Lecture 1/19

Friday, January 19, 2018 10:55 AM

MA 782 - Hypothesis Testing

Office Hours: MW: 2-3:30 MCS 229

<u>Topics</u>: All aspects of hypothesis testing.

- Preforming tests
- Quality of tests
- Properties of tests.
- UMP - Neyman Fisher

Course Layout

- HW - Biweekly 30%
- Midterm (mid March) 30%
- Final 40%

Statistical Decision Theory

Statistical Decision Theory

<u>Goal</u>: Make inferences about θ based on $(X_1, \ldots, X_n)$ where θ contains information about F_X.

Here θ can be either a finite dimensional or <u>infinite dimensional</u> and the $X_i \sim$ iid RV.

<u>Ex</u>: $\theta = F_X$ or $\theta = f_X$.

<u>Ex</u>: $X_i \sim \text{iid } N(\mu, \sigma^2)$ and $\theta = (\mu, \sigma^2)$

If we view the estimation problem jointly, one question we need to consider is

$$\mathrm{Cov}(\hat{\mu}, \hat{\sigma}^2)$$

Helpful for testing things like

$$H_0: \mu = 0, \sigma = 1$$

Ex: $X_i \sim iid\ f$ and $\theta = f$

$\hat{f}$ = kernel density / histogram estimator

Major Goal in Statistics

(1) Estimation

(2) Testing

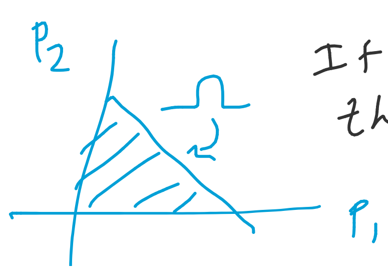
Ex: Suppose $\theta \in \Omega$ and we want to know if $\theta \in \Omega_0 \subseteq \Omega$.

(e.g.) Let $\mu \in \mathbb{R}$ and $H_0: \mu \leq 1$ then

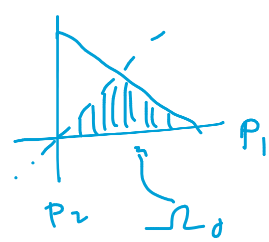
$\Omega_0 = (-\infty, -1]$

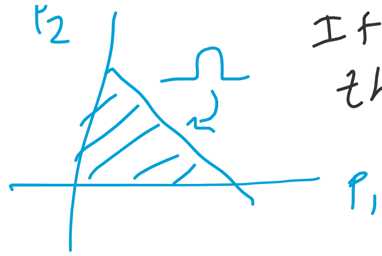
(e.g.) p_1, p_2 with $p_1 + p_2 \leq 1$

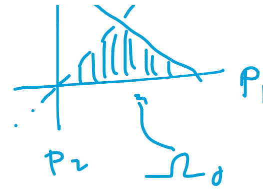
then



If $H_0: p_1 > p_2$ then



then 

If $H_0: p_1 > p_2$ then 

Another way to solve this:

Estimate $\mathbb{1}\{\theta \in \Omega_0\}$

<u>Next time</u>: Loss functions, Bayesian decision Theory, and Neymann-Fisher Lemma

Lecture 1/22

Monday, January 22, 2018 11:13 AM

Estimation Evaluation & Comparison

Suppose we estimate θ by $\delta(X)$ where X represents the data.

Consider a loss function $L(\theta, \delta(X))$

Satisfying
$$\begin{cases} L(\cdot,\cdot) \geq 0 \\ L(a,a) = 0 \end{cases}$$

Ex: Suppose $X \sim \text{Binomial}(n,p)$ with $n=100$. How do we estimate p?

One reasonable choice is $\frac{1}{n}\sum X_i$.

Why? Consider the loss given by

$$L_1(\delta(X),\theta) = [\delta - \theta]^2$$

$$L_2(\delta(X,\theta) = |\delta - \theta|$$

This gives

$$L_1(\hat{p}, p) = \left(\frac{1}{n}\sum x_i - p\right)^2$$

$$L_2(\hat{p}, p) = \left|\frac{1}{n}\sum x_i - p\right|$$

Now consider $\hat{p} = \frac{\sum x_i + 3}{n}$

Now for $p = 1/2$ $\sum x_i = 47$. Then

$\frac{\sum x_i + 3}{n}$ is the better estimator.

So the estimator depends both on X and p. So we need to evaluate these even moreso.

One way to do this through

Risk functions

Removes variability in $\delta(x)$

$$R(\delta, \theta) = \mathbb{E}(L(\delta, \theta))$$

Ex: Now continuing our example, with squared loss.

$$\mathbb{E}\left[\left(\frac{1}{n}\sum x_i - p\right)^2\right] < \mathbb{E}\left[\left(\frac{1}{n}\sum x_i + 3 - p\right)^2\right]$$

- Unbiased
- Identical Variance

- Biased
- Identical Variance

What about $\dfrac{\sum x_i + 3}{106}$?

Exercise: Show ↗ sometimes has a smaller Risk than $\hat{p} = \frac{1}{n}\sum x_i$

"Best" estimator still relies on unknown p.

We first got rid of Randomness in X by taking expectation. Now we use

Bayesian Risk $\quad \int R(\delta, \theta)\, p(\theta)\, d\theta$

$\int R(\delta, \theta) p(\theta) d\theta$

So all randomness removed - but how to choose $p(\theta)$? No clear answer...

Bayesian Classification Problem

Suppose $\Omega = \{\theta_0, \theta_1\}$ and we

bad ↑ ↑ good

estimate it by $\delta(x)$. We're going to use zero-one loss

$$L(\delta, \theta) = \begin{cases} 0, & \theta = \theta_1 \text{ and } \delta = \theta_1 \\ 0, & \text{" } \theta_0 \text{" } \quad \text{" } \theta_0 \text{"} \\ 1, & \text{" } \theta_1 \text{" } \quad \text{" } \theta_0 \text{"} \\ 1. & \text{" } \theta_0 \text{" } \quad \text{" } \theta_1 \text{"} \end{cases}$$

Now calculating Risk...

$E(L(\delta, \theta)) = P(\delta(x) = \theta \mid \theta = \theta_1)$

$$+ \; P(\delta(x) = \theta \mid \theta = \theta_0)$$

Suppose $\theta \sim \begin{cases} \theta_0 & \pi_0 \\ \theta_1 & \pi_1 \end{cases}$ which gives the

Bayes Risk as

$$P(\delta(x) = \theta_1)\pi_1 + P(\delta(x) = \theta_0)\pi_0$$

We now look to minimize this

Risk

$$\hat{\delta}_{Bayes}(x) = \int_{\delta=\theta_0} f(x \mid \theta = \theta_1)\pi_1 \, dx + \int_{\delta=\theta_1} f(x \mid \theta = \theta_0)\pi_0 \, dx$$

$$= \int \left[f(x \mid \theta = \theta_0)\pi_0 \mathbb{1}_{\delta=\theta_1} + f(x \mid \theta = \theta_1) \mathbb{1}_{\delta=\theta_0} \right] dx$$

Since we look to minimize we

always choose

$$\hat{\delta}_{Bayes} = \min\left(f(x|\theta=\theta_0)\pi_0,\ f(x|\theta=\theta_1)\pi_1\right)$$

That is if

$$f(x|\theta=\theta_1)\pi_0 > f(x|\theta=\theta_1)\pi_1$$

then we set $\hat{\delta}_{Bayes} = \theta_0$

Otherwise $\hat{\delta}_{Bayes} = \theta_1$.

Why does this not really help? $f(x|\theta=\theta_1)$ and $f(x|\theta=\theta_0)$ is the posterior distribution.

So we base our inference/
decisions on the posterior
distribution.

Lecture 1/24

Wednesday, January 24, 2018 11:07 AM

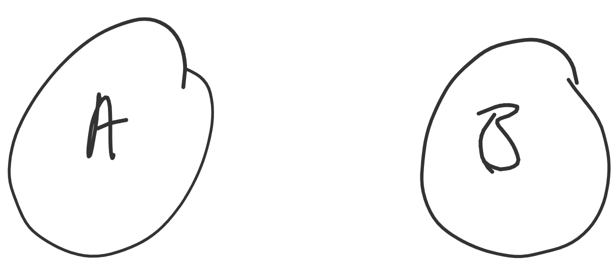
Notation: What does $\mathbb{E}_{\theta=\theta_0}$ mean?

$$X \sim N(\mu, \sigma^2)$$

$\mathbb{E}(X) = \mu$. implicitly assume μ is fixed.

If μ is random then we really mean

$$\mathbb{E}_X(x) = \mu.$$



μ_A μ_B

If $\underline{X}$ is a mixture of points from A and B then

$\mathbb{E}(X) = \mu_a$? $\mathbb{E}(X) = \mu_B$?

So we typically write

$\mathbb{E}_{\mu=\mu_A}(X) = \mu_A$

$= \mathbb{E}(X \mid G = A)$ for G indicating which group X comes from.

Measure Theory Based

Measure Theory Based

Expectation

$$\mathbb{E}(X) = \int x f(x) \underbrace{dx}$$

Same weight for each x value
$\approx$ uniform distribution

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}\left[\int x f(x|y) dx\right]$$

$$= \int\int x \underline{f(x|y)} dx \underline{f(y)} dy$$

product of densities

product of densities

that another strange

density over a

uniform ($dy\,dx = dz$)

Hypothesis Testing

Goal: Want to infer

$\theta \in \Omega_0 \subseteq \Omega$. based

on the observed $\underline{X}$.

and we have

$H_0: \theta \in \Omega_0$

$H_A: \theta \in \Omega_1$

where $\Omega_0 \cap \Omega_1 = \emptyset$ but it is possible that $\Omega_0 \cup \Omega_1 \subset \Omega$.

Ex: $H_0: \mu = 0$ for $\mu \in \mathbb{R}$.

$H_A: \mu > 0$.

Lecture 1/26

Friday, January 26, 2018 11:11 AM

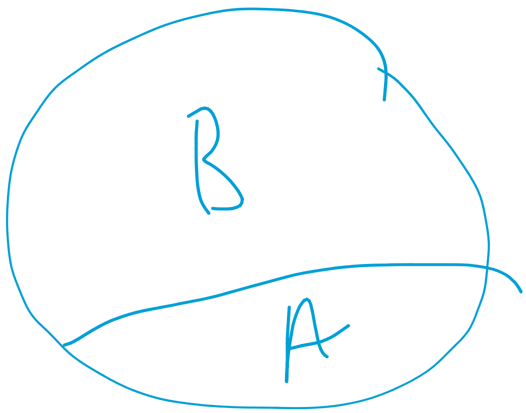
Motivating Example

Suppose that $\underline{X} \sim f(x)$

then $\mathbb{E}(X) = \int x f(x) dx$

but the information in $f(x)$ may be incorrect.

Say for example that the population is pantioned as



$Y = 1_A$

$X \sim Y f(x) + (1-Y) g(x)$

for group B having dist.

$g(x)$.

In this case $\mathbb{E}(X) = \mathbb{E}(X \mid Y=1)$.

but what we want is

$\mathbb{E}(X) = P_A \, \mathbb{E}(X \mid Y=1) + P_B \, \mathbb{E}(X \mid Y=0)$

· We always want to be able to change our understanding of f.

Measure Theory

Notation: $\mu : \mathcal{A} \mapsto [0,1]$

for $\mathcal{A}$ a σ-field over

Ω.

Suppose X is a R.V. then

X induces a measure μ_X

where $\mu_X(A) = \mathbb{P}(X \in A)$.

Then

$$\mathbb{E}(X) = \int x f(x) dx$$

$$= \int x \mu_X(dx)$$

When $\mu_X(dx) = \mathbb{P}(X \in (x, x+dx))$

$$\simeq f(x)\,dx.$$

Ex: $\mathbb{E}(X \mid Y = y)$

$$= \int x\, f^*(x|y)\,dx$$

$$= \int x\, f^*(x) \cdot \frac{f^*(x|y)}{f^*(x)}\,dx$$

$$= \int x\, f^*(x)\, \nu_y(dx)$$

when $\dfrac{\nu_y(dx)}{dx} = \dfrac{f^*(x|y)}{f^*(x)}$

$+ \ (x)$

Conditional expectation is just the expetation of X just over a different measure.

Hypothesis Testing

Goal: Test

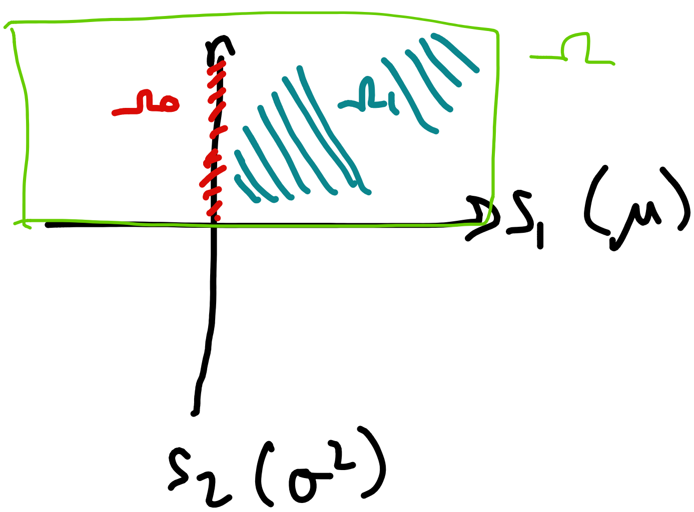
$H_0: \theta \in \Omega_0 \qquad \Omega_0 \cap \Omega_1 = \emptyset.$

$H_A: \theta \in \Omega_1$

Ex: $\underline{X} \sim N(\mu, \sigma^2)$

$H_0: \mu = 0 \qquad \Omega_0 = \left\{ (s_1, s_2) : s_1 = 0, s_2 > 0 \right\}$

$H_A: \mu > 0 \qquad \Omega_1 = \left\{ (s_1, s_2) : s_1 > 0, s_2 \geq 0 \right\}$



Critical/Rejection Region (S)

- Let $\delta(\underline{x}) = 1\ \{ \underline{x} \in S \}$

then if $\delta(\underline{x}) = 1$

We reject H_0.

We reject ...

<u>Def:</u> The **power function** is given by

$$\beta(\theta) = \mathbb{E}_\theta(\delta(\underline{X})) = \mathbb{P}_\theta(\delta(\underline{X}) = 1)$$

$$= \mathbb{P}_\theta(\text{reject } H_0).$$

<u>Goal:</u> Maximize $\beta(\theta)$ when $\theta \in \Omega_1$ and minimize $\beta(\theta)$ when $\theta \in \Omega_0$.

How do we do both simultaneously.

Lecture 1/29

Monday, January 29, 2018 11:14 AM

Hypothesis Testing

Recall we defined the

decision rule

$$\delta(\underline{X}) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{o.w.} \end{cases}$$

and the power function

$$\beta(\theta) = \mathbb{E}_\theta(\delta(\underline{X})) = \mathbb{P}_\theta(\text{reject } H_0)$$

Significance level

$$\alpha = \sup \beta(\theta)$$

$$\alpha = \sup_{\theta \in \Omega_0} \beta(\theta)$$

Goal: Minimize $\beta(\theta)$ for $\theta \in \Omega_0$ and maximize $\beta(\theta)$ for $\theta \in \Omega_1$

Hard to do simultaneously.

Strategy: Controll type I error at α and maximize β over the remaining Ω_1.

Ex: $H_0 : \mu \leq 0$ $\longleftrightarrow$ $H_0 : \mu = 0$

$H_1: \mu > 0$ $\quad H_1: \mu > 0$

because

$$\alpha = \sup_{\theta \in \Omega_0} \beta(\theta)$$

is maximized at α s.t. $\mu = 0$.

<u>Ex:</u> Suppose $\underline{X} \sim$ iid $N(\mu, 1)$

We wish to test

$H_0: \mu \leq 0$

$H_1: \mu > 0$

In this case consider

In this case consider

$$\delta(\underline{X}) = 1_{\{\bar{X} > \lambda\}}$$

for some threshold λ.

then for $\bar{X}_n \sim N(\mu, 1/n)$

$$\beta(\theta) = \mathbb{P}_\theta(\bar{X}_n > \lambda)$$

$$= \mathbb{P}_\theta\left(\frac{\bar{X}_n - \mu}{\sqrt{1/n}} > \frac{\lambda - \mu}{\sqrt{1/n}}\right)$$

$$= \mathbb{P}_\theta\left(Z > \sqrt{n}(\lambda - \mu)\right)$$

$$= 1 - \Phi\left[\sqrt{n}(\lambda - \mu)\right]$$

We want to controll the Type I error

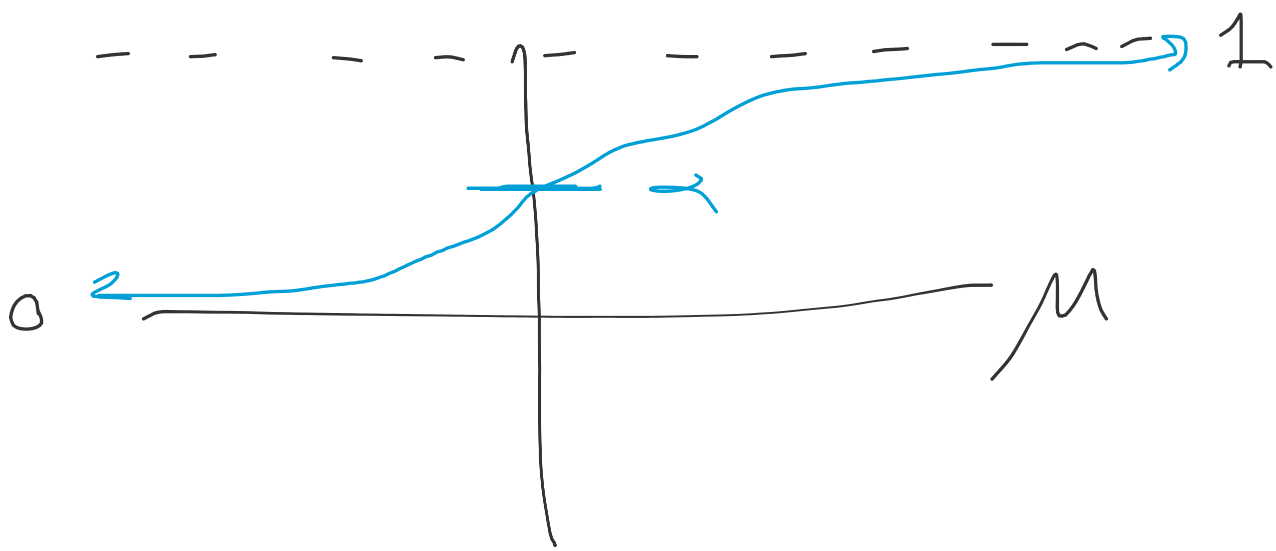
$$\sup_{\mu \le 0} \beta(\theta) = \sup_{\mu \le 0} \left(1 - \Phi(\sqrt{n}(\lambda - \mu))\right)$$

$$= 1 - \Phi(\sqrt{n}\lambda) \equiv \alpha$$

So our theshold will be given by

$$\lambda = \frac{1}{\sqrt{n}} \underbrace{\Phi^{-1}(1-\alpha)}_{(1-\alpha) \text{ quantile}}$$

of $N(0,1)$



<u>Def</u>: A **critical function** is a function $\phi(x)$ s.t.

$\phi(x)$ = prob. of rejection H_0 given $\underline{X} = x$.

- $\delta(\underline{x})$ a type of critical function.

<u>Proposition:</u> For two critical functions, ϕ_1, ϕ_2, the linear combination $\lambda \phi_1 + (1-\lambda)\phi_2$ is also a critical function if $0 \le \lambda \le 1$.

<u>Neyman Pearson Lemma.</u>

For testing $H_0: \theta = \theta_0$ $H_A: \theta = \theta_1$ for $\theta_0 \neq \theta_1$

<u>Goal:</u> Find test with best power performance.

For any $\phi(x)$,

For any $\phi(x)$,

$$\beta(\theta) = \mathbb{E}_\theta[\phi(\underline{X})]$$

<u>Prop</u>: (Lagrange mult.) Suppose $k \geq 0$ and ϕ^* maximizes

$$\mathbb{E}_{\theta_1}[\phi(\underline{X})] - k\,\mathbb{E}_{\theta_0}[\phi(\underline{X})]$$

among all critical func.

Let $\alpha = \mathbb{E}_{\theta_0}[\phi^*(\underline{X})]$ then

ϕ^* maximizes $\mathbb{E}_{\theta_1}(\phi(\underline{X}))$

over all critical functions

ϕ s.t. $\mathbb{E}_\theta[\phi(\underline{X})] = \alpha$.

Lecture 1/31

Wednesday, January 31, 2018 11:15 AM

Lagrange Multipliers

Suppose $k \geq 0$ and

ϕ^* maximizes

$$\mathbb{E}_{\theta_1} \phi(\underline{X}) - k\, \mathbb{E}_{\theta_0} \phi(\underline{X})$$

then ϕ^* maximizes

$\mathbb{E}_{\theta_1} \phi(\underline{X})$ among all

ϕ s.t.

$$\mathbb{E}_{\theta_0} \phi(\underline{X}) \leq \mathbb{E}_{\theta_0} \phi^*(\underline{X}) = \alpha$$

Pf: For any critical function $\tilde{\phi}$ s.t.

$$\mathbb{E}_{\theta_0} \tilde{\phi}(\underline{X}) \le \mathbb{E}_{\theta_0} \phi^*(\underline{X})$$

we want

$$\mathbb{E}_{\theta_1} \tilde{\phi}(\underline{X}) \le \mathbb{E}_{\theta_1} \phi^*(\underline{X})$$

Side note:

$\mathbb{E}_{\theta_1}(\phi(\underline{X})) \leftarrow$ 1 - Type II error

$\mathbb{E}_{\theta_0}[\phi(\underline{X})] \leftarrow$ Type I error

$$\mathbb{E}_{\theta_0}[\phi(\underline{X})] \leftarrow \text{Type I error}$$

By assump. we have

$$\mathbb{E}_{\theta_1}\phi^*(\underline{X}) - k\,\mathbb{E}_{\theta_0}\phi^*(\underline{X})$$

$$\geq$$

$$\mathbb{E}_{\theta_1}\tilde{\phi}(\underline{X}) - k\,\mathbb{E}_{\theta_0}\tilde{\phi}(\underline{X})$$

$$\Longrightarrow$$

$$\mathbb{E}_{\theta_1}\phi^*(\underline{X}) \geq \mathbb{E}_{\theta_1}\tilde{\phi}(\underline{X}) + k\left(\mathbb{E}_{\theta_0}\phi^*(\underline{X}) - \mathbb{E}_{\theta_0}\tilde{\phi}(\underline{X})\right)$$

$$\geq \mathbb{E}_{\theta_1}(\tilde{\phi}(\underline{X}))$$

$\blacksquare$

So focusing on our linear optimization function we can get the form

$$\mathbb{E}_{\theta_1}(\phi(\underline{X})) - k\,\mathbb{E}_{\theta_0}\,\phi(\underline{X})$$

$$= \int \phi(x)\, p_{\theta_1}(x)\,\mu(dx)$$

$$- k \int \phi(X)\, p_{\theta_0}(x)\,\mu(dx)$$

where $\underline{X} \sim p_\theta$ w.r.t. μ.

$$= \int \left[p_{\theta_1}(x) - k\, p_{\theta_0} \right] \phi(x)\, \mu(dx)$$

From here we see to maximize the objective function wrt $\phi(\underline{x})$.

We choose

$$\phi(x) = \begin{cases} 1 & p_{\theta_1}(x) > k\, p_{\theta_0}(x) \\ 0 & p_{\theta_1}(x) < k\, p_{\theta_0}(x) \end{cases}$$

Here P_{θ_0} = likelihood under the null

P_{θ_1} = likelihood under alternative

Here we can see α can be specified via k and vica versa.

Neyman-Pearson Lemma

For testing

$H_0: \mathbb{P} = P_0$ v.s. $H_1: \mathbb{P} = P_1$

where P_0 and P_1 have

densities p_1 & p_2

wrt measure μ

(i) $\forall\, 0 < \alpha < 1$

$\exists\, \phi^*$ and k s.t.

$\mathbb{E}_0\, \phi^*(X) = \alpha$ and

$$\phi^*(x) = \begin{cases} 1 & p_1(x) > k\, p_0(x) \\ 0 & < \\ x \in [0,1] & \end{cases}$$

$(\gamma \in [0,1])$ o.w.

ii) ϕ^* maximize $E_1 \phi^*$

among all tests that satisfy

$E_0 \phi(X) \leq \alpha$.

iii) ϕ^* unique almost everywhere.

Lecture 2/2

Friday, February 2, 2018 11:14 AM

Neyman-Pearson Lemma

For $H_0: \mathbb{P} = \mathbb{P}_0$ vs. $H_1: \mathbb{P} = \mathbb{P}_1$

where $\mathbb{P}_0$ and $\mathbb{P}_1$ have densities p_0 and p_1 wrt some measure μ, then

(i) $\forall\, 0 < \alpha < 1\ \exists k$ and γ

s.t. $\mathbb{E}_0\, \phi^*(\underline{X}) = \alpha$ where

$$\phi^*(\underline{X}) = \begin{cases} 1 & \text{if } p_1(x) > k\, p_0(x) \\ 0 & p_1(x) < k\, p_0(x) \\ \gamma \in [0,1] & p_1(x) = k\, p_0(x) \end{cases}$$

(ii) ϕ^* above maximizes $\mathbb{E}_1\, \phi(\underline{X})$ over all tests ϕ with $\mathbb{E}_0\, \phi(x) \le \alpha$

(iii) If $\tilde{\phi}$ also maximizes

$E_1 \phi(\underline{X})$ over all tests ϕ with

$E_0[\phi(\underline{X})] \le \alpha$ then

$\tilde{\phi} = \phi^*$ on the set

$$B = \{x : p_1(x) \neq k p_0(x)\}$$

$\underline{\text{Pf}}$: (i) Let $z = \dfrac{p_1(\underline{X})}{p_0(\underline{X})}$. Then

using our definition from above

$$E_0 \phi^*(x) = P_0\{z > k\} + \gamma P_0\{z = k\}$$

Let
$$\boxed{\begin{aligned} k &= F^{-1}(1-\alpha) \\ &= \inf\{z : F(z) \ge 1-\alpha\} \end{aligned}}$$

where F is the CDF of

z wrt μ.

(a) If F is continuous at k

then $P(z \le k) = F(k) = 1-\alpha$

and $P(z = k) = 0$ So

choose $\gamma = 0$.

(b) If F is not cont. at k

$F(k) \ge 1-\alpha$ and

$P(z < k) \le 1-\alpha$.

Now set

$$\boxed{\gamma = \frac{P(z \le k) - (1-\alpha)}{P(z = k)}}$$

So $0 \le \gamma \le 1$.

Pf of (ii): See Lagrange multiplier lemma.

Pf of iii:

$$\phi = \begin{cases} 1 & p_1 > k p_0 \\ 0 & < \\ \gamma & = \end{cases}$$

$$\phi^* - \tilde{\phi} > 0 \Rightarrow \phi^* > 0$$
$$\Rightarrow p_1(x) - k p_0(x) \geq 0$$

$$\phi^* - \tilde{\phi} < 0 \Rightarrow \phi^* < 1$$
$$\Rightarrow p_1(x) - k p_0(x) \leq 0$$

$$(\phi^* - \tilde{\phi})(p_1(x) - k p_0(x)) \geq 0$$

$$\Rightarrow \int (\phi^* - \tilde{\phi})(p_1(x) - k p_0(x))\mu(dx) \geq 0$$

$$\int \phi^* p_1(x) \geq \int \tilde{\phi} p_1(x) + k\left(\int \phi^* p_0 - \int \tilde{\phi} p_0\right)$$

$$\int \tilde{\phi} P_1(x) + h\left(\alpha - \mathbb{E}_0 \tilde{\phi}\right)$$

$$(\geq 0)$$

$$\geq \int \tilde{\phi} P_1(x) = \mathbb{E}_{\theta_1}(\tilde{\phi})$$

Lecture 2/5

Monday, February 5, 2018 11:09 AM

Neyman Pearson

Part iii Pf:

$$\{\phi^*(x) - \tilde{\phi}(x)\}\{p_1(x) - k p_0(x)\} \geq 0$$

and

$$0 \leq \int \{\phi^*(x) - \tilde{\phi}(x)\}\{p_1(x) - k p_0(x)\}$$

$$= \mathbb{E}_1[\phi^*(x)] - \mathbb{E}(\tilde{\phi}(x))$$

$$- k\{\mathbb{E}_0 \phi^*(x) - \mathbb{E}_0 \tilde{\phi}(x)\}$$

which implies that

$$\mathbb{E}_1 \phi^*(x) \geq \mathbb{E}_1(\tilde{\phi}(x)) + k\{\underbrace{\mathbb{E}_0(\phi^*(x))} - \underbrace{\mathbb{E}_0(\tilde{\phi}(x)}$$

$$\underbrace{\alpha \qquad \leq \alpha}_{\geq 0}$$

$$\geq E_1 \tilde{\phi}(X)$$

But recall that $E_1 \tilde{\phi}(X)$ is maximal by assumption so

$$\phi^* = \tilde{\phi}$$

Going back to the integral

$$\int \underbrace{(\phi^* - \tilde{\phi})(p_1(x) - k p_0(x))}_{\geq 0} = 0$$

$\Longrightarrow$

$$\{\underbrace{(\phi^* - \tilde{\phi})}_{A}\underbrace{(p_1(x) - k p_0(x))}_{B}\} \overset{a.e.}{=} 0$$

If $A \neq 0$ then $B = 0$

If $A = 0$ then $B \neq 0$

So $\phi^* = \tilde{\phi}$ on the set

$$B = \{x : p_1(x) \neq k p_0(x)\}.$$

Rmk: Most powerful tests need not be unique by N-P (iii).

Comment: $\mathbb{E}_0 \tilde{\phi}(x) = \alpha$ unless $k = 0$. What does that mean?

The test always rejects for $p_1(x) > 0$.

So really our M.P. test can be improved by increasing α.

Ex: $\mathbb{P}_0$: Unif$(0,1)$

$\mathbb{P}_1$: Unif$(1,2)$.

$$\tilde{\phi}(x) = 1_{\{x>1\}}.$$ our test

$$\phi^*(x) = \begin{cases} 1 & p_1(x) > k p_0(x) \\ 0 & p_1(x) < k p_0(x) \\ \gamma & p_1(x) = k p_0(x) \end{cases}$$

$$= \begin{cases} 1 \text{ if } & p_1(x) > 0 \\ \alpha & \text{o.w.} \end{cases}$$

why α?

So under the different hypotheses.

$E_0 \phi^* = \alpha$

$E_1 \phi^* = 1$

$E_1 \phi^* = 1$ $\quad E_0 \phi^* = \alpha$

$E_1 \tilde{\phi} = 1$ $\quad E_0 \tilde{\phi} = 0$

So our test is actually better.

Rmk: When we have perfect type II error rates, it is possible to have more reduction in Type I error.

But NP-tests require

$$\mathbb{E}_0(\phi^*(x)) = \alpha$$

So it is possible to get a better test in these cases $(k=0)$.

<u>Ex</u>: Suppose $X \sim Bin(n, \theta)$ $n=2$

$H_0: \theta = 1/2$ $\quad H_A: \theta = 3/4$.

By NP lemma

$$\phi^*(x) = \begin{cases} 1 & p_1(x) > k p_0(x) \\ 0 & p_1(x) < k p_0(x) \\ \gamma & p_1(x) = k p_0(x) \end{cases}$$

Where

$$P_\theta(x) = \binom{2}{x} \theta^x (1-\theta)^{2-x}$$

$$\Longleftrightarrow$$

$$\frac{P_1(x)}{P_0(x)} = \frac{(3/4)^x (1/4)^{2-x}}{(1/2)^x (1-1/2)^{2-x}} = \frac{\frac{3^x}{16}}{(1/2)^2} = \frac{3^x}{4}.$$

So we can reformulate our test.

$$\phi^*(x) = \begin{cases} 1 & \frac{3^x}{4} > k \\ 0 & < \\ \gamma & = \end{cases}$$

$$\mathbb{E}_0\, \phi^*(x) = \left\{ P_0\left(\frac{3^x}{4} > k\right) + \gamma\, P\left(\frac{3^x}{4} = k\right) \right\}$$

<u>set</u> α

set α.

Now note that under the null

x	0	1	2
$p(Y=x)$	1/4	1/2	1/4

$$\Rightarrow k = \frac{3^2}{4} \Rightarrow \gamma = \alpha \times 4 = 0.2$$

(based on $\alpha = 0.05$

So from here we can build the entire NP test.

Lecture 2/7

Wednesday, February 7, 2018 11:06 AM

Ex:

$H_0: \theta = 1/2$

$H_A: \theta = 3/4$

$$\phi^*(x) = \begin{cases} 1 & \frac{3^x}{4} > 9/4 \\ 0 & < \\ \gamma = .2 & = \end{cases}$$

$E_0 \phi^*(\underline{X}) = \alpha = 0.05$

If we want to translate from likelihood ratios to rejection regions

$$\phi^*(x) = \begin{cases} 1 & x > 2 \\ 0 & x < 2 \\ \gamma = 0.2 & x = 2 \end{cases}$$

Never reject with probability one. Reject when $X=2$ with probability γ.

Example of a random test.

To now calucate the power

$$\mathbb{E}_1(\phi^*) = 0.2\,\mathbb{P}(X=2)$$

$$= 0.2\left(\frac{3}{4}\right)^2 = \frac{9}{80}$$

Exercise: $X \sim \text{Binom}(n, \theta)$

Find N-P test and calculate its power.

<u>Corollary</u>: If $\mathbb{P}_0 \neq \mathbb{P}_1$ and ϕ^* is from NP with level $\alpha \in (0,1)$ then $\mathbb{E}_1 \phi^* > \alpha$.

<u>Pf</u>: Consider $\tilde{\phi}(\underline{x}) = \alpha$. Then if $\mathbb{E}_1 \phi^* = \alpha$ then by uniqueness $\tilde{\phi} = \phi^*$ on

$$B = \{x : p_1(x) \neq k p_0(x)\}$$

k is independent of $\hat{\phi}$

$$\mu\left[\{\phi^* = \hat{\phi}\} \cap B\right] = \mu(B)$$ since $\phi^* \neq \hat{\phi}$

$$\Rightarrow \mu\left[\{\phi^* = \hat{\phi}\}^c \cap B\right] \overset{*}{=} 0$$

So on B $\phi^* \neq \hat{\phi}$

hence $\mu(B) = 0$. and

$\mu(B^c) = 1$ or

$p_1(x) = k p_0(x)$ occurs with probability one. So probability one. So

$$\int p_1(x) = 1 = k \int p_0 = 1$$

So $k = 1$. and

$$p_1 \overset{a.s.}{=} p_0$$

but we assumed that

$$\mathbb{P}_0 \neq \mathbb{P}_1$$

Here we define

$$p_0 = \frac{d\mathbb{P}_0}{\mu(dx)}$$

$\mu(dx)$

Ex:

$$p_\theta = \begin{cases} \theta e^{-\theta x} & x > 0 \\ 0 & \text{o.w.} \end{cases}$$

$H_0: \theta = 1 \qquad H_A: \theta = \theta_1$

Find the NP test for $\theta_1 > 1$.

Lecture 2/9

Friday, February 9, 2018 11:18 AM

Ex: $p_\theta = \begin{cases} \theta e^{-\theta x} & x > 0 \\ 0 & \text{o.w.} \end{cases}$

$H_0: \theta = 1 \quad H_1: \theta = \theta_1$

$\theta_1 > 1$ Find NP test.

$$\phi^*(x) = \begin{cases} 1 & p_1(x) > k p_0(x) \\ 0 & < \\ \gamma & = \end{cases}$$

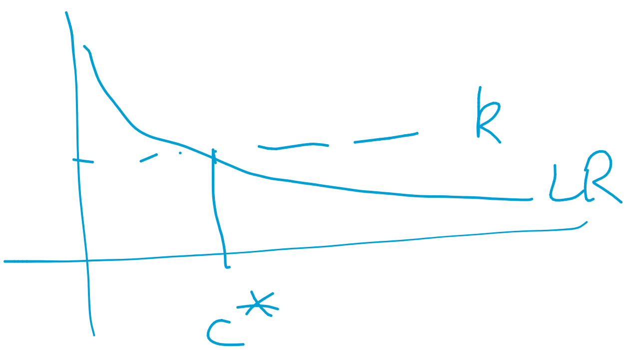
$$\frac{p_1(x)}{p_0(x)} = \frac{\theta_1 e^{-\theta_1 x}}{e^{-x}}$$

$$= \theta_1 e^{-(\theta_1 - 1)x}$$

$$\phi^*(x) = \begin{cases} 1 & \theta_1 e^{-(\theta_1 - 1)x} > k \\ 0 & < \end{cases}$$

$$\phi(x) = \begin{cases} 0 & \\ \gamma & < \\ & = \end{cases}$$

The likelihood ratio is decreasing



So we can again reformulate

$$\phi^* = \begin{cases} 1 & x < c^* \\ 0 & x > c^* \\ \gamma & x = c^* \end{cases}$$

$$\mathbb{E}_0 \phi^* = \mathbb{P}_0(X < c) + \gamma \underbrace{\mathbb{P}(X = c)}$$

θ_0

$$= \mathbb{P}_0 (X < c)$$

γ can be whatever.

$$= \int_0^c e^{-x} dx = -e^{-x} \Big|_0^c$$

$$= 1 - e^{-c} \overset{\text{set}}{=} \alpha$$

$$\Rightarrow 1 - \alpha = e^{-c}$$

$$\Rightarrow \boxed{c = -\log(1-\alpha)}$$

- α fixes k
- k fixes c
- so c independent of θ_1

So our final test is given by

$$\phi^*(x) = \begin{cases} 1 & x < -\log(1-\alpha) \\ 0 & x \geq -\log(1-\alpha) \end{cases}$$

Test does not depend on θ_1 so this test is most powerful for the test

$$H_0: \theta = 1 \quad H_1: \theta > 1.$$

- We will now begin to consider tests for sets not just points.

Monotone LR

The parametric family $\{P_\theta, \theta \in \Theta\}$ where P_θ

$\{P_\theta, \theta \in \Omega\}$ where P_θ

is the density w.r.t. μ

is said to be MLR

in $T(x)$ if

$\forall \theta_1 < \theta_2$ $\dfrac{P_{\theta_2}(x)}{P_{\theta_1}(x)}$ is

a nondecreasing function

of $T(x)$ for x at

which at least one

of $P_{\theta_1}, P_{\theta_2} > 0$

- Think of $T(x)$

as a s.s. for the family $\{P_\theta, \theta \in \Omega\}$.

<u>Thrm</u>: An exponential family

$$P_\theta = h(x) \exp\{y(\theta) T(x) - \beta(\theta)\}$$

where $y(\theta)$ is increasing

θ then MLR in

$T(x)$

<u>Pf</u>: $\forall\ \theta_2 > \theta_1$

$$\frac{P_{\theta_2}(x)}{P_{\theta_1}(x)} = \exp\{(y(\theta_2) - y(\theta_1)) T(x) - (\beta(\theta_2) - \beta(\theta_1))\}$$

increasing in $T(x)$ iff

$$y(\theta_2) - y(\theta_1) > 0$$

<u>Def</u>: Consider

$H_0: \theta \in \Omega_0 \quad H_A: \theta \in \Omega_1$

then

ϕ^* is UMP test iff

$\forall \theta \in \Omega_1$

$\mathbb{E}_\theta \phi^* \geq \mathbb{E}_\theta \phi$ holds

for all tests ϕ s.t.

$$\sup_{\theta \in \Omega_0} \mathbb{E}_\theta \phi \leq \sup_{\theta \in \Omega_0} \mathbb{E}_\theta \phi^*$$

Rmk: Difference in size and level.

size $\lesssim$ level.

Thm: Suppose $\{P_\theta\}$ is MLR in $T(X)$ and consider

$H_0: \theta \le \theta_0$ vs. $H_1: \theta > \theta_0$

then

(i) $\forall \alpha \in (0,1)$ $\exists \phi^*$ of the form

$$\phi^* = \begin{cases} 1 & T(x) > c \\ 0 & \quad < \\ \gamma \in [0,1] & \quad = \end{cases}$$

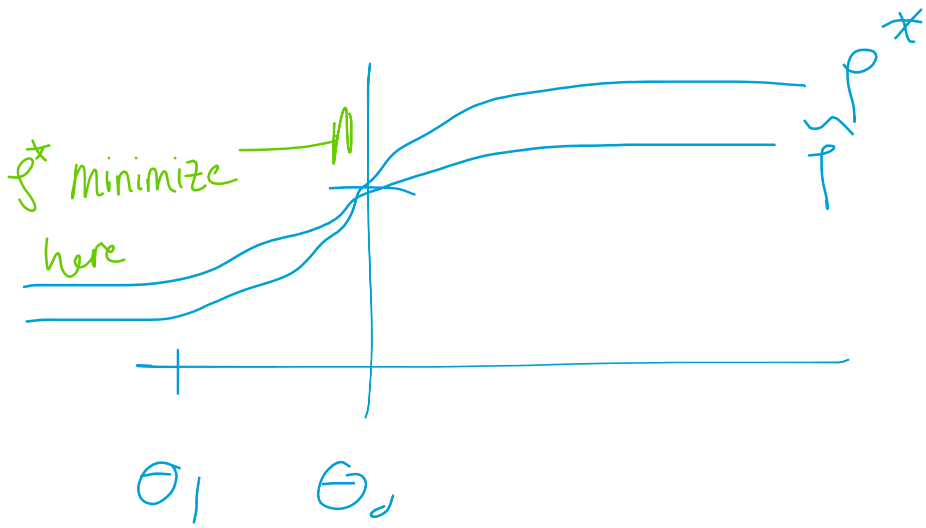
s.t. $\mathbb{E}_{\theta_0} \phi^* = \alpha$

(ii) ϕ^* is UMP at level α

(iii) The power function $\beta(\theta) = \mathbb{E}_\theta \phi^*$ is non-decreasing in θ and strictly increasing if $0 < \beta(\theta) < 1$

(iv) If $\theta_1 < \theta_0$ then ϕ^* minimizes $\mathbb{E}_{\theta_1} \phi$ among all ϕ s.t. $\mathbb{E}_{\theta_0} \phi = \alpha$

Exercise: Show (i) and (ii).



$\hat{\rho}^*$
$\hat{\rho}$
ρ^* minimize
here
θ_1
θ_d

Lecture 2/12

Monday, February 12, 2018 11:18 AM

Thrm: Suppose that

$P_\theta(x)$ is MLR in $T(x)$

and we want to test

$$H_0: \theta \le \theta_0 \qquad H_1 \quad \theta > \theta_0$$

$$\phi^*(x) = \begin{cases} 1 & T(x) > c \\ 0 & \quad < \\ \gamma & \quad = \end{cases}$$

(i) $\forall \alpha \in (0,1)$ $\exists c$ and γ s.t.

$E_{\theta_0} \phi^* = \alpha$

(ii) ϕ is UMP at level α

(iii) $\beta(\theta) = E_\theta \phi^*$ nondecreasing in θ

Pf: (i)

$$\mathbb{E}_{\theta_0} \phi^* = \mathbb{P}_{\theta_0}\{T(X) > c\} + \gamma \mathbb{P}_{\theta_0}(T(X) = c)$$

Same arguement as N.P.

(ii) $\forall \theta_1 > \theta_0$ consider the L.R.

$$\frac{p_{\theta_1}(x)}{\quad} = g(T(X))$$

$$\frac{\quad}{p_{\theta_0}(x)} = g(T(x))$$

where g is nondecreasing in $T(x)$. Define $k = g(c)$.

then $T(x) > c$ which implies $\phi^* = 1$.

If

$$\frac{p_{\theta_1}(x)}{p_{\theta_0}(x)} < k \Rightarrow T(x) < c \Rightarrow \phi^* = 0$$

So together ϕ^* maximize

$$\int \phi(x)\{p_{\theta_1}(x) - k p_{\theta_0}(x)\} d\mu.$$

By the NP lemma we have

ϕ^* M.P. $H_0: \theta = \theta_0$ $H_1: \theta = \theta_1$

$$\phi^*(x) = \begin{cases} 1 & p_{\theta_1}(x) > k p_{\theta_0}(x) \\ 0 & < \\ \begin{cases} 1 & T(x) > c \\ 0 & T(x) < c \\ \gamma & = \end{cases} & = \end{cases}$$

$(\; (\, \gamma \qquad\qquad =$

there because $g(\cdot)$ need

not be 1-1.

So $\forall \phi$ s.t. $\mathbb{E}_{\theta_0} \phi \le \alpha$ then

$\mathbb{E}_{\theta_1} \phi \le \mathbb{E}_{\theta_1} \phi^*$. but

θ_1 was arbitrary.

So ϕ^* is UMP across

Ω_1. $\blacksquare$

Pf: (iii) By the same argument

If (iii) by the same arg.

as above $\forall \theta_1 > \theta_2$

$$\phi^* = \arg\max_{\phi} \int (p_{\theta_1} - k p_{\theta_2}(x)) \phi \, d\mu.$$

and gives the MP test for $\theta_1 > \theta_2$.

then $\mathbb{E}_{\theta_1} \phi^* \geq \underbrace{\mathbb{E}_{\theta_1} (\mathbb{E}_{\theta_2} \phi^*)}_{\text{constant test.}}$

We want to show

$E_{\theta_1} \phi^* \geq \beta(\theta_2)$ Take

$\hat{\phi} := \beta(\theta_2)$ then

$\mathbb{E}_{\theta_1} \phi^* \geq \beta(\theta_2)$

Lecture 2/14

Wednesday, February 14, 2018 11:13 AM

Suppose $\{p_\theta(x)\}$ has MLR in $T(x)$ then

$$\phi^* = \begin{cases} 1 & T(x) > c \\ 0 & < \\ \gamma & = \end{cases}$$

satisfies all the results that we derived earlier for the NP test, when being used for

$H_0: \theta = \theta_1 \qquad H_1: \theta = \theta_2$

for any $\theta_1 < \theta_2$.

- NP - LR based but here $T(X)$ based on something more subtle.

<u>Pf</u>: Do as exercise.

<u>Rmk</u>: Because $\beta(\theta)$ is strictly monotone on $(0,1)$ then

for $H_0: \theta \le \theta_0$ $H_A: \theta > \theta_0$

the level is attained at

θ_0.

$\forall \tilde{\phi}$ s.t. $\sup_{\theta \in \Omega_0} \mathbb{E}_\theta \tilde{\phi} \le \alpha$

$\Rightarrow \mathbb{E}_{\theta_0} \tilde{\phi} \le \alpha$

$\Rightarrow \mathbb{E}_{\theta_1} \tilde{\phi} \le \mathbb{E}_{\theta_1} \hat{\phi}_1$ by

MP applied to $H_0: \theta = \theta_0$

$H_1: \theta = \theta_1$

$\Rightarrow \phi^*$ U.M.P.

<u>Pf (iv):</u>

Statement: If $\theta_1 < \theta_0$

then ϕ^* minimizes

$\mathbb{E}_{\theta_1} \phi$ among all ϕ s.t.

$\mathbb{E}_{\theta_1} \phi = \alpha$.

<u>Pf</u>: By proposition

ϕ^* maximizes $\mathbb{E}_{\theta_0} \phi - k \, \mathbb{E}_{\theta_1} \phi$, ,

ϕ^* maximizes $\mathbb{E}_{\theta_0}$ ...

when being viewed as a test

$H_0: \theta = \theta_1 \qquad H_A: \theta = \theta_0$

$$\mathbb{E}_{\theta_0} \phi^* - k\,\mathbb{E}_{\theta_1} \phi^* \geq \mathbb{E}_{\theta_0} \phi - k\,\mathbb{E}_{\theta_1} \phi$$

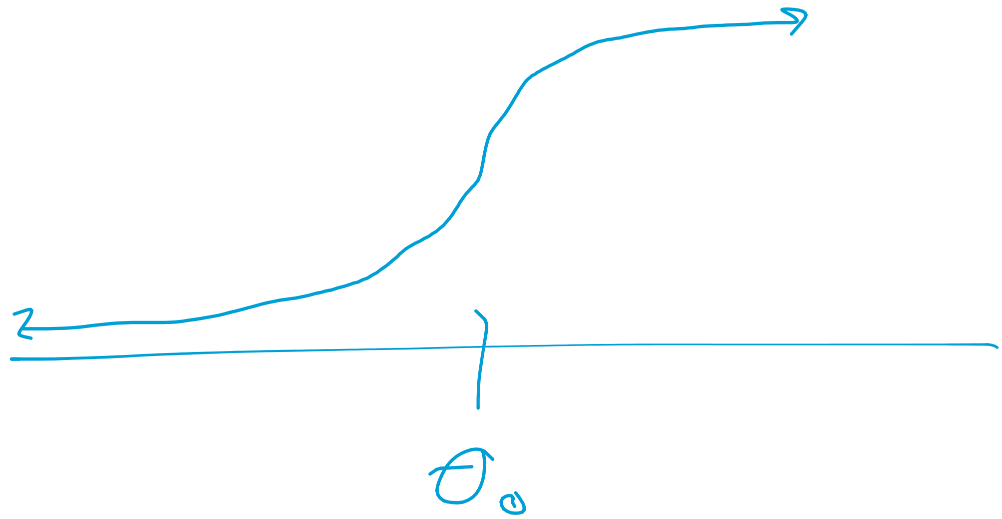
$$\Longrightarrow k\left\{\mathbb{E}_{\theta_1} \phi - \mathbb{E}_{\theta_1} \phi^*\right\}$$

$$\geq$$

$$\mathbb{E}_{\theta_0} \phi - \mathbb{E}_{\theta_0} \phi^*$$

$$= 0$$

So $\mathbb{E}_{\theta_1} \phi \geq \mathbb{E}_{\theta_1} \phi^*$ for arbitrary ϕ.



But we have yet to consider k.

Recall that

$$k = \left. \frac{P_{\theta_0}(x)}{P_{\theta_1}(x)} \right|_{T(x)=c}$$

$$k = 0 \implies \frac{p_{\theta_0}(x)}{p_{\theta_1}(X)} = 0 \quad T(x) = c$$

$$\underset{\text{MLR}}{\xrightarrow{\text{by}}} \quad \frac{p_{\theta_0}(x)}{p_{\theta_1}(x)} = 0 \quad T \le c$$

$$\Rightarrow \frac{p_{\theta_0}(x)}{p_{\theta_1}(x)} > 1 \quad \text{only if } T(x) > c$$

$\Rightarrow \mathbb{P}_{\theta_0}$ puts all mass on $\{T(x) > c\}$

$$\Rightarrow \mathbb{P}_{\theta}\{\phi^* = 1\} = 1.$$

$$\Rightarrow \quad \alpha = \mathbb{E}_{\theta_1}(\phi^*) = 1$$

$$\nexists\, \alpha \in (0,1).$$

$$k = \infty \Rightarrow$$

$$\frac{p_{\theta_0}(x)}{p_{\theta_1}(x)} = +\infty \text{ at } T(x) = c.$$

MLR

$$\Rightarrow \frac{p_{\theta_0}}{p_{\theta_1}} = \infty \qquad T(X) \geq c$$

Assuming the likelihood's are nice.

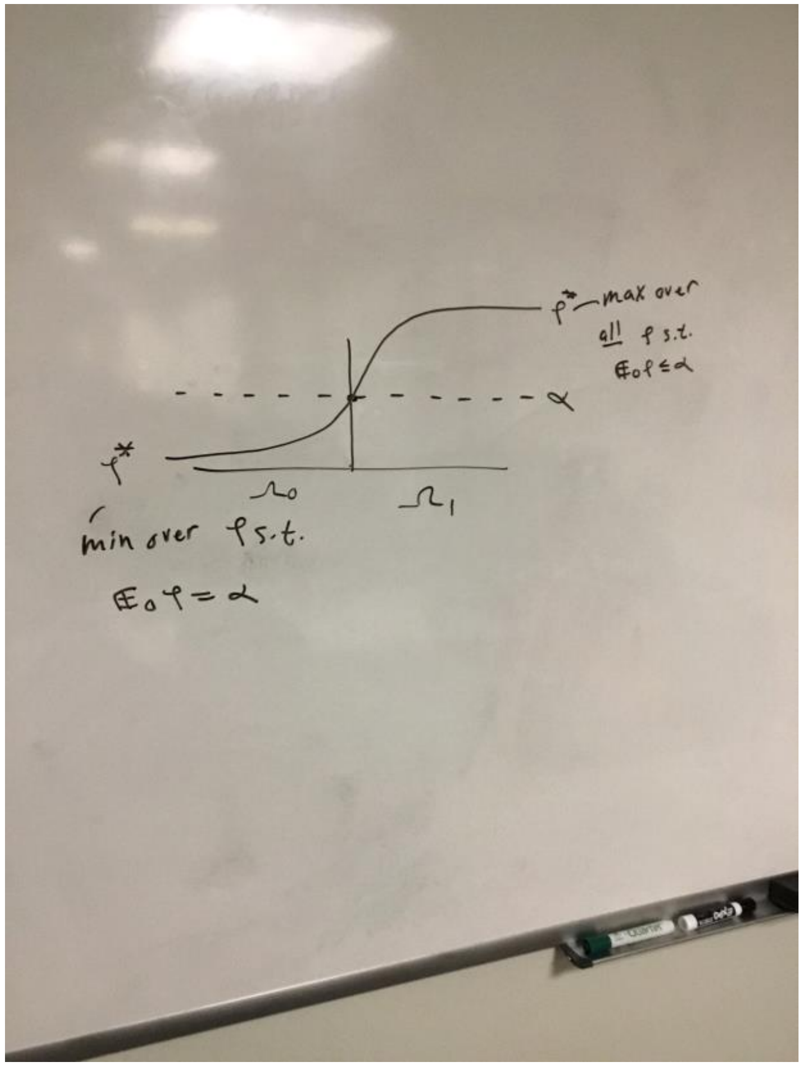
$$\Rightarrow p_{\theta_1} = 0 \qquad T(x) \geq c$$

$\Rightarrow P_{\theta_1}$ puts all mass on $\{T(X) < c\}$

$\Rightarrow P_{\theta_1}(\phi^* = 0) = 1$

$\Rightarrow \mathbb{E}_{\theta_1}(\phi^*) = 0$

which is minimal power.


φ^* — max over
all φ s.t.
$\mathbb{E}_0\varphi \le \alpha$
α
φ^*
Ω_0
Ω_1
min over φ s.t.
$\mathbb{E}_0\varphi = \alpha$

Lecture 2/16

Friday, February 16, 2018 11:15 AM

Ex: $X_1, \ldots, X_n \sim \text{Unif}(0, \theta)$

$\theta > 0$. Find UMP for

$H_0: \theta \leq 1 \quad H_1: \theta > 1$

$\forall\, \theta_1 < \theta_2$

$$\frac{p_{\theta_2}(x)}{p_{\theta_1}(x)} = \frac{\left(\frac{1}{\theta_2}\right)^n \mathbb{1}\{X_{(n)} < \theta_2\}}{\left(\frac{1}{\theta_1}\right)^n \mathbb{1}\{X_{(n)} < \theta_1\}}$$

$$= \left(\frac{\theta_1}{\theta_2}\right)^n \frac{\mathbb{1}\{X_{(n)} < \theta_2\}}{\mathbb{1}\{X_{(n)} < \theta_1\}}$$

$$= \begin{cases} \infty & \theta_1 \le X_{(n)} < \theta_2 \\ (\theta_1/\theta_2)^n & X_{(n)} < \theta_1 \\ ??? & X_{(n)} \ge \theta_2 \end{cases}$$

but by def we only consider points where

$$P_{\theta_1} \neq 0 \quad P_{\theta_2} \neq 0$$

So LR increasing in $X_{(n)}$. $\Rightarrow$ MLR in $X_{(n)}$.

So the UMP test is given by

$$\phi^* = \begin{cases} 1 & X_{(n)} > C \\ 0 & X_{(n)} < C \\ \gamma & X_{(n)} = C \end{cases}$$

$$E_{\theta=1}(\phi^*) = \mathbb{P}_{\theta=1}(X_{(n)} > C) + \gamma\, \mathbb{P}_{\theta=1}(X_{(n)} = C)$$

$$= \mathbb{P}_{\theta=1}(X_{(n)} > C)$$

$$= 1 - \left[\mathbb{P}(X_i \le C)\right]^n$$

$$= 1 - \left[\frac{c}{\theta}\right]^n = 1 - (c)^n = \alpha$$

So $c = (1-\alpha)^{1/n}$

and γ doesn't matter

in this case.

$$\varphi^* = \begin{cases} 1 & X_{(n)} > (1-\alpha)^{1/n} \\ 0 & X_{(n)} < (1-\alpha)^{1/n} \end{cases}$$

Now calculating the power func.

$$\mathbb{E}_\theta \varphi^* = 1 - \left(\tfrac{1}{\theta}\right)^n \quad \theta \geq 1$$

$$= 1 - \frac{1-\alpha}{\theta^n}$$

Now consider the test of the form

$$\tilde{\phi} = \begin{cases} 1 & X_{(n)} \geq 1 \\ \alpha & X_{(n)} < 1 \end{cases}$$

$$\mathbb{E}_\theta \tilde{\phi} = \mathbb{P}_\theta\{X_{(n)} \geq 1\}$$

$$+ \alpha \, \mathbb{P}(X_{(n)} < 1\}$$

Assumes $\theta \geq 1$

$$= 1 - \frac{1}{\theta^n} + \alpha \frac{1}{\theta^n}$$

$$= 1 - \frac{1-\alpha}{\theta^n}$$

So this is also U.M.P. which gives a much more natural huerist test. But what happens if $\theta < 1$.

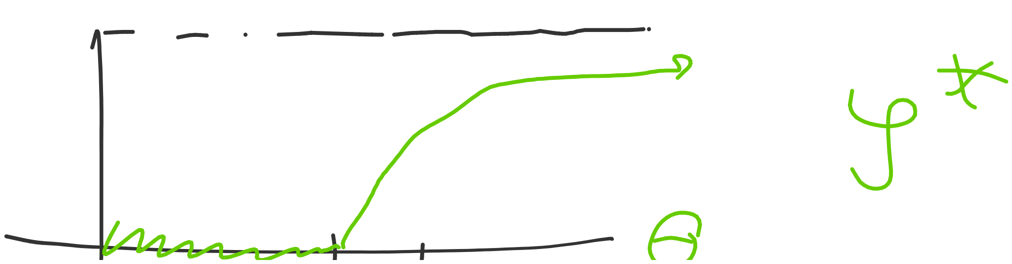
$$\mathbb{E}_\theta(\tilde{\varphi}) = \alpha$$

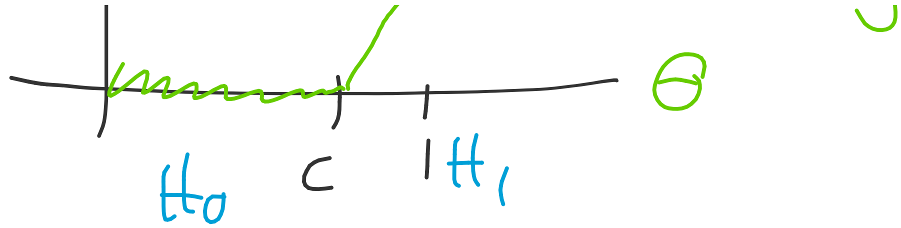
$$\mathbb{E}_\theta(\varphi^*) = 1 - \mathbb{P}(X_{(n)} \leq c)$$

$$= \begin{cases} 0 & \theta < c \\ 1 - \left(\frac{c}{\theta}\right)^n & c < \theta < 1 \end{cases}$$

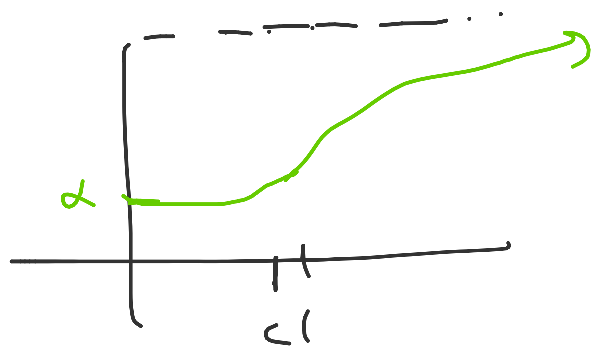
So all together

$$\mathbb{E}_\theta(\varphi^*) = \begin{cases} 1 - \frac{1-\alpha}{\theta^n} & \theta > (1-\alpha)^{1/n} \\ 0 & \theta < (1-\alpha)^{1/n} \end{cases}$$





θ
H_0
c
H_1



α
c

Lecture 2/20

Tuesday, February 20, 2018 11:15 AM

Ex: (Cauchy)

$$f_\theta(x) = \frac{1}{\pi} \cdot \frac{1}{1+(x-\theta)^2}$$

If $\{f_\theta : \theta \in \Theta\}$ is MLR in same $T(x)$ then

$\exists\, x \neq x'$ s.t. $T(x) < T(x')$

then $\forall\, \theta_1 < \theta_2$ by assumption of MLR.

$$\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} \leq \frac{f_{\theta_2}(x')}{f_{\theta_1}(x')}$$

- $$\frac{1+(x-\theta_1)^2}{1+(x-\theta_2)^2} \le \frac{1+(x'-\theta_1)^2}{1+(x'-\theta_2)^2}$$

$$\frac{1+(x-\theta_1)^2}{1+(x'-\theta_1)^2} \le \frac{1+(x-\theta_2)^2}{1+(x'-\theta_2)^2}$$

Consider

$$g(\theta) = \frac{1+(x-\theta)^2}{1+(x'-\theta)^2}$$

By assumption we require $g(\theta)$ to be nondecreasing. But

nondecreasing. But $g(\cdot)$ is decreasing in the neighborhood of x.

So $g(\cdot)$ is <u>not</u> nondecreasing.

Hence $\{f_\theta : \theta \in \Theta\}$ is not MLR.

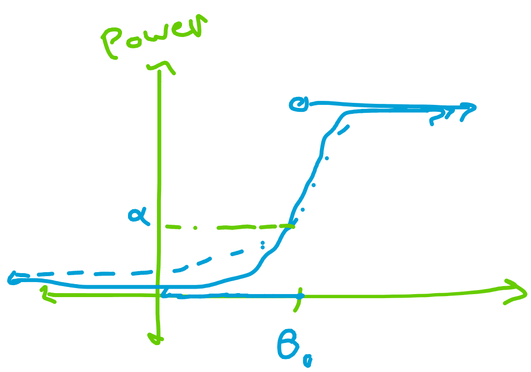
If the family is not MLR, what guarentees can we achieve about

our tests?

Locally Best Tests

Consider the test

$H_0: \theta \leq \theta_0 \qquad H_A: \theta > \theta_A$



- Our original optimization problem was

$$\min_{\theta \in \Omega_1} \beta(\theta) \text{ and } \max_{\theta \in \Omega_2} \beta(\theta)$$

· But really we look to maximize the derivative of $\beta(\theta)$ subject to $\beta(\theta_0) = \alpha$

We will use Lagrange to maximize

$$\max\left\{\beta'(\theta)\Big|_{\theta=\theta_0} - k\beta(\theta_0)\right\}$$

$$\beta(\theta) = \int \delta(x) f_\theta(x)\, dx$$

$$\beta'(\theta) = \int \delta(x) \frac{\partial f_\theta(x)}{\partial \theta} dx$$

So our maximization becomes

$$\max \int \delta(x) \left\{ \frac{\partial f_\theta(x)}{\partial \theta}\Big|_{\theta=\theta_0} - k f_\theta(x) \right\} dx$$

$$\Rightarrow \delta(x) = \begin{cases} 1 & \frac{\partial f_\theta(x)}{\partial x}\Big|_{\theta=\theta_0} > k f_{\theta_0}(x) \\ 0 & < \\ \gamma & = \end{cases}$$

Note: Here we assumed

$\beta(\theta)$ is diff. w.r.t. θ at $\theta=\theta_0$

Also note that

$$\frac{\partial f_\theta(x)}{\partial\theta} \Big/ f_\theta(x)$$

$$= \frac{\partial}{\partial\theta}\left(\log f_\theta(x)\right)$$

So our test can be reformulated as

$$\boxed{\delta(x) = \begin{cases} 1 & \partial/\partial\theta\left(\log f_\theta(x)\right)\Big|_{\theta=\theta_0} > k \\ 0 & < \\ \gamma & = \end{cases}}$$

Ex: (...)

Ex: (Cauchy)

$$\log f_\theta(x) = \log(1/x) + \log\left(\frac{1}{1+(x-\theta)^2}\right)$$

$$= -\ln(x) - \ln(1+(x-\theta)^2)$$

$$\frac{\partial \log f_\theta(x)}{\partial \theta} = \frac{2(x-\theta)}{1+(x-\theta)^2}$$

So the L.B.T.

$$\delta(x) = \begin{cases} 1 & \dfrac{2(x-\theta_0)}{1+(x-\theta_0)^2} > k \\ 0 & < \\ \gamma & = \end{cases}$$

Ex: Let

Ex: Let

$X_1, \dots, X_n \sim \text{iid } f_\theta(x) = \frac{1}{2} e^{-|x-\theta|}$

$H_0: \theta \leq \theta_0 \quad H_1: \theta > \theta_0$

Q: Find B.L.T.

Q: Find U.M.P.

Lecture 2/21

Wednesday, February 21, 2018 11:17 AM

Ex: $\frac{1}{2} e^{-|x-\theta|}$

$$\mathcal{L} = (1/2)^n \exp\left\{-\sum |X_i - \theta|\right\}$$

$$\log(\mathcal{L}) = n \log 1/2 - \sum_{i=1}^{n} |X_i - \theta|$$

$$\frac{\partial}{\partial \theta} \log \mathcal{L} = \sum_{i=1}^{n} \left\{ 1\{x_i > \theta\} - 1\{x_i < \theta\} \right\} \Big|_{\theta = \theta_0}$$

$$= \sum_{i=1}^{n} \left(1\{X_i > \theta_0\} - 1\{X_i < \theta_0\} \right)$$

$$= \sum_{i=1}^{n} \operatorname{sign}(x_i - \theta_0)$$

- Since the derivative

exists on $A \subseteq \Omega$

with $\mathbb{P}(A) = 1$.

This gives the L.B.T.

$$\phi^* = \begin{cases} 1 & \sum \text{sign}(x_i - \theta_0) > k \\ 0 & < \\ \gamma & = \end{cases}$$

Under the null

$$Y_i := \text{sign}(X_i - \theta_0) = \begin{cases} 1 & \mathbb{P}_0(X_i > \theta_0) \\ -1 & \mathbb{P}_0(X_i < \theta_0) \end{cases}$$

$$\mathbb{P}_0(X_i > \theta_0) = \int_{\theta_0}^{\infty} \frac{1}{2} e^{-|x - \theta_0|} dx = 1/2$$

So

$$Y_i = \begin{cases} 1 & 1/2 \\ -1 & 1/2 \end{cases}$$ and

$$\frac{Y_i+1}{2} \sim \text{Bern}(1/2)$$

So $\sum \frac{Y_i+1}{2} \sim \text{Binom}(n, 1/2)$

So for $W \sim \text{Binom}(n, 1/2)$

$$\sum Y_i \overset{D}{=} 2W - n$$

From here we can find γ and k.

Ex: (UMP) $\forall \theta_1 < \theta_2$

$$\frac{f_{\theta_2}(x_1, \ldots, x_n)}{f_{\theta_1}(x_1, \ldots, x_n)} = \exp\left\{\sum_{i=1}^{n} |x_i - \theta_1| - |x_i - \theta_2|\right\}$$

MLR hard to find if possible at all.

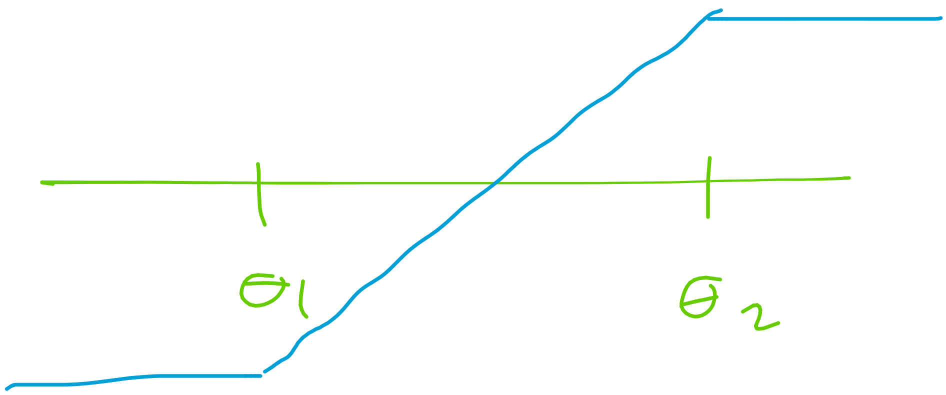
Special case $n=1$.

$$\frac{f_{\theta_2}}{f_{\theta_1}} = \exp\{|x-\theta_1| - |x-\theta_2|\}$$

$$\theta_1 - x - (\theta_2 - x) = \theta_1 - \theta_2 < 0$$

$$x - \theta_1 - (\theta_2 - x) = 2x - \theta_1 - \theta_2$$

$$x - \theta_1 - (x - \theta_2) = \theta_2 - \theta_1 > 0$$



So MLR in x.

and our test is

$$\phi^* = \begin{cases} 1 & x > c \\ 0 & x < c \\ \gamma & = \end{cases} = \begin{cases} 1 & T(x) > h \\ 0 & T(x) < h \\ \gamma & = \end{cases}$$

Two Sided Tests

Ex: $X \sim N(\theta, 1)$

$H_0: \theta = 0 \quad H_1: \theta \neq 0$

Consider $H_0: \theta = 0 \quad H_1: \theta > 0$

$H_0: \theta = 0 \quad H_1: \theta < 0$

Lecture 2/22

Friday, February 23, 2018 11:13 AM

Ex: $X \sim N(\theta, 1)$

$H_0: \theta = 0 \quad H_1: \theta \neq 0$

If ϕ^* is ump in the two sided case then it's certainly true that ϕ^* is ump for

$H_0: \theta = 0 \quad H_1: \theta > 0$

$$\phi^* = \begin{cases} 1 & f_1(x) > c\, f_0(x) \\ \gamma & = \\ 0 & < \end{cases}$$

$f_{\theta_1}(x)$

$$\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)} = \exp\left\{-\frac{1}{2}(x-\theta_1)^2 + \frac{1}{2}(x-\theta_0)^2\right\}$$

$$= \exp\left\{(\theta_1-\theta_0)x + \frac{1}{2}(\theta_0^2 - \theta_1^2)\right\}$$

This is MLR in x.

So the test is given by

$$\phi^* = \begin{cases} 1 & x > c \\ 0 & x < c \end{cases}$$

By uniqueness we can exclude "=" case.

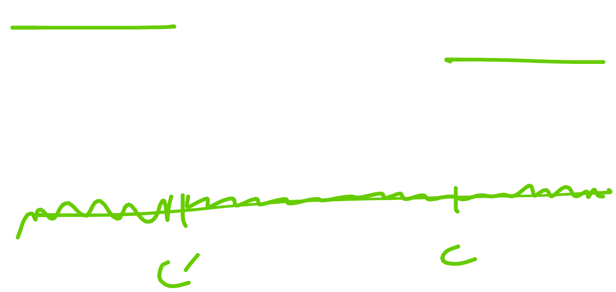
For the test

$H_0: \theta = 0 \quad H_1: \theta < 0$

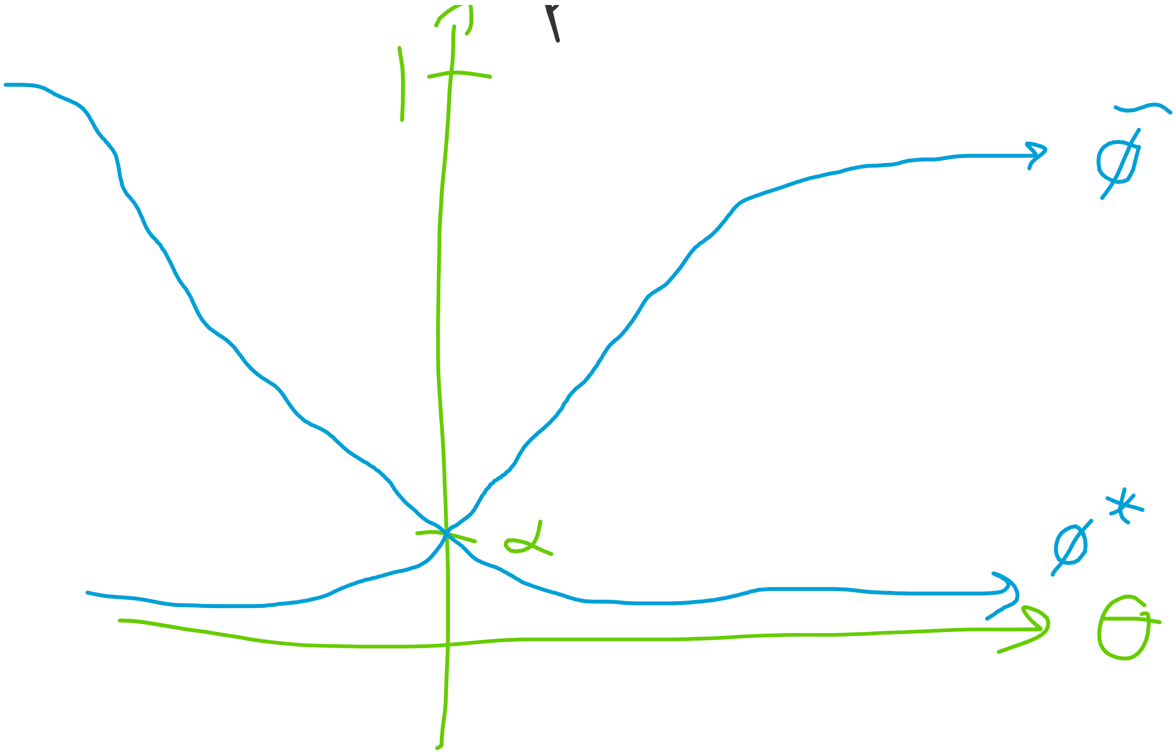
the UMP is given by

$$\tilde{\phi} = \begin{cases} 1 & x < c' \\ 0 & x > c' \end{cases}$$

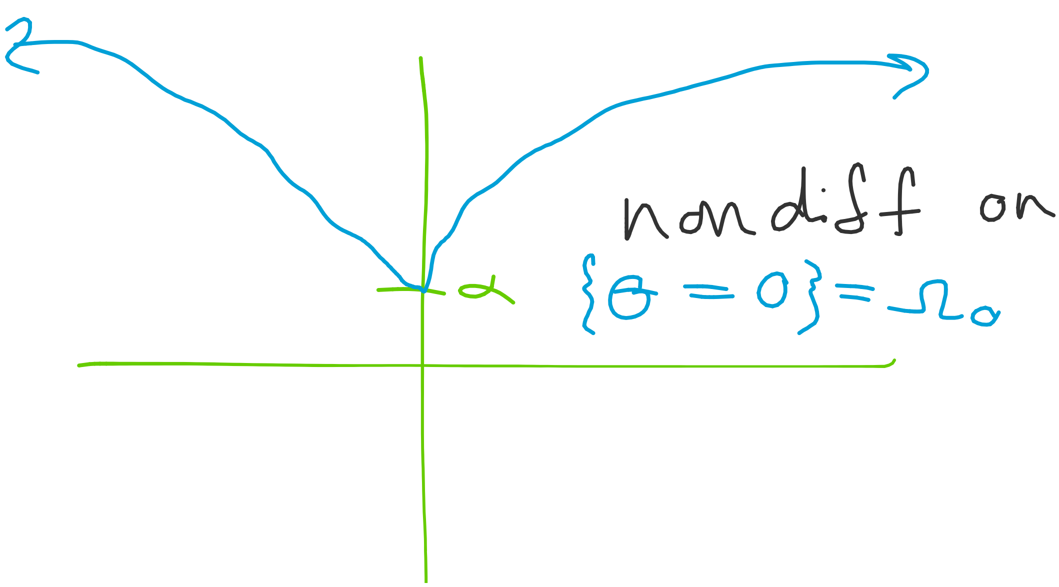
There is no function that exists such that that satisfies both of these equations.



If power



An optimal type test may look like



Unbiased Tests

<u>Def:</u> A test for $H_0: \theta \in \Omega_0$ vs $H_1: \theta \in \Omega_1$ is said to have <u>level</u> α if $\sup_{\theta \in \Omega_0} \beta(\phi) \leq \alpha$ and <u>size</u>

$$\sup_{\theta \in \Omega_0} \beta(\theta) = \alpha$$

<u>Def:</u> A test is said to be <u>unbiased</u> iff for $H_0: \theta \in \Omega_0$ vs $H_1: \theta \in \Omega_1$

if $\inf_{\theta \in \Omega_1} \beta(\theta) \geq \alpha$

<u>Def:</u> An <u>unbiased</u> test ϕ with size α is said to be UMPU iff

Def: An ...

is said to be UMPU iff for any unbiased test $\tilde{\phi}$ with size α we have

$$\beta_{\phi}(\theta) \geq \beta_{\tilde{\phi}}(\theta) \quad \forall\, \theta \in \Omega_1$$

Thrm: Assume $p_\theta(x) = h(x)e^{c(\theta)T(x) - B(\theta)}$ where $c(\theta)$ is strictly increasing.

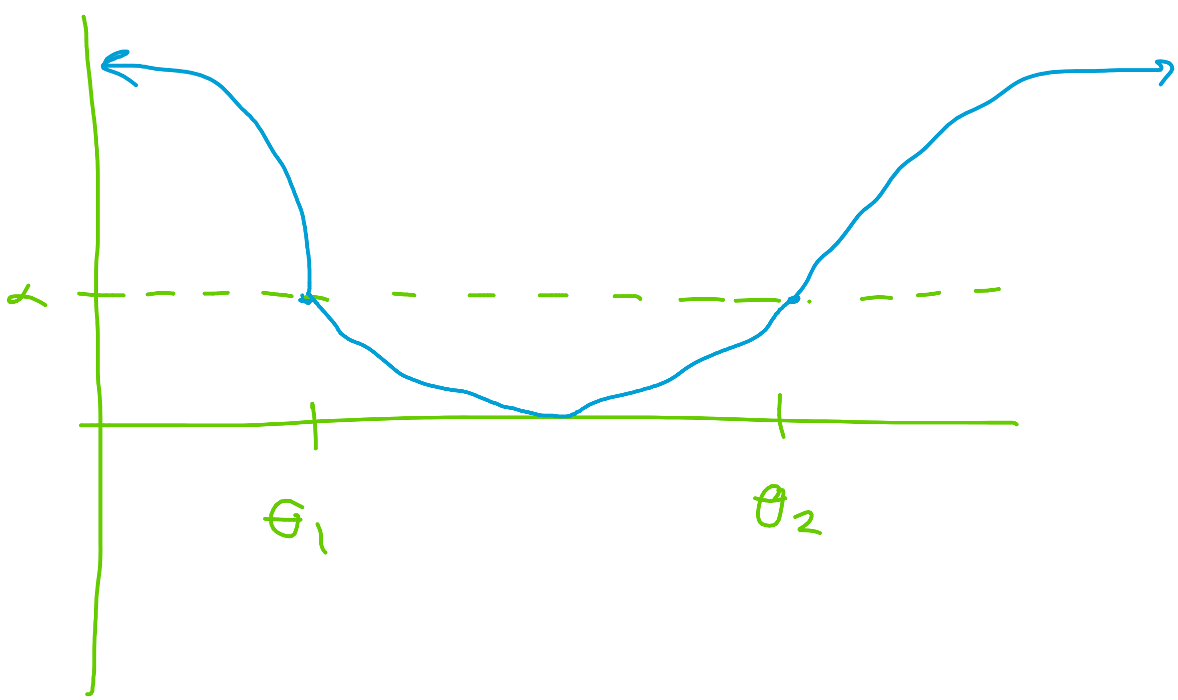
Then for the test $H_0: \theta_1 \leq \theta \leq \theta_2$

$H_1: \theta < \theta_1$ or $\theta > \theta_2$ Let

$$\phi^*(x) = \begin{cases} 1 & T(x) < c_1 \text{ or } T(x) > c_2 \\ 0 & c_1 < T(x) < c_2 \\ \gamma_1 & T(x) = c_1 \\ \gamma_2 & T(x) = c_2 \end{cases}$$

where $c_1, c_2, \gamma_1, \gamma_2$ are chosen where $c_1, c_2, \gamma_1, \gamma_2$ are chosen such that $\beta_{\phi^*}(\theta_1) = \beta_{\phi^*}(\theta_2) = \alpha$.

Then ϕ^* is UMPU with size α.



Lecture 2/26

Monday, February 26, 2018 11:10 AM

For two sided hyp. thesis tests the UMPU is given by

$$\phi^*(x) = \begin{cases} 1 & T(x) < c_1 \quad T(x) > c_2 \\ 0 & c_1 < T(x) < c_2 \\ \gamma_1 & T_1 = c_1 \\ \gamma_2 & T_2 = c_2 \end{cases}$$

s.t. $\beta_{\phi^*}(\theta_1) = \beta_{\phi^*}(\theta_2) = \alpha \in (0,1)$

<u>Pf:</u> By using Lagrange multipliers we maximize

$$\int \phi(x)\left\{ p_{\theta_3}(x) - k_1 p_{\theta_1}(x) - k_2 p_{\theta_2}(x) \right\} \mu(dx)$$

for any $\theta_3 \in \Omega_1$.

The optimizer is of the form

The optimizer is of the form

1 if $p_{\theta_3}(x) - k_1 p_{\theta_1}(x) - k_2 p_{\theta_2}(x) > 0$

0 if $\quad$ —— // —— < 0

γ if $\quad$ —— // —— $= 0$

Recall we assumed

$$p_\theta(x) = h(x)\exp\{c(\theta)T(x) - B(\theta)\}$$

So we just need to show the Lagrange solution matches the statement in the theorm.

$$p_{\theta_3}(x) - k_1 p_{\theta_1}(x) - k_2 p_{\theta_2}(x) > 0$$

$$e^{c(\theta_3)T(x) - B(\theta_3)} > k_1 e^{c(\theta_1)T(x) - B(\theta_1)} + k_2 e^{\overset{\theta_2}{\text{—//—}}}$$

$$1 > \frac{k_1 e^{(c(\theta_1) - c(\theta_3))T(x)}}{e^{\beta(\theta_1) - \beta(\theta_3)}} + \frac{k_2 e^{(c(\theta_2) - c(\theta_3))T(x)}}{e^{\beta(\theta_2) - \beta(\theta_3)}}$$

Assume that $\theta_3 < \theta_1$

then

$$a_1 e^{b_1 T(x)} + a_2 e^{b_2 T(x)} < 1$$

for $b_2 > b_1 > 0$

If $a_1 \leq 0$ $a_2 \leq 0$ then $1 < -a_1 e^{b_1 T(x)} - a_1 e^{b_2 T(x)}$

which violates $\alpha \in (0, 1)$.

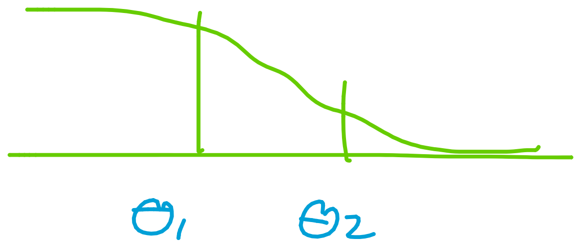
If $a_1 > 0$ $a_2 > 0$ then

the function is Monotone

in $T(x)$. So the power

function will look something

function will look something like



But $\beta(\theta_1) = \beta(\theta_2) = \alpha$

which can not occur.

When a_1, a_2 have different signs. (Suppose $a_1 < 0$ $a_2 > 0$)

Then

$$\frac{d}{dT(x)}\left(a_1 e^{\beta_1 T(x)} + a_2 e^{\beta_2 T(x)}\right)$$

$$= a_1 b_1 e^{b_1 T(x)} + a_2 b_2 e^{b_2 T(x)}$$

$$= e^{b_1 T(x)} \left\{ a_1 b_1 + a_2 b_2 e^{(b_2 - b_1) T(x)} \right\}$$

increasing or decreasing in $T(x)$.

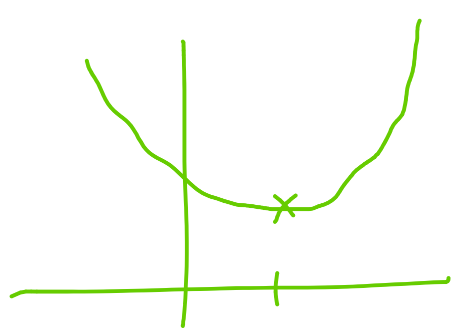
So our function looks like

(Fill in details from book). $\square$

We can use this test for

$H_0: \theta = \theta_0 \quad H_A: \theta \neq \theta_0$

For UMPU the power looks like



$\theta_0 = \Omega_0$

So our optimizer is given by

$$\arg\max_{\phi} \int \phi \left\{ p_{\theta_3}(x) - k_1 p_{\theta_0} - k_2 p'_{\theta_0} \right\}$$

Pr: $X \sim Exp(\theta)$

$H_0 : \theta = 1 \qquad H_A : \theta \neq 1.$

Lecture 2/28

Wednesday, February 28, 2018 11:16 AM

$X \sim \theta e^{-\theta x}$ $H_0: \theta = 1$ $H_1: \theta \neq 1$

$T(x) = -x$ so the UMPU

is given by

$$\phi^*(x) = \begin{cases} 1 & -x < c_1 \quad -x > c_2 \\ 0 & c_1 \leq -x \leq c_2 \\ \gamma_1 & -x = c_1 \\ \gamma_2 & -x = c_2 \end{cases}$$

Continuous so γ_1 γ_2 don't really matter

$$\phi^* = \begin{cases} 1 & -x < c_1 \text{ \& } -x > c_2 \\ 0 & c_1 \leq -x \leq c_2 \end{cases}$$

$$= \begin{cases} 1 & x < c_1' \quad x > c_2' \\ 0 & c_1' \le x \le c_2' \end{cases}$$

$$\mathbb{E}_0(\varphi^*) = \mathbb{P}_{\theta=1}(x < c_1' \text{ or } x > c_2)$$

$$= 1 - \mathbb{P}_{\theta=1}(c_1' \le x \le c_2')$$

$$= 1 - \int_{c_1'}^{c_2'} e^{-x} dx$$

$$= 1 - e^{-c_1} + e^{-c_2} = \alpha$$

And we also put a constaint of the derivative of the power function.

$$\frac{d\,\mathbb{E}_0(\theta)}{d\theta} = \frac{d}{d\theta}\left(e^{-c_2\theta} - e^{-c_1\theta}\right)$$

$$= c_1 e^{-c_1\theta} - c_2 e^{-c_2\theta} = 0$$

which give

$c_1 = 0.04$ for a given

$c_2 = 4.76$ $\alpha = 0.05$

Thrm: Assume

$$p_\theta(x) = h(x) e^{c(\theta)T(x) - B(\theta)}$$

where $c(\theta)$ is strictly increasing.

Let

$$\varphi^*(x) = \begin{cases} 1 & c_1 < T(x) < c_2 \\ 0 & T(x) < c_1,\ T(x) > c_2 \\ \gamma_i & \ldots \end{cases}$$

$$\varphi^*(x) = \begin{cases} 0 & T(x) < c_1,\ T(x) > c_2 \\ \gamma_1 & T(x) = c_1 \\ \gamma_2 & T(x) = c_2 \end{cases}$$

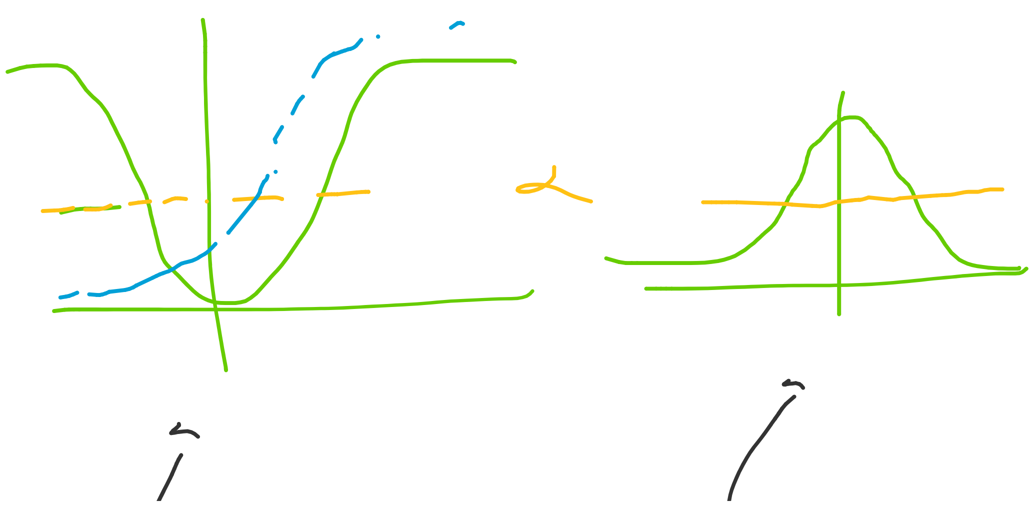
where $c_1, c_2, \gamma_1, \gamma_2$ are chosen such that

$$\beta_{\phi^*}(\theta_1) = \beta_{\phi^*}(\theta_2) = \alpha$$

for $0 < \alpha < 1$. Then φ^* is UMP for

H_0: $\theta \le \theta_1$ or $\theta \ge \theta_2$

H_1: $\theta_1 < \theta < \theta_2$



exists test that are better on one side

no such tests exist.

So UMPU $\Rightarrow$ UMP

Testing with Nusiance Param.

$X \sim N(\mu, \sigma^2)$

nusiance parameter.

$H_0: \mu \geq 0 \quad H_1: \mu < 0.$

Lecture 3/2

Friday, March 2, 2018 11:16 AM

Testing Under Nusiance

Suppose $p_\theta(t) = h(t)e^{t^T\theta - B(t)}$

for $t \in \mathbb{R}^n$

where $(\theta_1, \ldots, \theta_p)^T$ and

$t = (t_1, \ldots, t_n)$ and

we want to test

$H_0: \theta_1 \leq 0 \quad H_1: \theta_1 > 0$

($\theta_2, \ldots, \theta_p$ are nusiance parameters).

Def: Let $\overline{\Omega}_0$ and $\overline{\Omega}_1$ be

the closures of Ω_0, Ω_1

and set

$$\Omega_B = \overline{\Omega}_0 \cap \overline{\Omega}_1$$

A test is said to be

<u>α-similar</u> if $\beta(\theta) = \alpha$

$\forall \theta \in \Omega_B$.

<u>Thrm:</u> Suppose $\beta_\varphi(\theta)$

is continuous in θ

for all φ then if

φ^* is UMP, α-similar

with size α then

φ^* is UMPU.

<u>Pf:</u> We need to show

UMP + U.

UMP: Consider $\tilde{\varphi} = \alpha$

then $\hat{\varphi}$ is α-similar.

So φ^* UMP α-similar so

$\beta_{\varphi^*}(\theta) \geq \beta_{\hat{\varphi}}(\theta) = \alpha \ \forall \theta \in \Omega_1$

So φ^* is UMP.

u: For any unbiased test φ^0 with size α

$\Rightarrow$ φ^0 is α-similar

$\Rightarrow$ $\beta_{\varphi^*}(\theta) \geq \beta_{\varphi^0}(\theta) \ \forall \theta \in \Omega_1$

$\Rightarrow$ φ^* UMPU ∎

Is the power function always continuous?

always continuous?

$$\beta(\theta) = \int \varphi(x) h(x) e^{T(x)c(\theta) - \beta(\theta)} \mu(dx)$$

So if we can interchange.

$$\lim \int \# = \int \lim \#$$

then it suffices to check

Cont. of

$$h(x) e^{T(x)c(\theta) - B(\theta)}$$

$$\Longleftrightarrow$$

need to check cont. of

$$c(\theta) \ \& \ B(\theta)$$

But note that

$$e^{B(\theta)} = \int h(x) e^{T(x) c(\theta)} \mu(dx)$$

So really only need to check cont. of $c(\theta)$ given $\lim \int \# = \int \lim \#$

<u>Thrm:</u> UMPU test is given by

$$\varphi^*(\vec{T}) = \begin{cases} 1 & T_1 > g(T_2, \ldots, T_p) \\ 0 & < \\ \gamma & = \end{cases}$$

s.t. g, γ are chosen s.t.

$$\mathbb{E}_{\theta_1 = 0}(\varphi^* \mid T_2, \ldots, T_p) = \alpha$$

Pf:

$$T_1 | T_2 \cdots T_p = \frac{h(t) e^{t^T\theta - \beta(\theta)}}{\int h(t) e^{t^T\theta - \beta(\theta)} dt_1}$$

$$= \frac{h(t) e^{t^T\theta}}{\int h(t) e^{t^T\theta} dt_1}$$

$$= \frac{h(t) e^{t_1\theta_1} e^{t_2\theta_2} \cdots e^{t_p\theta_p}}{\int h(t) e^{t_1\theta_1} e^{t_2\theta_2} \cdots e^{t_p\theta_p} dt_1}$$

$$= \frac{h(t)}{\int h(t) e^{t_1\theta_1} dt_1}$$ independent of $T_2, \ldots, T_p$, $\theta_2, \ldots, \theta_p$

$\Rightarrow \varphi^*$ UMP

among all $\hat{\varphi}$ s.t.

$$\mathbb{E}_{\theta_1=0}(\hat{\varphi} \mid T_2, \dots, T_p) = \alpha$$

by the same proof for UMP on nonconditional tests.

For any $\hat{\varphi}$ that is α similar we have

$$\mathbb{E}_{\theta=0}\left\{\mathbb{E}_{\theta_1=0}(\hat{\varphi} \mid T_2, \dots, T_p)\right\} = \alpha$$

Lecture 3/12

Monday, March 12, 2018 11:19 AM

Consider

$$P_{\vec{\theta}}(\vec{t}) = h(\vec{t})e^{\vec{t}^T\vec{\theta} - B(\vec{\theta})}$$

and we wish to test

$H_0: \theta_1 \leq 0 \quad H_1: \theta_1 > 0$

with nusiance parameters

$\theta_2, \dots, \theta_p$ We use

$$\varphi^*(T) = \begin{cases} 1 & T_1 > g(T_2, \dots, T_p) \\ 0 & < \\ \gamma & = \end{cases}$$

Goal: find dist of $T_1|_{(T_2, \dots, T_p)}$

and hope it doesn't

depend on $T_2, \dots, T_p$.

$$f_{T_1|T_2,\ldots,T_p} = \frac{f_{T_1,\ldots,T_p}}{\int f_{T_1,\ldots,T_p}\, dT_1}$$

$$= \frac{h(\vec{t})\, e^{t_1\theta_1+\cdots+t_p\theta_p - B(\vec{\theta})}}{\int h(\vec{t})\, e^{t_1\theta_1+\cdots+t_p\theta_p - B(\vec{\theta})}\, dT_1}$$

$$= \frac{h(t)\, e^{t_1\theta_1}\, e^{t_2\theta_2+\cdots+t_p\theta_p - B(\theta)}}{e^{t_2\theta_2+\cdots+t_p\theta_p - B(\theta)} \int h(t)\, e^{t_1\theta_1}\, dt_1}$$

$$= \frac{h(t)\, e^{t_1\theta_1}}{\int h(t)\, e^{t_1\theta_1}\, dt_1}$$

← not a function of $(T_2,\ldots,T_p)$

From here we can find UMP,

From here

tests for all conditional tests.

Q: How can we show it is UMPU?

Sufficient to show UMP for α-similar tests.

For any α-similar test $\breve{\varphi}$ we have $\mathbb{E}_{\vec{\theta}=0}\,\breve{\varphi} = \alpha$

This implies by tower property

$$\mathbb{E}_{\theta_1=0}\left[\mathbb{E}(\breve{\varphi} \mid T_2, \dots, T_p)\right] = \alpha$$

If $\mathbb{E}(\tilde{\varphi} \mid T_2, \dots, T_p) = \alpha$

and φ^* is UMP by conditional tests we see

$$\mathbb{E}_1(\tilde{\varphi}) \leq \mathbb{E}_1(\varphi^*)$$

Therefore it suffices to show

$$\mathbb{E}(\tilde{\varphi} \mid T_2, \dots, T_p) = \alpha$$

Need to show

$$\mathbb{E}(\mathbb{E}(\ast)) = \alpha \Rightarrow \mathbb{E}(\ast) = \alpha$$

The density of $T_2, \dots, T_p$ is given by

$$\vec{t}^{\,T}\theta - \beta(\theta) \quad \dots \quad t_1\theta_1$$

$$e^{\vec{t}^T\theta - \beta(\theta)} \int h(\vec{t}) e^{t_1 \theta_1} \, dt_1$$

when $\theta_1 = 0$ the density of $T_2, \dots, T_p$ has an exponential family and is $(T_2, \dots, T_p)$ complete for $(\theta_2, \dots, \theta_p)$

By def of completeness

$$\mathbb{E}_{\theta_1}\left[\mathbb{E}(\varphi \mid T_2, \dots, T_p)\right] = \alpha$$

$$\mathbb{E}_{\theta_1}[\alpha] = \alpha$$

then by completeness

$$\mathbb{E}(\varphi \mid T_2, \dots, T_p) = \alpha. \quad \blacksquare$$

Ex: $X_1, \dots, X_n \sim \text{iid } N(\mu, \sigma^2)$

$H_0: \mu \le 0 \quad H_1: \mu > 0$

Applying the theorem, we write the joint density as

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}$$

$$= (\sqrt{2\pi}\sigma)^{-n} \exp\left\{\frac{-1}{2\sigma^2} \sum (x_i-\mu)^2\right\}$$

$$= (\sqrt{2\pi}\sigma)^{-n} \exp\left\{\frac{-1}{2\sigma^2} \sum (x_i - 2\mu x_i + \mu^2)\right\}$$

$$= \exp\left\{\frac{-1}{2\sigma^2}\left\{\sum x_i - 2\mu \sum x_i + n\mu^2 - n\log(\sqrt{2\pi}\sigma)\right]\right.$$

Lecture 3/14

Wednesday, March 14, 2018 11:16 AM

Suppose $x_i \sim iid\ N(\mu, \sigma^2)$

and we want to test

$H_0: \mu \leq 0 \quad H_A: \mu > 0$

then we write the joint

likelihood.

$f(x_1, \dots, x_n)$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-n\mu^2/2\sigma^2} e^{\mu/\sigma^2 \sum x_i - \frac{1}{2\sigma^2}\sum x_i^2}$$

So

$T_1(x) = \sum x_i$ sufficient for $\frac{\mu}{\sigma^2}$

$T_2(x) = \sum x_i^2$ sufficient for $\frac{-1}{2\sigma^2}$

Then the form of the

UMPU is given by

UMPU is given by

$$\varphi^*(x) = \begin{cases} 1 & T_1(x) > g(T_2(x)) \\ 0 & < \\ \gamma & = \end{cases}$$

where g, γ are chosen

s.t. $\mathbb{E}_{\theta_1 = 0}(\varphi^* \mid T_2) = \alpha$.

As $\sum X_i$ $\sum X_i^2$ are

not independent $T_1 > g(T_2)$

hard to analyze. But

by Basu's Thm $\bar{X} \perp S^2$.

When $\theta_1 = \dfrac{\mu}{\sigma^2} = 0$

$\Longrightarrow$ $\mu = 0$ and we have

$X_1 \ldots X_n \sim$ iid $N(0, \sigma^2)$.

$X_1, \ldots, X_n \sim$ iid $N(0, \sigma^2)$.

then $T_2 =$

couldn't figure it out...

Making a guess for $g(\cdot)$

Conjecture: $g(T_2) = c\sqrt{T_2}$

for c independent of σ^2.

$$\sum (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2$$
$$= T_2 - nT_1^2$$

So we know

$$\frac{T_1}{\sqrt{T_2 - nT_1^2}} \sim T$$

We know

$$c' > 0 \Rightarrow T_1^2 > (c')^2 \left(T_2 - nT_1^2\right)$$

$$T_1^2 > \frac{(c')^2 T_2}{1 + n(c')^2}$$

So

$$\frac{T_1}{\sqrt{T_2}} > c \iff \frac{T_1}{\sqrt{T_2 - nT_1}} > c'$$

for $c = \dfrac{(c')^2}{1 + n(c')^2}$

Lecture 3/19

Monday, March 19, 2018 11:17 AM

Testing Under Nusiance

$X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

$H_0: \sigma^2 = 1 \quad H_1: \sigma^2 \neq 1$

Writing the joint likelihood

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{ - \frac{\sum (x_i - \mu_i)^2}{2\sigma^2} \right\}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{\left\{ - \underbrace{\sum x_i^2}_{T_2} \underbrace{/2\sigma^2}_{\theta_2} + \underbrace{\mu/\sigma^2}_{\theta_1} \underbrace{\sum x_i}_{T_1} - \frac{\mu^2}{2} \right\}}$$

$$= \exp\left\{\theta_1 T_1 + \theta_2 T_2 - B(\theta_1, \theta_2)\right\}$$

This is an exponential family so the UMPU is given by

$$\phi^*(X_1, \ldots, X_n) = \begin{cases} 1 & T_2 > h(T_1) \quad T_2 < g(T_1) \\ 0 & \text{o.w.} \end{cases}$$

where h, g chosen s.t.

$$\mathbb{E}_{\theta_2 = -1/2}\left(\phi^* \mid T_1\right) = \alpha$$

$$\frac{\partial}{\partial \theta_2} \mathbb{E}\left(\phi^* \mid T_1\right)\Big|_{\underbrace{\theta_2 = -1/2}_{\text{Boundary}}} = 0$$

We can change the form of the test to be based on

$$T_2 - \frac{T_1^2}{n} > \tilde{h}(T_1)$$

$$\sum x_i^2 - \frac{(\sum x_i)^2}{n} > \tilde{h}(T_1)$$

$$= \sum (x_i - \bar{x})^2 > \tilde{h}(T_1)$$

independent of T_1 } makes the boundary calculations a lot easier

$\mathbb{E}_{\theta_2} = -1/2 \; (t^*(T_1)$

$$= \mathbb{E}_{\theta_2=-1/2}(\varphi^*)$$

and the UMP is given by

$$\varphi^* = \begin{cases} 1 & T_2 - \frac{T_1^2}{n} > \tilde{h} \quad T_2 - \frac{T_1^2}{n} < \tilde{g} \\ 0 & \text{o.w.} \end{cases}$$

Moreover we see

$$\frac{1}{\sigma^2}\left(T_2 - \frac{T_1^2}{n}\right) = \underbrace{\sum \underbrace{\frac{(X_i - \bar{X})^2}{\sigma^2}}_{Z}}_{\chi^2_{n-1}}$$

Using this

$$\mathbb{E}_{\theta_2=-1/2}(\varphi^*) = 1 - \int_{\tilde{g}}^{\tilde{h}} f_{\chi^2_{n-1}} = 1-\alpha$$

$$\frac{d}{d\sigma^2}\left(\int_{\hat{n}}^{+\infty} + \int_{-\infty}^{\bar{q}} f_{\sigma^2 \chi^2_{n-1}} \right)\Bigg|_{\sigma^2=1} = 0$$

- First fomula controlls the size of the rej. region
- Second constaint ensures they are symmetric

Lecture 3/23

Friday, March 23, 2018 11:16 AM

For tests in an exponential family we do our tests in terms of $T(x)$

$$\beta_\varphi(\theta) = \mathbb{E}_\theta(\varphi(T(x)))$$

$$= \int \varphi(t) h(t) e^{t\theta - B(\theta)} dt$$

$$\Rightarrow \beta'_\varphi(\theta) = \int \varphi(t) h(t) e^{t\theta - B(\theta)} \{t - B'(\theta)\}$$

$$= \mathbb{E}_\theta(T\varphi(T)) - B'(\theta)\mathbb{E}_\theta(\varphi(T))$$

$B(\theta)$ is cumulant so

$$B'(\theta) = \mathbb{E}(T)$$

With the criterion

$$\mathbb{E}_{\theta_0}(\varphi) = \alpha$$

$$\mathbb{E}_{\theta_0}(T(\varphi(T))) = \mathbb{E}_{\theta_0}(T)\underbrace{\mathbb{E}_{\theta_0}(\varphi(T))}_{\alpha}$$

not constant wrt θ

Ex: Suppose $X_1, \dots, X_n \overset{iid}{\sim} N(\mu_X, \sigma_X^2)$

$$Y_1, \dots, Y_n \overset{iid}{\sim} N(\mu_Y, \sigma_Y^2)$$

$H_0: \sigma_X^2 = \sigma_Y^2 \qquad H_1: \sigma_X^2 \neq \sigma_Y^2$

Writing the joint likelihood

$$\left(\frac{1}{2\pi\sigma_x\sigma_y}\right)^n e^{-\sum\frac{(x_i-\mu_x)^2}{2\sigma_x^2} - \sum\frac{(y_i-\mu_y)^2}{2\sigma_y^2}}$$

$$= \left(\frac{1}{2\pi\sigma_x\sigma_x}\right)^n e^{-\frac{\sum x_i^2}{2\sigma_x^2} + \frac{\sum x_i\mu_x}{\sigma_x^2}}$$

$$\times\, e^{-\frac{\sum y_i^2}{2\sigma_y^2} + \frac{\sum y_i\mu_y}{\sigma_y^2} + f(\mu_x,\mu_y,\sigma_x,\sigma_y)}$$

$$T_1(\underline{x}) = \sum x_i$$

$$T_2(\underline{x}) = \sum x_i^2$$

$$S_1(\underline{y}) = \sum y_i$$

$$S_2(\underline{y}) = \sum y_i^2$$

$$\mathcal{L} = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right)^n e^{\frac{\mu_x}{\sigma_x}T_1} e^{\frac{\mu_y}{\sigma_y}S_1} \underbrace{e^{-\frac{1}{2\sigma_x^2}T_2}}$$

$$\underbrace{e^{-1/2\sigma_x^2 S_2}}$$

Combine those

$$= \# \, e^{-\underbrace{\left(\frac{1}{2\sigma_x^2} - \frac{1}{2\sigma_y^2}\right)}_{\theta} T_2}$$

$$e^{-1/2\sigma_y^2 (\underbrace{S_2 + T_2}_{u})}$$

So the UMPU test is given by

$$\begin{cases} 1 & T_2 > g(T_1, S_1, u) \\ & T_2 < h(T_1, S_1, u) \end{cases}$$

$$\varphi(\underline{X},\underline{Y}) = \begin{cases} 1 & T_2 < g(\ldots) \\ & T_2 < h(T_1, S, u) \\ 0 & g(T_1, S_1, u) \le T_2 \le h(T_1, S_1, u) \\ \gamma_1 & T_2 = g \\ \gamma_2 & T_2 = h \end{cases}$$

Focusing on $T_2 > g$

$$T_2 - \frac{T_1^2}{n} > g - \frac{T_1^2}{n}$$

$$= \bar{g}(T_1, u, S_1)$$

So $T_2 - \frac{T_1^2}{n} \perp\!\!\!\perp S_1, T_1$

Need independence of u

Well

$$T_2 - \frac{T_1^2}{n} > \tilde{g}\left(T_1, S_1, u \pm \frac{T_1^2}{n} \pm \frac{S_1^2}{n}\right)$$

$$= \hat{g}\left(T_1, S_1, u - \frac{T_1^2}{n} - \frac{S_1^2}{n}\right)$$

Recall

$$T_2 - \frac{T_1^2}{n} \sim \sigma_x^2 \chi^2_{n-1}$$

$$S_2 - \frac{S_1^2}{n} \sim \sigma_x^2 \chi^2_{n-1}$$

$$u - \frac{T_1^2}{n} - \frac{S_1^2}{n}$$

$$= \left(T_2 - \frac{T_1^2}{n}\right) + \left(S_2 - \frac{S_1^2}{n}\right)$$

So

$$\frac{T_2 - \frac{T_1}{n}}{U - \frac{T_1^2}{n} - \frac{S_1^2}{n}} > g^\circ\left(T_1, S_1, U - \frac{T_1^2}{n} - \frac{S_1^2}{n}\right)$$

Top and bottom ind. of T_1, S_1. Suffices to show ratio ind. of $U - \frac{T_1^2}{n} - \frac{S_1^2}{n}$.

$W \sim \chi^2_{n-1}$ then

$V \sim \chi^2_{n-1}$

$$W + V \perp\!\!\!\perp \frac{W}{W+V} \sim F$$

As $t_2 - T_1^2 \sim \sigma^2 \chi^2$

As $t_2 - \frac{T_1^2}{n} \sim \sigma_x^2 \chi_{n-1}^2$

$$U - \frac{T_1^2}{n} - \frac{S_1^2}{n} = \left(T_2 - \frac{T_1^2}{n}\right) + \left(S_2 + \frac{S_1^2}{n}\right)$$

$$\sim \sigma_x^2 \chi_{n-1}^2 + \sigma_y^2 \chi_{n-1}^2$$

Hence we see that

$$\frac{T_2 - \frac{T_1^2}{n}}{U - \frac{T_1^2}{n} - \frac{S_1^2}{n}} \perp\!\!\!\perp U - \frac{T_1^2}{n} - \frac{S_1^2}{n}$$

So our statistic is the F and is UMPU.

Lecture 3/26

Monday, March 26, 2018 11:16 AM

Suppose we have data

B\A	Yes	No
Yes	N_{11}	N_{12}
No	N_{21}	N_{22}

$$(N_{11}, N_{12}, N_{21}, N_{22}) \sim \text{Multi}(n, p_{11}, p_{12}, p_{21}, p_{22})$$

Goal: Test if A and B are independent

Under the assumption of independence (the null here)

$$\mathbb{P}(A=1, B=1) = \mathbb{P}(A=1)\,\mathbb{P}(B=1)$$

$$\mathbb{P}(A=1, B=1) = \mathbb{P}(A=1)\mathbb{P}(B=1)$$

$$(p_{11}+p_{12})(p_{11}+p_{21}) = p_{11}$$

$$p_{11}(p_{11}+p_{12}+p_{21}) + p_{12}p_{21} = p_{11}$$

$$p_{11}(p_{11}+p_{12}+p_{21}) + p_{21}p_{12} = p_{11}\left(\sum p_{ij}\right)$$

$$p_{12}p_{21} = p_{11}p_{22}$$

So our test is given by

$H_0: \; p_{12}p_{21} = p_{11}p_{22}$

$H_A: \; p_{12}p_{21} \neq p_{11}p_{22}$

The likelihood is then given by

$$\binom{n}{N_{11}\, N_{12}\, N_{21}\, N_{22}} p_{11}^{N_{11}} p_{12}^{N_{12}} p_{21}^{N_{21}} p_{22}^{N_{22}}$$

$$= \frac{n!}{N_{11}!\, N_{12}!\, N_{21}!\, N_{22}!} e^{\sum N_{ij} \log p_{ij}}$$

So we have an expotential family

$$= \frac{n!}{N_{11}!\, N_{21}!\, N_{12}!\, N_{22}!} e^{N_{11} \log \frac{p_{11}}{p_{22}} + N_{12} \log \frac{p_{12}}{p_{22}}}$$

$$\times e^{N_{21} \log \frac{p_{21}}{p_{22}} + n \log p_{22}}$$

So we need to change the test

the test

$$H_0: \frac{p_{11} p_{22}}{p_{21} p_{12}} = 1 \quad \text{or}$$

$$H_0: \log\left(\frac{p_{11} p_{22}}{p_{21} p_{12}}\right) = 0$$

So rewriting the likelihood further

$$= \frac{n!}{N_{11}! N_{12}! N_{21}! N_{22}!} e^{N_{11}\left\{\log \frac{p_{11} p_{21}}{p_{12} p_{21}} + \log \frac{p_{12}}{p_{22}} + \log \frac{p_{21}}{p_{22}}\right\} + N_{12} \log \frac{p_{12}}{p_{22}} + \cdots}$$

Same as before

$$= \frac{n!}{N_{11}! N_{21}! N_{12}! N_{22}!} e^{\underbrace{N_{11}}_{T_1} \underbrace{\log\left(\frac{p_{11} p_{22}}{p_{12} p_{21}}\right)}_{\theta_1}}$$

$$e^{\underbrace{(N_{11}+N_{12})}_{T_2}\underbrace{\log\left(\frac{p_{12}}{p_{22}}\right)}_{\theta_2}+\underbrace{(N_{11}+N_{21})}_{T_3}\underbrace{\log\frac{p_{21}}{p_{22}}}_{\theta_3}}$$

$$e^{n\log p_{22}}$$

UMPU given by

$$\phi^*(N_{11}, N_{21}, N_{12}, N_{22})$$

$$= \begin{cases} 1 & T_1 > g(T_2,T_3) \text{ or } T_1 < h(T_2,T_3) \\ 0 & g < T_1 < h \\ \gamma_1 & T_1 = g \\ \gamma_2 & T_1 = h \end{cases}$$

$$\mathbb{P}_0(T_1 = t_1 \mid T_2 = t_2, T_3 = t_3)$$

$$= \frac{\dfrac{n!}{N_{11}!\,N_{12}!\,N_{21}!\,N_{22}!}\; e^{\sum \theta_i T_i + B(\theta_1, \theta_2, \theta_3)}}{\sum \dfrac{n!}{\cdot}\; e^{\sum \theta_i T_i + B(\theta_1, \theta_2, \theta_3)}}$$

$$\sum_{t_1} \frac{n!}{N_{11}!\,N_{12}!\,N_{21}!\,N_{22}!} e^{\sum \theta_i T_i + B(\theta_1, \theta_2, \theta_3)}$$

Under the null $\theta_1 = 0$
and we get

$$= \frac{\dfrac{n!}{t_1!\,(t_2-t_1)!\,(t_3-t_1)!\,(n-t_2-t_3+t_1)!}}{\displaystyle\sum_{t_1} \frac{n!}{t_1!\,(t_2-t_1)!\,(t_3-t_1)!\,(n-t_2-t_3+t_1)!}}$$

$$= \frac{\binom{t_2}{t_1}\binom{n-t_2}{t_3-t_1}}{\displaystyle\sum_{t_1} \binom{t_2}{t_1}\binom{n-t_2}{t_3-t_1}}$$

Lecture 3/28

Wednesday, March 28, 2018 11:19 AM

Recall we had

$$\mathbb{P}(T_1 = t_1 \mid T_2 = t_2, T_3 = t_3)$$

$$= \frac{\binom{t_2}{t_1}\binom{n-t_2}{t_3-t_1}}{\sum_{t_1} \binom{t_2}{t_1}\binom{n-t_2}{t_3-t_1}}$$

where

$$T_1 = N_{11}$$

$$T_2 = N_{11} + N_{12}$$

$$T_3 = N_{11} + N_{21}$$

Turns out

$$\sum_{t_1} \binom{t_2}{t_1}\binom{n-t_2}{t_3-t_1} = \binom{n}{t_3}$$

Argue via hypergeometric.

So

$$P(T_1 = t_1 \mid T_2 = t_2, T_3 = t_3)$$

$$= \frac{\binom{t_2}{t_1}\binom{n-t_2}{t_3-t_1}}{\binom{n}{t_3}}$$

Generalized LR Tests

$H_0: \theta \in \Omega_0 \quad H_1: \theta \in \Omega_1$

$X_1, \dots, X_n \sim \text{iid } f_\theta(x)$

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} f_\theta(x)$$

$$\hat{\theta}_0 = \sup_{\theta \in \Omega_0} \mathcal{L}(\theta)$$

Consider

$$\frac{\sup_{\theta \in \Omega_0} \mathcal{L}(\theta)}{\sup_{\theta \in \Omega_1} \mathcal{L}(\theta)}$$

If $\Omega_0 = \{w_0\}$ $\Omega_1 = \{w_1\}$

we get MP-test.

we get MP-test.

Since we could get some strange irregularities in Ω_1 we normally consider

$$\lambda(\theta) = \frac{\sup_{\theta \in \Omega_0} L(\theta)}{\sup_{\theta \in \Omega_0 \cup \Omega_1} L(\theta)}$$

Reject H_0 if $\lambda(\theta)$ is small.

<u>Ex:</u> $X_1, \dots, X_n \sim \text{iid } N(\mu_1, \sigma^2)$
$Y_1, \dots, Y_m \sim \text{iid } N(\mu_2, \sigma^2)$

$H_0: \mu_1 = \mu_2 \qquad H_1: \mu_1 \neq \mu_2$

We can write the joint likelihood as

$$\prod^{n} 1 \quad -\frac{(x_i-\mu_1)^2}{2\sigma^2}$$

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu_1)}{2\sigma^2}}$$

$$\prod_{i=1}^{m} \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{(y_i-\mu_2)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{m+n} e^{-\frac{1}{2\sigma^2}\left\{\sum^{n} + \sum^{m}\right\}}$$

Under H_0

$$\hat{\mu}_1 = \hat{\mu}_2 = \frac{\sum x_i + \sum y_i}{n+m}$$

$$\hat{\sigma}^2 = \frac{\sum (x_i-\hat{\mu}_1)^2 + \sum (y_i-\hat{\mu}_2)^2}{n+m}$$

Under $\Omega_0 \cup \Omega_1$

$$\hat{\mu}_1 = \frac{1}{n}\sum x_i \quad \hat{\mu}_2 = \frac{1}{m}\sum y_i$$

$$\hat{\sigma}^2 = \frac{\sum (x_i-\hat{\mu}_1)^2 + \sum (y_i-\hat{\mu}_2)^2}{n+m}$$

Finding cutoff is hard to find problistically.

Goal:

1. Finding limiting dist. of λ

2. Find convergenec rate

Suppose $\theta \in \Omega \subseteq \mathbb{R}^k$

$\Omega_0: \theta = g(\varphi), \; g: \mathbb{R}^{k-r} \longmapsto \mathbb{R}^k$

If $\dim(\Omega) = k$

$\dim(\Omega_0) = k-r$ then

under some regularity

conditions

$$-2 \log \lambda_n \xrightarrow{D} \chi^2_r$$

- So GLR tests are asymptotic tests.

Lecture 3/30

Friday, March 30, 2018 11:18 AM

Recall from last time

We stated

$\theta \in \Omega \subseteq \mathbb{R}^k \quad k \geq 1$

$\Omega_0 : \theta = g(\rho)$ where

$g : \mathbb{R}^{k-r} \longmapsto \mathbb{R}^k$

If $\dim(\Omega) = k$

$\dim(\Omega_0) = k - r$ then

$-2 \log \lambda \xrightarrow{D} \chi^2_r$

where

$$\lambda = \frac{\sup_{\theta \in \Omega_0} \mathcal{L}}{\sup_{\theta \in \Omega} \mathcal{L}}$$

Ex: $(\mu_1, \sigma^2), (\mu_2, \sigma^2)$

$H_0: \mu_1 = \mu_2$

$\Theta = (\mu_1, \mu_2, \sigma_2)$

$\varphi = (\mu_2, \sigma^2)$

$g(\varphi) = g(\mu_2, \sigma^2) = (\mu_1, \mu_2, \sigma^2)$

$g: \mathbb{R}^2 \longrightarrow \mathbb{R}^3$

So $k = 3$ $r = \underline{1}$

Ex: $\underline{X} \sim Pois(\lambda_1)$

$\underline{Y} \sim Pois(\lambda_2)$

$\underline{Z} \sim Pois(\lambda_3)$

$H_0: \lambda_1 = \sqrt{\lambda_2 \lambda_3}, \quad \lambda_1 + \lambda_3 = \lambda_2$

r = dimension of data reduction

So here $r = 2$

$\varphi = \lambda_2$ $g(\varphi) = (\lambda_1, \lambda_2, \lambda_3)$

but as we can write
all parameters as
functions of λ_2.

Proof sketch:

$$\log \frac{\mathcal{L}_1}{\mathcal{L}_0} = \log \mathcal{L}_1 - \log \mathcal{L}_0$$

$$\overset{\text{Taylor}}{=} \frac{\sum (z - \chi^2 + \dots)}{\sum z - \chi^2 + \dots)}$$

$\vdots$ $\blacksquare$

Wald Test

$H_0 : R(\theta) = \theta, \; R : \mathbb{R}^k \to \mathbb{R}^r \; r<k$

let $c(\theta) = \dfrac{\partial R(\theta)}{\partial \theta}$

then

$$W_n(\hat{\theta}_n) = n R(\hat{\theta}_n) \left\{ c(\hat{\theta}_n) I_k^{-1}(\hat{\theta}_n) c^T(\hat{\theta}_n) \right\} R^T(\hat{\theta}_n)$$

$$\xrightarrow{D} \chi_r^2$$

$1 \times r$

where $R(\theta) \in \mathbb{R}^{1 \times r}$

where $\hat{\theta}$ is MLE.

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, I^{-1}(\theta))$$

Pf: Under certain regularity condition

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, I^{-1}(\theta))$$

By delta method

$$\sqrt{n}(R(\hat{\theta}_n) - R(\theta)) \xrightarrow{D} N(0, c(\theta) I^{-1}(\theta) c^T)$$

So standardizing

let $V = c(\theta) I^{-1}(\theta) c^T(\theta)$

Then under the assumption that

V p.s.d.

$$V^{-1/2} \sqrt{n}(R(\hat{\theta}_n) - R(\theta)) \xrightarrow{D} N(0, I)$$

So take ℓ_2- norm

$$\| V^{-1/2} \sqrt{n} (R(\tilde{\theta}_n) - R(\theta)) \|_2 \xrightarrow{D} \chi^2_r$$

hence under the null

$R(\theta) = 0$.

Lecture 4/2

Monday, April 2, 2018 11:16 AM

Nonparametric Testing

Suppose that

$X_1, \ldots, X_n \sim \text{iid } F(x)$

and we want to test

$H_0: F(x) = G_\theta(x) \qquad \theta \in \mathbb{R}^p$

$H_1: F(x) \neq G_\theta(x)$

where G_θ is known given θ.

In order to begin testing we need to find an estimator $\hat{F}(x)$ and compare it to the null value.

We'll use the histogram $\tilde{F}(x)$ as our estimator.

as our estimator $\hat{F}(x)$.

<u>Empirical CDF</u>

$$\hat{F}_n(x) = \frac{1}{n} \underbrace{\sum_{i=1}^{n} 1(X_i \le x)}_{\text{sum of iid R.V.}}$$

$$\xrightarrow[a.s.]{P} \mathbb{E}\left[1(X_i \le x)\right] = F(x)$$

by (S/W)LLN.

By the CLT

$$\sqrt{n}\left(\hat{F}_n(x) - F(x)\right) \xrightarrow{D} N\left(0, F(x)(1-F(x))\right)$$

From here we can do pointwise tests but no global test.

Consider

$$\sqrt{n} \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)|$$

$$\mathbb{P}\left(\sup_{x\in\mathbb{R}} |\hat{F}_n(x) - F(x)| \not\to 0\right)$$

$$= \mathbb{P}\left(|\hat{F}_n(x_0) - F(x_0)| > \varepsilon\right)$$

$= 0$ by def. of a.e.

Conv.

<u>Thrm</u>: $\mathbb{P}\left(\sqrt{n} \sup_{x\in\mathbb{R}} |\hat{F}_n(x) - F(x)| \geq u\right)$

$$\longrightarrow 2 \sum_{j=1}^{\infty} (-1)^{j-1} e^{-2j^2u^2}$$

as $n \longmapsto +\infty$

So for asymptotic tests

set

$$\alpha = 2\sum_{j=1}^{\infty}(-1)^{j-1} e^{-2j^2u^2}$$

and solve for u.

Mostly done numerically.

An upper bound is given by

$$2\sum_{j=1}^{\infty}(-1)^{j+1} e^{-2j^2u^2} \leq 2e^{-2u^2}$$

We can also compare different norms of the functional estimates.

Lecture 4/4

Wednesday, April 4, 2018 11:17 AM

Recall from last time

$$\sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)| \xrightarrow{a.s.} 0$$

and

$$\sqrt{n}\left(\hat{F}_n(x) - F(x)\right) \xrightarrow[D]{} N(0, F(1-F))$$

Moreover

$$\sqrt{n} \sup_{x \in \mathbb{R}} |\hat{F}_n - F| \xrightarrow{D} \cancel{N(\quad)}$$

$$\|\hat{F}_n - F\|_2^2 \xrightarrow{D} \cancel{N(\quad)}$$

By Donsker's Thrm.

$$\sqrt{n}(\hat{F}_n - F) \Longrightarrow \mathbb{B}(F(x))$$

Brownian Bridge

$$\sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow{D} N$$

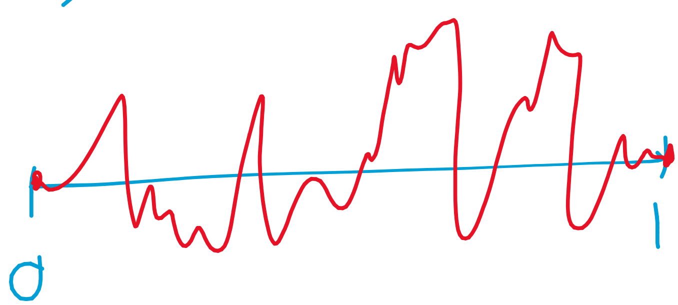
$$\left\{\sqrt{n}\,(\bar{X}_{[nt]} - \mu),\ t \in [0,1]\right\}$$

$$\Longrightarrow W(t)$$

$$B(t) = W(t) - t\,W(1)$$

$$B(0) = W(0) = 0$$

$$B(1) = W(1) - W(1) = 0$$



So we see

$$\sqrt{n} \sup_{x \in \mathbb{R}} |\hat{F}_n - F| \approx \sup_{x \in \mathbb{R}} B(F(x))$$

hypothesis testing via

Do hypothesis testing via Brownian Bridge.

Histogram Density Estimator

$$\hat{f}_n(x) = \frac{\hat{F}_n(x+b) - \hat{F}_n(x-b)}{2b}$$

$$= \frac{\#\{x-b \le X_i \le x+b\}}{2nb}$$

$$= \frac{\sum_{i=1}^{n} \mathbb{1}\{x-b \le X_i \le x+b\}}{2nb}$$

$$\mathbb{E}(\hat{f}_n(x)) = \frac{n}{2nb} \mathbb{P}(x-b \le X_i \le x+b)$$

$$= \frac{1}{2b}(F(x+b) - F(x-b))$$

now as $b \downarrow 0$

now as

$$\mathbb{E}(\hat{f}_n(x)) \longrightarrow F'(x) = f(x)$$

Now if f exists and is differentiable.

$$\mathbb{E}(\hat{f}_n(x)) = \frac{\int_{x-b_n}^{x+b_n} f(u)\,du}{2b_n}$$

$$= \frac{1}{2b_n}\int_{x-b_n}^{x+b_n} \Big\{ f(x) + (u-x) f'(x) +$$

$$+ \frac{(u-x)^2}{2} f''(x) + \dots \Big\} dx$$

$$= f(x) + \frac{1}{4b_n} f''(x) \int_{x-b_n}^{x+b_n} (u-x)^2 du \dots$$

$$= f(x) + O(b_n^2)$$

Lecture 4/6

Friday, April 6, 2018 11:16 AM

Suppose that

$X_1, \dots, X_n \sim \text{iid } f(x)$

$$\hat{f}_n(x) = \frac{\hat{F}_n(x+b_n) - \hat{F}_n(x-b_n)}{2b_n}$$

$$= \frac{1}{2nb_n} \sum_{i=1}^{n} \mathbb{1}\{x-b_n \le X_i \le x+b_n\}$$

$$\mathbb{E}(\hat{f}_n(x)) = \frac{1}{2b_n} \int_{x-b_n}^{x+b_n} f(u)\,du$$

$$= \frac{1}{2b_n} \int_{x-b_n}^{x+b_n} \{f(x) + (u-x)f'(x) + \dots\}\,du$$

$$= f(x) + \frac{1}{2b_n} \int_{x-b_n}^{x+b_n} \frac{(u-x)^2}{2}\,dx + \dots$$

$$= f(x) + \frac{1}{2b_n} f''(x) \int_{x-b_n}^{x+b_n} \frac{(u-x)^2}{2}\,du$$

$$+ o(b_n^2)$$

$$= f(x) + \frac{f''(x) b_n^2}{6} + o(b_n^2)$$

why one side is worse...

$$\mathbb{E}(\hat{f}_n) = \mathbb{E}\left(\frac{\hat{F}_n(x+b_n) - \hat{F}_n(x)}{b_n}\right)$$

$$= \frac{1}{b_n} \mathbb{E}(\hat{F}_n(x+b_n)) - \frac{1}{b_n} \mathbb{E}(\hat{F}_n(x))$$

$$= \frac{1}{b_n} \mathbb{E}\left(1(x \le X_i \le x+b_n)\right)$$

$$= \frac{1}{b_n} \mathbb{P}(x \le X_i \le x+b_n)$$

$$= \frac{1}{b_n} \int_x^{x+b_n} f_X(u)\, du$$

$$= \frac{1}{b_n} \int_x^{x+b_n} f_X(x) + (u-x) f'(x) + o(b_n) \, du$$

$$= f_X(x) + \frac{f'(x)}{b_n} \int_x^{x+b_n} (u-x) + o(b_n)$$

$$z = u - x \qquad dz = du$$

$$= f(x) + \frac{f'(x)}{b_n} \int_0^{b_n} z\, dz + o(b_n^2)$$

$$= f(x) + \frac{f'(x)}{2} b_n + o(b_n)$$

<u>Variance</u>

Bernoulli

$$\mathrm{Var}(\hat{f}_n) = \frac{1}{4n^2 b_n^2} \sum_{i=1}^{n} \mathrm{Var}(1_{\{<<\}})$$

$$= \frac{1}{4n b_n^2} \{F(x+b_n) - F(x-b_n)\}$$

$$[1 - F(x+b_n) + F(x-b_n)\}$$

Taylor

$$= \frac{1}{\cdot} [2\{f(x) + o(1)\} b_n]$$

$$= \frac{1}{4nb_n^2}\left[2\{f(x)+o(1)\}b_n\right] \quad |$$

$$\left[1-\underbrace{2\{f(x)+o(1)\}b_n}_{o(1)}\right]$$

$$= \frac{1}{2nb_n} f(x)\{1+o(1)\}$$

So $Var \to 0$ if

$nb_n \to +\infty$.

So variance converges to zero at a rate of

$$\boxed{\frac{1}{nb_n}}$$

Lecture 4/9

Monday, April 9, 2018 11:19 AM

Recall from last time

$$\hat{f}_n(x) = \frac{1}{2b_n} \sum_{i=1}^{n} \mathbb{1}_{(x \pm b_n)}(X_i)$$

$$\mathbb{E}(\hat{f}_n(x)) = f(x) + \frac{b_n^2}{6} f''(x) + o(b_n^2)$$

$$\longrightarrow f(x) \text{ as } \begin{array}{l} n \to \infty \\ b_n \to 0 \end{array}$$

$$Var(\hat{f}_n) = \frac{1}{2nb_n} f(x) \{1 + o(1)\}$$

$$\longrightarrow 0 \text{ as } nb_n \to \infty$$

So we need the bin size to shrink st.

$$\boxed{\begin{array}{l} nb_n \longmapsto \infty \\ b_n \longmapsto 0 \end{array}}$$

$$MSE(\hat{f}_n) = (Bias)^2 + Var$$

$$= \left\{ \frac{1}{36} f''(x)^2 b_n^4 + \frac{f(x)}{2nb_n} \right] \left[1 + o(1) \right]$$

$$= \left\{ \frac{1}{36} f''(x)^2 b_n^4 + \frac{f(x)}{2nb_n} \right\} \left[1 + o(1) \right]$$

- Good example of bias - variance tradeoff.

- We will select b_n by minimizing the asymptotic M.S.E.

$$g(u) = au^4 + \frac{b}{u}$$

$$\frac{\partial g(u)}{\partial u} = 4au^3 - \frac{b}{u^2} = 0$$

$$\Rightarrow 4au^3 = \frac{b}{u^2}$$

$$\Rightarrow u^5 = \frac{b}{4a}$$

$$\Rightarrow u = \left(\frac{b}{4a} \right)^{1/5}$$

$$\Rightarrow u = \left(\frac{f(x)/2n}{4f''(x)^2/36} \right)^{1/5}$$

$$= \left(\frac{36}{8n} \frac{f(x)}{f''(x)} \right)^{1/5}$$

$$= \left(\frac{9f(x)}{2nf''(x)} \right)^{1/5}$$

$$= \left(\frac{9}{2}\right)^{1/5} \left(\frac{f(x)}{f''(x)} \right)^{1/5} n^{-1/5}$$

Bandwidth still a function of x so we will instead use integrated MSE (MISE)

Kernel Density Estimates

$$\hat{f}_n(x) = \frac{1}{nb_n} \sum_{i=1}^{n} K\left(\frac{X_i - x}{b_n}\right)$$

where $K(\cdot)$ is a kernel function and b_n is a bandwidth satisfying

$$\boxed{b_n \to 0 \qquad nb_n \to +\infty}$$

If $K(\cdot)$ is continuous then $\hat{f}_n$ is also cont.

If

$$K(v) = \frac{1\{-1 \le v \le 1\}}{2}$$

then $\hat{f}_n$ is histogram.

Now consider

$$\lim_{x \to x_0} \hat{f}_n(x) = \lim_{x \to x_0} \frac{1}{nb_n} \sum_{i=1}^{n} K\left(\frac{X_i - x}{b_n}\right)$$

$$= \frac{1}{nb_n} \sum_{i=1}^{n} K\left(\frac{X_i - x_0}{b_n}\right)$$

$$= \frac{1}{n b_n} \sum_{i=1} K\left(\frac{\quad}{b_n} \right)$$

$= \hat{f}_n(x_0)$ So $K(\cdot)$ cont.

is sufficient for $\hat{f}_n$ to be cont.

Now for the bias calculatn.

$$\mathbb{E}(\hat{f}_n) - f_n = \frac{1}{b_n} \int K\left(\frac{u-x}{b_n}\right) f(u)\,du - f(x)$$

$$= \int K(v) f(x + v b_n)\,dv - f(x)$$

$$= \int K(v) \{ f(x+b_n) - f(x) \}\,dv$$

b/c $\int K(v)\,dw = 1$

$$= \int K(v) \left\{ f'(x)\, b_n v + b_n^2 \frac{f''(x)}{2} v^2 + o(b_n^2) \right\} dv$$

Now if K is symmetric about 0

$$= f'(x) \int K(v)\, v b_n\, dv \longrightarrow 0$$

$$+ \frac{f''(x)}{2} b_n^2 \int K(v) v^2 \, dv$$

$$+ o(b_n^2)$$

$$= \frac{f''(x)}{2} \|K\|_2^2 b_n^2 + o(b_n^2)$$

Lecture 4/13

Friday, April 13, 2018 11:13 AM

Recall from last time

$\hat{b}_n = cn^{-1/5}$

Today's focus

$$\left[\hat{f}_n(x) - \mathbb{E}(\hat{f}_n(x))\right]$$

well

$$\hat{f}_n = \frac{1}{n}\sum_{i=1}^{n} \underbrace{\frac{1}{b_n} K\left(\frac{X_i - x}{b_n}\right)}_{Y_i}$$

So by the iid ness of

the Y_i

$$\frac{\sum_{i=1}^{n} Y_{in} - \mathbb{E}(Y_{i,n})}{\sqrt{\mathrm{Var}(\sum Y_{in})}} \xrightarrow{D} N(0,1)$$

$\hat{f}_n(x) - \mathbb{E}(\hat{f}_n)$ $-$

$$\frac{\hat{f}_n(x) - \mathbb{E}(\hat{f}_n)}{\sqrt{\frac{1}{nb_n} f(x) \int K(u)^2 du \, \{1 + o(1)\}}} \to Z$$

$$\frac{\hat{f}_n - \mathbb{E}(\hat{f}_n)}{\sqrt{\frac{1}{nb_n} f(x) \int K(u)^2 du}} \to Z$$

$$\sqrt{nb_n} \left\{ \hat{f}_n(x) - \mathbb{E}(\hat{f}_n) \right\}$$

$$\xrightarrow{D} N\left(0, f(x) \int K(u)^2 du\right)$$

Now if bias $\to 0$

$$\sqrt{nb_n} \frac{b_n^2 f''(x)}{2} \to 0$$

we can write

$$\sqrt{nb_n} (\hat{f}_n - f(x)) \to N\left(0, f(x) \int K(u)^2\right)$$

but note that for

but note that for

$$\hat{b}_n = cn^{-1/5}$$

$$\sqrt{n\hat{b}_n}\,\hat{b}_n^2 \not\to 0$$

So we don't get a C.L.T. if we use bandwidth of the optimal MSE.

People use <u>pointwise</u> C.I. based on the CLT variance.

Issues

- Developed pointwise not the entire curve
- optimal bandwidth

- Optimal bandwidth doesn't even give ns this variance

If

$$\sqrt{nb_n}\left(g_{\hat{\theta}_n}(x) - f(x)\right) \xrightarrow{P} 0$$

then

$$\sqrt{nb_n}\left(\hat{f}_n(x) - g_{\hat{\theta}_n}(x)\right) \xrightarrow{D} N\left(0, f(x)\int K(u)^2 du\right)$$

$$\sqrt{n}\left(b_n^{5/2}\right) \to 0.$$

So for $b_n = cn^{-k}$

$$n^{1/2}\left(c\, n^{-5k/2}\right)$$

$$= c\, n^{-5k+1/2} < 0 \quad \boxed{k > 1/5}$$

$$\sqrt{n}\left(\sqrt{b_n}\right) b_n^2$$

$$= \sqrt{n}\left(n^{-2/5}\, n^{-2/5}\right)$$

$$= n^{1/2}\left(n^{-4/5}\right)$$

$$= n^{1/5} \longmapsto \infty.$$

Suppose...

H_0: $f(x_0) = g_\theta(x_0)$

$f(x_1) = g_\theta(x_1)$

Lecture 4/18

Wednesday, April 18, 2018 11:15 AM

$$\hat{f}_n(x) = \frac{1}{nb_n} \sum_{i=1}^{n} K\left(\frac{x_i - x}{b_n}\right)$$

$$\sqrt{nb_n}\left(\hat{f}_n(x) - \underbrace{\mathbb{E}(\hat{f}_n(x))}\right) \to N\left(0, f(x)\int K^2\right)$$

Can be replaced by $f(x)$ if $nb_n^5 \to 0$

We want to Consider tests of the following type.

$x_0 \in \mathbb{R}$

$H_0: f(x_0) = a$ vs. $H_1: f(x_0) = a$

$H_0: f(x_0) = g_\theta(x_0)$ $H_1: f(x_0) = g_\theta(x_0)$

where g_θ has a known form expeet the parameter vector θ

<u>Ex:</u> $X_1, \dots, X_n \sim$ iid $f(x)$ with mean μ and variance σ^2. Let $g_{(\mu,\sigma^2)}(x)$ be the normal density. Test

$$H_0: f(x_0) = g_{(\mu,\sigma^2)}(x_0)$$

$$H_1: f(x_0) \neq g_{(\mu,\sigma^2)}(x_0)$$

Well

$$\sqrt{n b_n}\left(\hat{f}_n - f\right) \xrightarrow{D} N\left(0, f(x)\int K^2\right)$$

under regularity conditions.

Under regularity conditions.

Or

$$\frac{\sqrt{nb_n}(\hat{f}_n - f)}{\sqrt{f \int K^2}} \xrightarrow{D} Z$$

For the variance we can use the plugin estimator

$$\frac{\sqrt{nb_n}(\hat{f}_n - f)}{(\hat{f} \int K^2)^{1/2}} \xrightarrow{D} Z$$

Let $\hat{\mu}_n = \bar{X}$ and $\hat{\sigma}^2 = S_n^2$ then under H_0

$$\{q_{(\hat{\mu}, \hat{\sigma}^2)}(x_0) - f(x_0)\} = O_p\left(\frac{1}{\sqrt{n}}\right)$$

where O_p means

$\forall \varepsilon > 0 \ \exists M > 0$ s. t. $\forall N$

$$\mathbb{P}\left(\sqrt{n}\left|g_{(\hat{\mu}_n, \hat{\sigma}^2)}(x_0) - f(x_0)\right| > M\right) < \varepsilon$$

Under the null

$$\frac{\sqrt{nb_n}\left(\hat{f}_n(x_0) - g_{(\hat{\mu}, \hat{\sigma}_n^2)}(x_0)\right)}{\sqrt{\tilde{f}_n(x_0)\int K(v)^2 dv}} \xrightarrow{D} Z$$

Define

$$Z_n := \frac{\sqrt{nb_n}\left(\hat{f}_n(x_0) - g_{(\hat{\mu}, \hat{\sigma}_n^2)}(x_0)\right)}{\sqrt{\hat{f}_n(x_0)\int K(v)^2 dv}}$$

then the asymptotic test can be completed as

$$|Z_n| > z_{1-\alpha/2}$$

Multipoint Testing

Lecture 4/20

Friday, April 20, 2018 11:17 AM

Suppose we have

$$\hat{f}_n(x) = \frac{1}{nb_n} \sum_{i=1}^{n} K\left(\frac{X_i - x}{b_n}\right)$$

$$\sqrt{nb_n}\left\{\hat{f}_n(x) - E\left(\hat{f}_n(x)\right)\right\} \xrightarrow{D} N\left(0, f(x)\int K^2(u)\,du\right)$$

Consider $x_1 \neq x_2$

and we want to analyze the behavior of

$$\sqrt{nb_n}\begin{Bmatrix} \hat{f}_n(x_1) \\ \hat{f}_n(x_2) \end{Bmatrix} = \frac{\sqrt{nb_n}}{n} \sum_{i=1}^{n} \begin{Bmatrix} \frac{1}{b_n} K\left(\frac{X_i - x_1}{b_n}\right) \\ \frac{1}{b_n} K\left(\frac{X_i - x_2}{b_n}\right) \end{Bmatrix}$$

So by the multivariate CLT we have this converges to a joint normal

Converges to a joint normal

$$\xrightarrow{D} N\left\{ \begin{pmatrix} \mathbb{E}(\hat{f}_n(x_1)) \\ \mathbb{E}(\hat{f}_n(x_2)) \end{pmatrix}, \Sigma \right\}$$

where

$$\Sigma = \begin{pmatrix} f(x_1)\int K^2(u)\,du & ? \\ ? & f(x_2)\int K^2(u)\,du \end{pmatrix}$$

The covariance is given by the limit of

$$n b_n \operatorname{Cov}\left(\hat{f}_n(x_1), \hat{f}_n(x_2)\right)$$

$$= \frac{1}{n b_n} \operatorname{Cov}\left\{ \sum K\left(\frac{X_i - x_1}{b_n}\right), \sum K\left(\frac{X_i - x_2}{b_n}\right) \right\}$$

$$\overset{\text{ind}}{=} \frac{1}{n b_n} \sum_{i=1}^{n} \operatorname{Cov}\left(K\left(\frac{X_i - x_1}{b_n}\right), K\left(\frac{X_i - x_2}{b_n}\right) \right)$$

$$= \frac{1}{nb_n} \sum_{i=1} \mathrm{Cov}\left(K\left(\frac{X_i - x_1}{b_n}\right), K\left(\frac{X_i - x_2}{b_n}\right)\right)$$

id

$$= \frac{1}{b_n} \mathrm{Cov}\left(K\left(\frac{X_n - x_1}{b_n}\right), K\left(\frac{X_n - x_2}{b_n}\right)\right)$$

focusing on the covariance term only

$$\mathbb{E}\left(K\left(\frac{X_n - x_1}{b_n}\right) K\left(\frac{X_n - x_2}{b_n}\right)\right)$$

$$- \mathbb{E}\left[K\left(\frac{X_n - x_1}{b_n}\right)\right] \mathbb{E}\left[K\left(\frac{X_n - x_2}{b_n}\right)\right]$$

$$= \int K\left(\frac{u - x_1}{b_n}\right) K\left(\frac{u - x_2}{b_n}\right) f(u)\,du$$

$$- \int K\left(\frac{u - x_1}{b_n}\right) f(u)\,du \int K\left(\frac{u - x_2}{b_n}\right) f(u)\,du$$

We know by bias calculations

$$\mathbb{E}\left(K\left(\frac{X_n - x_1}{b_n}\right)\right) = b_n f(x_1)\{1 + o(1)\}$$

So

$$\frac{1}{b_n}\mathbb{E}(*)\,\mathbb{E}(*) = b_n f(x_1) f(x_2) \longrightarrow 0$$

So we need to only focus on the first term.

$$\frac{1}{b_n}\int K\left(\frac{u - x_1}{b_n}\right) K\left(\frac{u - x_2}{b_n}\right) f(u)\,du$$

$$= \frac{1}{b_n}\int K(v) K\left(v + \frac{x_1 - x_2}{b_n}\right) f(x_1 + b_n)\,du$$

letting $n \to \infty$

we see by L.D.C.T.

$$\lim_{n\to\infty} K\left(v + \frac{x_1 - x_2}{b_n}\right) = 0$$

So $\int K(v) K\left(v + \frac{x_1 - x_2}{b_n}\right) f(x_1 + b_n$

$\longrightarrow 0$

So the MVCLT

$$\begin{Bmatrix} \hat{f}_n(x_1) \\ \hat{f}_n(x_2) \end{Bmatrix} \xrightarrow{D} N\left(\begin{matrix} \mathbb{E}(\hat{f}_n(x_1)) \\ \mathbb{E}(\hat{f}_n(x_2)) \end{matrix}, \Sigma \right)$$

$$\Sigma = \begin{pmatrix} f(x_1)\int K^2(v)\,dv & 0 \\ 0 & f(x_2)\int K^2(v)\,dv \end{pmatrix}$$

So $X_i = \hat{f}_n(x_i) - \mathbb{E}(\hat{f}_n(x_i))$

$$\left(\frac{X_1}{\sigma^2_{x_1}}\right)^2 + \left(\frac{X_2}{\sigma^2_{x_2}}\right)^2 \xrightarrow{D} \chi^2_2$$

and in general

$$\sum_{i=1}^{n}\left(\frac{X_i}{\sigma^2_{X_i}}\right)^2 \xrightarrow{D} \chi^2_n$$

We can use this result for testing hypothesis

$$H_0: f(x_1) = g_0(x_1)$$

$$f(x_2) = g_0(x_2)$$

Infinite sample testing

Want to analyze

$$\sup_{x \in S}\left|\frac{\hat{f}_n(x) - f(x)}{\sqrt{f(x)}}\right|$$

and

$$\int_{x \in S} \frac{\left(\hat{f}_n(x) - f(x)\right)^2}{f(x)} \, dx$$

where S is compact.

Lecture 4/23

Monday, April 23, 2018 11:16 AM

Consider the nonparametric regression model

$$Y_i = g(x_i) + e_i \quad i = 1, \dots, n$$

If we assume g is smooth and use OLS optimization criterion

$x \in \mathbb{R}$ estimate $g(x)$ by

$$\sum_{i=1}^{n} (Y_i - g(x))^2$$

gives the mean

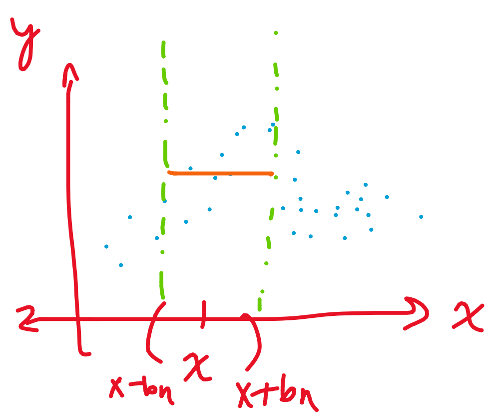
Instead weight the x_i by their distance to x. and by their distance to x. and

minimize

$$\sum_{i=1}^{n} (Y_i - g(x)) K\left(\frac{X_i - x}{b_n}\right)$$

Ex: rectangular

$$K\left(\frac{X_i - x}{b_n}\right) = \frac{1}{2} \mathbb{1}\{|X_i - x| \le b_n\}$$



We can view kernel regression

as a weighted least squares.

$$\sum (Y_i - \hat{y}_x) w_i = 0$$

$$\hat{y}_x = \frac{\sum Y_i w_i}{\sum w_i}$$

So

$$\boxed{\hat{g}(x) = \frac{\sum_{i=1}^{n} Y_i K\left(\frac{x_i - x}{b_n}\right)}{\sum_{i=1}^{n} K\left(\frac{x_i - x}{b_n}\right)}}$$

Does $\hat{g}(x) \to g(x)$

$$\hat{g}(x) = \frac{\frac{1}{nb_n} \sum Y_i K\left(\frac{x_i - x}{b_n}\right)}{\frac{1}{nb_n} \sum K\left(\frac{x_i - x}{b_n}\right)}$$

We know the bottom

$\longrightarrow f(x) \quad O_p(n^{-1/5})$

For the top factor as follows

$$\underbrace{\frac{1}{nb_n} \sum g(X_i)\, K(\#)}_{} + \frac{1}{nb_n} \sum e_i K(\#)$$

$$= \frac{1}{nb_n} \sum \left\{ g(x) + (X_i - x) g'(x) + \dots \right\} K(\#)$$

$$\simeq g(x) \left\{ f(x) + o_p(1) \right\}$$

$$\longrightarrow g(x) f(x)$$

So plugging in

$$\longrightarrow \frac{g(x) f(x) + 0}{f(x)} = g(x)$$

because the sum of iid R.V.

$$\frac{1}{n} \sum e_i \frac{1}{b_n} K(\#)$$

by SLLN

$$\to \mathbb{E}\left(e_i \frac{1}{b_n} K(\#)\right)$$

$$= \mathbb{E}(e_i)\, \mathbb{E}\left(\frac{1}{b_n} K(\#)\right)$$

$$= 0$$

Let

$$U_n(x) = \frac{1}{n b_n} \sum Y_i K(\#)$$

$$V_n(x) = \frac{1}{n b_n} \sum K(\#)$$

Now let's consider

$$U_n(x) - \mathbb{E}(U_n(x))$$

$$= \frac{1}{nb_n} \sum \left[g(X_i) K - \mathbb{E}\left(g(X_i) K \right) \right]$$

$$+ \frac{1}{nb_n} \sum e_i K$$

The first term suggests a C.L.T. Finding the rate...

$$\mathrm{Var}\left(\frac{g(x_i) K(\#)}{b_n} \right)$$

$$= \mathbb{E}\left\{ \frac{g^2(x_i) K^2(\#)}{b_n^2} \right\} - \left[\mathbb{E}\left\{ \frac{g(x_i) K}{b_n} \right\} \right]^2$$

$$= \frac{1}{bh^2} \int g^2(u) K_u^2 f(u) du$$

$$- \left(\frac{1}{b_n} \int g(u) K_u f(u) du \right)^2$$

Lecture 4/25

Wednesday, April 25, 2018 11:15 AM

Let

$$\hat{g}_i(x_i\cdot) = \frac{\frac{1}{nb_n}\sum Y_i K(\#)}{\frac{1}{nb_n}\sum K(\#)}$$

$$= \frac{T_n(x)}{\hat{f}_n(x)}$$

We want to find a CLT type result.

$$\hat{g}_n(x) - g(x) = \frac{T_n(x) - g(x)\hat{f}_n(x)}{\hat{f}_n(x)}$$

Since

$$T_n(x) - g(x)\hat{f}_n(x)$$

$$= \frac{1}{nb_n}\sum_{j=1}^{n}(Y_i - g(x))K(\#)$$

By the CLT

$$U_n(x) = T_n(x) - g(x)\hat{f}_n(x)$$

$$\frac{U_n(x) - \mathbb{E}(U_n(x))}{\sqrt{Var(U_n(x))}} \xrightarrow{D} Z$$

Then we have

$$\mathbb{E}(U_n(x)) = \frac{1}{b_n}\mathbb{E}\Big(\big(Y_i - g(X_i)\big)K(\#)\Big)$$

$$= \frac{1}{b_n}\mathbb{E}\Big(\big(g(X_i) - g(x)\big)K(\#)\Big)$$

$$= \frac{1}{b_n}\int\{g(u) - g(x)\}K\left(\frac{u-x}{b_n}\right)f(u)du$$

$$= \int\underbrace{\{g(x+b_n v) - g(x)\}}_{\longrightarrow 0 \text{ by LDCT}}K(v)f(x+b_n v)\,dv$$

We also can about

rates here. Taylor expanding the limiting terms.

$$= \int \left\{ g'(x) b_n v + \frac{g''(x)}{2} b_n^2 v^2 \right\}$$

$$\left\{ f(x) + f'(x) b_n v + \ldots \right\} K(v)\, dv$$

$$\simeq \int \left\{ g'(x) f'(x) + \frac{g''(x)}{2} f(x) \right\} b_n^2 v^2$$

$$K(v)\, dv$$

$$\mathbb{E}(u_n(x)) = O(b_n^2)$$

$$Var(u_n(x)) = \frac{1}{n b_n^2} \Big(Var\Big(g(x_i) - g(x) \Big)$$

$$K\left(\frac{x_i - x}{b_n} \right) \Big) + Var(e_i K(x)) \Big)$$

$$+ O$$

$$= \frac{1}{n b_n^2} \left\{ \int (g(u) - g(x))^2 K(t)^2 f \, du \right.$$

$$- \left[\int (g(u) - g(x))^2 K(t) f \, du \right]^2$$

$$\left. + \mathbb{E}(e_i^2)\, \mathbb{E}(K^2(t)) \right\}$$

$$= \frac{1}{n b_n^{\cancel{2}}} \left\{ \cancel{b_n} \int (g(x + b_n v) - g(x))^2 K(v) \right.$$

$$f(x + b_n v)\, dv - b_n^2 \left[\int \{ g(x + b_n v) K(v) \right.$$

$$\left. \left. f(x + b_n v)\, dv \right]^2 + \sigma^2 \cancel{b_n} \int K(v)^2 f(x + b_n v)\, dv \right\}$$

$$= \frac{1}{n b_n} \sigma^2 f(x) \| K(v) \|_2^2 + O(b_n^{??})$$

So

$$\frac{U_n(x) - \mathbb{E}(U_n(x))}{\sqrt{\mathrm{Var}(U_n(x))}} =$$

$$\sqrt{\operatorname{var}(U_n(x))}$$

$$\frac{\sqrt{nb_n}\left(U_n(x) - b_n^2(x)\left(g'(x)f'(x)\right) + \frac{g''(x)}{2}f(x)\|K\|_2^2\right)}{\sqrt{\sigma^2 f(x)\|K\|_2^2}} \xrightarrow{D} Z$$

If $nb_n^5 \to 0$ then

$$\sqrt{nb_n}\, U_n(x) \xrightarrow{D} N\left(0, \sigma^2 \|K\|_2^2 f(x)\right)$$

So by Slutsky

$$\frac{\sqrt{nb_n}\, U_n(x)}{\hat{f}_n(x)} \xrightarrow{D} N\left(0, \sigma^2 \|K\|_2^2\right)$$

And hence

$$\sqrt{nb_n}\left(\hat{g}_n(x) - g(x)\right) \xrightarrow{D} N\left(0, \sigma^2 \|K\|_2^2\right)$$

Lecture 4/27

Friday, April 27, 2018 11:18 AM

False Discovery Rate

Suppose we have null hypothesis $\{H_i\}_i^m$ each with p-values $p_1, \dots, p_m$ and we consider

	Nonsig.	Sig	Tot
True Null	U	V	m_0
False	T	S	m_1
	$m-r$	r	m

Bonferroni Correction:

Significant if $p < \frac{\alpha}{m}$

$$\mathbb{P}(\text{at least one false discovery})$$

$$= \mathbb{P}(\text{at least one Type I})$$

$$= \mathbb{P}\left(\bigcup_{i=1}^{m} \text{Type I on } i\right)$$

$$\leq \sum_{i=1}^{m} \mathbb{P}(\text{Type I on } i)$$

$$= \sum_{i=1}^{m} \frac{\alpha}{m} = \alpha$$

FDR: (Benjamini & Hochberg)

1. order p-values into

$$P_{(1)} \leq P_{(2)} \leq \cdots \leq P_{(m)}$$

2. Let

$$k = \max\left\{ i : P_{(i)} \leq \frac{i}{m} \alpha \right\}$$

then we reject

$H_{(1)}, \dots, H_{(k)}$

- Less strict the Bonferroni
- More strict than constant

α

Lecture 4/30

Monday, April 30, 2018 11:15 AM

Recall the B-H method

$k = \max\{ i : P_{(i)} \leq i/m \, \tau \}$

Reject $H_{(1)}, \ldots, H_{(k)}$

Back to B.F.

$\mathbb{P}(\text{at least one false disc.})$

$= \mathbb{P}(\bigcup \text{false discovery})$

$= 1 - \mathbb{P}(\bigcap \text{no false discovery})$

now assuming independence of the tests

of the tests

$$= 1 - \prod_{i=1}^{m} \mathbb{P}(\text{no false on } \#i)$$

$$= 1 - (1-\underbrace{\beta}_{\text{marginal type 1.}})^m = \alpha$$

Solving for β

$$\beta = 1 - (1-\alpha)^{1/m}$$

$$\approx \frac{\alpha}{m} \text{ by Taylor.}$$

for $m=2$ $\alpha = 0.05$

$$1 - (1 - 0.05)^{1/2} = 0.0253$$

$$1-(1-.05)^{1/2}=0.0253$$

Def:

FDP (false discovery prop.)

$$FDP=\frac{V}{\max\{R,1\}}$$

V ← # false discovery

R ← # discovery

FDR (false discovery rate)

$$\boxed{FDR=\mathbb{E}(FDP)}$$

Thrm: Assume that

(i) Indep. test stat.

(ii) Cont. dist of T.S.

then B-H controlls

FDR at α

Pf: (only consider $m = m_0$)

$$FDP = \begin{cases} 1 & V > 0 \\ 0 & \text{o.w.} \end{cases}$$

$$\mathbb{E}(FDP) = \mathbb{P}\Big(V > 0\Big)$$

$$= \mathbb{P}\Big(P_{(i)} \le \frac{i}{m}\alpha \text{ for some } 1 \le i \le m\Big)$$

$= \underset{\cdots}{\text{induction}} = \alpha$

induction

$m=1$: $\mathbb{P}\big(P_{(1)} \le \frac{1}{m}\alpha\big)$

$$= \mathbb{P}(P_1 \le \alpha) = \alpha$$

Aside: $p = \mathbb{P}(X > T)$

R.V. with $X \overset{D}{=} T$

test stat

$$p = 1 - F_T(X) = 1 - F_T(T)$$

$$\mathbb{P}(p \le a) = \mathbb{P}(1 - F_T(T) \le a)$$

$$= \mathbb{P}(F_T(T) \ge 1 - a)$$

$$= \mathbb{P}(T \ge q_{1-a}) = a$$

Therefore $p \sim \mathrm{Unif}(0,1)$

Assume true for m

then

$$\mathbb{P}\left(P_{(i)} \le \frac{i}{m+1}\alpha \quad \text{for some } 1 \le i \le m+1 \right)$$

$$= \mathbb{P}\left(P_{(m+1)} \le \alpha,\ P_{(i)} \le \frac{i}{m+1}\alpha \right)$$

$$+ \mathbb{P}\left(P_{(m+1)} > \alpha,\ P_{(i)} \le \frac{i}{m+1}\alpha \right)$$

$$= \mathbb{P}\left(P_{(m+1)} \le \alpha \right) + \mathbb{P}\left(P_{(m+1)} > \alpha, \# \right)$$

$$= \alpha^{m+1} + \mathbb{E}\left(\mathbb{1}\left(P_{(m+1)} > \alpha\right) \mathbb{1}\left(P_{(i)} \le \frac{\alpha}{m+1}\right) \right)$$

$$= \alpha^{m+1} + \mathbb{E}\left\{ \mathbb{E}\left[\# \mid P_{(m+1)} \right] \right\}$$

$$= \alpha^{m+1} + \mathbb{E}\left[\mathbb{1}(P_{(m+1)} > \alpha)\ \alpha \right]$$

$$= \alpha^{m+1} + \alpha\, \mathbb{E}\left\{ \mathbb{1}\left(\frac{P_{(i)}}{P_{(m+1)}} \le \frac{i}{(m+1)} \frac{\alpha}{P_{(m+1)}} \right) \right\}$$

Claim: $\dfrac{P_{(i)}}{P_{(m+1)}} \,\Big|\, P_{(m+1)} \sim \text{Unif}(0, P_{(m+1)})$

then we have

$$= \alpha^{m+1} + \mathbb{E}\left\{ \mathbb{1}(P_{(m+1)} > \alpha)\ \mathbb{E}\left(\# \right) \Big| P_{(m+1)} \right\}$$

$$= \alpha^{m+1} + \mathbb{E}\left\{ \mathbb{1}(P_{(m+1)} > \alpha)\ \frac{m}{m+1} \frac{\alpha}{P_{(m+1)}} \right\}$$

$$= \alpha^{m+1} + \int_{\alpha}^{1} \frac{m}{m+1} \frac{\alpha}{x} (m+1) x^{m} \, dx$$

$$= \alpha^{m+1} + m\alpha \int_{\alpha}^{1} x^{m-1}\, dx$$

$$= \alpha^{m+1} + m\alpha \left[\frac{x^m}{m} \right]_{\alpha}^{1}$$

$$= \alpha^{m+1} + \alpha - \alpha^{m+1} = \alpha$$

Lecture 5/2

Wednesday, May 2, 2018 11:16 AM

Let $X_1, \dots, X_n \sim \text{Unif}(0, \theta)$

$$f_{X_{(1)} \cdots X_{(n-1)} \mid X_{(n)}}(u_1, \dots, u_{n-1} \mid u_n)$$

$$= \frac{f_{X_{(1)} \cdots X_{(n)}}(u_1, \dots, u_n)}{f_{X_{(n)}}(u_n)}$$

$$\mathbb{P}\left((X_{(1)}, \dots, X_{(n)}) \in A \right)$$

$$= \mathbb{P}\left(\bigcup_{p \in \text{Perm}} (X_{p_1}, \dots, X_{p_n}) \in A \right)$$

$$= \sum_{p} \mathbb{P}\left((X_{p_1}, \dots, X_{p_n}) \in A \right)$$

$$= \sum_{P} \int_{A} f(x_1, \dots, x_n)$$

$$= n! \int_{A} f(x_1, \dots, x_n)$$

So we have

$$f_{X_{(1)} \dots X_{(n)}} = n! \, \theta^{-n} \mathbb{1}(u_i < u_{i+1})$$

and back to the joint...

$$= \frac{n! \, \theta^{-n} \mathbb{1}(u_i \le u_{i+1})}{n \, u_n^{n-1} \, \theta^{-n} \mathbb{1}(0 < u_n < \theta)}$$

$$= \frac{(n-1)!}{u_n^{\,n-1}} \mathbb{1}(0 < u_1 < \dots < u_n)$$

So this cond. dist

is the order statics

of uniform on

$(0, u_n)$