IE613: Online Machine Learning **Jan-Apr 2018**

Lecture 12: February, 16 2018

Lecturer: M. K. Hanawal *Scribes: Prashant Trivedi*

Recap

In last class we discussed about the EXP3.P algorithm and the regret bounds in term of expectation and in terms of high probability and proved that both of them have same order, thus one question that we can ask is, Is it possible to come up with any other algorithm which will give us a better regret bounds? This we are going to see in this class and will prove that there is no such algorithm which give us better regret bounds.

12.1 Introduction

In EXP3 algorithm we only focused on exploitation whereas in EXP3.P we also considered exploration along with exploitation. In last class we argued that the exploration term is essentially $\frac{\gamma}{K}$ which is given by the uniform distribution. Also we noticed that both EXP3 and EXP3.P gave us same regret bound in terms of order which is $O(\sqrt{TKlogK})$, also we argued that in EXP3 we can show the bounds in expectations whereas in EXP3.P we can show the bounds in expectation as well as in high probability. Now the question is can we get this bound in Expectation and in high probability only by doing the exploration and the answer to this question is yes! we can do this and the algorithm that we have for this EXP3-IX.

Algorithm 1 EXP3-IX

Parameters: $\eta_t > 0$, $\gamma_t > 0$.
Initialization: $w_{1,i} = 1$.
for $t = 1, 2, \ldots, T$**, repeat**

1. $p_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^{K} w_{t,j}}$.
2. Draw $I_t \sim \boldsymbol{p}_t = (p_{t,1}, \ldots, p_{t,K})$.
3. Observe loss $\ell_{\{t,I_t\}}$.
4. $\tilde{\ell}_{t,i} \leftarrow \frac{\ell_{t,i}}{p_{t,i}+\gamma_t}\mathbb{I}_{\{I_t=i\}}$ for all $i \in [K]$.
5. $w_{t+1,i} \leftarrow w_{t,i}e^{-\eta_t\tilde{\ell}_{t,i}}$ for all $i \in [K]$.

Clearly, the loss estimates are defined as

$$\tilde{\ell}_{t,i} = \frac{\ell_{t,i}}{p_{t,i} + \gamma_t}\mathbb{I}_{\{I_t=i\}}, \tag{12.1}$$

for all i and an appropriately chosen $\gamma_t > 0$. This technique of defining loss is referred to as as *Implicit eXploration*, or, in short, IX.

12.2 High-probability regret bounds via implicit exploration

Following theorem states high-probability bound on the regret of EXP3-IX.

Theorem 12.1 *Fix an arbitrary $\delta > 0$. With $\eta_t = 2\gamma_t = \sqrt{\frac{2\log K}{KT}}$ for all t, EXP3-IX guarantees*

$$\tilde{R}_T \leq 2\sqrt{2KT\log K} + \left(\sqrt{\frac{2KT}{\log K}} + 1\right)\log(2/\delta)$$

with probability at least $1-\delta$. Furthermore, setting $\eta_t = 2\gamma_t = \sqrt{\frac{\log K}{Kt}}$ for all t, the bound becomes

$$\tilde{R}_T \leq 4\sqrt{KT\log K} + \left(2\sqrt{\frac{KT}{\log K}} + 1\right)\log(2/\delta).$$

To prove this theorem we need a lemma (one of its version we proved in the previous lecture). We will use the similar lemma (stated below) for EXP3-IX also.

Lemma 12.2 *Let (γ_t) be a* fixed *non-increasing sequence with $\gamma_t \geq 0$ and let $\alpha_{t,i}$ be nonnegative $\mathcal{F}_{t-1}$-measurable random variables satisfying $\alpha_{t,i} \leq 2\gamma_t$ for all t and i. Then, with probability at least $1-\delta$,*

$$\sum_{t=1}^{T}\sum_{i=1}^{K}\alpha_{t,i}\left(\tilde{\ell}_{t,i} - \ell_{t,i}\right) \leq \log(1/\delta)$$

A particularly important special case of the above lemma is the following:

Corollary 12.3 *Let $\gamma_t = \gamma \geq 0$, then $\forall\ \delta \in (0,1)$, with probability atleast (1-δ)*

$$\sum_{t=1}^{T}\left(\tilde{\ell}_{t,i} - \ell_{t,i}\right) \leq \frac{\log(K/\delta)}{2\gamma} \tag{12.2}$$

simultaneously holds for all $i \in [K]$

Now we will proceed to prove the theorem(12.1)

Proof: Let us fix an arbitrary $\delta' \in (0,1)$. Following the standard analysis of EXP3 in the loss game and non increasing learning rates , we can obtain the bound (see appendix.)

$$\sum_{t=1}^{T}\left(\sum_{i=1}^{K}p_{t,i}\tilde{\ell}_{t,i} - \tilde{\ell}_{t,j}\right) \leq \frac{\log K}{\eta_T} + \sum_{t=1}^{T}\frac{\eta_t}{2}\sum_{i=1}^{K}p_{t,i}\left(\tilde{\ell}_{t,i}\right)^2 \tag{12.3}$$

for any j.

Now observe that

$$\sum_{i=1}^{K}p_{t,i}\tilde{\ell}_{t,i} = \sum_{i=1}^{K}\mathbb{I}_{\{I_t=i\}}\frac{\ell_{t,i}(p_{t,i}+\gamma_t)}{p_{t,i}+\gamma_t} - \gamma_t\sum_{i=1}^{K}\mathbb{I}_{\{I_t=i\}}\frac{\ell_{t,i}}{p_{t,i}+\gamma_t} = \ell_{t,I_t} - \gamma_t\sum_{i=1}^{K}\tilde{\ell}_{t,i} \tag{12.4}$$

Substituting this back in equation(12.3) we have:

$$\sum_{t=1}^{T}\left[\left(\ell_{t,I_t}-\gamma_t\sum_{i=1}^{K}\tilde{\ell}_{t,i}\right)-\ell_{t,j}\right]\leq\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\frac{\eta_t}{2}\sum_{i=1}^{K}p_{t,i}\left(\tilde{\ell}_{t,i}\right)^2 \tag{12.5}$$

Also notice that, $\sum_{i=1}^{K}p_{t,i}\tilde{\ell}_{t,i}^2\leq\sum_{i=1}^{K}\tilde{\ell}_{t,i}$ holds by the boundedness of the losses.

$$\sum_{t=1}^{T}\left(\ell_{t,I_t}-\tilde{\ell}_{t,j}\right)\leq\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)\sum_{i=1}^{K}\tilde{\ell}_{t,i} \tag{12.6}$$

Adding and subtracting one $\sum_{t=1}^{T}\ell_{t,j}$ on R.H.S. we have

$$\sum_{t=1}^{T}\left(\ell_{t,I_t}-\sum_{i=1}^{K}\tilde{\ell}_{t,j}\right)+\sum_{t=1}^{T}\ell_{t,j}-\sum_{t=1}^{T}\ell_{t,j}\leq\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)\sum_{i=1}^{K}\tilde{\ell}_{t,i}$$

$$\Rightarrow\sum_{t=1}^{T}\left(\ell_{t,I_t}-\ell_{t,j}\right)\leq\sum_{t=1}^{T}\left(\tilde{\ell}_{t,j}-\ell_{t,j}\right)+\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)\sum_{i=1}^{K}\tilde{\ell}_{t,i}$$

Thus, we get that

$$\begin{aligned}\sum_{t=1}^{T}(\ell_{t,I_t}-\ell_{t,j})\leq&\sum_{t=1}^{T}\left(\tilde{\ell}_{t,j}-\ell_{t,j}\right)+\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)\sum_{i=1}^{K}\tilde{\ell}_{t,i}\\ \leq&\frac{\log(K/\delta')}{2\gamma}+\frac{\log K}{\eta}+\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)\sum_{i=1}^{K}\ell_{t,i}+\log(1/\delta')\end{aligned}$$

holds with probability at least $1-2\delta'$, where the last line follows from an application of lemma(12.2) with $\alpha_{t,i}=\eta_t/2+\gamma_t$ for all t,i and taking $j=arg\,\underset{i}{min}\;L_{T,i}$ and $\delta'=\delta/2$, and using the boundedness of the losses,

$$\tilde{R}_T\leq\frac{\log(2K/\delta)}{2\gamma_T}+\frac{\log K}{\eta_T}+K\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)+\log(1/\delta).$$

The statements of the theorem then follow immediately, noting that $\sum_{t=1}^{T}1/\sqrt{t}\leq2\sqrt{T}$. ■

12.3 Appendix

1. Proof of inequality 12.3

Proof: For Exp3-IX algorithm the probability is given by :

$$\begin{aligned}p_{t,i}&=\frac{e^{-(\eta\sum_{j=1}^{t-1}\tilde{\ell}_{j,i})}}{\sum_{i=1}^{K}e^{-(\eta\sum_{j=1}^{t-1}\tilde{\ell}_{j,i})}}\\&=\frac{e^{-(\eta\tilde{L}_{t-1,i})}}{\sum_{i=1}^{K}e^{-(\eta\tilde{L}_{t-1,i})}}\end{aligned}$$

where $\tilde{L}_{t-1,i}$ denotes the cumulative loss upto time $t-1$ Now consider $\frac{1}{\eta}\log\sum_{i=1}^{K} p_{t,i}e^{-\eta\tilde{\ell}_{t,i}}$ Putting the above probability $p_{t,i}$ in this we have

$$\begin{aligned}
\frac{1}{\eta}\log\sum_{i=1}^{K} p_{t,i}e^{-\eta\tilde{\ell}_{t,i}} &= \frac{1}{\eta}\log\sum_{i=1}^{K}\frac{e^{-(\eta\sum_{j=1}^{t-1}\tilde{\ell}_{j,i})}}{\sum_{i=1}^{K}e^{-(\eta\sum_{j=1}^{t-1}\tilde{\ell}_{j,i})}}\cdot e^{-\eta\tilde{\ell}_{t,i}} \\
&= \frac{1}{\eta}\log\sum_{i=1}^{K}\frac{e^{-(\eta\sum_{j=1}^{t}\tilde{\ell}_{j,i})}}{\sum_{i=1}^{K}e^{-(\eta\sum_{j=1}^{t-1}\tilde{\ell}_{j,i})}} \qquad \textit{(combining exponentials in numerator)}
\end{aligned}$$

therefore, we have

$$\frac{1}{\eta}\log\sum_{i=1}^{K} p_{t,i}e^{-\eta\tilde{\ell}_{t,i}} = \frac{1}{\eta}\left(\log\sum_{i=1}^{K}e^{-\eta\tilde{L}_{t,i}} - \log\sum_{i=1}^{K}e^{-\eta\tilde{L}_{t-1,i}}\right) \tag{12.7}$$

Taking sum over all $t = 1, 2, \ldots, T$ we have,

$$\begin{aligned}
\sum_{t=1}^{T}\frac{1}{\eta}\log\sum_{i=1}^{K} p_{t,i}e^{-\eta\tilde{\ell}_{t,i}} &= \sum_{t=1}^{T}\frac{1}{\eta}\left(\log\sum_{i=1}^{K}e^{-\eta\tilde{L}_{t,i}} - \log\sum_{i=1}^{K}e^{-\eta\tilde{L}_{t-1,i}}\right) \\
&= \frac{1}{\eta}\left(\sum_{t=1}^{T}\log\sum_{i=1}^{K}e^{-\eta\tilde{L}_{t,i}} - \sum_{t=1}^{T}\log\sum_{i=1}^{K}e^{-\eta\tilde{L}_{t-1,i}}\right) \\
&= \frac{1}{\eta}\left(\log\sum_{i=1}^{K}e^{-\eta\tilde{L}_{T,i}} - \log\sum_{i=1}^{K}e^{-\eta\tilde{L}_{0,i}}\right) \qquad \textit{(summing over all } T\textit{)}
\end{aligned}$$

also we know that $\quad \tilde{L}_{0,i} = 0, \quad e^{-\eta\tilde{L}_{0,i}} = 1$ *and* $\sum_{i=1}^{K}e^{-\eta\tilde{L}_{0,i}} = K$

therefore we now have,

$$\sum_{t=1}^{T}\frac{1}{\eta}\log\sum_{i=1}^{K} p_{t,i}e^{-\eta\tilde{\ell}_{t,i}} = \frac{1}{\eta}\log\sum_{i=1}^{K}e^{-\eta\tilde{L}_{T,i}} - \frac{\log K}{\eta} \tag{12.8}$$

Now we will use the following inequalities to get the bounds:

$$\begin{aligned}
\log(x) &\leq x - 1 \\
e^{-x} &\leq 1 - x + \frac{x^2}{2}
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\eta}\log\sum_{i=1}^{K} p_{t,i}e^{-\eta\tilde{\ell}_{t,i}} &\leq \frac{1}{\eta}\left(\sum_{i=1}^{K}p_{t,i}e^{-\eta\tilde{\ell}_{t,i}} - 1\right) \\
&\leq \frac{1}{\eta}\left(\sum_{i=1}^{K}p_{t,i}\left(1 - \eta\tilde{\ell}_{t,i} + \frac{\eta^2\tilde{\ell}_{t,i}^2}{2}\right) - 1\right) \\
&\leq \frac{1}{\eta}\left(-\sum_{i=1}^{K}p_{t,i}\eta\tilde{\ell}_{t,i} + \sum_{i=1}^{K}p_{t,i}\frac{\eta^2\tilde{\ell}_{t,i}^2}{2}\right) \qquad \textit{(because} \sum_{i=1}^{K}p_{t,i} = 1)
\end{aligned}$$

therefore we have,

$$\frac{1}{\eta}\log\sum_{i=1}^{K} p_{t,i}e^{-\eta\tilde{\ell}_{t,i}} \leq \frac{\eta}{2}\sum_{i=1}^{K}p_{t,i}\tilde{\ell}_{t,i}^2 - \sum_{i=1}^{K}p_{t,i}\eta\tilde{\ell}_{t,i} \tag{12.9}$$

from inequalities(12.9) and equation(12.8) we have,

$$\begin{aligned}
\frac{1}{\eta}\log\sum_{i=1}^{K} e^{-\eta\tilde{L}_{T,i}} - \frac{\log K}{\eta} &\leq \sum_{t=1}^{T}\left(\frac{\eta}{2}\sum_{i=1}^{K} p_{t,i}\tilde{\ell}_{t,i}^2 - \sum_{i=1}^{K} p_{t,i}\eta\tilde{\ell}_{t,i}\right) \\
\sum_{t=1}^{T}\sum_{i=1}^{K} p_{t,i}\tilde{\ell}_{t,i} + \frac{1}{\eta}\log\sum_{i=1}^{K} e^{-\eta\tilde{L}_{T,i}} &\leq \sum_{t=1}^{T}\frac{\eta}{2}\sum_{i=1}^{K} p_{t,i}\tilde{\ell}_{t,i}^2 + \frac{\log K}{\eta}
\end{aligned}$$

now applying the following:

$$\begin{aligned}
\tilde{L}_{T,i} \geq \min_i \tilde{L}_{T,i} &\geq \eta \min_i \tilde{L}_{T,i} \\
e^{-\tilde{L}_{T,i}} &\leq e^{-\eta \min \tilde{L}_{T,i}} \\
therefore\ \ \tilde{L}_{T,i} &\geq \eta\tilde{L}_{T,i} \\
e^{-\tilde{L}_{T,i}} &\leq e^{-\eta\tilde{L}_{T,i}} \\
But\ \ \sum_{i=1}^{K} e^{-\tilde{L}_{T,i}} &\leq e^{-\eta\tilde{L}_{T,i}} \quad \forall i \in [K]
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{1}{\eta}\log\sum_{i=1}^{K} e^{-\tilde{L}_{T,i}} &\geq \frac{1}{\eta}\log\left(e^{-\tilde{L}_{T,i}}\right) \\
&= -\frac{1}{\eta}\eta\tilde{L}_{T,i} = -\tilde{L}_{T,i}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{t=1}^{T}\sum_{i=1}^{K} p_{t,i}\tilde{\ell}_{t,i} - \tilde{L}_{T,i} &\leq \frac{\log K}{\eta} + \sum_{t=1}^{T}\frac{\eta}{2}\sum_{i=1}^{K} p_{t,i}\tilde{\ell}_{t,i}^2 \\
\Rightarrow \sum_{t=1}^{T}\left(\sum_{i=1}^{K} p_{t,i}\tilde{\ell}_{t,i} - \tilde{\ell}_{t,j}\right) &\leq \frac{\log K}{\eta} + \sum_{t=1}^{T}\frac{\eta}{2}\sum_{i=1}^{K} p_{t,i}\left(\tilde{\ell}_{t,i}\right)^2
\end{aligned}$$

■

12.4 Lower Bound

So far we had seen all the algorithms for finding the regret bounds in both the setting viz. in expectation as well as in high probability and we found that for all the algorithms that we have

$$\tilde{R}_T \leq O(\sqrt{KT\log K}) \ \ and \ \ \bar{R}_T \leq O(\sqrt{KT\log K})$$

Now we would like to find the lower bound on these quantities and we will thus make the following claim:
Claim: $\bar{R}_T \leq \Omega(\sqrt{KT})$
In other words we would like to proof the following theorem:

Theorem 12.4 *There exists a sequence of losses such that for any T, "pseudo regret" is* $\Omega(\sqrt{KT})$.

This means that the adversary can come up with an strategy/ sequence such that the regret is at least $\Omega(\sqrt{KT})$, and this also means that in this sense all the algorithms that we had studied so far are optimal. We can write the above theorem mathematically as:

Theorem 12.5

$$\inf\ \sup\left(\mathbb{E}\sum_{t=1}^{T}\ell_{t,I_t}-\min_{i\in[K]}\mathbb{E}\sum_{t=1}^{T}\ell_{t,i}\right)\geq\frac{1}{20}\sqrt{TK}$$

where inf is over all the learner's strategy and the sup is over all the adversary strategies.

The proof of this theorem requires the stochastic argument. This theorem says that however the learner choose his strategy he is going to incur atleast $\frac{1}{20}\sqrt{TK}$ loss.

In order to emphasize that our losses are stochastic (in particular, Bernoulli random variables), we use $l_{t,i}\in\{0,1\}$ to denote the loss incur by playing arm i at time t.

Since $\mathbb{E}\sum_{t=1}^{T}\ell_{t,I_t}-\min_{i\in[K]}\mathbb{E}\sum_{t=1}^{T}\ell_{t,i}=\bar{R_T}\leq\mathbb{E}R_T$, Theorem 12.5 immediately entails a lower bound on the regret.

The general idea of the proof goes as follows. Since at least one arm is pulled less than T/K times, for this arm one cannot differentiate between a Bernoulli of parameter 1/2 and and a Bernoulli of parameter $1/2+\sqrt{K/T}$. Thus, if all arms are Bernoulli of parameter 1/2 but one, whose parameter is $1/2+\sqrt{K/T}$, then the forecaster should incur a regret of order $T\sqrt{K/T}=\sqrt{TK}$. In order to formalize this idea, we use the **Kullback-Leibler divergence**, and in particular **Pinsker's inequality**, to compare the behaviour of a given forecaster against: (1) the distribution where all arms are Bernoulli of parameter 1/2; (2) the same distribution where the parameter of one arm is increased by ϵ.

We start by proving a more general lemma, which could also be used to derive lower bounds. The proof of Theorem 12.5 then follows by a simple optimization over ϵ.

Lemma 12.6 *Let $\epsilon\in[0,1)$. For any $i\in[K]$ let $\mathbb{E}_i$ be the expectation against the joint distribution of losses where all arms are i.i.d. Bernoulli of parameter $\frac{1-\epsilon}{2}$ but arm i, which is i.i.d. Bernoulli of parameter $\frac{1+\epsilon}{2}$. Then,*

$$\min_{i\in[K]}\mathbb{E}_i\sum_{t=1}^{T}\left(l_{t,I_t}-l_{t,i}\right)\geq T\epsilon\left(1-\frac{1}{K}-\sqrt{\epsilon\ln\frac{1+\epsilon}{1-\epsilon}}\sqrt{\frac{T}{2K}}\right).$$

Proof: We provide a proof in five steps by lower bounding $\frac{1}{K}\sum_{i=1}^{K}\mathbb{E}_i\sum_{t=1}^{n}(Y_{i,t}-Y_{I_t,t})$. This implies the statement of the lemma because a -min is larger than a mean.

First step: Empirical distribution of plays.

We start by considering a deterministic forecaster. Let $q_T=\left(q_{1,T},\ldots,q_{K,T}\right)$ be the empirical distribution of plays over the arms defined by $q_{i,T}=\frac{T_i(T)}{T}$ where $B_i(T)$ denotes the number of times arm i was selected in the first T rounds. Let J_T be drawn according to q_T. We denote by $\mathbb{P}_i$ the law of J_T against the distribution where all arms are i.i.d. Bernoulli of parameter $\frac{1-\epsilon}{2}$ but arm i, which is i.i.d. Bernoulli of parameter $\frac{1+\epsilon}{2}$ (we call this the i-th stochastic adversary). Recall that $\mathbb{P}_i(J_T=j)=\mathbb{E}_i fracB_j(T)n$, hence

$$\mathbb{E}_i\sum_{t=1}^{T}\left(l_{t,I_t}-l_{t,i}\right)=\epsilon T\sum_{j\neq i}\mathbb{P}_i(J_T=j)=\epsilon n\left(1-\mathbb{P}_i(J_T=i)\right)$$

which implies

$$\frac{1}{K}\sum_{i=1}^{K}\mathbb{E}_i\sum_{t=1}^{T}\big(l_{t,I_t}-l_{t,i}\big)=\epsilon T\left(1-\frac{1}{K}\sum_{i=1}^{K}\mathbb{P}_i(J_T=i)\right)\ . \tag{12.10}$$

Second step: Pinsker's inequality.

Let $\mathbb{P}_0$ be the law of J_T for the distribution where all arms are i.i.d. Bernoulli of parameter $\frac{1-\epsilon}{2}$. Then Pinsker's inequality immediately gives $\mathbb{P}_i(J_T=i)\le\mathbb{P}_0(J_T=i)+\sqrt{\frac{1}{2}\mathbb{K}(\mathbb{P}_0,\mathbb{P}_i)}$, and so

$$\frac{1}{K}\sum_{i=1}^{K}\mathbb{P}_i(J_T=i)\le\frac{1}{K}+\frac{1}{K}\sum_{i=1}^{K}\sqrt{\frac{1}{2}\mathbb{K}(\mathbb{P}_0,\mathbb{P}_i)}\ . \tag{12.11}$$

Third step: Computation of $\mathbb{K}(\mathbb{P}_0,\mathbb{P}_i)$.

Since the forecaster is deterministic, the sequence of losses $l^T=(l_1,\dots,l_T)\in\{0,1\}^T$ received by the forecaster uniquely determines the empirical distribution of plays q_T. In particular, the law of J_T conditionally to l^T is the same for any i-th stochastic adversary. For each $i=0,\dots,K$, let $\mathbb{P}_i^T$ be the law of l^T against the i-th adversary. Then one can easily show that $\mathbb{K}(\mathbb{P}_0,\mathbb{P}_i)\le\mathbb{K}(\mathbb{P}_0^T,\mathbb{P}_i^T)$. Now we use the chain rule for **Kullback-Leibler divergence** iteratively to introduce the laws $\mathbb{P}_i^t$ of $l^t=(l_1,\dots,l_t)$. More precisely, we have

$$\begin{aligned}
&\mathbb{K}(\mathbb{P}_0^T,\mathbb{P}_i^T)\\
&=\mathbb{K}(\mathbb{P}_0^1,\mathbb{P}_i^1)+\sum_{t=2}^{T}\sum_{y^{t-1}}\mathbb{P}_0^{t-1}(y^{t-1})\,\mathbb{K}\big(\mathbb{P}_0^t(\cdot\mid y^{t-1}),\mathbb{P}_i^t(\cdot\mid y^{t-1})\big)\\
&=\mathbb{K}(\mathbb{P}_0^1,\mathbb{P}_i^1)+\sum_{t=2}^{T}\left(\sum_{y^{t-1}\,:\,I_t=i}\mathbb{P}_0^{t-1}(y^{t-1})\,\mathbb{K}\left(\tfrac{1-\epsilon}{2},\tfrac{1+\epsilon}{2}\right)\right.\\
&\qquad\qquad\left.+\sum_{y^{t-1}\,:\,I_t\neq i}\mathbb{P}_0^{t-1}(y^{t-1})\,\mathbb{K}\left(\tfrac{1+\epsilon}{2},\tfrac{1+\epsilon}{2}\right)\right)\\
&=\mathbb{K}\left(\tfrac{1-\epsilon}{2},\tfrac{1+\epsilon}{2}\right)\mathbb{E}_0\,B_i(T)\ .
\end{aligned}\tag{12.12}$$

Fourth step: conclusion for deterministic forecasters.

By using that the square root is concave, and combining $\mathbb{K}(\mathbb{P}_0,\mathbb{P}_i)\le\mathbb{K}(\mathbb{P}_0^T,\mathbb{P}_i^T)$ with (12.12), we deduce that

$$\begin{aligned}
\frac{1}{K}\sum_{i=1}^{K}\sqrt{\mathbb{K}(\mathbb{P}_0,\mathbb{P}_i)}&\le\sqrt{\frac{1}{K}\sum_{i=1}^{K}\mathbb{K}(\mathbb{P}_0,\mathbb{P}_i)}\\
&\le\sqrt{\frac{1}{K}\sum_{i=1}^{K}\mathbb{K}\left(\tfrac{1-\epsilon}{2},\tfrac{1+\epsilon}{2}\right)\mathbb{E}_0\,B_i(T)}\\
&=\sqrt{\frac{T}{K}\mathbb{K}\left(\tfrac{1-\epsilon}{2},\tfrac{1+\epsilon}{2}\right)}\ .
\end{aligned}\tag{12.13}$$

We conclude the proof for deterministic forecasters by applying (12.11) and (12.13) to (12.10), and observing that $\mathbb{K}\left(\frac{1-\epsilon}{2}, \frac{1+\epsilon}{2}\right) = \epsilon \ln \frac{1+\epsilon}{1-\epsilon}$.

Fifth step: randomized forecasters via Fubini's Theorem.

Extending previous results to randomized forecasters is easy. Denote by $\mathbb{E}_r$ the expectation with respect to the forecaster's internal randomization. Then **Fubini's Theorem** implies

$$\frac{1}{K}\sum_{i=1}^{K}\mathbb{E}_i\sum_{t=1}^{T}\mathbb{E}_r\big(l_{t,I_t} - l_{t,i}\big) = \mathbb{E}_r\frac{1}{K}\sum_{i=1}^{K}\mathbb{E}_i\sum_{t=1}^{T}\big(l_{t,I_t} - l_{t,i}\big) \ .$$

Now the proof is concluded by applying the lower bound on $\frac{1}{K}\sum_{i=1}^{K}\mathbb{E}_i\sum_{t=1}^{T}\big(l_{t,I_t} - l_{t,i}\big)$, which we proved in previous steps, to each realization of the forecaster's random bits. ■