UNC BIOS 761

Summary

Creating my own project

Author:
Mingwei Fei

Special colaborator:
Bonnie

Jan, 2023

Contents

Chapter 1

Loss Function and Optimization

Chapter 2

Hypothesis Test

2.1 UMP

Neyman - Pearson Lemma: Uniformly max power test

The theory mainly compares the probability distribution and find a critical region that is a uniformly max power test. To learn a theorem, we need to understand the proof and also an associated example.

Suppose that $P_0 = \{p_0\}$ and $P_1 = \{p_1\}$. Let f_j be the p.d.f.

2.1.1 Proof

To prove one test ϕ is the uniform max power test $(E_{\theta_0}(\phi) = \alpha)$, then we assume another test ϕ^N with type I error $\leq \alpha$, and prove that the power of this test is bigger than the other test.

(i) Existence of UMP test For every α, there exists a UMP test of size α, which is

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

where $\gamma \in [0, 1]$ and $k \geq 0$ are some constants chosen so that $E[\phi] = \alpha$ when $P = P_0$ ($k = \infty$ is allowed).

(ii) Uniqueness of UMP test Assume another test ϕ^N, such that $E_{P_0}[\phi^N] \leq \alpha$. We need to prove that $\beta_\phi \geq \beta_{\phi^N}$. only when the set $B = \{x : f_1(x) = k f_0(x)\}$, and $v(B) = 0$, then we have a unique nonrandomized UMP; otherwise UMP tests are randomized on the set B and the randomization is necessary for UMP tests to have the given size α.

$$\phi^N(x) = \begin{cases} 1 & f_1(x) > k f_0(x) \\ 0 & f_1(x) < k f_0(x) \end{cases}$$

(iii) Proof

We can see ϕ^N has the following characteristics: (montone) If $\phi^N - \phi > 0$, then $\phi^N > 0$ and $f_1(x) \geq k f_0(x)$. If $\phi^N - \phi < 0$, then $\phi^N < 1$ and $f_1(x) \leq k f_0(x)$.

In any case,

$$[\phi^N - \phi][f_1(x) - k f_0(x)] \geq 0$$

and therefore,

$$\int [\phi^N - \phi][f_1(x) - k f_0(x)] dv \geq 0$$

$$\int [\phi^N - \phi] f_1(x) dv \geq \int [\phi^N - \phi] k f_0(x) dv$$

The left-hand side is $E_1(\phi^N) - E_1(\phi)$ and the right-hand side

$$k(E_0[\phi^N] - E_0[\phi]) = k(\alpha - E_0[\phi]) \geq 0$$

2.1.2 Example

The example just gives two distributions: Suppose that X is a sample of size 1, $P_0 = \{p_0\}$ and $P_1 = \{p_1\}$, where P_0 is $N(0,1)$ and P_1 is the double exponential distribution $DE(0,2)$ with the p.d.f. $4^{-1} e^{-|x|/2}$

Since $P\left(f_1(x) = c f_0(x)\right) = 0$, there is a unique nonrandomized UMP test.

$$\phi(x) = \begin{cases} 1 & \left(4^{-1} e^{-|x|/2}\right)^2 > k^2 (2\pi)^{-\frac{1}{2} \times 2} \exp\left(-x^2\right) \\ 0 & f_1(x) < k f_0(x) \end{cases}$$

which is $|x| > t$ or $|x| < 1 - t$ for some $t > \frac{1}{2}$. Suppose that $\alpha < \frac{1}{3}$, to determine t, we use

$$\alpha = E_0[\phi] = P_0[|x| > t] + P_0[|x| < 1 - t]$$

(i) t should be larger than 1.

If $t \leq 1$, then $P_0(|x| > t) \geq P_0(|x| > 1) = 0.3374 > \alpha$

4

(ii) So the probability simplified

$$\alpha = P_0(|x| > t) = \Phi(-t) + 1 - \Phi(t)$$

thus, $t = \Phi^{-1}(1 - \alpha/2)$ and $\phi = I_{(t,\infty)}(|X|)$. Note, it is not necessary to find out what c is.

Another common example is in binomial distribution, which we will have randomized UMP test.

An interesting phenomenon is that ϕ is a test that does not depend on P_1. So it is the range of parameters in H_1.

2.2 Likelihood Ratio Test

Likelihood ratio test is associated with the Neyman-Pearson UMP test, it realizes on the monotone likelihood ratio and an extension of UMP test.

2.2.1 Definition

Let $l(\theta) = f_\theta(X)$ be the likelihood function. For testing $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$, a likelihood ratio (LR) test is any test that rejects H_0 if and only if $\lambda(X) < c$, where $c \in [0, 1]$ and $\lambda(X)$ is the likelihood ratio defined by

$$\lambda(X) = \frac{\sup\limits_{\theta \in \Theta_0} l(\theta)}{\sup\limits_{\theta \in \Theta} l(\theta)}$$

Question:

(i) Why we need to use the SUP here? Is it related to the monotone likelihood ratio? We could see that $\sup\limits_{\theta \in \Theta} l(\theta)$ is a fixed number, comparing with $\sup\limits_{\theta \in \Theta_1} l(\theta)$. While in the nominator, any likelihood would be less than $\sup\limits_{\theta \in \Theta_0} l(\theta)$, so if we rejected the H_0 by the sup, any other θ_0 would be rejected as well.

(ii)

2.2.2 Relationship between LR test and UMP test

Why do we need to study this?

(i) Why do we use LR test while not UMP test?

One requirement for UMP test is to have monotone likelihood ratio, however ie. when $H_0 : a < \theta < b$, the likelihood increases and decreases, so the monotone LR doesn't hold here. However, it won't stop us using the LR test. As the LR test only uses the Sup of likelihood, which does not depend on the monotone ratio.

(ii) Here we compare the rejection region that the two tests give. UMP test won't give a rejection region that go to different direction, as it requires the power uniformly increase to max.

Suppose that X has a p.d.f. in a one-parameter exponential family:

$$f_\theta(x) = \exp\{\eta(\theta)Y(x) - \xi(\theta)\}h(x)$$

w.r.t. a $\sigma-$ finite measure v, where η is a strictly increasing and differentiable function of θ.

(i) For testing $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$ there is an LR test whose rejection region is the same as that of the UMP test given in Theorem 6.2.

proof: monotone increasing of $l(\theta_0)/l(\hat{\theta})$, and when $\lambda = l(\theta_0)/l(\hat{\theta}) < c$, is equal to $\theta_0 < \hat{\theta}$

Let $\hat{\theta}$ be the MLE of θ. Note that $l(\theta)$ is increasing when $\theta < \hat{\theta}$ and decreasing when $\theta > \hat{\theta}$. Thus

$$\lambda = \begin{cases} 1 & \theta \geq \hat{\theta} \\ \frac{l(\theta)}{l(\hat{\theta})} & \theta < \hat{\theta} \end{cases}$$

Then $\lambda(X) < c$ is the same as $\hat{\theta} > \theta_0$ and $\frac{l(\theta)}{l(\hat{\theta})} < c$.

From the property of exponential families, $\hat{\theta}$ is a solution of the likelihood equation. For any $\theta_0 \in \Theta$, $log(l(\hat{\theta})) - logl(\theta_0)$ is strictly increasing in Y when $\hat{\theta} > \theta_0$, and strictly decreasing in Y when $\hat{\theta} < \theta_0$.

Hence, for any $d \in R, \hat{\theta} > \theta$ and $l(\theta_0)/l(\hat{\theta}) < c$ is equivalent to $Y > d$ for some $c \in (0,1)$.

Here, we need to pay attention to the monotone increasing or decreasing is respective to Y. When Y increases, the LR increases or decreases.

(ii) For testing $H_0 : \theta \leq \theta_1, \theta \geq \theta_2$ versus $H_1 : \theta_1 < \theta < \theta_2$, there is an LR test whose rejection region is the same as that of the UMP test T given in Theorem 6.3.

(iii) For testing the other two-sided hypotheses, there is an LR test whose rejection region is equivalent to $Y(X) < c_1$ or $Y(X) > c_2$ for some constants c_1 and c_2.

2.2.3 Example

We would like to show the connection between LR test and UMP test.

(i) Consider the testing problem $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ based on i.i.d. $X_1, \ldots, X_n$ from the uniform distribution $U(0, \theta)$.

This problem we can do in both UMP and LR test. The lebesgue p.d.f. used in UMP test is the same as the likelihood function Let

(ii)

2.3 Generalized NP lemma

Let $f_1, \ldots, f_{m+1}$ be real-valued, μ- integrable functions defined on a Euclidean space X. Suppose that for given constants $c_1, \ldots, c_m$ there exists a critical function ϕ satisfying

$$\int \phi f_i d\mu = c_i, \qquad i = 1, ..m \tag{2.1}$$

Let C be the class of critical functions ϕ for which (1.1) holds

(i) Among all members of C there exists one that maximizes $\int \phi f_{m+1} d\mu$

Note that, $\int \phi f_{m+1} d\mu$ is the power of the test, as the critical function ϕ is to $f_1, ..f_m$.

(ii) A sufficient condition for a member ϕ_0 of C to maximize $\int \phi f_{m+1} d\mu$ (over C) is the existence of constants $k_1, \ldots, k_m$ such that

$$\phi_0(x) = \begin{cases} 1 & f_{m+1}(x) > \sum_{i=1}^{m} k_i f_i(x) \\ 0 & f_{m+1}(x) < \sum_{i=1}^{m} k_i f_i(x) \end{cases} \tag{2.2}$$

Proof: it is always a good practice to transform the complicated and general case into simple case, in this problem, we will try to transform into two parameter case.

Take $\phi \in C$. Note that $\int (\phi_0 - \phi)(f_{m+1} - \sum_{i=1}^{m} k_i f_i) d\mu \geq 0$ since the integrand is ≥ 0. [this is the same as in the UMP proof].

$$\int (\phi_0 - \phi) f_{m+1} d\mu \geq \sum_{i=1}^{m} k_i \int (\phi_0 - \phi) f_i d\mu \geq 0,$$

$$\int \phi_0 f_{m+1} d\mu \geq \int \phi f_{m+1} d\mu$$

(iii) If a member of C satisfies (1.2) with $k_1, \ldots k_m \geq 0$, then it maximizes $\int \phi f_{m+1} d\mu$ among all critical functions satisfying $\int \phi f_i d\mu \leq c_i$, for $i = 1, ..m$.

(iii) The set

$$M := \left(\int \phi f_1 d\mu, .., \int \phi f_m d\mu \right)$$

where ϕ is a critical function. It is convex and closed. If $(c_1, ..c_m)$ is an interior point of M, then there exists constants $k_1, .., k_m$ and a test ϕ_0 satisfying (1.1) and (1.2). And a necessary condition for a member of C to maximize $\int \phi f_{m+1} d\mu$ is that (1.2) holds a.e. μ.

2.3.1 Example

How do we use the general NP lemma? The probability could be all different

Suppose that $X_1, \ldots X_n$ are i.i.d. from the Cauchy location family $p_\theta(x) = \frac{1}{\pi} \frac{1}{1+(x-\theta)^2}$, for $x \in R$. Consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$. Can we find a test ϕ of size α such that ϕ maximizes

$$\frac{d}{d\theta} \beta_\phi(\theta_0) = \frac{d}{d\theta} E_\theta[\phi(X)]\Big|_{\theta=\theta_0}$$

For any test ϕ the power is given by

$$\beta_\phi(\theta) = E_\phi[\phi(X)] = \int \phi(X) p(x, \theta) dx$$

where $p(x, \theta)$ is the joint density of the model. So, if the interchange of differentiation and integration is justifiable, then

$$\beta'_\phi(\theta) = \int \phi(X) \frac{\partial}{\partial\theta} p(x, \theta) dx$$

Thus, by the generalized N-P lemma, a test of the form

$$\phi_)(x) = \begin{cases} 1 & \frac{\partial}{\partial\theta} p(X, \theta_0) > kp(X, \theta_0) \\ 0 & \frac{\partial}{\partial\theta} p(X, \theta_0) < kp(X, \theta_0) \end{cases}$$

maximizes $\beta'_\phi(\theta_0)$ among all ϕ with $E_{\theta_0}\phi(X) = \alpha$. This test is said to be locally most powerful of size α.

Also the likelihood function and density function are correlated.

$$\frac{\partial}{\partial\theta} p(X, \theta_0) > kp(X, \theta_0) \rightarrow \frac{\partial}{\partial\theta} logp(X, \theta_0) > k$$

$$S_n(\theta_0) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{l}_{\theta_0}(X_i) > k'$$

The point of doing that way is to simplify the equation, so that we be able to get the statistics involving X.

2.4 Unbiased tests

2.4.1 Definition

(i) Unbiased test (Constrain by enforcing unbiasedness):

A test ϕ for $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$ with level α is unbiased if its power $\beta_\phi(\theta) := E_\theta[\phi(X)]$ satisfies

$$\beta_\phi(\theta) \leq \alpha, \qquad \theta \in \Theta_0$$
$$\beta_\phi(\theta) \geq \alpha, \qquad \theta \in \Theta_1$$

If there is a UMP test ϕ^*, then it is automatically unbiased because $\beta_{\phi^*} \geq \beta(\theta)$, for all $\theta \in \Theta_1$, where ϕ is the degenerate test, which equals α regardless of the observed data.

Unbiasedness enforces the appealing property that the probability of rejection is greater under any alternative distribution than it is under any null distribution. A uniformly most powerful test is always unbiased if it exists.

While this notion of unbiasedness differs from the definition we encountered when discussing point estimation, we can check that this is actually a special case of risk unbiasedness when the loss function L is such that $L(\theta_0, reject) = 1 - \alpha, L(\theta_1, accept) = \alpha$

(ii) UMP Unbiased test:

we will focus on finding uniformly most powerful unbiased (UMPU) tests in settings in which UMP tests do not exist. These tests often exist for testing $\theta_1 \leq \tilde{\theta}$ vs $\theta_1 > \tilde{\theta}$ in the presence of nuisance parameters $(\theta_2, \ldots \theta_n)$ and for testing $\theta = \tilde{\theta} vs \theta \neq \tilde{\theta}$.

A UMP unbiased test (UMPU) level α is a test ϕ_0 for which

$$\beta_{\phi_0}(\theta) \geq \beta_\phi(\theta), \qquad \theta \in \Theta_1$$

for all unbiased level α tests ϕ.

This definition says that, among all the unbiased test, the UMPU test has higher power.

Generally, if the power function $\theta \to \beta_\phi(\theta)$ is continuous in θ (as is the case for any canonical form exponential family on the natural parameter space), then ϕ unbiased and of level α implies that $\beta_\phi(\theta) = \alpha$ for all $\theta \in \omega$. We have a name for tests that match the level on the boundary.

Proof. Firstly, because ϕ_0 is UMP α-similar tests, it is at least as powerful as $\phi_\alpha(X) \equiv \alpha$, and the power of ϕ_0 on Ω_1 is therefore $\geq \alpha$. Hence, ϕ_0 is unbiased.

Secondly, an unbiased level-α test must, by definition, have expectation value $\leq \alpha$ for $\theta \in \Omega_0$ and $\geq \alpha$ for $\theta \in \Omega_1$. By continuity such a test must have expectation α

on the common boundary. Therefore, the set of unbiased level-α tests is a subset of α-similar level-α tests, amongst which ϕ_0 is most powerful. Hence, ϕ_0 is also as powerful as any unbiased level-α test. ϕ_0 is UMPU.

2.4.2 Two sided testing without Nuisance parameters

It is based on the hypothesis that, the null hypothesis has only one θ_0, while the H_1 has two different direction of range. So based on the UMPU test definition, we look at the power function as a function of θ. And it reaches the minimum at θ_0.

Remember that UMP test is just based on the type I error rate and power, it is not important to find out what the likelihood ratio is. And the same applies to UMPU test. We only need to focus on the power function, and reaches the minimum at θ_0. Furthermore, we can find out the critical function ϕ.

Let us test $H_0 : \theta = \theta_0 vs. H_1 : \theta \neq \theta_0$, when X is distributed according some member of the one-dimensional exponential family

$$p_\theta(x) = h(x) \exp\left(\theta T(x) - A(\theta)\right)$$

We have seen that no UMP test exists in the normal case. Our goal here is to find a UMPU test.

Since we are working with an exponential family, the power function is continuous, and, by Lemma 1, it suffices to find a UMP level α test amongst α-similar tests. Since $\omega = \Omega_0$, any UMP α-similar test ϕ has

$$\beta_\phi(\theta_0) = E_{\theta_0}\phi(X) = \alpha$$
$$\beta_\phi(\theta_0) \leq \beta_\phi(\theta), \qquad \theta \in R$$

since $\phi_\alpha(x) \equiv \alpha$ is also α-similar.

Since θ_0 minimizes β_ϕ, and β_ϕ is differentiable with derivative

$$\beta'_\phi(\theta) = \int \phi(x) \frac{d}{d\theta} p_\theta(x) d\mu(x)$$

We have the constraint

$$0 = \beta'_\phi(\theta_0) = \int \phi(x) \frac{d}{d\theta} p_{\theta_0}(x) d\mu(x)$$

2.5 Multiple Constraints - Method of Undetermined Multipliers (MoUM)

In this setting, $H_0 : \theta = \theta_0 vs. H_1 : \theta = \theta'$. We will fix a simple alternative $\theta = \theta'$ and hope that our best test has no θ' dependence. We would like to maximize power subject to

$$\int \phi p_{\theta_0}(x)d\mu(x) = \alpha$$

$$\int \phi \frac{d}{d\theta}p_{\theta_0}(x)d\mu(x) = 0$$

For a 1-parameter exponential family, we have

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

$$\frac{d}{d\theta}p_\theta(x) = h(x)e^{\theta T(x) - A(\theta)}(T(x) - A'(\theta)) = p_\theta(x)\Big[T(x) - E_\theta(T(X))\Big]$$

Applying the reasoning from the previous section, we find that a most powerful test has rejection region defined by

$$p_{\theta'}(x) > k_1 p(\theta_0)(x) + k_2 \frac{d}{d\theta}p(\theta_0)(x)$$

for some values of k_1 and k_2, which is equivalent to

$$\frac{e^{(\theta' - \theta_0)T(x)}}{k_1' + k_2'T(x)} > const$$

Now consider the set of values of T(x) satisfying this constraint. Because the constraint is that an exponential function exceeds a linear function, the set of values of T(x) satisfying this constraint is either a one-sided interval

2.6   Likelihood Ratio Test

Derive the likelihood ratio test for the hypothesis in part (e) and derive its asymptotic distribution under H_0. From part (e), we have the parameter estimates under H_0. While under alternative hypothesis, we have $\mu_{ij} = n_{ij}$.

$$LRT_n = 2(LR(\pi_{H_1}) - LR(\pi_{H_0})) = 2\left(\sum_{i=1}^{I}\sum_{j=1}^{J}n_{ij}log\pi_{ij} - \sum_{i=1}^{I}\sum_{j=1}^{J}n_{ij}log\pi_{i+}\pi_{+j}\right)$$

$$= 2\left(\sum_{i=1}^{I}\sum_{j=1}^{J}n_{ij}log\frac{\pi_{ij}}{\pi_{i+}\pi_{+j}}\right)$$

$$= 2\left(\sum_{i=1}^{I}\sum_{j=1}^{J}n_{ij}log\frac{n_{ij}n}{n_{i+}n_{+j}}\right) \sim \chi^2_{(I-1)(J-1)}$$

Note that the full model has $(IJ - 1)$ parameters, and the null hypothesis has $(I - 1) + (J - 1)$ parameters.

$$df = I \times J - 1 - (I - 1) - (J - 1)$$
$$= (I - 1)(J - 1)$$

Bibliography

[1] Armstrong, M.A. *Basic Topology.* England: Editorial Board, 2000.

[2] Coxeter, H.S.M. *Introduction to Geometry.* Toronto: John Wiley and Sons Inc, 1969.

[3] Chronister, James. *Blender Basics.* Estados Unidos: 2004.

[4] Engelking, Ryszard. *General Topology.* Berlin: Hildermann, 1989.

[5] Engelking, Ryszard y Karol Sieklucki. *Topology: A Geometric Approach.* Berlin: Hildermann, 1992.

[6] Hatcher, Allen. *Algebraic Topology.* Cambridge: Cambridge University Press, 2002.

[7] Hoffman, Kenneth y Ray Kunze. *Linear Algebra, second edition.* New Jersey: Prentice Hall. Inc, 1961.

[8] Kackzynski, T., K. Mischainkow y M. Mrozen. *Algebraic Topology: A Computational Approach.* 2000.

[9] Kosniowski, Czes. *Topología Algebraica.* Barcelona: Editorial Revert, 1986.

[10] Lefschetz, Solomon. *Algebraic Topology.* Rhode Island: American Mathematical Society, 1991.

[11] May, J. P. *A Concise Course in Algebraic Topology.* Chicago.

[12] Mill, J. Van. *Infinite-Dimensional Topology.* Amsterdam: Elsevier Science Publications B. V, 1989.

[13] Munkres, James R. *Topology.* New York: Prentice Hall Inc, 2000.

[14] Roseman, Dennis. *Elementary Topology.* New Jersey: Prentice Hall Inc, 1999.

[15] Sato, Hajime. *Algebraic Topology: An Intuituve Approach.* Rhode Island: American Mathematical Society, 1996.

[16] Zomorodian, Afra J. *Topology for Computing.* United States: Cambridge University Press, 2005.