به نام خدا

علی قاسم زاده ۴۰۱۱۰۶۳۳۹

تمرین ۱ RL

۱.۱ a) داریم که $\mathbb{E}\{x\}=0$ , $a \leq x \leq b$ است، $s>0$ و می‌خواهیم ثابت کنیم که:

$$\mathbb{E}\{e^{sx}\} \leq e^{s^2(b-a)^2/8}$$

حل: داریم که،

$$x = \lambda b + (1-\lambda) a \quad \text{where} \quad \lambda = \frac{x-a}{b-a}$$

می‌دانیم که exp تابعی convex است ⟸ خواهیم داشت که:

$$e^{sx} \leq \lambda e^{sb} + (1-\lambda) e^{sa} = \frac{x-a}{b-a} e^{sb} + \frac{b-x}{b-a} e^{sa}$$

$$\longrightarrow \mathbb{E}\{e^{sx}\} \leq \mathbb{E}\left\{\frac{x-a}{b-a} e^{sb} + \frac{b-x}{b-a} e^{sa}\right\} = \underbrace{\frac{-a}{b-a} e^{sb} + \frac{b}{b-a} e^{sa}}$$

log این را $\psi(s)$ می‌نامیم

$$\Rightarrow \mathbb{E}\{e^{sx}\} \leq e^{sa}\left(\frac{b}{b-a} - \frac{a}{b-a} e^{s(b-a)}\right)$$

$$\Rightarrow \psi(s) = \operatorname{Log}\left(\frac{-a}{b-a} e^{sb} + \frac{b}{b-a} e^{sa}\right), \quad \psi(0) = 0$$

$$\psi'(0) = \mathbb{E}\{X\} = 0 \longleftarrow$$ از خواص Moment generating function می‌دانیم که

$$\operatorname{var}_{P_s}(X), \quad P_s(x) = \frac{e^{sx} P(x)}{\mathbb{E}\{e^{sx}\}}$$

$$* \longrightarrow \psi''(s) \leq \frac{(b-a)^2}{4} \xrightarrow{\text{دوبار انتگرال می‌گیریم}} \psi(s) \leq \frac{s^2(b-a)^2}{8}$$

$$\longrightarrow \underline{\mathbb{E}\{e^{sx}\} \leq e^{s^2(b-a)^2/8}}$$

* هر متغیر تصادفی در رنج $[a,b]$ در بیشترین حالت واریانسش $\left(\frac{b-a}{2}\right)^2$ هست
زمانی هست که داده ها دو سر a و b باشند و بقیه جاها نباشند.

(b نامساوی هافدینگ را باید نشان دهیم. ابتدا نشان می‌دهیم که

$$P\left(\frac{1}{n}\sum (z_i - \mathbb{E}\{z_i\}) \geq t\right) \leq \exp\left(\frac{-2nt^2}{(b-a)^2}\right)$$

سپس طرف دیگر هم مشابه آن اثبات می‌شود.

حل، در نظر می‌گیریم

$$X_i = Z_i - \mathbb{E}\{Z_i\} \longrightarrow \ , \mathbb{E}\{X_i\} = 0$$

$$X_i \in [a - \mathbb{E}\{X_i\},\ b - \mathbb{E}\{X_i\}]$$

ابتدا هم از بخش a داریم که،

$$E\{e^{sX_i}\} \leq e^{s^2(b-a)^2/8}$$

از طرفی داریم که $Z_i$ها مستقل بودند $\Longleftarrow$

$$\mathbb{E}\{e^{s\sum X_i}\} = \prod_{i=1}^{n} \mathbb{E}\{e^{sX_i}\} \leq e^{ns^2(b-a)^2/8}$$

از باند چرنوف داریم که،

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq t\right) = \mathbb{P}\left(\sum X_i \geq nt\right) \leq \frac{\mathbb{E}\{e^{s\sum X_i}\}}{e^{snt}}$$

$$\leq e^{ns^2(b-a)/8 - snt}$$

حالا برای کمینه کردن توان مشتق می‌گیریم،

$$\frac{ns(b-a)^2}{4} - nt = 0$$

$$\longrightarrow s = \frac{4t}{(b-a)^2} \longrightarrow P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq t\right) \leq e^{-2nt^2/(b-a)^2}$$

برای tail پایینی هم کافیست $X_i$ را $-X_i$ قرار بدهیم و روابط بالا را از اول بنویسیم به همان می‌رسیم.

اثبات باند چرنوف ۱

$$S_n = \sum X_i$$

$$P(S_n \geq E\{S_n\}+nt') \;,\; t>0$$

از نامساوی مارکوف داریم که ؛

$$P(S_n\geq E\{S_n\}+nt') = P(e^{sS_n} \geq e^{s(E\{S_n\}+nt')}) \leq \frac{E\{e^{s\sum_{i}^{n} X_i}\}}{e^{s(E\{S_n\}+nt')}}$$

حالا کافیه $t$ را در نظر بگیریم به طوری که $nt = nt' + E\{S_n\}$

$t = t' + \frac{\sum E\{X_i\}}{n}$ یا به عبارتی $t = t' + \frac{E\{S_n\}}{n}$ $\leftarrow$

$\Leftarrow$

$$P(S_n \geq nt) \leq \frac{E\{e^{s\sum X_i}\}}{e^{s(nt)}}$$

---

۲.۱ (a ۱- حل: داریم که از نامساوی چرنوف ؛

$$P(X-\mu > t) = P(e^{\lambda(X-\mu)} \geq e^{\lambda t}) \leq \frac{E\{e^{\lambda(X-\mu)}\}}{e^{\lambda t}}$$

$$E\{e^{\lambda(X-\mu)}\} \leq e^{\lambda^2\sigma^2/2} \longleftarrow \text{sub-guassian}$$

$$P(X-\mu>t) \leq e^{\lambda^2\sigma^2/2 - \lambda t}$$

به ازای هر $\lambda$ ای باید برقرار باشد پس مشتق می‌گیریم و $\lambda$ ای را قرار می‌دهیم که کران بالا را کمینه کند ؛

$$\lambda\sigma^2 - t = 0 \longrightarrow \lambda = \frac{t}{\sigma^2} \longrightarrow$$

$$P(X-\mu>t) \leq e^{-t^2/2\sigma^2}$$

۲- اینبار چرنوف را برای $X$ - اجرا می‌کنیم، خواهیم داشت که:

$$\mathbb{P}(X-\mu<-t)=\mathbb{P}(e^{-\lambda(X-\mu)}>e^{\lambda t})\leq\frac{\mathbb{E}\{e^{-\lambda(X-\mu)}\}}{e^{\lambda t}}$$

از زیرگاوسی باز داریم که:

$$\mathbb{E}\{e^{-\lambda(X-\mu)}\}\leq e^{\lambda^2\sigma^2/2}$$

$$\Rightarrow P(X-\mu<-t)\leq e^{\lambda^2\sigma^2/2-\lambda t}$$

دوباره مشتق از توان ← و داریم که:

$$\mathbb{P}(X-\mu\leq-t)\leq e^{-t^2/2\sigma^2}$$

۳- حالا داریم که:

$$\mathbb{P}(|X-\mu|\geq t)=\mathbb{P}(X-\mu\geq t)+\mathbb{P}(X-\mu\leq-t)\leq 2e^{-t^2/2\sigma^2}$$

b) ابتدا تغییر پارامتر می‌دهیم: $Y_i=X_i-\mu_i$ ← $\mathbb{E}\{Y_i\}=0$ و زیرگاوسی هست ⇐

$$\mathbb{E}\{e^{\lambda Y_i}\}\leq e^{\lambda^2\sigma_i^2/2}$$

$S_n=\sum Y_i$

$$\mathbb{E}\{e^{\lambda S_n}\}=\prod_{i=1}^{n}\mathbb{E}\{e^{\lambda Y_i}\}\leq\prod_{i=1}^{n}e^{\lambda^2\sigma_i^2/2}=e^{\lambda^2\sum\sigma_i^2/2}$$

← $S_n$ یک زیرگاوسی با پارامتر $\sqrt{\sum\sigma_i^2/2}$ است.

$$\Rightarrow P(S_n\geq t)\leq\exp\left(-\frac{t^2}{2\sum\sigma_i^2}\right)$$

← از بخش a

Scanned with CamScanner

و به طور مشابه بدست می آید برای طرف دیگر که

$$P(S_n \le -t) \le \exp\left(\frac{-t^2}{2\sum \sigma_i^2}\right)$$

و برای دو طرفه باز داریم که

$$P(|S_n| \ge t) \le 2\exp\left(\frac{-t^2}{2\sum \sigma_i^2}\right)$$

c) حال $X_1 \ldots X_n$ ، متغیرهای iid زیرگاوسی اند با واریانس $\sigma^2$ , $\mu = \mathbb{E}(X)$

خواهیم داشت که

$$\bar{X}_n = \frac{1}{n}\sum X_i \to \bar{X}_n - \mu = \frac{1}{n}\sum (X_i - \mu)$$

چون $X_1 \ldots X_n$ متغیرهای زیرگاوسی iid هستند، جمعشان هم متغیر زیرگاوسی خواهد بود با میانگین $\mu$ و واریانس $\sigma^2/n$

$$\mathbb{E}\{e^{\lambda(\bar{X}_n - \mu)}\} \le e^{\lambda^2\sigma^2/2n}$$

از باند زیرگاوسی داریم که

$$\Rightarrow P(\bar{X}_n - \mu \ge \varepsilon) \le e^{-n\varepsilon^2/2\sigma^2}$$

~~حالا می خواهیم~~ حال $\varepsilon$ را برابر با $\sqrt{\frac{2\sigma^2\log(1/\delta)}{n}}$ قرار می دهیم، خواهیم داشت که

$$P(\bar{X}_n - \mu \ge \varepsilon) \le \delta$$

$$\Rightarrow P(\bar{X}_n - \mu < \varepsilon) \ge 1 - \delta$$

## ۲- UCB :

### ۱.۲

a

$$R_n = \mathbb{E}\left\{\sum_{t=1}^{n} \mu^* - X_{A_t}\right\} \quad , \quad \mathbb{E}\left\{\sum_{t=1}^{n} X_{A_t}\right\} = \mathbb{E}\left\{\sum_{a\in A}\sum_{t=1}^{n} X_a \,\mathbb{I}\{A_t = a\}\right\} =$$

$$= \sum_{a\in A} \mathbb{E}\left\{\sum_{t=1}^{n} X_a \,\mathbb{I}\{A_t=a\}\right\} = \sum_{a\in A} \mathbb{E}\{T_a(n)\}\,\mu_a$$

$T_a(n)$ تعداد باری است که تا زمان $n$ ، arm انتخابی ما $a$ بوده است.

جمع $T_a(n)$ ها برابر است با $n$

$$\Rightarrow R_n = n\mu^* - \sum_{a\in A} \mu_a\, \mathbb{E}\{T_a(n)\} = \sum_{a\in A} (\mu^* - \mu_a)\,\mathbb{E}\{T_a(n)\}$$

$$\Rightarrow R_n = \sum_{a\in A} \Delta_a\, \mathbb{E}\{T_a(n)\}$$

b) فرض کنیم که دوتا arm داریم، arm1 بهینه باشد، arm2 غیربهینه باشد، از آنجایی که ریواردها تصادفی است در ابتدا هم $UCB_1$ و $UCB_2$ مقدارشان $\infty$ است، ابتدا هر دو یک دور اجرا می‌شوند، حالا فرض کنیم که arm1 کمتر از میانگین arm2 بشه دور اول، و arm2 در طول training حول میانگین اش حرکت کنه، و زیاد تغییر نکنه اون موقع اتفاقی که می‌افته اینه که arm1 دیگه تغییر نخواهد کرد مقدار تخمینی اش چون هیچ وقت انتخاب نمی‌شود برابر آبیست؛

$$UCB_1 = \sqrt{\frac{2\log(1/\delta)}{T_1(t-1)}} + \hat{\mu}_1 \quad , \quad UCB_2 = \sqrt{\frac{2\log(1/\delta)}{T_2(t-1)}} + \hat{\mu}_2$$

یعنی یه جوری پیش بره که $R_2 \le UCB_2$ باشه ولی در دور اول $UCB_1 < R_1$ اتفاق بیفته ←

این طوری هیچ وقت ① انتخاب نمیشه، و regret خطی خواهد شد.

c) اگر $G_i$ اتفاق بیفتد خواهیم داشت که، arm1 بهینه است و UCB هم کمتر مساوی $\mu_1$ است

و از طرفی $\mu_1 \le UCB_1$ است. حالا فرض کنیم که $T_i(n) > u_i$ است یعنی تعداد اتفاق‌ها (arm i تا زمان n از $u_i$ بیشتر باشد $\Leftarrow$ در نظر می‌گیریم که:

$$t^* = \min_{t} \{ t \le n : T_i(t) = u_{i+1} \}$$

اولین جایی که $u_{i+1}$ بار arm i انتخاب شده است، پس در این هنگام $UCB_i$ از $UCB_1$ بیشتر بوده است:

$$\underbrace{UCB_i(t^*-1, \delta)}_{\text{arm i اندیس}} \ge \underbrace{UCB_1(t^*-1, \delta)}_{\text{arm 1 اندیس}}$$

در زمان $t^*-1$، arm i به تعداد $u_i$ بار انتخاب شده

از شرط دوم $G_i$ داریم که

$$UCB_1(t,\delta) \ge \mu_1 \ , \ UCB_i(t^*-1,\delta) = \hat{\mu}_{i,u_i} + \sqrt{\frac{2\ln(1/\delta)}{u_i}} < \mu_1$$

به ازای $t \le n$ به ازای $t = t^*-1$ هم اوله یعنی $UCB_1(t^*-1,\delta) \ge \mu_1$

از اینجا نتیجه می‌شود که $UCB_1(t^*-1,\delta) > UCB_i(t^*-1,\delta)$ ولی بالاتر داشتیم که

$$UCB_i(t^*-1,\delta) \ge UCB_1(t^*-1,\delta) \quad * \quad \Rightarrow \quad T_i(n) \le u_i$$

d) از بالا داشتیم که $T_i(n) \le u_i$ حالا این وقتی بوده که $G_i$ برقرار باشه برابر وقتی که برقرار نباشه؟ خواهیم داشت که:

$$\mathbb{E}\{T_i(n)\} = \underbrace{\mathbb{E}\{T_i(n) \mathbb{1}_{G_i}\}} + \mathbb{E}\{T_i(n) \mathbb{1}_{G_i^c}\}$$

$$\le u_i \mathbb{1}_{G_i} \Rightarrow u_i \Pr(G_i) \le u_i$$

$$T_i(n) \le n \rightarrow \mathbb{E}\{T_i(n) \mathbb{1}_{G_i^c}\} \le n\Pr(G_i^c)$$

$$\Rightarrow \mathbb{E}\{T_i(n)\} \le u_i + n\Pr(G_i^c)$$

e) داریم که،

$$\Delta_i = \mu_1 - \mu_i > 0 \quad \text{on the condition} \quad \Delta_i - \sqrt{\frac{2\operatorname{Log}(1/\delta)}{u_i}} \geq c\Delta_i, \quad c \in (0,1)$$

شرط ، $\sqrt{\frac{2\operatorname{Log}(1/\delta)}{u_i}} \leq (1-c)\Delta_i \longleftrightarrow \frac{2\operatorname{Log}(1/\delta)}{u_i} \leq (1-c)^2 \Delta_i^2 \longleftrightarrow$

$$\operatorname{Ln}(1/\delta) \leq \frac{u_i (1-c)^2 \Delta_i^2}{2}$$

حالا داریم که،

$$\left\{\hat{\mu}_{i,u_i} + \sqrt{\frac{2\operatorname{Ln}(1/\delta)}{u_i}} \geq \mu_1\right\} = \left\{\hat{\mu}_{i,u_i} - \mu_i \geq \underbrace{\mu_1 - \mu_i}_{\Delta_i} - \sqrt{\frac{2\operatorname{Ln}(1/\delta)}{u_i}}\right\}$$

طبق شرط صورت سوال داریم که،

$$\underbrace{\left\{\hat{\mu}_{i,u_i} - \mu_i \geq \Delta_i - \sqrt{\frac{2\operatorname{Ln}(1/\delta)}{u_i}}\right\}}_{1} \subseteq \underbrace{\left\{\hat{\mu}_{i,u_i} - \mu_i \geq c\Delta_i\right\}}_{2} \Rightarrow$$

$$\Pr(1) \leq \Pr(2)$$

حالا میاییم هافدینگ را برایش می نویسیم، خواهیم داشت که،

$$\Pr(\hat{\mu}_{i,u_i} - \mu_i \geq \varepsilon) \leq \exp(-2u_i \varepsilon^2) \quad \text{for } \varepsilon > 0$$

$$\Rightarrow \Pr(\hat{\mu}_{i,u_i} - \mu_i \geq c\Delta_i) \leq \exp(-2u_i (c\Delta_i)^2) = \exp(-2c^2 u_i \Delta_i^2)$$

$$\Rightarrow \boxed{\Pr\left(\hat{\mu}_{i,u_i} + \sqrt{\frac{2\operatorname{Ln}(1/\delta)}{u_i}} \geq \mu_1\right) \leq \exp\left(-\frac{u_i c^2 \Delta_i^2}{2}\right)}$$

۴)

اجتماع دوتا ترم $\longrightarrow Pr(A\cup B) \le Pr(A) + Pr(B)$

$$Pr(G_i^c) \le Pr\left(\exists t\le n : UCB_1(t,\delta)\le \mu_1\right) + \overbrace{\underbrace{Pr\left(\hat{\mu}_{i,u_i} + \sqrt{\frac{2\ln(1/\delta)}{u_i}} \ge \mu_1\right)}_{\le \exp\left(-\frac{u_i c^2 \Delta_i^2}{2}\right)}}^{B}$$

$$UCB_1(t,\delta) = \hat{\mu}_1(T_1(t)) + \sqrt{\frac{2\ln(1/\delta)}{T_1(t)}} \quad , \quad T_1(t) > 0$$

برای اینکه $UCB_1(t,\delta)$ از $\mu_1$ کمتر باشد، باید داشته باشیم:

$$\hat{\mu}_1(T_1(t)) \le \mu_1 - \sqrt{\frac{2\ln(1/\delta)}{T_1(t)}} \Longrightarrow$$

هافدینگ:

$$Pr\left(\hat{\mu}_1(s) \le \mu_1 - \sqrt{\frac{2\ln(1/\delta)}{s}}\right) \le \delta$$

وقتی arm 1 به تعداد s انتخاب شده باشد.

$$\Rightarrow Pr\big(\underbrace{\exists t\le n : UCB_1(t,\delta) \le \mu_1}_{A}\big) \le \sum_{t=1}^{n} Pr\left(UCB_1(t,\delta)\le\mu_1\right) \le n\delta$$

$$\Longrightarrow Pr(A) \le n\delta$$

با ترکیب دو بخش قبل داریم که،

$$Pr(G_i^c) \le P(A) + P(B) \le n\delta + \exp\left(-\frac{u_i c^2 \Delta_i^2}{2}\right)$$

۵) با استفاده از نتایج بخش‌های قبل داریم که،

$$\mathbb{E}\{T_i(n)\} \le u_i + n\left(n\delta + \exp\left(-\frac{u_i c^2\Delta_i^2}{2}\right)\right)$$

با انتخاب $\delta = 1/n^2$، حالا می‌خواهیم $u_i$ را انتخاب کنیم که به کران مورد نظر برسیم،

قرار می‌دهیم $u_i = \frac{2\ln(n)}{c^2\Delta_i^2} \Leftarrow$ خواهیم داشت که،

$$\exp\left(-\frac{u_i c^2\Delta_i^2}{2}\right) = \exp(-\ln(n)) = \frac{1}{n}$$

$$\Rightarrow \mathbb{E}\{T_i(n)\} \leq \frac{۲\ln(n)}{c^۲\Delta_i^۲} + n\left(\frac{1}{n}+\frac{1}{n}\right) = ۲ + \frac{۲\ln(n)}{c^۲\Delta_i^۲} \leq ۳ + \frac{14\ln(n)}{\Delta_i^۲}$$

کافیه $c$ را هم بگیریم ۱/۲

h) اگر $\delta = 1/n^۲$ باشد، آنگاه داریم که

$$R_n \leq ۳\sum_{i=1}^{K}\Delta_i + \sum_{i:\Delta_i>0}\frac{14\ln(n)}{\Delta_i} \qquad ???$$

داشتیم: $R_n = \sum_i \Delta_i\, \mathbb{E}\{T_i(n)\} \leq \sum \Delta_i\left(۳ + \frac{14\ln(n)}{\Delta_i^۲}\right) =$

$$۳\sum_{i=1}^{K}\Delta_i + \sum_{i:\Delta_i>0}\frac{14\ln(n)}{\Delta_i}$$

$$\Rightarrow R_n \leq ۳\sum_{i=1}^{K}\Delta_i + \sum_{i:\Delta_i>0}\frac{14\ln(n)}{\Delta_i}$$

i) در نظر می‌گیریم $\Delta_* = ۴\sqrt{\frac{K\log(n)}{n}}$ حالا براساس $\Delta_*$ میاییم دو تا مجموعه می‌سازیم:

$$L = \{i : \Delta_i \geq \Delta_*\}, \qquad S = \{i,\ 0<\Delta_i<\Delta^*\}$$

$$\sum_{i\in L}\frac{14\log(n)}{\Delta_i} \leq \frac{14\log(n)}{\Delta_*}|L| \leq \frac{14\log(n)\,|L|}{۴\sqrt{\frac{K\log(n)}{n}}} = ۴\sqrt{nK\log n}$$

$$\sum_{i\in S}\Delta_i T_i(n) \leq n\Delta^* = ۴n\sqrt{\frac{K\log(n)}{n}}$$

$$\Rightarrow R_n \leq ۳\sum_{i=1}^{K}\Delta_i + ۸n\sqrt{\frac{K\log(n)}{n}} = \boxed{۳\sum_{i=1}^{K}\Delta_i + ۸\sqrt{nK\log(n)}}$$

۲.۲ اینبار هم مثل حالت قبل داریم که:

$$R_n = \mathbb{E}\left\{\sum_{t=1}^{n}\sum_{k=1}^{2^t} \mu^* - X_{A_t}^{(k)}\right\}, \quad \mathbb{E}\left\{\sum_{t=1}^{n}\sum_{k=1}^{2^t} X_{A_t}^{(k)}\right\} =$$

$$\mathbb{E}\left\{\sum_{a\in A}\sum_{t=1}^{n}\sum_{k=1}^{2^t} X_a^{(k)}\, \mathbb{I}(A_t = a)\right\} =$$

$$\sum_{a\in A} \mathbb{E}\left\{\sum_{t=1}^{n}\sum_{k=1}^{2^t} X_a^{(k)}\, \mathbb{I}(A_t = a)\right\} = \sum_{a\in A} \mu_a\, \mathbb{E}\{T_a(n)\}$$

$T_a(n)$ هم تعداد باری است که یک arm انتخاب شده است. اینبار نسبت به قبل $2^\ell$ تایی زیاد می‌شود.

$$\Rightarrow R_n = \sum_{a\in A} \Delta_a\, \mathbb{E}\{T_a(n)\}$$

حالا $G_i$, $G_i^c$ مثل قبلاً در نظر باید بگیریم. فقط قبلش داریم که:

چون finite horizon هستیم داریم که

$$L_{max} = \lceil \log_2(n) \rceil$$

از طرفی داریم که،

$$\sum_{\ell=1}^{L_{max}} 2^{\ell-1} = 2^{L_{max}} - 1 \geq n$$

$$T_i(t_{\ell-1}) = 1 + \sum_{\substack{1\leq m<\ell \\ A_m = i}} 2^{m-1}$$

فرض کنیم epoch های ما $1, \dots, \ell-1$ هستند.

$$\Rightarrow UCB_i(t_{\ell-1}, \delta) = \hat{\mu}_i(t_{\ell-1}) + \sqrt{\frac{2\ln(1/\delta)}{T_i(t_{\ell-1})}}$$

$$A_\ell = \underset{1\leq i\leq k}{\operatorname{argmax}}\ UCB_i(t_{\ell-1}, \delta)$$

$A_\ell$ هم $2^{\ell-1}$ بار انجام میشه پشت سرهم.

Good، حالا داریم که ای arm ماکزیمم $u_i$ بار نهایتاً انتخاب شده است ⇐

دور آخری که arm اون انتخاب شده است عبارت است از: $L_i = \lceil \log_2(u_i) \rceil + 1$

طبق تعریف سقف داریم که:

$$2^{L_i - 1} \le 2^{\lceil \log_2(u_i) \rceil} \le 2u_i$$

$$\rightarrow 2^{L_i} \le 4u_i$$

حالا خواهیم داشت که:

$$\underline{T_i(n) \le 2^{L_i} \le 4u_i}$$

برای حالت Good بالا.

حالا داریم که:

$$\mathbb{E}\{T_i(n)\} \le 4u_i + n\Pr(G_i^c)$$

فقط این بخش کوچکتر مساوی اش مثل بخش غیر اختیاری هست اثباتش و ~~از ... آن خود ...~~ جلوتر بعضی اشاره میشه ***

حالا میاییم $u_i$ بهینه را مثل بخش قبل انتخاب می‌کنیم: $\left\lceil \frac{16\ln(n)}{\Delta_i^2} \right\rceil$

$$u_i \ge \frac{16\ln(n)}{\Delta_i^2} \Rightarrow \sqrt{\frac{4\ln(n)}{u_i}} \le \sqrt{\frac{4\ln(n)}{16\ln(n)/\Delta_i^2}} = \frac{\Delta_i}{2}$$

~~good event~~

حالا $G_i$ good event را در نظر بگیریم:

$$E_\ell^{(1)} = \left\{ \hat{\mu}_1(t_{\ell-1}) + \sqrt{\frac{4\ln(n)}{T_1(t_{\ell-1})}} \ge \mu^* \right\}, \left\{ \ell = 1 \_\_\_ L_{max} \right\}$$

$$E_i^{(2)} = \left\{ \hat{\mu}_i(u_i) + \sqrt{\frac{4\ln(n)}{u_i}} < \mu^* \right\}$$

حالا داریم که،

$$G_i = \left( \bigcap_{\ell=1}^{L_{max}} E_\ell^{(1)} \right) \cap E_i^{(2)}$$

حالا باز داریم که،

$$Pr(G_i^c) \leq Pr\Big( \underbrace{\bigcup_{\ell=1}^{L_{max}} (E_\ell^{(1)})^c}_{A} \Big) + Pr\Big( \underbrace{(E_i^{(2)})^c}_{B} \Big)$$

$$Pr\left( \hat{\mu}(t_{\ell-1}) + \sqrt{\frac{4\ln(n)}{T_1(t_{\ell-1})}} < \mu^* \right) \longrightarrow$$

هافدینگ میگفت، $Pr(\hat{\mu}_1(t) \leq \mu^* - \varepsilon) \leq \exp(-\frac{t\varepsilon^2}{2})$

حالا $\varepsilon$ را $\sqrt{\frac{4\ln(n)}{t}}$ می‌گیریم، خواهیم داشت که

$$Pr\left( \hat{\mu}_1(t_{\ell-1}) + \sqrt{\frac{4\ln(n)}{T_1(t_{\ell-1})}} < \mu^* \right) \leq 1/n^2$$

حالا از آنجایی که $L_{max} = \lceil \log_2(n+1) \rceil$ خواهیم داشت که،

$$A = Pr(\exists \ell \in \{1, \ldots, L_{max}\} : E_\ell^{(1)} \text{ fails}) \leq L_{max} \cdot \frac{1}{n^2} = \frac{\lceil \log_2(n+1) \rceil}{n^2}$$

$$\Rightarrow P(A) \leq \frac{\lceil \log_2(n+1) \rceil}{n^2}$$

$$B : \left\{ \hat{\mu}_i(u_i) \geq \mu^* - \sqrt{\frac{4\ln(n)}{u_i}} \right\} = \left\{ \hat{\mu}_i(u_i) \geq \mu_i + \Delta_i - \sqrt{\frac{4\ln(n)}{u_i}} \right\}$$

داشتیم که

$$u_i \geq \frac{16\ln(n)}{\Delta_i^2} \Rightarrow \sqrt{\frac{4\ln(n)}{u_i}} \leq \sqrt{\frac{4\ln(n)}{\frac{16\ln(n)}{\Delta_i^2}}} = \frac{\Delta_i}{2}$$

$$\Longrightarrow \Delta_i - \sqrt{\frac{4\ln(n)}{u_i}} \geq \frac{\Delta_i}{2} \longrightarrow B \subseteq \left\{ \hat{\mu}_i(u_i) \geq \mu_i + \frac{\Delta_i}{2} \right\}$$

حالا داریم که،

$$\Pr\left(\hat{\mu}_i(u_i)-\mu_i \ge \frac{\Delta_i}{2}\right) \le \exp\left(-\frac{u_i\left(\frac{\Delta_i}{2}\right)^2}{2}\right) = \exp\left(-\frac{u_i\Delta_i^2}{8}\right)$$

$$\le \exp\left(-\frac{\Delta_i^2}{8}\cdot\frac{16\ln(n)}{\Delta_i^2}\right) \le \exp(-2\ln(n)) = \frac{1}{n^2}$$

$$\Rightarrow \Pr(G_i^c) \le \frac{\lceil \log_2(n+1)\rceil}{n^2} + \frac{1}{n^2}$$

*** فرض کنیم $G_i$ برقرار باشه، نشون میدیم که arm i فقط در $\ell=1,\dots,L_i$ می‌تونه انتخاب بشه.

$L_i = \lceil \log_2(u_i)\rceil + 1$ and thus $\ell \le L_i \Rightarrow 2^{\ell-1} \le 2^{L_i-1} < 2u_i$

فرض کنیم در $\ell \ge \lceil \log_2(u_i)\rceil$ ، arm i ام انتخاب شود خواهیم داشت که،

در ابتدای $\ell$ ام epoch ام UCB برای arm i کمتر از $\mu^*$ خواهد بود درحالیکه UCB برای arm 1 حداقل $\mu^*$ است ⇐ اول argmax نیاد i امین arm را انتخاب کند.

در شروع اپاک $\ell$ ام:

$$\hat{\mu}_1(t_{\ell-1}) + \sqrt{\frac{4\ln(n)}{T_1(t_{\ell-1})}} \ge \mu^*$$

$$UCB_1(t_{\ell-1},\delta) \ge \mu^*$$

درحالیکه روی $G_i$ میدانیم که وقتی $T_i(t)$ به $u_i$ میرسه ما داریم که،

$$\hat{\mu}_i(u_i) + \sqrt{\frac{4\ln(n)}{u_i}} < \mu^* \quad \text{for any time } t \ge u_i \longrightarrow$$

$$\hat{\mu}_i(t) + \sqrt{\frac{4\ln(n)}{t}} \le \hat{\mu}_i(u_i) + \sqrt{\frac{4\ln(n)}{u_i}} < \mu^*$$

اگر در ابتدای epoch $\ell$ داشته باشیم که ، $T_i(t_{\ell-1}) \geq u_i$

$$\Rightarrow UCB_i(t_{\ell-1}, \delta) = \hat{\mu}_i(t_{\ell-1}) + \sqrt{\frac{4\ln(n)}{T_i(t_{\ell-1})}} < \mu^*$$

$$T_i(t_{\ell-1}) = \sum_{\substack{1 \leq m < \ell \\ A_m = i}} 2^{m-1} \geq 2^{(\ell-1)-1} = 2^{\ell-2}$$

وقتی به اول epoch $\ell$ میرسیم تعداد کل pull ها از arm i حداقل است.

همچنین داریم که $2^{m-1} \geq 2^{\ell-2}$. چون m نهایتاً $\ell-1$ است و $m-1$ نهایتاً $\ell-2$ است.

$$\Rightarrow \forall \ell, \quad \ell > \lceil \log_2(u_i) \rceil + 1 \rightarrow$$ ایندکس اولین epoch در دور $\ell$ ام.

$$\Rightarrow \ell-2 \geq \log_2(u_i) + 1 - 1 \Rightarrow 2^{\ell-2} \geq u_i$$

$$T_i(t_{\ell-1}) \geq 2^{\ell-2} \geq u_i \Longrightarrow$$

$$UCB_i(t_{\ell-1}) < \mu^* \leq UCB_1(t_{\ell-1})$$

$$L_i = \lceil \log_2(u_i) \rceil + 1 \longrightarrow$$ طرف تا این دور می‌تونه بره.

$$T_i(n) \leq \sum_{\ell=1}^{L_i} 2^{\ell-1} = 2^{L_i} - 1 \leq 2\left(2^{\lceil \log_2(u_i) \rceil}\right) - 1 \leq 4u_i - 1 \leq 4u_i$$

$$\Rightarrow \boxed{\text{if } G_i \text{ holds } T_i(n) \leq 4u_i}$$

حالا ترکیب می‌کنیم، خواهیم داشت که،

$$G_i:\ T_i(n) \le 4u_i$$

$$\Pr(G_i^c) \le \boxed{\frac{\lceil \log_2(n+1) \rceil}{n^2} + \frac{1}{n^2}}$$

$$\Rightarrow \mathbb{E}\{T_i(n)\} \le 4u_i + n\Pr(G_i^c) \le 4u_i + n.(\quad)$$

قرار می‌دهیم، $u_i = \left\lceil \frac{16 \ln(n)}{\Delta_i^2} \right\rceil \rightsquigarrow u_i \le \frac{16\ln(n)}{\Delta_i^2} + 1 \Rightarrow$

$$4u_i + \frac{\lceil \log_2(n+1) \rceil + 1}{n} \le 4\left(\frac{16\ln(n)}{\Delta_i^2} + 1\right) + \frac{\lceil \log_2(n+1) \rceil + 1}{n}$$

$$= \frac{64 \ln(n)}{\Delta_i^2} + 1 + \frac{\lceil \log_2(n+1) \rceil + 1}{n}$$

$$\Longrightarrow R_n = \sum_{\Delta_i>0} \Delta_i \mathbb{E}\{T_i(n)\} \le \sum_{\Delta_i>0} \Delta_i\left(\frac{64\ln(n)}{\Delta_i^2} + \dots\right)$$

$$= 64 \sum_{i:\Delta_i>0} \frac{\ln(n)}{\Delta_i} + \frac{\lceil \log_2(n+1) \rceil + 1}{n} \sum_{i:\Delta_i>0} \Delta_i$$

$$\Rightarrow R_n \le \left(64 \sum_{i:\Delta_i>0} \frac{\ln(n)}{\Delta_i}\right) + \left(\frac{\lceil \log_2(n+1) \rceil + 1}{n} \sum_{\Delta_i>0} \Delta_i\right)$$

حالا مثل بخش آخر مسئله/ قبل پیش میریم، خواهیم داشت،

$$\bar{\Delta} = 4\sqrt{\frac{K \ln(n)}{n}}$$

Scanned with CamScanner

، $\Delta_i \geq \bar{\Delta}$

$$\Rightarrow \sum_{\Delta_i>0} \frac{\ln(n)}{\Delta_i} \leq K \frac{\ln(n)}{\bar{\Delta}} = \frac{\sqrt{Kn\ln(n)}}{4}$$

$$\rightsquigarrow 64 \sum_{\Delta_i>0} \frac{\ln(n)}{\Delta_i} \leq 64 \frac{\sqrt{Kn\ln(n)}}{4} = 16\sqrt{Kn\ln(n)}$$

$0 < \Delta_i < \bar{\Delta}$ خواهیم داشت که،

$$\cancel{\sum \Delta_i \leq \sum \bar{\Delta}}$$

$$\sum \Delta_i T_i(n) \leq n\bar{\Delta} \leq 4n\sqrt{\frac{K\ln(n)}{n}} = 4\sqrt{nK\ln(n)}$$

$$\Rightarrow \boxed{R_n \leq \frac{\lceil \log_2(n+1)\rceil + 1}{n} \sum_{\Delta_i>0} \Delta_i + 20\sqrt{Kn\ln(n)}}$$

Scanned with CamScanner

Randomized weighted majority Algo ۱.۳

a) داریم که $S_t = \sum_{i=1}^{N} w_i(t)$ ، $\tilde{m}_t$ معلوم می‌کند که اشتباه رخ داده یا نه،

حالا داریم که:

$$i_t \text{ is wrong}, \quad S_{t+1} = S_t - \varepsilon w_{i_t}(t)$$

$$i_t \text{ is true}, \quad S_{t+1} = S_t$$

$$\Rightarrow E\{S_{t+1} | S_t\} = S_t - \varepsilon . E\{w_{i_t}(t)\, \tilde{m}_t\}$$

همچنین احتمال انتخاب $i_t$ برابر است با $\frac{w_i(t)}{S_t}$ ، $\tilde{m}_t = 1$ اگر $i_t$ غلط باشد.

$$\Rightarrow E\{w_{i_t}(t)\, \tilde{m}_t\} = \sum_{i=1}^{N} \frac{w_i(t)}{S_t} . \tilde{m}_t(i) . w_i(t)$$

همچنین داریم که:

$$P(\tilde{m}_t = 1) = \text{احتمال اینکه expert انتخابی اشتباه کند} = \sum_{i=1}^{N} \frac{w_i(t)}{S_t} \tilde{m}_t(i)$$

$$\Rightarrow E\{S_{t+1} | S_t\} = S_t(1 - \varepsilon P(\tilde{m}_t = 1))$$

حالا روی $S_t$ ها میانگین $E$ می‌گیریم،

$$\boxed{E\{S_{t+1}\} = E\{S_t\}(1 - \varepsilon P(\tilde{m}_t = 1))}$$

b) داریم که از بخش الف)

$$E\{S_{T+1}\} = E\{S_T\}(1 - \varepsilon P(\tilde{m}_t = 1)) = \cdots$$

$$E\{S_1\} \prod_{t=1}^{T} (1 - \varepsilon . P(\tilde{m}_t = 1))$$

وزن اولیه همه 1 است $\leftarrow$ $S_1 = N$ پس خواهیم داشت که:

$$\mathbb{E}\{S_{T+1}\} = N \prod_{t=1}^{T} (1 - \varepsilon P(\tilde{m}_t = 1))$$

می‌دانیم از ریاضی 1 که $1 - x \leq e^{-x}$ پس خواهیم داشت که

$$\boxed{\mathbb{E}\{S_{T+1}\} \leq N \exp\left(\varepsilon \sum_{t=1}^{T} P(\tilde{m}_t = 1)\right)}$$

c) وزن اولیه $w_i$ ها همگی 1 اند و هر بار که غلط داریم $(1-\varepsilon)$ برابر می‌شود $\Leftarrow$

در نهایت خواهیم داشت که اگر $M_i$ بار غلط داشتیم $\Leftarrow$

$$w_i(T+1) = w_i(0)(1-\varepsilon)^{M_i} = (1-\varepsilon)^{M_i}$$

چونکه $w_i(T+1)$ کوچکتر از جمعشان است $\Leftarrow$ $w_i(T+1) \leq S_{T+1}$

$$\Rightarrow (1-\varepsilon)^{M_i} \leq S_{T+1} \xrightarrow{\mathbb{E} \text{ می‌گیریم}}$$

$$(1-\varepsilon)^{M_i} \leq \mathbb{E}\{S_{T+1}\} \leq N \cdot \exp\left(-\varepsilon \sum_{t=1}^{T} P(\tilde{m}_t = 1)\right)$$

حالا $\log$ می‌گیریم، خواهیم داشت که:

$$M_i \ln(1-\varepsilon) \leq \ln N - \varepsilon\, \mathbb{E}\{M\} \rightarrow$$

$$\varepsilon\, \mathbb{E}\{M\} \leq \ln N - M_i \underbrace{\ln(1-\varepsilon)}$$

$$\Rightarrow \mathbb{E}\{M\} \leq -M_i \left(\frac{\ln(1-\varepsilon)}{\varepsilon}\right) + \frac{\ln N}{\varepsilon}$$

حالا برای $\varepsilon \leq \frac{1}{2}$ داریم که:

$$\ln\left(\frac{1}{1-\varepsilon}\right) \leq \varepsilon + \varepsilon^2$$

$$\Rightarrow \boxed{\mathbb{E}\{M\} \leq (1+\varepsilon) M_i + \frac{\ln N}{\varepsilon} \qquad \forall i \in [N]}$$

پس از بخش قبل داریم که:

$$\mathbb{E}\{M\} \leq (1+\varepsilon) M_i + \frac{\ln N}{\varepsilon}$$

حالا با انتخاب $\varepsilon = \sqrt{\frac{\ln N}{T}}$ خواهیم داشت که:

$$\mathbb{E}\{M\} \leq M_i + \sqrt{\frac{\ln N}{T}}\, M_i + \sqrt{T \ln N}$$

از طرفی $M_i \leq T$ است $\Leftarrow$

$$\mathbb{E}\{M\} \leq M_i + \sqrt{\frac{\ln N}{T}}\, T + \sqrt{T \ln N} = M_i + 2\sqrt{T \ln N}$$

این به ازای هر $i$ ای برقرار است $\leftarrow$ می‌توانیم مینیمم را هم قرار دهیم:

$$\Rightarrow \mathbb{E}\{M\} \leq \min_i M_i + 2\sqrt{T \ln N}$$

همچنین regret عبارت است از:

$$\text{Regret} = \mathbb{E}\{M\} - \min_i M_i \leq 2\sqrt{T \ln N}$$