \exp(\phi(x, i)' \theta)}.$$

(Or decision trees, or ...
Exp4 cannot be applied to an infinite $\Pi$ for computational (can’t maintain the $q_t$ distribution) and statistical ($\log |\Pi| = \infty$) reasons.

But the cardinality of $\Pi$ might not capture its complexity. A smaller class might be essentially the same. Consider the following approach:

1. Construct a finite approximation $\hat{\Pi}$ to $\Pi$.
2. Use Exp4 on $\hat{\Pi}$.
Infinite comparison classes

Consider an i.i.d. stochastic model: \((X_t, Y_t) \sim P\).

Suppose the approximation is such that, for every \(\pi \in \Pi\), there is a \(\hat{\pi} \in \Pi\) with
\[
\Pr (\pi(X_t) \neq \hat{\pi}(X_t)) \leq \epsilon,
\]
then for \(Y \in [0, 1]\),
\[
\mathbb{E} \left| Y_{\pi(X_t), t} - Y_{\hat{\pi}(X_t), t} \right| \leq \epsilon.
\]
Infinite comparison classes

\[ \overline{R}_n(\Pi) = \sup_{\pi \in \Pi} \mathbb{E} \sum_{t=1}^{n} Y_{\pi(X_t),t} - \mathbb{E} \sum_{t=1}^{n} Y_{I_t,t} \]

\[ = \sup_{\pi \in \Pi} \mathbb{E} \sum_{t=1}^{n} Y_{\pi(X_t),t} - \sup_{\hat{\pi} \in \hat{\Pi}} \mathbb{E} \sum_{t=1}^{n} Y_{\hat{\pi}(X_t),t} \]

\[ + \sup_{\hat{\pi} \in \hat{\Pi}} \mathbb{E} \sum_{t=1}^{n} Y_{\hat{\pi}(X_t),t} - \mathbb{E} \sum_{t=1}^{n} Y_{I_t,t} \]

\[ = \sup_{\pi \in \Pi} \inf_{\hat{\pi} \in \hat{\Pi}} \mathbb{E} \sum_{t=1}^{n} (Y_{\pi(X_t),t} - Y_{\hat{\pi}(X_t),t}) \]

\[ + \sup_{\hat{\pi} \in \hat{\Pi}} \mathbb{E} \sum_{t=1}^{n} Y_{\hat{\pi}(X_t),t} - \mathbb{E} \sum_{t=1}^{n} Y_{I_t,t} \]

\[ \leq n\epsilon + \overline{R}_n(\hat{\Pi}). \]
Infinite comparison classes

A set \( \hat{\Pi} \) that can \( \epsilon \)-approximate \( \Pi \) in this way is called an \( \epsilon \)-cover of \( \Pi \) in the pseudometric

\[
\rho(\hat{\pi}, \pi) = \Pr(\pi(X_t) \neq \hat{\pi}(X_t)).
\]

The cardinality of the smallest \( \epsilon \)-cover of \( \Pi \) is called its \( \epsilon \)-covering number, and denoted \( \mathcal{N}_\Pi(\epsilon) \).
Infinite comparison classes

**Theorem:** Under the i.i.d. stochastic model: \((X_t, Y_t) \sim P\), strategy Exp4 on the class \(\hat{\Pi}\), which is a minimal \(\epsilon\)-cover of \(\Pi\), where \(\epsilon\) is chosen to minimize

\[
\epsilon + \sqrt{\frac{2k \log \mathcal{N}_\Pi(\epsilon)}{n}},
\]

gives pseudo-regret

\[
\overline{R}_n \leq n \min_{\epsilon \geq 0} \left( \epsilon + \sqrt{\frac{2k \log \mathcal{N}_\Pi(\epsilon)}{n}} \right).
\]
How could we construct an $\epsilon$-cover $\hat{\Pi}$ of $\Pi$?

If $\Pi$ is a parametric class, $\Pi = \{\pi_\theta : \theta \in \Theta\}$, where, for all $x \in \mathcal{X}$, the map $\theta \rightarrow \pi_\theta(x)$ is a Lipschitz map: $\rho(\pi_\theta, \pi_{\theta'}) \leq c \|\theta - \theta'\|$, and $\Theta$ is compact, then we can construct an $(\epsilon/c)$-cover $\hat{\Theta}$ of $\Theta$, and define

$$\hat{\Pi} = \left\{ \pi_{\hat{\theta}} : \hat{\theta} \in \hat{\Theta} \right\}.$$

(For instance, consider the parameterized class

$$\pi_\theta(j|x) = \frac{\exp(\phi(x,j)'\theta)}{\sum_i \exp(\phi(x,i)'\theta)}$$

with bounded features $\phi$ and bounded parameters $\theta$.)
Another example: Suppose that the shattering coefficient

$$S_{\Pi}(n) := \max_{x_1, \ldots, x_n \in \mathcal{X}} |\{(\pi(x_1), \ldots, \pi(x_n)) : \pi \in \Pi\}|$$

grows slowly with $n$ (much slower than exponential in $n$). Then we can use that to build a small cover.

High level idea:

1. Gather some data $X_1, \ldots, X_m$ (making arbitrary decisions $I_t$),

2. Construct $\hat{\Pi}$ containing one representative for each element of $\{(\pi(X_1), \ldots, \pi(X_m)) : \pi \in \Pi\}$. (So that $|\hat{\Pi}| \leq S_{\Pi}(m)$.)

3. Use Exp4 with $\hat{\Pi}$. 

- - - 

Infinite comparison classes
Infinite comparison classes
**Theorem:** Under the i.i.d. stochastic model: \((X_t, Y_t) \sim P\), with probability \(1 - \delta\), the \(\hat{\Pi}\) constructed in this way is an \(\epsilon\)-cover for \(\Pi\) of size no more than \(S_\Pi(m)\), for

\[
\epsilon = \frac{2}{m} \log_2 \left( \frac{2S_\Pi(2m)^2}{\delta} \right).
\]

Thus, the pseudo-regret of this strategy satisfies

\[
\overline{R}_n \leq m + (n - m)\delta + (n - m)\epsilon + \sqrt{2(n - m)k \log(S_\Pi(m))}.
\]

If \(S_\Pi(m) = O \left( (m/d)^d \right)\), setting \(m = \sqrt{nd \log(n/d)}\) and \(\delta = m/n\) gives

\[
\overline{R}_n = O \left( \sqrt{nk d \log \frac{n}{d}} \right).
\]
Infinite comparison classes

A symmetrization idea due to Vapnik and Chervonenkis, plus a simple counting argument shows that \( \hat{\Pi} \) is an \( \epsilon \)-cover:

**Lemma:** Given i.i.d. data \( D_n = \{X_1, \ldots, X_n\} \), and a set \( \mathcal{E} \) of events in \( \mathcal{X} \),

\[
P^n (\exists E \in \mathcal{E}, \ D \cap E = \emptyset, \ P(E) \geq \epsilon) \leq 2S_{\mathcal{E}}(2n)2^{-\epsilon n/2},
\]

where \( S_{\mathcal{E}}(n) \) is the shattering coefficient of \( \{1_E : E \in \mathcal{E}\} \).

Defining \( \mathcal{E} = \{\{x : \pi(x) = \hat{\pi}(x)\} : (\pi, \hat{\pi}) \in \Pi^2\} \), we have, with probability at least \( 1 - \delta \) over \( D_m \), the initial \( m \)-sample, for every \( \pi \in \Pi \) there is a \( \hat{\pi} \in \hat{\Pi} \) (the one that equals \( \pi \) on \( D_m \)) with

\[
Pr(\pi(X) \neq \hat{\pi}(X)) \leq \epsilon,
\]

that is, \( \hat{\Pi} \) is an \( \epsilon \)-cover for \( \Pi \).
Infinite comparison classes

When does $S_{\Pi}(n)$ grow slowly with $n$?

**Definition:** A class $\Pi \subseteq \{0, 1\}^X$ **shatters** $\{x_1, \ldots, x_d\} \subseteq X$ means that $|\Pi(x_1^d)| = 2^d$.

The Vapnik-Chervonenkis dimension of $\Pi$ is

$$d_{VC}(\Pi) = \max \left\{ d : \text{some } x_1, \ldots, x_d \in X \text{ is shattered by } \Pi \right\}$$

$$= \max \left\{ d : S_{\Pi}(d) = 2^d \right\}.$$
Vapnik-Chervonenkis dimension: “Sauer’s Lemma”

**Theorem:** [Vapnik-Chervonenkis] \( d_{VC}(F) \leq d \) implies

\[
S_\Pi(n) \leq \sum_{i=0}^{d} \binom{n}{i}.
\]

If \( n \geq d \), the latter sum is no more than \( \left( \frac{en}{d} \right)^d \).

So the VC-dimension is a single integer summary of the shatter coefficients: either it is finite, and \( S_\Pi(n) = O(n^d) \), or \( S_\Pi(n) = 2^n \). No other growth is possible.

\[
S_\Pi(n) \begin{cases} 
= 2^n & \text{if } n \leq d, \\
\leq \left( \frac{e}{d} \right)^d n^d & \text{if } n > d.
\end{cases}
\]
Vapnik-Chervonenkis dimension: “Sauer’s Lemma”

Stronger than this: finiteness of the VC-dimension is necessary. If the VC-dimension is infinite, then there are distributions for which competing with \( \Pi \), even in the full information case, is impossible: for every strategy, there is a probability distribution such that with high probability, the regret grows linearly.

(And it’s the same story for \( k \)-valued functions, modulo \( \log k \) factors.)
Consider a parameterized class of $k$-valued functions,

$$\Pi = \{ x \mapsto f(x, \theta) : \theta \in \mathbb{R}^p \},$$

where $f : \mathbb{R}^m \times \mathbb{R}^p \to \{1, \ldots, k\}$.

Suppose that $f$ can be computed using no more than $t$ operations of the following kinds:

1. arithmetic ($+,-,\times,/)$,
2. comparisons ($>,$ $=,$ $<$),
3. output a constant in $\{1, \ldots, k\}$

**Theorem:** $d_{VC}(F) = O(pt \log k)$.

(And a similar story applies, with a worse dependence on $t$, if we include the exponential function in the set of operations.)
Competing with infinite $\Pi \subseteq \{1, \ldots, k\}^X$:

- If we want to compete with an infinite $\Pi$ for all distributions on $\mathcal{X} \times [0, 1]^k$, $S_\Pi(n)$ must have polynomial growth, say $O(n^d)$.

- We can use i.i.d. data to build an $\epsilon$-cover of $\Pi$ of size $O(S_\Pi(n)) = O(n^d)$.

- Running Exp4 with this class of experts gives regret $\overline{R_n} = O(\sqrt{nkd \log n})$.

- The drawback is computational: $S_\Pi(n)$ is polynomial in $n$, but exponential in the dimension $d$. For example, for

$$\pi(x) = \arg \max_{j \in A} \phi(x, j)'\theta,$$
the computation grows exponentially with the number of features.