Lower Bounds

$(\Omega, \mathcal{F}, P) \qquad (\Omega, \mathcal{F}, Q)$

$$H_0: X_t \sim P = \mathcal{N}(0,1) \qquad t = 1, 2, \dots, T$$
$$H_1: X_t \sim Q = \mathcal{N}(\Delta,1)$$

How big does T need to be to decide whether H_0, or H_1 is true?

If we define $\hat{\mu}_T = \frac{1}{T}\sum_{t=1}^{T} X_t$ then w.p. $\geq 1-\delta$

$$|\hat{\mu}_T - \mathbb{E}[X_1]| \leq \sqrt{\frac{2\log(2/\delta)}{T}} = \Delta/2$$

Choose $T = 8\Delta^{-2}\log(2/\delta)$, compute $\hat{\mu}_T$

and decide H_1 if $\hat{\mu}_T > \Delta/2$, H_0 if $\hat{\mu}_T < \Delta/2$.

This is correct w.p. $\geq 1-\delta$.

Def The total Variation between P, Q is

$$TV(P,Q) = \sup_{A \in \mathcal{F}} |P(A) - Q(A)| = \int |p(x) - q(x)|\, dx$$

Def The Kullback-Liebler Divergence btwn P, Q is

$$KL(P|Q) = \int \log\left(\frac{dP}{dQ}(\omega)\right) dP(\omega) = \int \log\left(\frac{p(x)}{q(x)}\right) p(x)\, dx$$

Lemma (Pinsker's inequality) $TV(P,Q) \le \sqrt{KL(P\|Q)/2}$.

$\le \sqrt{1 - e^{-KL(P\|Q)/2}}$

Aside: likelihood of $X_1, \ldots, X_T$ under $\mathcal{N}(\mu, 1)$

is $\prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(X_t-\mu)^2}{2}\right)$ so likelihood

ratio of H_1 to H_0 is $\prod_{t=1}^{T} \frac{\exp\left(-\frac{(X_t-\Delta)^2}{2}\right)}{\exp\left(-\frac{X_t^2}{2}\right)} = \frac{q(\{X_t\}_{t=1}^T)}{q(\{X_t\}_{t=1}^T)}$

The likelihood ratio test (LRT) says

pick $q(H_1)$ if $\prod_{t=1}^{T} \frac{q(x_t)}{p(x_t)} > \frac{1}{\delta}$ and $p(H_0)$ otherwise.

Make an error under P (i.e. $X_t \sim P \ \forall t$)

w/ prob

$$\mathbb{P}_{X_t \sim P}\left(\prod_{t=1}^{T} \frac{q(x_t)}{p(x_t)} > \frac{1}{\delta}\right) \le \delta \cdot \mathbb{E}_P\left[\prod_{t=1}^{T} \frac{q(x_t)}{p(x_t)}\right] = \delta.$$

$X_1 \sim P$
$$\mathbb{E}_{X_1 \sim P}\left[\frac{q(X_1)}{P(X_1)}\right] = \int \cancel{p(x)} \cdot \frac{q(x)}{\cancel{p(x)}} \, dx = 1$$

$$\mathbb{P}_{X_t \sim q}\left(\prod_{t=1}^{T} \frac{q(x_t)}{p(x_t)} < \delta\right) \le \delta.$$

$$\mathbb{P}_{x_t \sim p}\left(\prod_{t=1}^{T} \frac{q(x_t)}{p(x_t)} > \frac{1}{\delta} \right) = \mathbb{P}_{x_t \sim p}\left(\sum_{t=1}^{T} \log\left(\frac{q(x_t)}{p(x_t)}\right) > \log(1/\delta) \right)$$

$$\mathbb{E}_{x_t \sim q}\left[\sum_{t=1}^{T} \log\left(\frac{q(x_t)}{p(x_t)}\right) \right] = T \, KL(Q|P)$$

Under Q, $\quad T \simeq \dfrac{\log(1/\delta)}{KL(Q|P)}$

$P = \mathcal{N}(0,1)$
$Q = \mathcal{N}(\Delta,1)$ $\qquad KL(Q,P) = \Delta^2/2$

Sequential Probability Ratio Test (SPRT)

$$\tau = \min\left\{ T : \sum_{t=1}^{T} \log\left(\frac{q(x_t)}{p(x_t)}\right) > \log(1/\delta) \right\}$$

$$\log(1/\delta) \simeq \underbrace{\mathbb{E}_q\left[\sum_{t=1}^{\tau} \log\left(\frac{q(x_t)}{p(x_t)}\right) \right] = \mathbb{E}_q[\tau] \, KL(q|P)}$$

Wald's Identity.

$$\mathbb{E}_q[\tau] \le \frac{\log(1/\delta)}{KL(q|P)}.$$

$$\mathbb{E}_p[\tau] \le \frac{\log(1/\delta)}{KL(p|q)}$$

Theorem For every $T \geq n$ $\exists$ bandit instance ν_θ where $\nu_{\theta,i} = \mathcal{N}(\theta_i, 1)$ such that for a policy π:

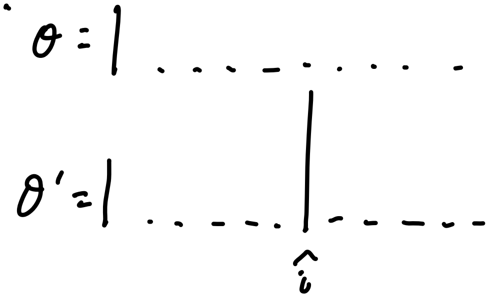
$$\inf_{\pi} \max_{\theta} \max_{i} \mathbb{E}_\theta \left[\sum_{t=1}^{T} X_{t,i} - X_{t,I_t} \right] = \inf_{\pi} \max_{\theta} \max_{i} \mathbb{E}_\theta \left[\sum_{t=1}^{T} \theta_i - \theta_{I_t} \right] \geq \sqrt{\frac{(n-1)T}{256}}.$$

Proof Let $\theta = (\Delta, 0, \ldots, 0)$. For any π

$\exists\ \hat{i} \in [n] \setminus \{1\}$ s.t. $\mathbb{E}_\theta[T_{\hat{i}}] \leq \frac{T}{n-1}$.

Define $\theta' = (\Delta, 0, \ldots, 2\Delta, 0, \ldots 0)$

$\uparrow \hat{i}$ location.



$\mathbb{P}(X > \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}$

$$\mathbb{E}_{\theta'}[R_T] = \sum_{i=1}^{n} (\theta'_{\hat{i}} - \theta'_i) \mathbb{E}_{\theta'}[T_i] \geq (\theta'_{\hat{i}} - \theta'_1) \mathbb{E}_{\theta'}[T_1]$$

$$\geq (\theta'_{\hat{i}} - \theta'_1) \mathbb{P}_{\theta'}(T_1 > T/2)\, T/2$$

$$= \Delta \mathbb{P}_{\theta'}(T_1 > T/2)\, T/2$$

$$\mathbb{E}_\theta[R_T] = \sum_{i=1}^{n} (\theta_1 - \theta_i) \mathbb{E}_\theta[T_i] = \sum_{i=2}^{n} \Delta \mathbb{E}_\theta[T_i] = \Delta \mathbb{E}_\theta\left[\sum_{i \neq 1} T_i\right]$$

$$\geq \Delta \mathbb{P}_\theta\left(\sum_{i \neq 1} T_i > T/2\right) T/2$$

$$= \Delta \mathbb{P}_\theta(T_1 < T/2)\, T/2$$

$$= \Delta \left(1 - \mathbb{P}_\theta(T_1 > T/2)\right) T/2$$

$$\max_{\mu \in \{\theta, \theta'\}} \mathbb{E}_\mu [R_T] \geq \frac{1}{2}\left(\mathbb{E}_{\theta'}[R_T] + \mathbb{E}_\theta [R_T] \right)$$

$$\geq \frac{\Delta T}{4}\left(\mathbb{P}_{\theta'}(T_1 > T/2) + 1 - \mathbb{P}_\theta (T_1 > T/2) \right)$$

$$\geq \frac{\Delta T}{4}\left(1 - \left| \mathbb{P}_\theta (T_1 > T/2) - \mathbb{P}_{\theta'}(T_1 > T/2) \right| \right)$$

$$\geq \frac{\Delta T}{4}\left(1 - TV(\mathbb{P}_\theta, \mathbb{P}_{\theta'}) \right)$$

$$\geq \frac{\Delta T}{4}\left(1 - \sqrt{KL(\mathbb{P}_\theta \| \mathbb{P}_{\theta'})/2} \right)$$

$$\geq \frac{\Delta T}{4}\left(1 - \sqrt{\Delta^2 \mathbb{E}_\theta [T_{\hat{i}}]} \right)$$

$$\geq \frac{\Delta T}{4}\left(1 - \sqrt{\Delta^2 \frac{T}{n-1}} \right) \qquad \Delta = \sqrt{\frac{n-1}{4T}}$$

$$\geq \sqrt{\frac{(n-1)T}{256}}$$

$\boxed{\text{Lemma}}$ Let $\nu = (\nu_1, \ldots, \nu_n)$ and ν' be two bandit instances, Fix policy π and let $\mathbb{P}_\nu$ be measure under ν, and $\mathbb{P}_{\nu'}$ under ν'. Then

$$KL(\mathbb{P}_\nu \| \mathbb{P}_{\nu'}) = \sum_{i=1}^{n} \mathbb{E}_\nu [T_i(T)] \, KL(\nu_i \| \nu_i')$$

Furthermore, if τ is a stopping time, the above holds w/ $T = \tau$.

$\boxed{\text{Proof}}$ $(I_1, X_1, I_2, X_2, I_3, X_3, \ldots, I_T, X_T) = Z$

$$d\mathbb{P}_\nu(z) = \prod_{t=1}^{T} \nu_{I_t}(X_t)\, \pi(I_t \mid \{I_s, X_s\}_{s<t})$$

$$KL(\mathbb{P}_\nu \mid \mathbb{P}_{\nu'}) = \mathbb{E}_\nu\left[\log\left(\frac{d\mathbb{P}_\nu(z)}{d\mathbb{P}_{\nu'}(z)}\right)\right]$$

$$= \mathbb{E}_\nu\left[\log\left(\frac{\prod_t \nu_{I_t}(X_t)}{\prod_t \nu'_{I_t}(X_t)}\right)\right]$$

$$= \mathbb{E}_\nu\left[\sum_{t=1}^{T} \log\left(\frac{\nu_{I_t}(X_t)}{\nu'_{I_t}(X_t)}\right)\right]$$

$$\mathbb{E}_\nu\left[\log\left(\frac{\nu_{I_t}(X_t)}{\nu'_{I_t}(X_t)}\right)\right] = \mathbb{E}_\nu\left[\mathbb{E}\left[\log\left(\frac{\nu_i(X_t)}{\nu_i(X_t)}\right) \middle| I_t = i\right]\right]$$

$$= \mathbb{E}_\nu\left[KL(\nu_{I_t} \mid \nu'_{I_t})\right]$$

$$= \mathbb{E}_\nu\left[\sum_{i=1}^{n}\sum_{t=1}^{T} \mathbb{1}\{I_t = i\}\, KL(\nu_i \mid \nu'_i)\right]$$

$$= \sum_{i=1}^{n} KL(\nu_i \mid \nu'_i)\quad \mathbb{E}_\nu\left[\sum_{t} \mathbb{1}\{I_t = i\}\right]$$

$$= \sum_i \mathbb{E}_\nu[T_i]\, KL(\nu_i \mid \nu'_i).$$

Best arm identification

for $t = 1, 2, \ldots, T$

  Player chooses $I_t \in [n]$

  Nature reveals $X_{t, I_t} \sim \mathcal{N}(\theta_{I_t}, 1)$

  If player stops:

    Output $\hat{i} \in [n]$

Def We say an algorithm is δ-correct if

$$\mathbb{P}_{\theta_*}\left(\hat{i} = \arg\max_i \theta_{*,i}\right) \geq 1 - \delta \quad \text{for all } \theta_* \in \mathbb{R}^d.$$

Recall Action Elim. alg identifies $i_* = \arg\max_i \theta_i$ w.p. $\geq 1-\delta$ after just $T = c \sum_{i \neq i_*} \Delta_i^{-2} \log\left(\frac{n \log(\Delta^{-2})}{\delta}\right)$ pulls $\Delta_i = \theta_{i_*} - \theta_i$.

Fix any $\theta \in \mathbb{R}^n$, assume WLOG $i_* = 1$

$$\nu_i = \mathcal{N}(\theta_i, 1) \quad \forall i$$

$$\nu_i^{(j)} = \begin{cases} \nu_i & \text{if } i \neq j \\ \mathcal{N}(\theta_i + \Delta_i + \varepsilon, 1) & \text{if } i = j \end{cases}$$

Lemma (Data Processing Inequality) Given a probability space $(\Omega, \mathcal{F})$ and dist. P, Q R.V. Z w/ push-forward measure P_Z, Q_Z

$$KL(P_Z \| Q_Z) \le KL(P \| Q)$$

$$\underbrace{KL\big(\underbrace{\mathbb{P}_\nu(\hat{i}=1)}_{\ge 1-\delta} \,\|\, \underbrace{\mathbb{P}_{\nu^{(j)}}(\hat{i}=1)}_{<\delta}\big)}_{} \le KL(\mathbb{P}_\nu \| \mathbb{P}_{\nu^{(j)}})$$

$\mathbb{P}_\nu(\hat{i}=1)=p \qquad \mathbb{P}_{\nu^{(j)}}(\hat{i}=1)=q$

$$= p\log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right)$$

$$\cancel{\ge (1-\delta)\log\left(\frac{1-\delta}{\delta}\right) + \delta\log\left(\frac{\delta}{1-\delta}\right)}$$

$$\ge \log\left(\frac{1}{2.4\delta}\right)$$

$$\Rightarrow \quad \log\left(\frac{1}{2.4\delta}\right) \le KL(\mathbb{P}_\nu \| \mathbb{P}_{\nu^{(j)}})$$

$$= \mathbb{E}_\nu[T_j(s)]\, \underbrace{KL(\nu_j \| \nu_j^{(j)})}_{}$$

$(\Delta_j+\varepsilon)^2/2 =$

$\mathcal{N}(\theta_j+\Delta_j+\varepsilon, 1)$

$$\Longrightarrow \quad \mathbb{E}_\nu[T_j(\tau)] \geq \frac{2}{(\Delta_j+\varepsilon)^2}\log\left(\frac{1}{2.4\delta}\right)$$

$$\mathbb{E}_\nu[\tau] \geq \sum_{j=2}^{n} \mathbb{E}_\nu[T_j(\tau)] \geq \sum_{j=2}^{n} \frac{2\log(1/2.4\delta)}{(\Delta_j+\varepsilon)^2}$$

$\varepsilon>0$ was arbitrary, take $\varepsilon \to 0$.

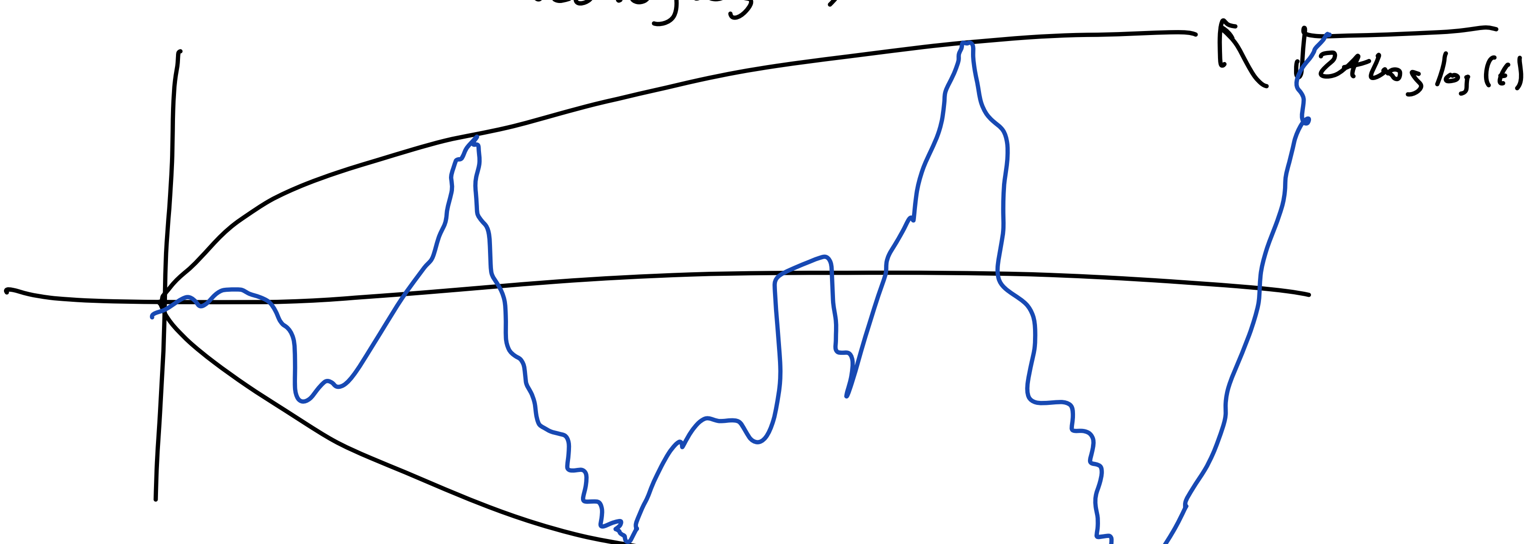
Thm Any δ-correct algorithm requires

$$\mathbb{E}_\nu[\tau] \geq \sum_{i\neq i_\nu} \frac{2\log(1/2.4\delta)}{\Delta_i^2}.$$

Lemma Law of the Iterated Logarithm (LIL)

Let $Z_t \sim \mathcal{N}(0,1)$. and $S_t = \sum_{s=1}^{t} Z_s$

Then $$\limsup_{t\to\infty} \frac{S_t}{\sqrt{2t\log\log(t)}} = 1 \quad \text{a.s.}$$



$$\mathbb{E}[|S_t|] \asymp \sqrt{t}$$

Using similar techniques one can show:

Thm Fix any alg s.t. $\max_\theta \mathbb{E}_\theta[R_T] \lesssim T^{1-\alpha}$ for any $\alpha > 0$

then

$$\lim_{T\to\infty} \frac{\mathbb{E}_\theta[R_T]}{\log(T)} \gtrsim \sum_{i \neq i_\theta} \Delta_i^{-1}$$