`MATH5905Inference` `hypothesis`

Week 8 Hypothesis Testing

6.0 Basic Concepts

- **Two types of errors:**
  - type I: P(reject H_0 | H_0 correct) = level of test (siginificance) **错判了好学生**
  - type II: P(accept H_0 | H_1 correct) = 1 - (power of the test) **放过了坏学生**
- H_0 : often an initial claim that is based on previous analyses or specialized knowledge.
- H_1 : alternative hypothesis is what you might believe to be true or hope to prove true.
- $\alpha = size = level\ of\ significance$: probability of rejecting H_0 when it is correct
- **power:** probability of accepting H_0 when H_1 is correct
- **test:** A test statistic measures the degree of agreement between a sample of data and the null hypothesis.

> What is a test statistic?
> Critical Values: Find a Critical Value in Any Tail
> Power Functions

6.3 Neyman Pearson Lemma

> The problem of constructing most-powerful tests for simple hypotheses is solved by the Neyman–Pearson lemma, according to which the likelihood-ratio test is a most-powerful test.
>
> If the competing hypotheses H_0 and H_1 are compound, the problem of devising most-powerful tests is formulated in terms of a uniformly most-powerful test, if such a test exists.

Suppose we have **simple case hypothesis** with $H_0 : \theta = \theta_0 \ and \ H_1 : \theta = \theta_1$

$$\varphi^*(x) = \begin{cases} 1, & \text{if } x \in S = \{x : \frac{L(x,\theta_1)}{L(x,\theta_1)} > C\} \ reject\ region \\ \gamma, & \text{if } x \in R = \{x : \frac{L(x,\theta_1)}{L(x,\theta_1)} = C\} \ hesitate \\ 0, & \text{if } x \in A = \{x : \frac{L(x,\theta_1)}{L(x,\theta_1)} < C\} \ accept\ region \end{cases} \tag{1}$$

Notes:

1. $E_{\theta_0}\varphi^* = \alpha$ can be interpreted as when $\theta = \theta_0$, power of test = size of test
2. First of all, we pick out test φ whose size is less than selected α (we want to minimize out type I error). Secondly, among all of these qualified tests, we want to find the one with the largest power to minimize the probability of committing type II error. We describe the "power" as $E_{\theta_1}\varphi^*$
3. φ^* is unique

6.6 Composite

6.6.1 Monotone likelihood ratio

Definition:
$\frac{L(x,\theta'')}{L(x,\theta')}$ is a non-decreasing function of $T(x) = T(x_1, x_2 \ldots, x_n)$

Note: For one-parameter exponential family, if $c(\theta)$ is monotone **increasing**, this family has MLR in $T(x) = \sum_{i=1}^n d(X_i)$

6.6.2 Blackwell & Girshick Theorem

$If\ X \sim L(x, \theta)$ is MLR in $T(x)$, for **testing** $H_0 : \theta \le \theta_0 \ versus \ H_1 : \theta > \theta_0$

$$\varphi^*(x) = \begin{cases} 1, & \text{if } T(x) > k \\ \gamma, & \text{if } T(x) = k \\ 0, & \text{if } T(x) < k \end{cases} \tag{2}$$

- $K :\ upper\ \alpha.100\%\ point\ of\ the\ P_{\theta_0} - distribution\ of\ T(X)$
- $E_\theta \phi^* :\ increasing\ power\ function$

- ϕ^* : $UMP\ \alpha - test$

Note:

- To determine the threshold k, we need to solve the equation
  $E_{\theta_0}\varphi^* = P(T(x) > k \mid \theta = \theta_0) = \alpha$
- Power function:
  $Power(t) = E_\theta\varphi^* = P(reject\ H_0 \mid \theta_1\ is\ true) = P_\theta(T(x) > k = \alpha \mid \theta = t)$
- If we have $H_0 : \theta \geq \theta_0\ versus\ H_1 : \theta < \theta_0$, our test should be constructed as:

$$\varphi^*_{variant}(x) = \begin{cases} 1, & \text{if } T(x) < k \\ \gamma, & \text{if } T(x) = k \\ 0, & \text{if } T(x) > k \end{cases} \tag{3}$$

What if UMP test does not exist?

6.7 unbiased UMP test

Definition:

Test φ of $H_0 : \theta \in \Theta_0(\Theta_0 \in \Theta)$ versus $H_1 : \theta \in \Theta \setminus \Theta_0$ is an unbiased size $\alpha - test$ if:

- $E_\theta\varphi \leq \alpha$ for all $\theta \in \Theta_0$
- $E_\theta\varphi \geq \alpha$ for all $\theta \in \Theta \setminus \Theta_0$

6.9 Locally most powerful test

6.10 Likelihood ratio test

Confusion

1. **In which situation there is no UMP exists?**
2. **What is critical region?**
   Answer:
   the UMPCR is the region that gives the smallest chance of making a Type I or II error
3. **What is simple hypothesis? What is composite?**
   Answer:
   Composite hypotheses have multiple options for solutions. For example, H0: $\sigma > 8$ is a composite hypothesis because it doesn't specify a value for σ; The solution could be anything over 8.

4. **In NP lemma, what does γ stand for?**

   Answer:

   When the likelihood ratio equals to the threshold, we doubt whether reject or accept null hypothesis.

   We make a random decision for the alternative with a probability $\gamma \in (0, 1)$

5. **Can you interpret power function graph? The probability of committing type II error decrease as θ increases. Is θ our observed sample mean?**

6. α and k, what is their difference?

7. Illustration of different type of H_0