Hypothesis Testing

· Def: A hypothesis is a statement about a population parameter.

Two complementary hypothesis in one H-T problem are called null hypo: H_0 and alternative hypothese H_A. (or H_1)

e.g. $A \subseteq \Theta_0$. $H_0 = \theta \in A$. $H_1 = \theta \in A^c$.

(1) Method of Finding Tests:

· The procedure of hypothesis testing is:

After observing the samples taken, make a decision on accepting H_0 or H_1. The subset of "H_0 is rejected" of sample space is rejection region R.

Remark: If $H_0 = \theta = 1.5$, $H_1 = \theta > 0.5$. We will not simply calculate the estimate of θ, i.e. $\hat{\theta}$. Compare $\hat{\theta}$ with θ. Since if $\hat{\theta} = 0.500001$, it's vague to say $\hat{\theta}$ falls into H_0 or H_1.

→ That's also why we put $\theta = \theta_0$ on H_0. A little deviation can be tolerated

Moreover, H_0 and H_1 are asymmetric. We're pretending to protect H_0.

① Likelihood Ratio Test:

Def: Likelihood Ratio Test Statistic for $H_0: \theta \in \Theta_0$ vs. $H_1: \theta \in \Theta_0^c$ is

$$\lambda(\vec{x}) = \frac{\sup_{\Theta_0} L(\theta|\vec{x})}{\sup_{\Theta} L(\theta|\vec{x})}$$

Likelihood Ratio Test is any test has rejection region of form $\{\vec{x} \mid \lambda(\vec{x}) \le c\}$. $0 \le c \le 1$

A computational simplification:

Thm. $T(\vec{X})$ is sufficient statistic for θ. $T(\vec{X}) \sim g(t|\theta)$

$$\lambda^*(t) = \frac{\sup_{\Theta_0} g(t|\theta)}{\sup_{\Theta} g(t|\theta)}.$$ Then $\lambda(\vec{x}) = \lambda^*(T(\vec{x}))$, $\forall \vec{x} \in \mathcal{X}$.

pf: By Factorization Thm.

② Bayesian Tests:

Given $\theta \sim \pi(\cdot)$. $X \sim f(x|\theta)$. Calculate $\pi(\theta|x)$

Then we have: $P(\theta \in \Theta_0|\vec{x})$. $P(\theta \in \Theta_0^c|\vec{x})$

since these two prob have sum 1.

We can set rejection region: $R = \{\vec{x} \mid P(\theta \in \Theta_0^c|\vec{x}) > c\}$

for some const. $c \in (0,1)$

③ Union-Intersection Test and Intersection-Union Test:

· Test for complicated null hypotheses can be developed from simpler null hypotheses.

i) U-I Method:

· For $H_0: \theta \in \bigcap_{\gamma \in I} \Theta_\gamma$. We can separate it:

Test: $H_{0\gamma}: \theta \in \Theta_\gamma$ v.s. $H_{1\gamma}: \theta \in \Theta_\gamma^c$ with

rejection region R_γ

So. the rejection region of H_0 is $R = \bigcup_{\gamma \in I} R_\gamma$

ii) I-U Method:

Analogously. $H_0: \theta \in \bigcup_{\gamma \in I} \Theta_\gamma$. separate:

$H_{0\gamma}: \theta \in \Theta_\gamma$ v.s. $H_{1\gamma}: \theta \in \Theta_\gamma^c$. with

rejection region R_γ.

Then $R = \bigcap_{\gamma \in I} R_\gamma$

(2) Methods of evaluting tests:

· Intuitively. hypothesis test are evaluted by comparing the prob. of making mistake.

① Error prob and Power function:

		Decision: Accept H_0	Decision: Reject H_0
Truth	H_0	Correct Decision	Type I error
Truth	H_1	Type II error	Correct Decision

$\therefore \begin{cases} p(\text{Type I error}) = p(\vec{X} \in R \mid \theta \in \Theta_0) \\ p(\text{Type II error}) = p(\vec{X} \in R^c \mid \theta \in \Theta_0^c) \end{cases}$

Def: The power function of hypothesis test with rejection Region R is: $\beta(\theta) = P_\theta(\vec{X} \in R)$

$\Rightarrow$ For $0<\alpha<1$. a test with power function $\beta(\theta)$ is "size" α test, if $\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$.

is a "level" α test if $\sup_{\theta \in \Theta_0} \beta(\theta) \le \alpha$.

Remark: It's usually impossible to make both error probs small. We limit on the class with small $p(\text{Type I})$ firstly, then minimize $p(\text{Type II})$

$\Rightarrow$ It's important to specify null and alternative hypotheses for controling $p(\text{Type I error})$

Methods of determing the null hypothese H_0

i) Choose simpler hypo (e.g. $\theta = 0.5$) on H_0. It's easy to find $p(\text{Type I error})$

ii) Put the hypo you trust more on H_0 (with stronger evidence, not mentally)

iii) Put the hypo which will lead to serious consequense owing to reject it on H_0.
e.g. put the hypo of some result of an experiment on H_1. H_1 is also called research hypothesis.

· Def: cutoff number z_α is $P(Z \geq z_\alpha) = \alpha$.

(Distinguish it from quantile, percentile)

· Def: A test with power function $\beta(\theta)$ is unbiased

if $\beta(\theta') \geq \beta(\theta'')$, $\forall \theta' \in \Theta_0^c$, $\theta'' \in \Theta_0$.

② The most powerful Test:

· As before, we will restrict on a class C of tests for $H_0: \theta \in \Theta_0$ v.s. $H_1: \theta \in \Theta_0^c$ with level α. Then a good test in such class should have a small prob of Type II error.

$\Rightarrow$ Def: A test in C with power function $\beta(\theta)$ is uniformly most powerful (UMP) if for any other test with $\beta'(\theta)$,

$\Rightarrow \beta(\theta) \geq \beta'(\theta)$, $\forall \theta \in \Theta^c$.

Remark: Specify the elements in C: (H_0 vs. H_1)

generalizedly randomized Tests:

$$\phi(\vec{X}) = \begin{cases} 1, & \vec{X} \in R \\ 0, & \vec{X} \in R^c \\ \lambda, & \vec{X} \in \partial R \end{cases}$$ (λ can be controled to set a size α test, exactly)

$C = \{\phi_i \mid P(\phi_i = 1 \mid \theta \in \Theta_0) \leq \alpha\}$.

To characterize UMP Test. begin with both simple hypothesis (e.g. $\theta = \lambda$. for some const. λ)

<u>Thm.</u> (Neyman - Pearson Lemma)

Consider $H_0: \theta = \theta_0$ v.s. $H_1: \theta = \theta_1$. $\vec{X} \sim f(\vec{x}|\theta)$

For a test with rejection Region R. st.

$\vec{x} \in R \Leftrightarrow f(x|\theta_1) \geq k f(x|\theta_0)$. for some $k \geq 0$.

(i.e. $\vec{x} \in R^c$ if $f(x|\theta_1) < k f(x|\theta_0)$).

Let $P_{\theta_0}(X \in R) = \alpha$.

i) Then it's the UMP test. of α level.

ii) Conversely. every UMP test of α level (if exist) satisfies the condition above. except a set A. where $P_{\theta_0}(\vec{X} \in A) = P_{\theta_1}(\vec{X} \in A) = 0$ When there exists a test satisfies the condition above, with replace "$k \geq 0$" with "$k > 0$".

<u>pf:</u> i) (Sufficient) → used to construct a UMP Test

Denote that test ϕ. For any other test ϕ' with rejection region R'. st. $P_{\theta_0}(\vec{x} \in R') \leq \alpha$

Then $\beta(\theta_1) - \beta'(\theta_1) = \int_R f(x|\theta_1) - \int_{R'} f(x|\theta)$

$= \int_{R/R'} f(x|\theta_1) - \int_{R'/R} f(x|\theta_1)$

$\geq k[\int_{R/R'} f(x|\theta_0) - \int_{R'/R} f(x|\theta_0)]$

$= k [\int_R f(x|\theta_0) - \int_{R'} f(x|\theta_0)] \geq 0$.

Replace "$\int$" with "Σ" for discrete case.

ii) (Necessary) Only for conti case.

Note i). "=" holds when $P_{\theta_0}(R/R') = P_{\theta_1}(R'/R) = 0$.

$\therefore$ ϕ, ϕ''s rejection region R, R' only differs a prob. measure 0 set.

→ Used to show nonexistence of UMP test. (Because its R has some kind of Uniqueness)

Remark: i) The point is that it transform "comparing power function (θ_1)" to "Comparing level" by the condition of rejection region.

ii) For $T(\vec{x})$ s.s. for θ, the condition can be reduced: (suppose test based on T has R_T = reject. region)

$$t \in R_T \iff g(t|\theta_1) \geq k\, g(t|\theta_0)$$

Next, we consider one of hypothese is composite hypo:

Def: A family of pdf's or pmf's $\{g(t|\theta) | \theta \in \Theta\}$ for univariate r.v. T has a monotone likelihood Ratio (MLR) iff for $\forall\ \theta_2 > \theta_1$, $\frac{g(t|\theta_2)}{g(t|\theta_1)}$ is mono on t.

Remark: $X \sim h(t)c(\theta)e^{w(\theta)t}$. w is monotone. Then, it has MLR!

Thm. (Karlin-Rubin)

Consider $H_0: \theta \leq \theta_0$ v.s. $H_1: \theta > \theta_0$. $T \sim g(t|\theta)$, which has a MLR of increasing. T is s.s. for θ. Then $\forall t_0$, "Rejects H_0 when $T \geq t_0$" is UMP $\alpha = P(R|H_0)$ test.

pf: For reducing to simple hypothesis

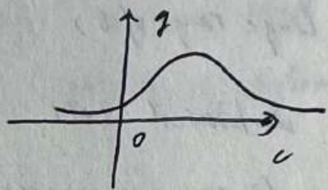
Fix $\theta' > \theta_0$. Consider $H_0': \theta = \theta_0$ v.s. $H_1': \theta = \theta'$.

Lemma. If $g(t|\theta)$ of T has an increasing type MLR. Then for $\theta_1 < \theta_2$, $P_{\theta_1}(T > c) \le P_{\theta_2}(T > c)$, $\forall c$.

Pf: Let $g(c) = P_{\theta_2}(T > c) - P_{\theta_1}(T > c)$

$g'(c) = g(c|\theta_1) - g(c|\theta_2) = g(c|\theta_1)\left(1 - \dfrac{g(c|\theta_2)}{g(c|\theta_1)}\right)$

Note $g(c|\theta_2)/g(c|\theta_1) \uparrow$. $g(\infty) = g(-\infty) = 0$.

$\Rightarrow g(c)$: 

$\therefore g(c) \ge 0$ $\forall c \in \mathbb{R}$.

$\Rightarrow \beta(\theta)$ is nondecreasing on θ.

i) $\sup_{\theta \le \theta_0} \beta(\theta) = \beta(\theta_0) = \alpha$. So it's level α test.

ii) Def: $k' = \inf_{t \in T} \dfrac{g(t|\theta')}{g(t|\theta_0)}$. $T = \{t > t_0,\ g(t|\theta') \text{ or } g(t|\theta_0) > 0\}$.

$\therefore T \ge t_0 \Leftrightarrow g(t|\theta') \ge k' g(t|\theta_0) \Leftrightarrow t \in R$.

$\therefore$ For any other α level test with $\beta^*(\cdot)$.

since $\beta^*(\theta_0) \le \sup_{\theta \le \theta_0} \beta^*(\theta) \le \alpha$. $\therefore \beta^*(\theta') \le \beta(\theta')$. $\forall \theta' > \theta_0$.

Since θ' is arbitrary. It holds for $H_0'': \theta = \theta_0$ v.s $H_1'': \theta > \theta_0$

By i). extend to $H_0: \theta \le \theta_0$ v.s. $H_1: \theta > \theta_0$.

Remark: i) For decreasing type. Let $T = -T$.

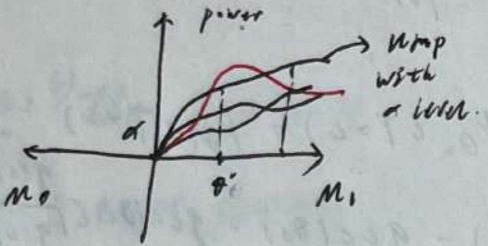
reverse R: $T < t_0$.

ii) For $H_0: \theta \geq \theta_0$ v.s. $H_1: \theta < \theta_0$. Similarly.

reverse $T \leq t_0$!

iii) We can operate hypothese from: simple v.s. simple.

$\Rightarrow$ composite v.s. simple $\Rightarrow$ composite v.s. composite.



(If the red curve exists. Then UMP test with α level doesn't exist.)

(Nonexistence of UMP test often shows up in two sides hypotheses for its large range of θ)

$\Rightarrow$ Restrict on unbiased test may result in finding UMP test!

③ Size of UI and IU tests:

i) UI Test:

· A relationship between LRT and UIT:

Thm. For $H_0: \theta \in \Theta_0$ v.s. $H_1: \theta \in \Theta_0^c$. $\Theta_0 = \bigcap_{\gamma \in I} \Theta_\gamma$

$\lambda_\gamma(\vec{x})$ is LRT for each $H_{0\gamma}$ vs. $H_{1\gamma}$. $\lambda(x)$ is LRT for H_0 v.s. H_1. $T(x) = \inf_{\gamma \in I} \lambda_\gamma(x)$

For test one: $R: \{T(x) \leq c\}$ with $\beta_T(\theta)$

test two: $R: \{\lambda(x) \leq c\}$. with $\beta_\lambda(\theta)$.

Then. $T(\vec{x}) \geq \lambda(\vec{x})$. $\forall \vec{x} \in \Omega$. Besides.

$\beta_T(\theta) \le \beta_\lambda(\theta)$. $\forall \theta \in \Theta$. so if LRT test is level α. then UIT is α level.

<u>Pf:</u> $\forall \gamma \in I$. $\lambda_\gamma(\vec{x}) \ge \lambda(\vec{x})$, since $\Theta_\gamma \supseteq \Theta_0$

From $T(\vec{x}) \ge \lambda(\vec{x})$ $\therefore \{T(\vec{x}) \le c\} \subseteq \{\lambda(\vec{x}) \le c\}$.

$\therefore \beta_T(\theta) \le \beta_\lambda(\theta)$. $\therefore \sup_{\theta\in\Theta_0} \beta_T(\theta) \le \sup_{\theta\in\Theta_0} \beta_\lambda(\theta) \le \alpha$.

<u>Remark:</u> LRT is more powerful than UIT. But we usually use UIT:

$\begin{cases} \text{i) UIT has smaller type I error prob.} \\ \text{ii) If } H_0 \text{ is rejected. we can look at } H_{0\gamma} \end{cases}$

to see why! (additional information!)

ii) <u>IU Test:</u>

<u>Thm.</u> α_γ is size of test of $H_{0\gamma}$ v.s. $H_{1\gamma}$. with rejection region R_γ. Then IUT with $R = \bigcap_{\gamma\in I} R_\gamma$ is level $\alpha = \sup_{\gamma\in I} \alpha_\gamma$ test.

<u>Pf:</u> $\forall \theta \in \Theta$. $P_\theta(X \in R) \le P_\theta(X \in R_\gamma)$. Since $R \subseteq R_\gamma$. $\forall \gamma \in I$.

<u>Remark:</u> The shortcoming is that it's conservative, Since α may be much larger than its size.

<u>Thm.</u> For $H_0: \theta \in \bigcup_1^k \Theta_i$. R_i is rejection region for $H_{0i}: \theta \in \Theta_i$ of level α. If for some i. $1 \le i \le k$. st. exists a seq of parameters $\{\theta_l\} \subseteq \Theta_i$. st.

$\begin{cases} \lim_{l\to\infty} P_{\theta_l}(\vec{X} \in R_i) = \alpha \\ \forall j \ne i. \lim_{l\to\infty} P_{\theta_l}(\vec{X} \in R_j) = 1 \end{cases} \Rightarrow$ IUT with $R = \bigcap_1^k R_i$ is size α test.

Pf: By Thm above. $\sup_{\theta\in\Theta_0} P_\theta(\vec{X}\in R) \le \alpha$.

$$\sup_{\theta\in\Theta_0} P_\theta(\vec{X}\in R) \ge \lim_{l\to\infty} P_{\theta_l}(\vec{X}\in R) = \lim_{l\to\infty} P_{\theta_l}(\vec{X}\in \bigcap_1^k R_i)$$

$$\ge \lim_{l\to\infty} \sum_1^k P_{\theta_l}(X\in R_j) - (k-1) = \alpha.$$

Since $P_{\theta_l}(\vec{X}\in\bigcap_1^k R_i) = 1 - P_{\theta_l}(\vec{X}\in \bigcup_1^k R_i^c)$

$$\ge 1 - \sum_1^k P_{\theta_l}(\vec{X}\in R_i^c) = \sum_1^k P_{\theta_l}(\vec{X}\in R_i) - (k-1) \quad \square$$

Remark: Particular case: $\sup_{\theta\in\Theta_i} \beta_i(\theta) = \lim_{l\to\infty} P_{\theta_l}(\vec{X}\in R_i) = \alpha$.

④ p-value:

α hasn't given out yet.

Then we can determine by giving "α" by ourselves. ← · Another way of reporting the result (Not use Accept or reject) of test is to report a statistic — p-value.

Def: p-value $p(\vec{X})$ is a test statistic s.t. $0\le p(\vec{x})\le 1$. $\forall\, \vec{x}\in\Omega$. Small $p(\vec{x})$ means it's skewed. giving evidence that H_1 is true.

(*) Because: $P_\theta(p(X)\le\alpha) = P_\theta(X\in R_1)$. naturally! ←

$\Rightarrow$ A p-value is valid if $P_\theta(p(\vec{X})\le\alpha)\le\alpha$. (*)

So use it to construct R. we can obtain a level α test.

The most common way to define a p-value:

Thm. $W(\vec{X})$ is a test statistic. s.t. large value of $W(\vec{x})$ gives evidence to H_1 is true. Def:

$p(\vec{x}) = \sup_{\theta\in\Theta_0} P_\theta(W(\vec{X}) \ge W(\vec{x}))$. for each $\vec{x}$.

Then $p(\vec{X})$ is valid p-value.

Pf: $P_\theta(W(\vec{X}) \geq W(\vec{x})) = P_\theta(-W(\vec{x}) \geq -W(\vec{X}))$

$= F_\theta(-W(\vec{x})) = P_\theta(\vec{x})$. F_θ is cdf of $-W(\vec{X})$.

$\therefore P_\theta(p(\vec{X}) \leq \alpha) \leq P_\theta(P_\theta(\vec{X}) \leq \alpha) = P_\theta(F_\theta(-W(x)) \leq \alpha)$

$= P_\theta(-W(x) \in A_\alpha) = F_\theta(-W_\alpha(x)) \leq \alpha$.

$A_\alpha = \{-W(x) \mid F_\theta(-W(x)) \leq \alpha\} = (-\infty, -W_\alpha(x)]$. (right-conti)

where $F_\theta(-W_\alpha(x)) \leq \alpha$. Since $-W_\alpha(x) \in A_\alpha$.

Remark: i) $p(\vec{x}) \downarrow$ means $W(\vec{X}) > W(\vec{x})$ has small prob.

$\Leftrightarrow W(\vec{x})$ is large. $\therefore$ gives support to H_1.

ii) An useful Lemma:

· X has cdf $F_X(\cdot)$. Then $p(F_X(X) \leq x) \leq x$.

Pf: $\{X \mid F_X(X) \leq x\} = (-\infty, t_x]$ or $(-\infty, t_x) \triangleq A_x$.

The second case can happen in discrete case.

$\therefore \lim_{t \uparrow t_x} F_X(t) = F_X(t_x) \leq x$. (By extending)

$\therefore p(F_X(X) \leq x) = p(X \in A_x) = p(X \leq t_x) =$

$F_X(t_x) \leq x$. Since $t_x \in A_x$.

iii) Another interpretation of p-value: (Observe significant level)

Under the condition in the Thm.

If C_α is a critical value chosen s.t. $\{\vec{x} \mid W(\vec{x}) \geq C_\alpha\}$

is a rejection region of size α test of H_0. Then

$p(\vec{x})$ is the smallest value of level that rejects H_0

Pf: $\alpha = \sup_{\theta \in \Theta_0} P_\theta(W(\vec{X}) \geq C_\alpha)$

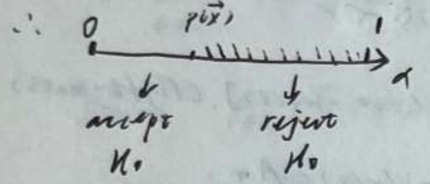
$p(\vec{x}) = \sup_{\theta \in \Theta_0} P_\theta(W(\vec{X}) \geq W(\vec{x}))$

If $p(\vec{x}) = \sup_{\theta \in \Theta_0} P_\theta(W(\vec{X}) \geq W(\vec{x})) \geq \alpha = \sup_{\theta \in \Theta_0} P_\theta(W(\vec{X}) \geq c_\alpha)$

$\therefore$ $W(\vec{x}) \geq c_\alpha$. (Argue by contradiction). reject H_0

Conversely if $p(\vec{x}) < \alpha$. we have $W(\vec{x}) < c_\alpha$. accept H_0

$\therefore$



when $p(\vec{x})$ is extreme. we need to support H_0 strongly to accept H_0.

Remark: i) Note that the smaller p-value is. the more chance we can reject H_0. That's because if we want to accept H_0. then the prob. of type I error should be small enough! (< p-value)

ii) p-value is the "Level" of experiment we have done.

Another method to Construct p-value:

· suppose $S(\vec{X})$ is a. s.s. for $\{f(\vec{x}|\theta) | \theta \in \Theta_0\}$

(not $\{f(x|\theta) | \theta \in \Theta\}$. $\Theta_0 \subseteq \Theta$). Then for each $\vec{x} \in \mathcal{X}$.

Def: $p(\vec{x}) = P(W(\vec{X}) \geq W(\vec{x}) \mid S = S(\vec{x}))$ (indept of $\theta \in \Theta_0$)

Then it's valid: for $\forall \theta \in \Theta_0$

$$P_\theta(p(\vec{X}) \leq \alpha) = \sum_s P(p(\vec{X}) \leq \alpha \mid S = s) P_\theta(S = s)$$

$$\leq \sum_s \alpha P_\theta(S = s) = \alpha.$$

Remark: p-value for IUT:

$H_0: \theta \in \bigcup_{j}^{k} \Theta_j$. p_j is p-value for H_{0j}. Then

$p(\vec{X}) = \max_{1 \leq j \leq k} p_j(\vec{X})$ is a valid p-value for H_0.

($P_\theta(p(\vec{X}) \leq \alpha) \leq P_\theta(p_j(\vec{X}) \leq \alpha) \leq \alpha$, Since $\exists j$.

$\theta \in \Theta_j$. $p_j(\vec{X}) \leq p(\vec{X})$. Analogous for UIT!)