بسم الله الرحمن الرحیم

# یادگیری برخط

## جلسه پانزدهم: بندیت زمینه‌ای

# مرور بندیت دشمنانه

# بندیت دشمنانه

For rounds $t = 1, 2, \ldots, n$:

Learner selects distribution $P_t \in \mathcal{P}_{k-1}$ and samples $A_t$ from $P_t$.

Learner observes reward $X_t = x_{tA_t}$.

# مسئله بندیت دشمنانه

سود

افق زمانی

$$R_n(\pi, x) = \max_{i \in [k]} \sum_{t=1}^{n} x_{ti} - \mathbb{E}\left[\sum_{t=1}^{n} x_{tA_t}\right]$$

تعداد دسته‌ها

انتخاب ما

$$R_n^*(\pi) = \sup_{x \in [0,1]^{n \times k}} R_n(\pi, x).$$

# الگوریتم EXP3

**Input:** $n$, $k$, $\eta$

Set $\hat{S}_{0i} = 0$ for all $i$

**for** $t = 1, \ldots, n$ **do**

Calculate the sampling distribution $P_t$:

$$P_{ti} = \frac{\exp\left(\eta \hat{S}_{t-1,i}\right)}{\sum_{j=1}^{k} \exp\left(\eta \hat{S}_{t-1,j}\right)}$$

Sample $A_t \sim P_t$ and observe reward $X_t$

Calculate $\hat{S}_{ti}$:

$$\hat{S}_{ti} = \hat{S}_{t-1,i} + \underbrace{1 - \frac{\mathbb{I}\left\{A_t = i\right\}\left(1 - X_t\right)}{P_{ti}}}_{\widehat{X}_{t,i}}$$

**end for**

$$\hat{S}_{ti} = \sum_{s=1}^{t} \hat{X}_{si} \qquad P_{ti} = \frac{\exp\left(\eta \hat{S}_{t-1,i}\right)}{\sum_{j=1}^{k} \exp\left(\eta \hat{S}_{t-1,j}\right)} \qquad W_t = \sum_{j=1}^{k} \exp\left(\eta \hat{S}_{tj}\right)$$

$$\exp(\eta \hat{S}_{ni}) \leq \sum_{j=1}^{k} \exp(\eta \hat{S}_{nj}) = W_n = W_0 \frac{W_1}{W_0} \cdots \frac{W_n}{W_{n-1}} = k \prod_{t=1}^{n} \frac{W_t}{W_{t-1}}.$$

$$\begin{aligned}
\frac{W_t}{W_{t-1}} &= \sum_{j=1}^{k} \frac{\exp(\eta \hat{S}_{t-1,j})}{W_{t-1}} \exp(\eta \hat{X}_{tj}) = \sum_{j=1}^{k} P_{tj} \exp(\eta \hat{X}_{tj}). \\
&\leq 1 + \eta \sum_{j=1}^{k} P_{tj} \hat{X}_{tj} + \eta^2 \sum_{j=1}^{k} P_{tj} \hat{X}_{tj}^2 \\
&\leq \exp\left(\eta \sum_{j=1}^{k} P_{tj} \hat{X}_{tj} + \eta^2 \sum_{j=1}^{k} P_{tj} \hat{X}_{tj}^2\right)
\end{aligned}$$

$$\exp\left(\eta \hat{S}_{ni}\right) \leq k \exp\left(\eta \hat{S}_n + \eta^2 \sum_{t=1}^{n} \sum_{j=1}^{k} P_{tj} \hat{X}_{tj}^2\right)$$

# تحلیل الگوریتم EXP3

سود دسته i

سود ما

$$\hat{S}_{ni} - \hat{S}_n \leq \frac{\log(k)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{j=1}^{k} P_{tj} (\hat{X}_{tj} - 1)^2$$

# تحلیل الگوریتم EXP3

سود دسته i

سود ما

$$\hat{S}_{ni} - \hat{S}_n \leq \frac{\log(k)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{n}\sum_{j=1}^{k} P_{tj}(\hat{X}_{tj} - 1)^2$$

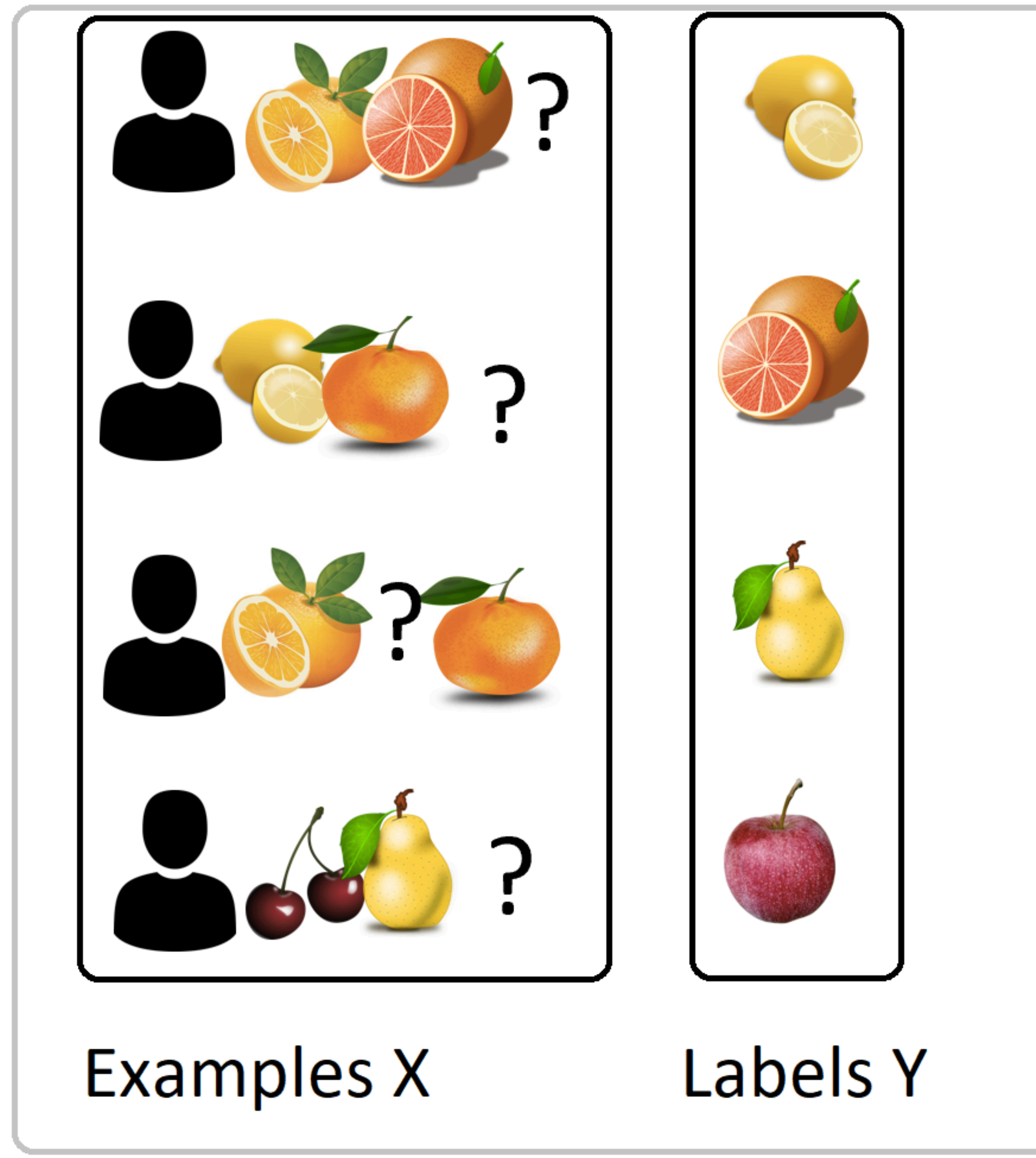
$$R_n(\pi, x) \leq 2\sqrt{nk\log(k)}.$$

$$
\begin{aligned}
\mathbb{E}\left[\sum_{j=1}^{k} P_{tj} \hat{X}_{tj}^2\right] &= \mathbb{E}\left[\sum_{j=1}^{k} P_{tj}\left(1 - \frac{\mathbb{I}\left\{A_t = j\right\} y_{tj}}{P_{tj}}\right)^2\right] \\
&= \mathbb{E}\left[\sum_{j=1}^{k} P_{tj}\left(1 - 2\,\frac{\mathbb{I}\left\{A_t = j\right\} y_{tj}}{P_{tj}} + \frac{\mathbb{I}\left\{A_t = j\right\} y_{tj}^2}{P_{tj}^2}\right)\right] \\
&= \mathbb{E}\left[1 - 2Y_t + \mathbb{E}_{t-1}\left[\sum_{j=1}^{k} \frac{\mathbb{I}\left\{A_t = j\right\} y_{tj}^2}{P_{tj}}\right]\right] \\
&= \mathbb{E}\left[1 - 2Y_t + \sum_{j=1}^{k} y_{tj}^2\right] \\
&= \mathbb{E}\left[(1 - Y_t)^2 + \sum_{j \neq A_t} y_{tj}^2\right] \\
&\leq k.
\end{aligned}
$$

# بندیت زمینه‌ای

## تعریف بندیت زمینه‌ای

## دشمنانه



# بندیت زمینه‌ای: انگیزش

- اطلاعات زمینه
  - مثال: پیشنهاد فیلم

# تعریف بندیت دشمنانه

For rounds $t = 1, 2, \ldots, n$:

Learner selects distribution $P_t \in \mathcal{P}_{k-1}$ and samples $A_t$ from $P_t$.

Learner observes reward $X_t = x_{tA_t}$.

# تعریف بندیت دشمنانه

Adversary secretly chooses rewards $(x_t)_{t=1}^n$ with $x_t \in [0,1]^k$

Adversary secretly chooses contexts $(c_t)_{t=1}^n$ with $c_t \in \mathcal{C}$

For rounds $t = 1, 2, \ldots, n$:

Learner selects distribution $P_t \in \mathcal{P}_{k-1}$ and samples $A_t$ from $P_t$.

Learner observes reward $X_t = x_{tA_t}$.

# تعریف بندیت دشمنانه

Adversary secretly chooses rewards $(x_t)_{t=1}^n$ with $x_t \in [0,1]^k$

Adversary secretly chooses contexts $(c_t)_{t=1}^n$ with $c_t \in \mathcal{C}$

For rounds $t = 1, 2, \ldots, n$:

- Learner observes context $c_t \in \mathcal{C}$ where $\mathcal{C}$ is an arbitrary fixed set of contexts.
- Learner selects distribution $P_t \in \mathcal{P}_{k-1}$ and samples $A_t$ from $P_t$.
- Learner observes reward $X_t = x_{tA_t}$.

# تعریف بندیت زمینه‌ای

Adversary secretly chooses rewards $(x_t)_{t=1}^n$ with $x_t \in [0,1]^k$

Adversary secretly chooses contexts $(c_t)_{t=1}^n$ with $c_t \in \mathcal{C}$

For rounds $t = 1, 2, \dots, n$:

Learner observes context $c_t \in \mathcal{C}$ where $\mathcal{C}$ is an arbitrary fixed set of contexts.

Learner selects distribution $P_t \in \mathcal{P}_{k-1}$ and samples $A_t$ from $P_t$.

Learner observes reward $X_t = x_{tA_t}$.

# چگونه از زمینه استفاده کنیم؟

- بدون زمینه:

$$R_n(\pi, x) \leq 2\sqrt{nk\log(k)}.$$

# چگونه از زمینه استفاده کنیم؟

- بدون زمینه:

$$R_n(\pi, x) \leq 2\sqrt{nk\log(k)}.$$

زمینه‌ای چه فایده‌ای دارد؟

# روش ۱ استفاده از زمینه: برخورد مستقل با هر زمینه

$$= \mathbb{E}\left[\sum_{c\in\mathcal{C}} \max_{i\in[k]} \sum_{t\in[n]:c_t=c} (x_{ti} - X_t)\right]$$

$$R_n(\pi, x, c) := \sum_{c\in\mathscr{C}} \max_{i\in[k]} \sum_{t:c_t=c} x_{ti} - \mathbb{E}\left[\sum_{t=1}^{n} X_t\right]$$

$$R_n^*(\pi) = \sup_{x\in[0,1]^{n\times k}, c\in\mathbb{C}} R_n(\pi, x, c)$$

# روش ۱: یک EXP3 برای هر زمینه

هر زمینه،
$n/|C|$ بار

# روش ۱: یک EXP3 برای هر زمینه

$$R_{nc} \leq 2\sqrt{k \sum_{t=1}^{n} \mathbb{I}\{c_t = c\} \log(k)},$$

هر زمینه،
$n/|C|$ بار

# روش ۱: یک EXP3 برای هر زمینه

$$R_{nc} \leq 2\sqrt{k\sum_{t=1}^{n}\mathbb{I}\left\{c_t = c\right\}\log(k)},$$

$$R_n = \sum_{c\in\mathcal{C}} R_{nc} \leq 2\sum_{c\in\mathcal{C}}\sqrt{k\log(k)\sum_{t=1}^{n}\mathbb{I}\left\{c_t = c\right\}}.$$

هر زمینه،
$n/|C|$ بار

# روش ۱: یک EXP3 برای هر زمینه

$$R_{nc} \leq 2\sqrt{k \sum_{t=1}^{n} \mathbb{I}\{c_t = c\} \log(k)},$$

$$R_n = \sum_{c \in \mathcal{C}} R_{nc} \leq 2 \sum_{c \in \mathcal{C}} \sqrt{k \log(k) \sum_{t=1}^{n} \mathbb{I}\{c_t = c\}}.$$

$$R_n \leq 2\sqrt{nk|\mathcal{C}| \log(k)}.$$

هر زمینه،
$n/|C|$ بار

# روش ۱: یک EXP3 برای هر زمینه

$$R_{nc} \leq 2\sqrt{k \sum_{t=1}^{n} \mathbb{I}\{c_t = c\} \log(k)},$$

$$R_n = \sum_{c \in \mathcal{C}} R_{nc} \leq 2 \sum_{c \in \mathcal{C}} \sqrt{k \log(k) \sum_{t=1}^{n} \mathbb{I}\{c_t = c\}}.$$

$$R_n \leq 2\sqrt{nk|\mathcal{C}|\log(k)}.$$

هر زمینه،
بار $n/|C|$

# چرا این مدل‌سازی خوب است؟

$$\mathbb{E}\left[\sum_{t=1}^{n} X_t\right] \geq \max_{i\in[k]} \sum_{t=1}^{n} x_{ti} - 2\sqrt{kn\log(k)}.$$

$$\mathbb{E}\left[\sum_{t=1}^{n} X_t\right] \geq \sum_{c\in\mathcal{C}} \max_{i\in[k]} \sum_{t\in[n]:c_t=c} x_{ti} - 2\sqrt{kn|\mathcal{C}|\log(k)}.$$

بسم الله الرحمن الرحیم

# یادگیری برخط

## جلسه شانزدهم: بندیت زمینه‌ای

# بندیت زمینه‌ای

## توابع محدودتر

# کمی واقعی‌تر!

$$R_n = \mathbb{E}\left[\sum_{c \in \mathcal{C}} \max_{i \in [k]} \sum_{t \in [n] : c_t = c} (x_{ti} - X_t)\right]$$

# کمی واقعی‌تر!

$$R_n = \mathbb{E}\left[\sum_{c\in\mathcal{C}} \max_{i\in[k]} \sum_{t\in[n]:c_t=c} (x_{ti} - X_t)\right]$$

$$\max_{\phi:\mathscr{C}\to[k]} \sum_{c\in\mathscr{C}} \sum_{t:c_t=c} x_{t,\phi(c_t)} - X_t$$

# کمی واقعی‌تر!

$$R_n = \mathbb{E}\left[\sum_{c\in\mathcal{C}} \max_{i\in[k]} \sum_{t\in[n]:c_t=c} (x_{ti} - X_t)\right]$$

$$\max_{\phi:\mathscr{C}\to[k]} \sum_{c\in\mathscr{C}} \sum_{t:c_t=c} x_{t,\phi(c_t)} - X_t$$

$$R_n = \mathbb{E}\left[\max_{\phi\in\Phi} \sum_{t=1}^{n} (x_{t\phi(c_t)} - X_t)\right]$$

# نمونه‌هایی از $\Phi$

$$\frac{1}{|\mathcal{C}|^2} \sum_{c,d \in \mathcal{C}} (1 - s(c,d)) \mathbb{I}\{\phi(c) \neq \phi(d)\}$$

# نمونه‌هایی از $\Phi$

- همه توابع (∘

$$\frac{1}{|\mathcal{C}|^2}\sum_{c,d\in\mathcal{C}}(1-s(c,d))\mathbb{I}\{\phi(c)\neq\phi(d)\}$$

# نمونه‌هایی از $\Phi$

- ۰) همه توابع
- ۱) به ازای یک افراز از C، توابعی که به هر زیرمجموعه از افراز یک عدد ثابت نسبت می‌دهند.

$$\frac{1}{|\mathcal{C}|^2} \sum_{c,d \in \mathcal{C}} (1 - s(c,d)) \mathbb{I}\{\phi(c) \neq \phi(d)\}$$

# نمونه‌هایی از $\Phi$

- ۰) همه توابع
- ۱) به ازای یک افراز از C، توابعی که به هر زیرمجموعه از افراز یک عدد ثابت نسبت می‌دهند.
- ۲) توابع با تغییرات کوچک

$$\frac{1}{|\mathcal{C}|^2}\sum_{c,d\in\mathcal{C}}(1-s(c,d))\mathbb{I}\{\phi(c)\neq\phi(d)\}$$

# نمونه‌هایی از $\Phi$

- ۰) همه توابع

- ۱) به ازای یک افراز از C، توابعی که به هر زیرمجموعه از افراز یک عدد ثابت نسبت می‌دهند.

- ۲) توابع با تغییرات کوچک

$$\frac{1}{|\mathcal{C}|^2}\sum_{c,d\in\mathcal{C}}(1-s(c,d))\mathbb{I}\{\phi(c)\neq\phi(d)\}$$

- ۳) توابع نامزد (مثلا: یادگیری غیربرخط)

$$\phi_1,\ldots,\phi_M:\mathcal{C}\to[k].$$

# نمونه‌هایی از $\Phi$

- ۰) همه توابع
- ۱) به ازای یک افراز از C، توابعی که به هر زیرمجموعه از افراز یک عدد ثابت نسبت می‌دهند.
- ۲) توابع با تغییرات کوچک

$$\frac{1}{|\mathcal{C}|^2}\sum_{c,d\in\mathcal{C}}(1-s(c,d))\mathbb{I}\{\phi(c)\neq\phi(d)\}$$

- ۳) توابع نامزد (مثلا: یادگیری غیربرخط)

$$\phi_1,\ldots,\phi_M:\mathcal{C}\to[k].$$

# بندیت زمینه‌ای

## راهنمایی متخصصین

$$\phi_1, \ldots, \phi_M : \mathcal{C} \to \mathbb{P}_k$$

$$R_n = \mathbb{E}\left[\max_{\phi \in \Phi} \sum_t \left( \sum_{i=1}^{k} \phi(c_t)_i \cdot x_{t,i} - X_t \right)\right]$$

$$R_n = \mathbb{E}\left[\max_{\phi\in\Phi}\sum_{t=1}^{n}(x_{t\phi(c_t)} - X_t)\right]$$

$$\phi_1, \ldots, \phi_M : \mathcal{C} \to \mathbb{P}_k$$

$$R_n = \mathbb{E}\left[\max_{\phi\in\Phi}\sum_{t}\left(\sum_{i=1}^{k}\phi(c_t)_i \cdot x_{t,i} - X_t\right)\right]$$

$$\phi_1, \ldots, \phi_M : \mathcal{C} \to [k].$$

$$R_n = \mathbb{E}\left[\max_{\phi \in \Phi} \sum_{t=1}^{n} (x_{t\phi(c_t)} - X_t)\right]$$

$$\phi_1, \ldots, \phi_M : \mathcal{C} \to \mathbb{P}_k$$

$$R_n = \mathbb{E}\left[\max_{\phi \in \Phi} \sum_{t} \left( \sum_{i=1}^{k} \phi(c_t)_i \cdot x_{t,i} - X_t \right)\right]$$

# صورت مسئله «بندیت با راهنمایی متخصصین»

Adversary secretly chooses rewards $x \in [0,1]^{n \times k}$

Experts secretly choose predictions $E^{(1)}, \dots, E^{(n)}$

For rounds $t = 1, 2, \dots, n$:

Learner observes predictions of all experts, $E^{(t)} \in [0,1]^{M \times k}$.

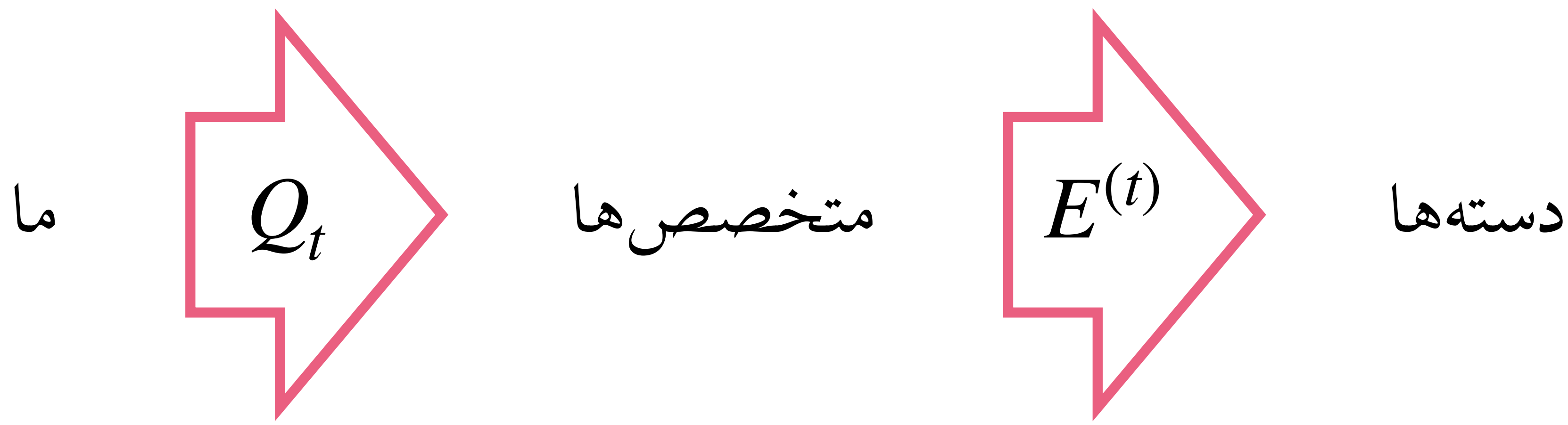
Learner selects a distribution $P_t \in \mathcal{P}_{k-1}$.

Action $A_t$ is sampled from $P_t$ and the reward is $X_t = x_{tA_t}$.

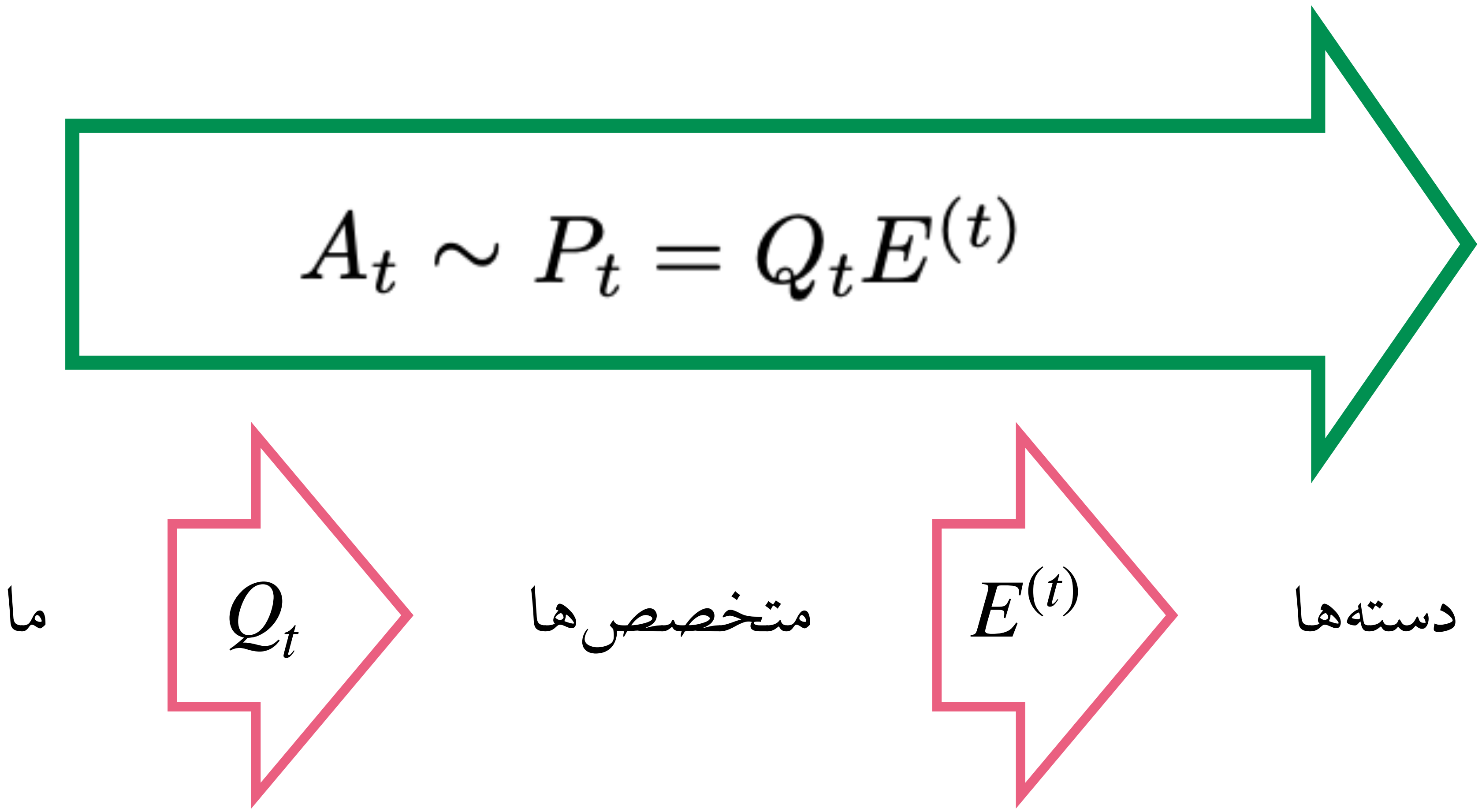
$$R_n = \mathbb{E}\left[\max_{m \in [M]} \sum_{t=1}^{n} E_m^{(t)} x_t - \sum_{t=1}^{n} X_t\right]$$

هدف

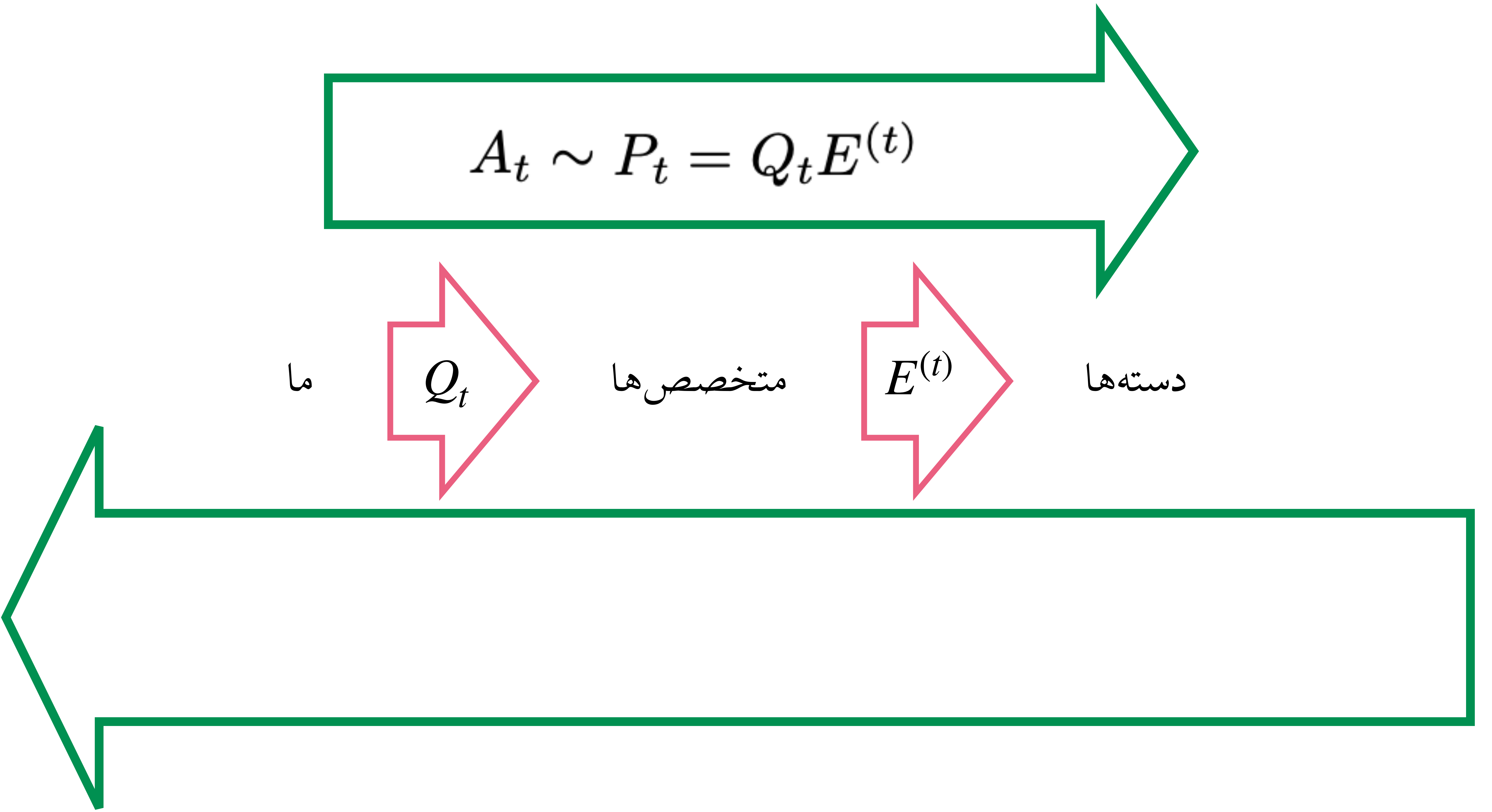
# ایده



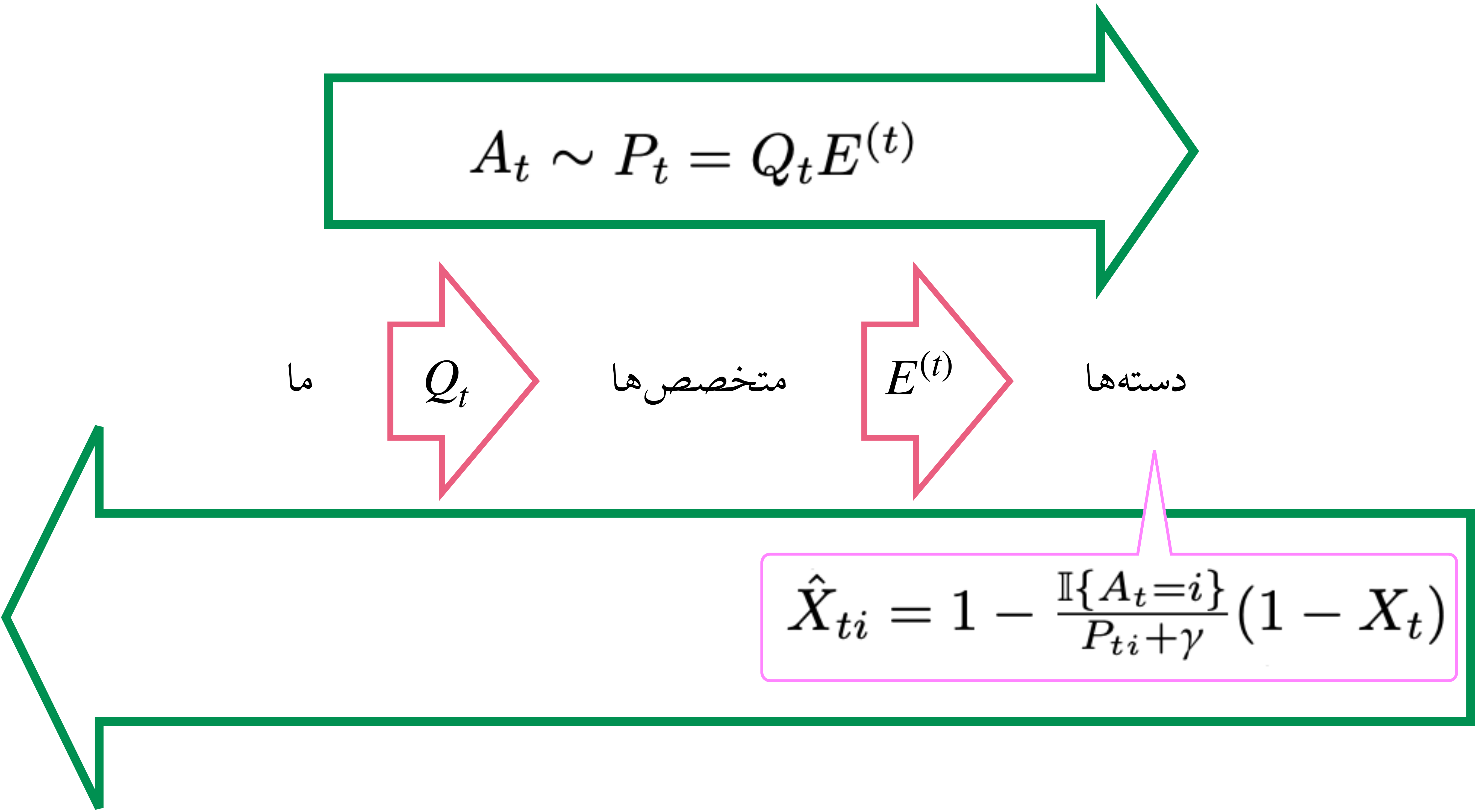
# ایده



# ایده



# ایده



$A_t \sim P_t = Q_t E^{(t)}$
دسته‌ها
$E^{(t)}$
متخصص‌ها
$Q_t$
ما
$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$

# ایده

$$A_t \sim P_t = Q_t E^{(t)}$$

ما $Q_t$ متخصص‌ها $E^{(t)}$ دسته‌ها

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

# ایده

$$A_t \sim P_t = Q_t E^{(t)}$$

ما $\quad Q_t \quad$ متخصص‌ها $\quad E^{(t)} \quad$ دسته‌ها

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

طبق تحلیل الگوریتم EXP3 داریم:

با احتمال $E^{(t)}$ $Q_t$ انتخاب دسته‌ها

**Lemma 18.2.** *For any $m^* \in [M]$, it holds that*

$$\sum_{t=1}^{n} \tilde{X}_{tm^*} - \sum_{t=1}^{n} \sum_{m=1}^{M} Q_{tm} \tilde{X}_{tm} \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} Q_{tm} (1 - \tilde{X}_{tm})^2.$$

طبق تحلیل الگوریتم
EXP3 داریم:

با احتمال $Q_t$ $E^{(t)}$ انتخاب دسته‌ها

**Lemma 18.2.** *For any $m^* \in [M]$, it holds that*

$$\sum_{t=1}^{n} \tilde{X}_{tm^*} - \sum_{t=1}^{n}\sum_{m=1}^{M} Q_{tm}\tilde{X}_{tm} \leq \frac{\log(M)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{n}\sum_{m=1}^{M} Q_{tm}(1-\tilde{X}_{tm})^2.$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{n}\sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1-\tilde{X}_{tm})^2\right].$$

$y_{ti} = 1 - x_{ti}$ دسته‌ها $\xleftarrow{E^{(t)}}$ متخصص‌ها $\xleftarrow{Q_t}$ ما

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\begin{aligned} & E^{(t)}(1 - \widehat{X}_t) \\ & = 1 - E^{(t)} \widehat{X}_t \\ & = 1 - \widetilde{X}_t \end{aligned}$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm}$$

$$\hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t$$

$$\hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\quad Q_t \quad$ متخصص‌ها $\quad E^{(t)} \quad$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\begin{aligned} E^{(t)}(1 - \widehat{X}_t) &= 1 - E^{(t)} \widehat{X}_t \\ &= 1 - \widetilde{X}_t \end{aligned}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

---

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها

$$y_{ti} = 1 - x_{ti}$$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\begin{aligned} E^{(t)}(1 - \widehat{X}_t) \\ = 1 - E^{(t)} \widehat{X}_t \\ = 1 - \widetilde{X}_t \end{aligned}$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm}$$

$$\hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t$$

$$\hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

---

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E^{(t)}_{mA_t} y_{tA_t}}{P_{tA_t}}\right)^2\right]$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right) \qquad \tilde{X}_t = E^{(t)} \hat{X}_t \qquad \hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$E^{(t)}(1 - \widehat{X}_t) = 1 - E^{(t)} \widehat{X}_t = 1 - \widetilde{X}_t$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

---

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E^{(t)}_{mA_t} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E^{(t)}_{mi} y_{ti}\right)^2}{P_{ti}}$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $y_{ti} = 1 - x_{ti}$

$\exp\left(\eta \widetilde{X}_{t,m}\right)$

$\tilde{X}_t = E^{(t)} \hat{X}_t$

$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$

$$E^{(t)}(1 - \widehat{X}_t) = 1 - E^{(t)} \widehat{X}_t = 1 - \widetilde{X}_t$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

---

$$\mathbb{E}_t[\tilde{Y}^2_{tm}] = \mathbb{E}_t\left[\left(\frac{E^{(t)}_{mA_t} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E^{(t)}_{mi} y_{ti}\right)^2}{P_{ti}} \le \sum_{i=1}^{k} \frac{E^{(t)}_{mi}}{P_{ti}}.$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\begin{aligned} E^{(t)}(1 - \widehat{X}_t) &= 1 - E^{(t)} \widehat{X}_t \\ &= 1 - \widetilde{X}_t \end{aligned}$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm}$$

$$\hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t$$

$$\hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

---

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E^{(t)}_{mA_t} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E^{(t)}_{mi} y_{ti}\right)^2}{P_{ti}} \leq \sum_{i=1}^{k} \frac{E^{(t)}_{mi}}{P_{ti}}.$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq \mathbb{E}\left[\sum_{m=1}^{M} Q_{tm} \sum_{i=1}^{k} \frac{E^{(t)}_{mi}}{P_{ti}}\right]$$

ما $\xleftarrow{Q_t}$ متخصص‌ها $\xleftarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$E^{(t)}(1 - \widehat{X}_t) = 1 - E^{(t)} \widehat{X}_t = 1 - \widetilde{X}_t$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

---

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E^{(t)}_{mA_t} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E^{(t)}_{mi} y_{ti}\right)^2}{P_{ti}} \le \sum_{i=1}^{k} \frac{E^{(t)}_{mi}}{P_{ti}}.$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \le \mathbb{E}\left[\sum_{m=1}^{M} Q_{tm} \sum_{i=1}^{k} \frac{E^{(t)}_{mi}}{P_{ti}}\right] = \mathbb{E}\left[\sum_{i=1}^{k} \frac{\sum_{m=1}^{M} Q_{tm} E^{(t)}_{mi}}{P_{ti}}\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right) \qquad \tilde{X}_t = E^{(t)} \hat{X}_t \qquad \hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\begin{aligned} E^{(t)}(1 - \widehat{X}_t) \\ = 1 - E^{(t)}\widehat{X}_t \\ = 1 - \widetilde{X}_t \end{aligned}$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{it} = \frac{\mathbb{I}[A_t = i] y_{ti}}{P_{ti}}$$

---

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E_{mA_t}^{(t)} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E_{mi}^{(t)} y_{ti}\right)^2}{P_{ti}} \leq \sum_{i=1}^{k} \frac{E_{mi}^{(t)}}{P_{ti}}.$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq \mathbb{E}\left[\sum_{m=1}^{M} Q_{tm} \sum_{i=1}^{k} \frac{E_{mi}^{(t)}}{P_{ti}}\right] = \mathbb{E}\left[\sum_{i=1}^{k} \frac{\sum_{m=1}^{M} Q_{tm} E_{mi}^{(t)}}{P_{ti}}\right] = k.$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1 - \tilde{X}_{tm})^2\right].$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq k$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1 - \tilde{X}_{tm})^2\right].$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq k$$

$$\leq \frac{\log(M)}{\eta} + \frac{\eta n k}{2}$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1 - \tilde{X}_{tm})^2\right].$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq k$$

$$\leq \frac{\log(M)}{\eta} + \frac{\eta n k}{2} = \sqrt{2nk\log(M)}.$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1 - \tilde{X}_{tm})^2\right].$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq k$$

$$\leq \frac{\log(M)}{\eta} + \frac{\eta n k}{2} = \sqrt{2nk\log(M)}.$$

THEOREM 18.1. *Let $\gamma = 0$ and $\eta = \sqrt{2\log(M)/(nk)}$, and denote by $R_n$ the expected regret of Exp4 defined in Algorithm 11 after $n$ rounds. Then,*

$$R_n \leq \sqrt{2nk\log(M)}. \tag{18.7}$$

$$R_n \leq \sqrt{2nk \log(M)}.$$

$$R_n \leq \sqrt{2nk\log(M)}.$$

- مثال: $\Phi$ = همه توابع از $C$ به $[k]$

$$R_n \leq \sqrt{2nk\log(M)}.$$

- مثال: $\Phi$ = همه توابع از C به [k]

$$M = k^C$$

$$R_n \leq \sqrt{2nk\log(M)}.$$

- مثال: $\Phi$ = همه توابع از C به [k]

$$M = k^{\mathcal{C}}$$

$$R_n \leq \sqrt{2nk|\mathcal{C}|\log(k)},$$

$$R_n \leq \sqrt{2nk\log(M)}.$$

- مثال: Φ = همه توابع از C به [k]

$$M = k^{\mathcal{C}}$$

$$R_n \leq \sqrt{2nk|\mathcal{C}|\log(k)},$$

- مثال: Φ = تعدادی تابع

$$R_n \leq \sqrt{2nk \log(M)}.$$

- مثال: $\Phi$ = همه توابع از C به [k]

$$M = k^{\mathcal{C}}$$

$$R_n \leq \sqrt{2nk|\mathcal{C}| \log(k)},$$

- مثال: $\Phi$ = تعدادی تابع

$$R_n \leq \sqrt{2nk \log(|\Phi|)}.$$

$$E_t^* = \sum_{s=1}^{t} \sum_{i=1}^{k} \max_{m \in [M]} E_{mi}^{(s)}$$

THEOREM 18.3. *Assume the same conditions as in Theorem 18.1, except let* $\eta_t = \sqrt{\log(M)/E_t^*}$. *Then there exists a universal constant* $C > 0$ *such that*

$$R_n \leq C\sqrt{E_n^* \log(M)}. \tag{18.14}$$

$$E_t^* = \sum_{s=1}^{t} \sum_{i=1}^{k} \max_{m \in [M]} E_{mi}^{(s)}$$

THEOREM 18.1. *Let $\gamma = 0$ and $\eta = \sqrt{2\log(M)/(nk)}$, and denote by $R_n$ the expected regret of Exp4 defined in Algorithm 11 after $n$ rounds. Then,*

$$R_n \leq \sqrt{2nk\log(M)}. \tag{18.7}$$

THEOREM 18.3. *Assume the same conditions as in Theorem 18.1, except let $\eta_t = \sqrt{\log(M)/E_t^*}$. Then there exists a universal constant $C > 0$ such that*

$$R_n \leq C\sqrt{E_n^* \log(M)}. \tag{18.14}$$

# پایان