Let $X_1, X_2, \ldots$ be a $\mathcal{F}_t$-adapted random sequence.

Consider hypothesis test, for known densities p_0, p_1:

$$H_0 : X_t \overset{iid}{\sim} p_0 \quad \forall t$$
$$H_1 : X_t \overset{iid}{\sim} p_1 \quad \forall t$$

Goal: define a stopping time $\tau \in \mathbb{N}$ such that at time τ declare hypothesis.

Define likelihood ratio @ time t, $L_t = \prod_{s=1}^{t} \frac{p_1(X_s)}{p_0(X_s)}$.

Let $\mathbb{E}_i[\cdot]$, $\mathbb{P}_i[\cdot]$ be expectation and probability laws under H_i.

Note L_t is an $\mathcal{F}_t$-adapted martingale under H_0:

$$\mathbb{E}_0[L_{t+1} \mid \mathcal{F}_t] = L_t \cdot \mathbb{E}_0\left[\frac{p_1(X_{t+1})}{p_0(X_{t+1})} \,\Big|\, \mathcal{F}_t\right]$$
$$= L_t \cdot \int_x \frac{p_1(x)}{\cancel{p_0(x)}} \cancel{p_0(x)}\, dx$$
$$= L_t.$$

We can apply maximal inequality for super-marting

$$\Longrightarrow \quad \mathbb{P}_0(\exists t \in \mathbb{N} : L_t > 1/\delta) \le \delta$$

But also note L_t^{-1} is a martingale under H_1.

$$\Longrightarrow \quad \mathbb{P}_1(\exists t \in \mathbb{N} : L_t^{-1} > 1/\delta) \le \delta$$
$$\Longrightarrow \quad \mathbb{P}_1(\exists t \in \mathbb{N} : L_t < \delta) \le \delta$$

$$\tau = \min\{t : L_t \notin [\delta, 1/\delta]\}$$

Example : Gaussian $\quad P_0 = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad , \quad P_1 = \frac{1}{\sqrt{2\pi}} e^{-(x-\Delta)^2/2}$

$\equiv \mathcal{N}(0,1) \qquad \equiv \mathcal{N}(\Delta,1)$

$$L_t = \prod_{s=1}^{t} \frac{P_1(X_s)}{P_0(X_s)} = \prod_{s=1}^{t} \exp\left(-(X_s-\Delta)^2/2 + X_s^2/2\right)$$

$$= \prod_{s=1}^{t} \exp\left(X_s\Delta - \Delta^2/2 \right)$$

$$= \exp\left(\Delta\left(\sum X_s\right) - t\Delta^2/2\right)$$

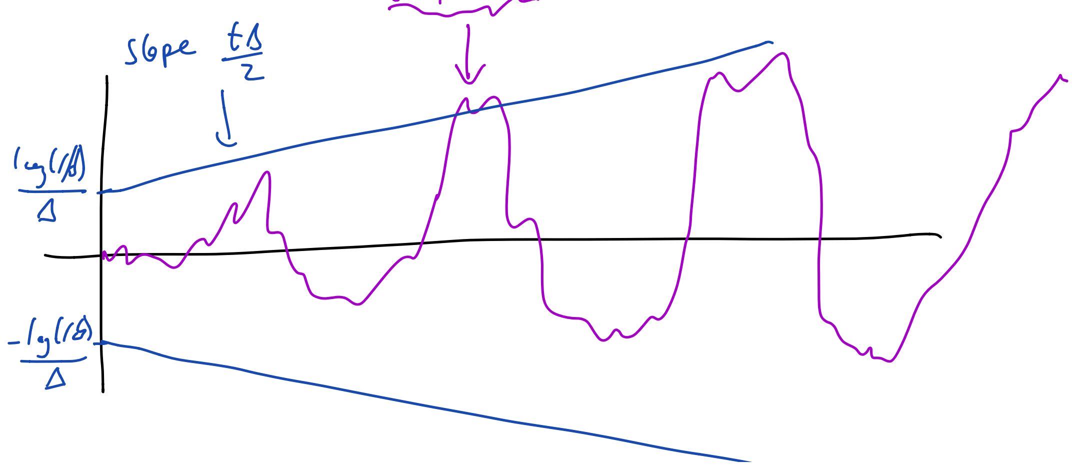
$$\mathbb{P}_0\left(\exists t \in \mathbb{N} : \sum X_s \geq \frac{t\Delta}{2} + \frac{\log(1/\delta)}{\Delta}\right)$$

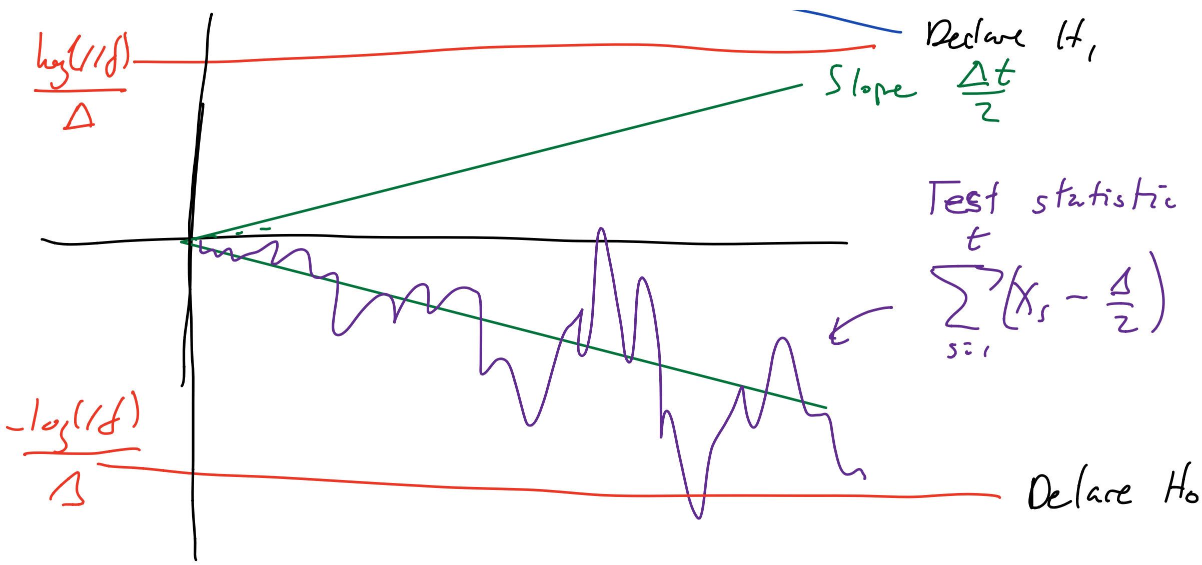
$$= \mathbb{P}_0\left(\exists t \in \mathbb{N} : L_t \geq 1/\delta\right) \leq \delta$$

$$\mathbb{P}_1\left(\exists t \in \mathbb{N} : \sum X_s \leq \frac{t\Delta}{2} - \frac{\log(1/\delta)}{\Delta}\right) \leq \delta$$

$$\tau = \min\{t : L_t \notin [\delta, 1/\delta]\}$$

$$= \min\left\{t : \underbrace{\sum_{s=1}^{t} X_s - \frac{t\Delta}{2}}_{} \notin \left[\frac{-\log(1/\delta)}{\Delta}, \frac{\log(1/\delta)}{\Delta}\right]\right\}$$





Conclude that test is correct in the sense that it will not output incorrect hypothesis w.p $\geq 1-\delta$.

When will test stop? That is, $\mathbb{E}_i[\tau]$?

Lemma (Wald's identity). Let X_t be iid w/ mean μ. Let $\tau \in \mathbb{N}$ be a stopping time with $\mathbb{E}[\tau] < \infty$. Then $\mathbb{E}\left[\sum_{t=1}^{\tau} X_t\right] = \mu \cdot \mathbb{E}[\tau]$.

Proof (sketch): if $\mu \in \{-\infty, \infty\}$ statement trivial so assume otherwise. Then we can conclude $\mathbb{E}[|X_t|] < c$ for some $c < \infty$. W/ $\mathbb{E}[\tau] < \infty$ we can apply Doob's optional stopping to martingale

$$M_t = \sum_{s=1}^{t} X_s - t\mu. \quad \mathbb{E}[M_\tau] = \mathbb{E}[M_0] = 0 \qquad \blacksquare$$

To apply Wald's inequality consider

$$\mathbb{E}_0[\operatorname{Log}(L_{\tau'})] = ✪$$

where $\tau' = \min\{t \in \mathbb{N} : L_t < \delta\}$

$$✪ = \mathbb{E}_0\left[\sum_{t=1}^{\tau'} \log\left(\frac{P_1(X_t)}{P_0(X_t)}\right)\right]$$

$$= \mathbb{E}_0[\tau']\, \mathbb{E}_0\left[\log\left(\frac{P_1(X_t)}{P_0(X_t)}\right)\right] = Ⓐ$$

But at τ', we have that

$L_{\tau'} \simeq \delta$ (ignoring the overshoot)

$$\Rightarrow \quad ✪ = \mathbb{E}_0[\log(L_{\tau'})] \simeq \log(\delta) = Ⓑ$$

$$\mathbb{E}_0[\tau'] = \frac{\log(\delta)}{\mathbb{E}_0\left[\log\left(\frac{P_1(X_1)}{P_0(X_1)}\right)\right]}$$

$$= \frac{\log(1/d)}{KL(P_0 | P_1)}.$$

Recall $KL(p|q) = \int p(x) \log\left(\frac{p(x)}{q(x)}\right) dx.$

Similar calculation leads to

$$\mathbb{E}_1[\tau''] = \frac{\log(1/\delta)}{KL(P_1|P_0)}$$

Recall: on HW1 you showed that <u>any</u> procedure that decides between $N(0,1)$ and $N(\Delta,1)$ requires at least $\frac{\log(1/\delta)}{KL(P_1|P_0)}$ samples.

$$= \frac{2\log(1/\delta)}{\Delta^2}$$

Conclude: This test of H_0 vs H_1 is <u>optimal</u> in the sense that the expected stopping time cannot be improved. Sequential probability ratio test (SPRT).

Method of mixture

Consider hypothesis test w/ $\mathbb{E}[X_t] = \mu$

$H_0 : \mu = 0$

$H_1 : \mu \neq 0$

Recall before, we had binary hypothesis test of $L_t = \prod_{s=1}^{t} \frac{p(X_s; \Delta)}{p(X_s; 0)}$. Now we don't know Δ. So we define prior belief over Δ. Let $h(\mu)$ be a density.

Then $$L_t = \int_\mu h(\mu) \prod_{s=1}^{t} \frac{p(X_s; \mu)}{p(X_s; 0)} \, d\mu$$

is a martingale.

Previously we mixed w/ $h \equiv$ Gaussian but this is arbitrary. In practice h can be prior knowledge on effectsize μ.

Fix $x_1, \dots, x_n \in \mathbb{R}^d$; and then observed

$$y_i = \langle \theta^*, x_i \rangle + \eta_i \quad \text{where } \eta_i \sim N(0,1).$$

Then $\hat{\theta} = (X^T X)^{-1} X^T y$. W.p. $\geq 1-\delta$, $z \in \mathbb{R}^d$

$$\frac{\langle \hat{\theta} - \theta^*, z \rangle}{\|z\|_{(X^T X)^{-1}}} \leq \sqrt{2\log(1/\delta)}$$

Using Martingale bound (which allowed for $x_t \in \mathcal{F}_{t-1}$, $y_t \in \mathcal{F}_t$)

$$\langle \hat{\theta} - \theta^*, z \rangle \leq \|z\|_{(X^T X)^{-1}} \cdot \|\hat{\theta} - \theta^*\|_{(X^T X)}$$

$$\leq \|z\|_{(X^T X)^{-1}} \cdot c\sqrt{d + \log(1/\delta)}$$

$$\frac{\langle \hat{\theta} - \theta^*, z \rangle}{\|z\|_{(X^T X)^{-1}}} \leq c\sqrt{d + \log(1/\delta)}$$

Fix $z = \mathbb{1}$ and then construct seq $(x_1, x_2, \dots$

where $x_t \in \mathcal{F}_t$ and $y_t = \langle \theta^*, x_t \rangle + \eta_t$

$$\mathbb{E}\left[\frac{|\langle \hat{\theta} - \theta^*, z \rangle|}{\|z\|_{(X^T X)^{-1}}}\right] \geq \sqrt{d}$$

$$\hat{\theta} = (X^T X)^{-1} X^T \eta = \left(\sum_{i=1}^{n} T_i e_i e_i^T\right)^{-1} \left(\sum_{t=1}^{T} e_{I_t} \eta_t\right)$$

$$\boxed{y_t = \langle \theta^*, x_t \rangle + \eta_t}$$

$\varepsilon_i \in \{-1, 1\}$ w.p. $1/2$. Fix N

$$\tau = \min\left\{N+1, \min\left\{t : \sum_{i=1}^{t} \varepsilon_i \geq 1\right\}\right\}$$

$$\mathbb{E}\left[\frac{1}{\tau}\sum_{i=1}^{\tau}\varepsilon_i\right] > 0$$

For a fixed $n \in \mathbb{N}$ $\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\right] = 0$

$$= \mathbb{E}\left[\sum_{t=1}^{N}\mathbb{1}\{\tau = t\}\frac{1}{\tau}\sum_{i=1}^{\tau}\varepsilon_i\right] + \underbrace{\mathbb{E}\left[\mathbb{1}\{\tau = N+1\}\frac{1}{\tau}\sum_{i=1}^{\tau}\varepsilon_i\right]}$$

$$\geq \mathbb{E}\left[\sum_{t=1}^{N}\mathbb{1}\{\tau = t\}\cdot\frac{1}{t}\right] - \mathbb{E}\left[\left|\frac{1}{N+1}\sum_{i=1}^{N+1}\varepsilon_i\right|\right]$$

$$\geq \mathbb{E}\left[\mathbb{1}\{\tau = 1\}\cdot\frac{1}{1}\right] - \sqrt{\frac{2}{\pi(N+1)}}$$

$$\geq 0.1 -$$

$$\geq c > 0$$

$$\phi = \mathbb{E}_{Z\sim N(0,1)}\left[|Z|\right] \quad \left(\text{I think it } \sqrt{\tfrac{2}{\pi}}\right)$$

$u = x^2/2$

$du = x$

$$2\int_{x\geq 0}\frac{1}{\sqrt{2\pi}}\, x\, e^{-x^2/2}\, dx$$

$$= 2\int_{\cancel{x}\, u=0}\frac{1}{\sqrt{2\pi}}\, e^{-u}\, du = \sqrt{\frac{2}{\pi}}$$