STA 3000 Lecture 4

$H_0 : \theta \in \Theta_0 \subseteq \Theta$

$H_1 : \theta \in \Theta_1 \subseteq \Theta$

$\Theta_0 \cap \Theta_1 = \emptyset$

$\Theta_0 \cup \Theta_1 = \Theta$.

"Critical function" — $\phi(X) = \begin{cases} 1 & \text{reject} \\ \pi(\theta, X) & \text{reject w.p. } \pi. \\ 0 & \text{do not reject} \end{cases}$

Significance level — $\alpha_\phi := \sup_{\theta \in \Theta_0} \mathbb{E}_\theta[\phi(X)]$.

Power — $\beta_\phi(\theta) = \mathbb{E}_\theta[\phi(X)]$ for $\theta \in \Theta_1$.

Goal: keep $\alpha_\phi \le \alpha$ while trying to maximize β.

Simple-vs-simple testing. $\Theta_0 = \{\theta_0\}$ $\Theta_1 = \{\theta_1\}$

$$\alpha_\phi = \int \phi(x) \cdot p_0(x) \, d\mu(x)$$

$$\beta_\phi = \int \phi(x) \, p_1(x) \, d\mu(x).$$

Likelihood ratio $L(x) = \dfrac{p_1(x)}{p_0(x)}$.

Def. (LRT) $\varphi^*(x) = \begin{cases} 1 & L(x) > c \\ \pi & L(x) = c \\ 0 & L(x) < c \end{cases}$.

Lemma (Neymann-Pearson)

(Assuming p_0 and p_1 has common support)

Given $\alpha \in (0,1)$. $\exists$ LRT φ^*_α w/ level $= \alpha$

and it is optimal.

Proof. Under dictionary order, φ^* is monotone dec in $(C, -\pi)$

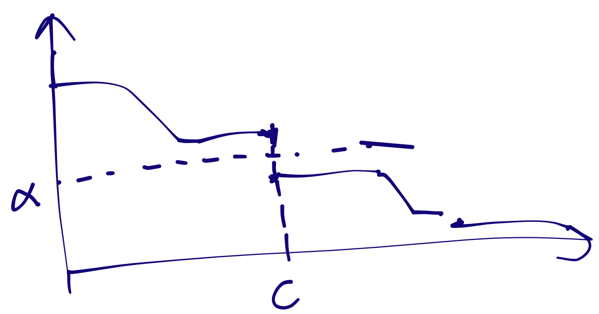
$(C=0, \pi=1) \Rightarrow$ level $=1$

$(C \to +\infty, \pi=0)$. level $= \int p_0(x) \cdot \mathbb{1}\{p_1(x) > c \cdot p_0(x)\} \, d\mu(x)$

$$\le \int \frac{p_1(x)}{c \cdot p_0(x)} \cdot p_0(x) \, d\mu(x)$$

$$= \frac{1}{c} \to 0.$$

$\exists (c_0, \pi_0)$ s.t. level $= \alpha$.



Optimality: For any other test φ w/ level $\le \alpha$.

$$\mathbb{E}_1[\varphi(X)] - c\,\mathbb{E}_0[\varphi(X)]$$

$$= \int_{p_1 \ge cp_0} |p_1 - cp_0| \cdot \phi \, d\mu - \int_{p_1 < cp_0} |p_1 - cp_0| \, \phi \, d\mu.$$

$$\le \int_{p_1 \ge cp_0} (p_1 - cp_0) \, d\mu = \mathbb{E}_1[\phi^*(X)] - c\,\mathbb{E}_0[\phi^*(X)].$$

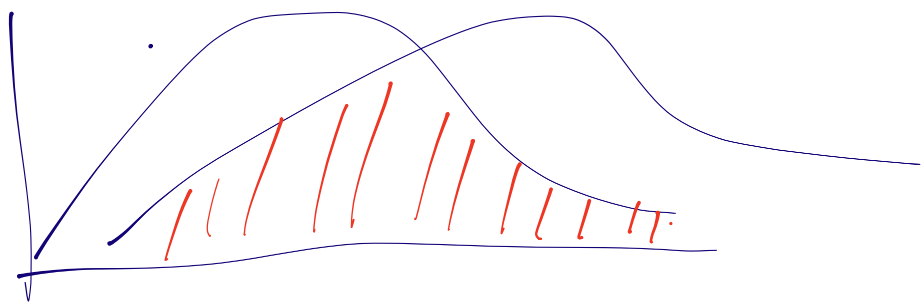
$$\mathbb{E}_1[\phi(X)] \le \mathbb{E}_1[\phi^*(X)] - c\cdot \underbrace{\mathbb{E}_0[\phi^*(X)]}_{=\alpha} + c\,\underbrace{\mathbb{E}_0[\phi(X)]}_{\le \alpha}$$

$$\le \mathbb{E}_1[\phi^*(X)].$$

$$\mathbb{E}_0[\phi(X)] + \mathbb{E}_1[1-\phi(X)]$$

$$= \int \left\{ p_0(x)\cdot\phi(x) + p_1(x)\cdot\left(1-\phi(x)\right)\right\} d\mu(x)$$

$$\ge \int \min\{p_0(x), p_1(x)\}\, d\mu(x).$$



$$= 1 - d_{TV}(P_1, P_0).$$

where $d_{TV}(P_1, P_0) := \frac{1}{2}\int |p_1(x) - p_0(x)|\, d\mu(x)$

$$= \sup_{f: \|f\|_\infty \le 1} \left| \mathbb{E}_{P_1}[f(X)] - \mathbb{E}_{P_0}[f(X)] \right|.$$

$= \mathbb{P}(X \ne Y)$ under optimal coupling of $X\sim P_0$, $Y\sim P_1$

Pinsker's ineq

$$d_{TV}(p,q) \leq \sqrt{\frac{1}{2} D_{KL}(p\|q)}$$

where $D_{KL}(p\|q) := \int p \log \frac{p}{q} \, d\mu(x)$.

$$d_{TV}(p,q) \leq \sqrt{\frac{1}{2} \chi^2(p\|q)}$$

where $\chi^2(p\|q) := \int p \cdot \left(\frac{p}{q} - 1\right) d\mu(x)$.

e.g. $X_1, X_2, \dots, X_n \overset{iid}{\sim} \mathcal{N}(\theta, 1)$.

$H_0: \quad \theta = 0$

$H_1: \quad \theta = \mu$.

$$d_{TV}\left(P_1^{\otimes n}, P_0^{\otimes n}\right) \leq \sqrt{\frac{1}{2} D_{KL}\left(P_1^{\otimes n} \| P_0^{\otimes n}\right)}$$

Tensorization of KL

$$D_{KL}(P_1 \times P_2 \times \cdots P_n \| Q_1 \times Q_2 \times \cdots Q_n) = \sum_{i=1}^{n} D_{KL}(P_i \| Q_i)$$

$$= \sqrt{\frac{n}{2} \cdot D_{KL}(P_1 \| P_0)}.$$

$$= \sqrt{\frac{n\mu^2}{2}}.$$

If $\mu = \frac{1}{\sqrt{n}}$.

Type I + Type II error $\geq 1 - \frac{1}{\sqrt{2}}$

for any possible test.

Def. (UMP test).

$\beta_{\phi^*}(\theta) \geq \beta_\phi(\theta)$ for any $\theta \in \Theta_1$

and any level-α test ϕ.

($\phi \equiv \varphi$).

"~~Uniformly~~ Most Powerful."

($\theta_1 > \theta_0$)

e.g. $X \sim N(\theta, 1)$ $\quad H_0: \theta = \theta_0$

$H_1: \theta = \theta_1$

$$L(x) = \exp\left((\theta_1 - \theta_0) \cdot x - \frac{(\theta_1^2 - \theta_0^2)}{2} \right)$$

$$\phi^*(x) = 1\{x \geq c\} = 1\{x \geq \theta_0 + z_\alpha\}$$

ϕ^* optimal for H_0 vs H_1 for any $\theta_1 > \theta_0$.

$H_0: \theta = \theta_0$
$H_1: \theta > \theta_0$. $\quad$ — $\quad \phi^*$ is UMP

Def. Monotone likelihood ratio (MLR).

1-D testing about θ, if there exists a stat

$T(X) \in \mathbb{R}$, s.t. $\forall \theta_1 < \theta_2$ we have

$p_{\theta_2}(x) / p_{\theta_1}(x)$ is non-dec function of T.

and $p_{\theta_1} \neq p_{\theta_2}$ for $\theta_1 \neq \theta_2$.

e.g. (exp family). $p_\theta(x) = \exp(\eta(\theta) \cdot T(x) - B(\theta)) \cdot h(x)$.

$$p_{\theta_2} / p_{\theta_1}(x) = \exp\left(\left(\eta(\theta_2) - \eta(\theta_1)\right) \cdot T(x) - B(\theta_2) + B(\theta_1)\right).$$

If η is monotonic in θ, MLR.

Thm: Consider $H_0: \theta \leq \theta_0$ vs. $H_1: \theta > \theta_0$, Assume MLR

$\exists$ UMP
$$\phi^* = \begin{cases} 0 & T(X) < c \\ \gamma & T(X) = c \\ 1 & T(X) > c \end{cases}$$
w/ $\mathbb{E}_{\theta_0}[\phi(X)] = \alpha$

Proof: ϕ^* is Neymann-Pearson test (LRT)

for $H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_1$

for any possible $\theta_1 > \theta_0$.

e.g. $X \sim \text{Binom}(n, \theta)$

$H_0: \theta \leq \theta_0$

$H_1: \theta > \theta_0$.

$T(X) = X$.

$$\phi(X) = \begin{cases} 0 & X < c \\ \gamma & X = c \\ 1 & X > c. \end{cases}$$

Need to make sure

$$\mathbb{E}_{\theta_0}(\phi(X)) = \mathbb{P}_{\theta_0}(X>c) + \gamma\cdot\mathbb{P}_{\theta_0}(X=c) = \alpha.$$

— Select c smallest integer s.t. $\mathbb{P}_{\theta_0}(X>c) \le \alpha$.

— $\gamma = \dfrac{\alpha - \mathbb{P}_{\theta_0}(X>c)}{\mathbb{P}_{\theta_0}(X=c)}$.

Two-sided testing

$H_0: \theta = \theta_0$

v.s. $H_1: \theta \neq \theta_0$.

e.g. $X \sim N(\theta, 1)$

unbiased

$$\phi(X) = \mathbb{1}\{|X-\theta_0| \ge z_{\alpha/2}\}.$$

$$\tilde{\phi}(X) = \mathbb{1}\left\{\begin{array}{l} X-\theta_0 > z_{\alpha/3} \\ \text{or } X-\theta_0 < -z_{2\alpha/3} \end{array}\right\}.$$

not unbiased.

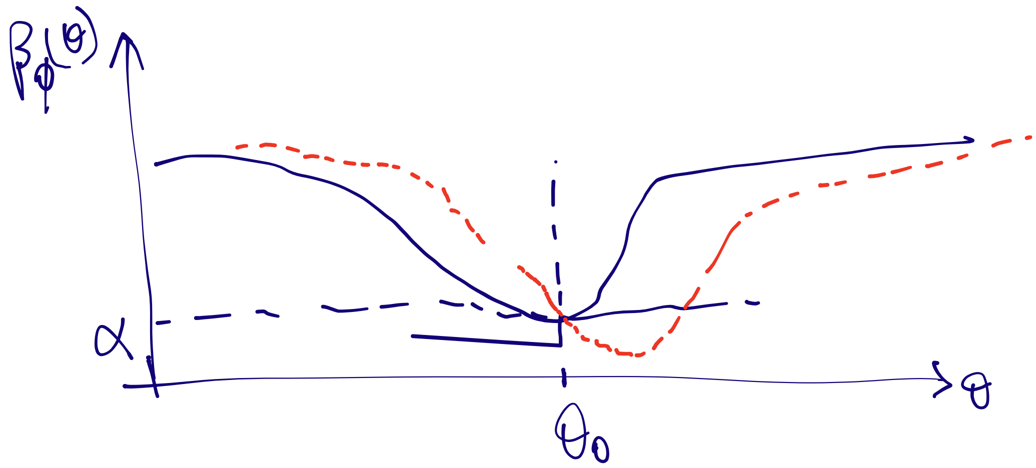
Def. ("Unbiased ~~test~~").

We call ϕ unbiased if $\mathbb{E}_\theta[\phi(X)] \ge \alpha$

$(\forall \theta \in \Theta_1)$.

(Why the name?

$\mathbb{E}_\theta[(\delta(X)-g(\theta))^2] \le \mathbb{E}_{\theta'}[(\delta(X)-g(\theta'))^2] \quad \forall \theta'$)



UMPU = Uniformly most powerful among unbiased tests.

Exp family. $p_\eta(x) = \exp(\eta \cdot T(x) - A(\eta))\, h(x)$.

Test $H_0: \eta = \eta_0$ vs. $H_1: \eta \neq \eta_0$.

F.O.C $\frac{d}{d\eta} \mathbb{E}_\eta[\phi(X)]$

$$= \int \frac{d}{d\eta}\left(\phi(x) \cdot e^{\eta T(x) - A(\eta)} \cdot h(x)\right) d\mu(x)$$

$$= \int \phi(x) \cdot (T(x) - A'(\eta)) \cdot p_\eta(x)\, d\mu(x)$$

$$\mathbb{E}[(X - \mathbb{E}(X)) \cdot Y]$$
$$= \mathbb{E}[(X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y))].$$

$$A'(\eta) = \mathbb{E}_\eta[T(X)]$$

So $\frac{d}{d\eta} \mathbb{E}_\eta[\phi(X)] = \mathrm{Cov}_\eta(T(X), \phi(X)).$

Need
$$\begin{cases} \mathbb{E}_{\eta_0}\left[T(X)\cdot(\phi^*(X)-\alpha)\right]=0 \\ \mathbb{E}_{\eta_0}\left[\phi^*(X)\right]=\alpha. \end{cases} \quad (*)$$

$$\phi^*(x)=\begin{cases} 0 & T(X)\in(C_1,C_2) \\ 1 & T(X)>C_2 \text{ or } T(X)<C_1 \\ \gamma_i & T(X)=C_i \quad \text{for } i=1,2. \end{cases}$$

Thm:- Suppose $\eta_0\in \text{int}\ \eta(\Theta)$. (exp family)
$\forall\alpha\in(0,1)$. $\exists$ two-sided level α test ϕ^*
of the form above satisfying. $(*)$
ϕ^* is UMPU.

Proof idea for existence

(C_1,γ_1), second equation $\Longrightarrow$ (C_2,γ_2).

$$\begin{cases} C_1\to-\infty & \beta_\phi'(\eta_0)\geq 0 \\ C_1\to+\infty & \beta_\phi'(\eta_0)\leq 0. \end{cases}$$

Def (Confidence ~~set~~).

$C(X)$ is $(1-\alpha)$ confidence set if $\mathbb{P}_\theta\big(g(\theta) \in C(X)\big) \geq 1-\alpha$ $(\forall \theta \in \Theta)$.

Duality: level-α test $\phi_{\theta_0}(x)$ for $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$.
valid $\forall \theta_0 \in \Theta$.

Then $C(X) = \{\theta \in \Theta : \phi_\theta(X) < 1\}$ is $(1-\alpha)$ CI.

Proof: $\mathbb{P}_\theta\big(\theta \notin C(X)\big) = \mathbb{P}\big(\phi_\theta(x) = 1\big) \leq \alpha.$

The other way around:
$C(X)$ $(1-\alpha)$ CI for θ

$\phi_\theta(x) = \mathbb{1}\{\theta_0 \notin C(X)\}$ is level-α test for $H_0: \theta = \theta_0$ v.s $H_1: \theta \neq \theta_0$.

Proof: $\mathbb{P}_{\theta_0}\big(\phi_{\theta_0}(x) = 1\big) = \mathbb{P}_{\theta_0}\big(\theta_0 \notin C(X)\big)$
$\leq \alpha$

Multivariate testing.

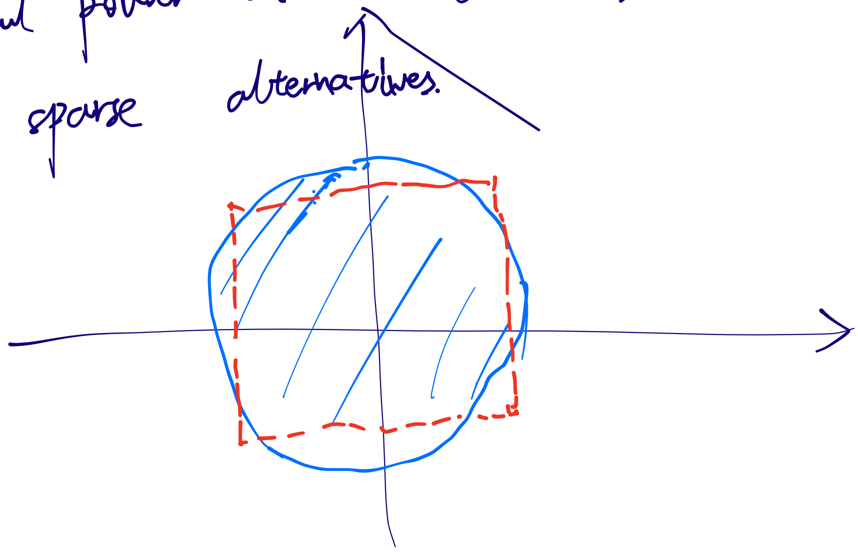
e.g. $X \sim N(\theta, I_d)$

$\phi_1(x) = \mathbb{1}\{\|X\|_2 \geq c_1\}$.

$\phi_2(x) = \mathbb{1}\{\|X\|_\infty \geq c_2\}$.

ϕ_1 : equal power at all directions.

ϕ_2 : favours sparse alternatives.



"Minimax testing".

$$H_0 : \theta \in \Theta_0$$

$$H_1 : \theta \in \Theta_1(\varepsilon)$$

$$:= \Theta_1 \cap \{\theta : dist(\theta, \Theta_0) \geq \varepsilon\}.$$

$$R_\varepsilon(\phi) = \sup_{\theta \in \Theta_0} \mathbb{E}[\phi(X)] + \sup_{\theta \in \Theta_1(\varepsilon)} \mathbb{E}[1 - \phi(X)].$$

Find ε s.t. $R_\varepsilon(\phi) < 1/10$.

e.g. Recall univariate normal.

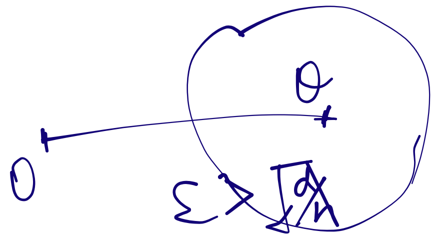
$$\varepsilon_n^* \asymp 1/\sqrt{n}.$$

How about $X_1, X_2 \cdots X_n \overset{iid}{\sim} N(\theta, I_d)$

$H_0 : \theta = 0$ $\quad$ $H_1(\varepsilon) : \|\theta\|_2 \geq \varepsilon$.

Optimal $\varepsilon_n^* \overset{?}{=} \sqrt{d/n}$.

since $\|\bar{X}_n - \theta\|_2 \le \frac{\sqrt{d}}{\sqrt{n}}$ w.h.p.



Thm. The minimax testing radius $\varepsilon_n^* \asymp \dfrac{d^{1/4}}{n^{1/2}}$.

Roadmap
- Construct a test that works at ε_n^*
- Information-theoretically, no better test (i.e. bounding d_{TV}).

Part I.

Test $\quad \phi(X) = \mathbb{1}\{\|\frac{1}{n}\sum_1^n X_i\|_2^2 \ge C\}$

$$Y = \frac{1}{n}\sum_{i=1}^n X_i \sim N(\theta, I_d/n).$$

$$\mathbb{E}_\theta\left[\|Y\|_2^2\right] = \frac{d}{n} + \|\theta\|_2^2$$

$\xi_j \overset{iid}{\sim} N(0,1)$ $(j = 1, 2, \dots, d)$

$$\mathrm{var}_\theta\left(\|Y\|_2^2\right) = \sum_{j=1}^d \mathrm{var}_\theta\left(Y_j^2\right)$$

$$= \sum_{j=1}^d \left\{ \mathbb{E}\left|\theta_j + \xi_j/\sqrt{n}\right|^4 - \left(\mathbb{E}\left|\theta_j + \xi_j/\sqrt{n}\right|^2\right)^2 \right\}$$

$$= \frac{4}{n}\|\theta\|_2^2 + \frac{2d}{n^2}$$

Fluctuation in $\|Y\|_2^2 \asymp \frac{\|\theta\|_2}{\sqrt{n}} + \frac{\sqrt{d}}{n}$

$$\ll \mathbb{E}[\|Y\|_2^2].$$

$$C = \mathbb{E}_0[\|Y\|_2^2] + a \cdot \sqrt{\mathrm{var}_0(\|Y\|_2^2)}.$$

$$\left(= \frac{d}{n} + a \cdot \sqrt{\frac{2d}{n^2}} \right).$$

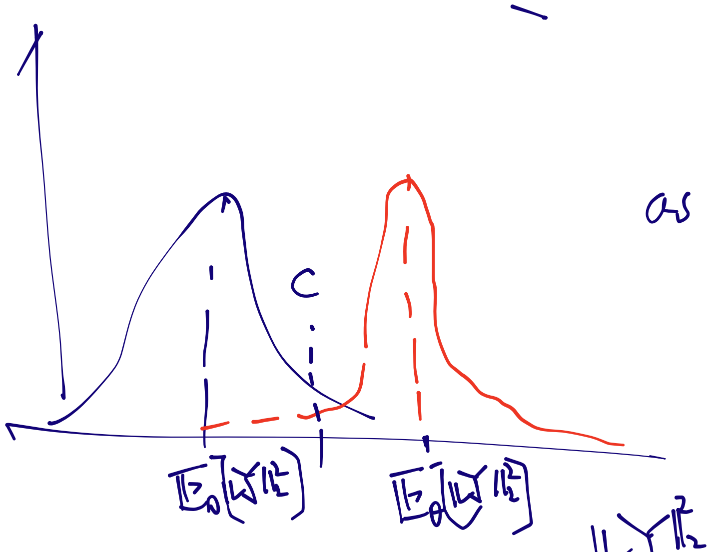
$$\mathbb{E}_0[\phi(X)] = \mathbb{P}_0\left(\|Y\|_2^2 - \mathbb{E}_0[\|Y\|_2^2] \geq a \cdot \sqrt{\mathrm{var}(\|Y\|_2^2)} \right)$$

$$\leq \frac{1}{a^2} \quad \text{(Chebyshev's ineq).}$$

$$\mathbb{E}_\theta[1 - \phi(X)] = \mathbb{P}_\theta\left(\|Y\|_2^2 < C \right).$$

$$= \mathbb{P}_\theta\left(\|Y\|_2^2 - \mathbb{E}_\theta[\|Y\|_2^2] \leq C - \mathbb{E}_\theta[\|Y\|_2^2] \right)$$

$$\leq \frac{\mathrm{var}_\theta(\|Y\|_2^2)}{\left(\mathbb{E}_\theta[\|Y\|_2^2] - C\right)^2}$$



as long as

$$\mathbb{E}_\theta[\|Y\|_2^2] > C.$$

Want $\mathbb{E}_\theta\left[1-\phi(X)\right] \leq 1/a^2$.

Requires

$$\frac{d}{n} + \|\theta\|_2^2 - a\sqrt{\frac{2d}{n} + \frac{4}{n}\|\theta\|_2^2} \geq \frac{d}{n} + \frac{a\sqrt{2d}}{n}.$$

$(a=3)$. $\Updownarrow$

$$\|\theta\|_2^2 \geq \frac{4\sqrt{d}}{n} + 6\sqrt{\frac{d}{n^2} + \frac{\|\theta\|_2^2}{n}}$$

Solve for θ,

holds true as long as $\|\theta\|_2 \geq C' \cdot \frac{d^{1/4}}{n^{1/2}}$.