Parametric Statistics

Hypothesis Testing - Part 2

Sofia Triantafillou

sof.triantafillou@gmail.com

University of Crete
Department of Mathematics and Applied Mathematics

December 18, 2023

Lecture Summary

9.1 Hypothesis Testing

Recap

Testing Hypotheses

- Hypothesis Space: $\theta \in \Omega_0$ vs $\theta \in \Omega_0$.
- Data space: $\mathbf{X} \in S_0$ vs $\mathbf{X} \in S_1$.
- Statistic space: $T \notin R$ vs $T \in R$, R: rejection region.

Power function: $P(T \in R \mid \theta)$

- Type I error: Falsely reject the null.
- Type II error: Falsely fail to reject the null.

For $\theta \in \Omega_0$: The power function is the probability of α Type I error.

Size/Level of a test

- $H_0 : \mu = 4 \quad \text{vs} \quad H_1 : \mu \neq 4.$
- Reject the null if $|\bar{x}_n - 4| > c$.
- Power function:

$$P\left(T \in R\right) = P\left(\bar{x}_n > c + \mu_0\right) + P\left(x_n < -c + \mu_0\right) =$$
$$= P\left(z > \frac{c + \mu_0 - \mu}{\sigma}\sqrt{n}\right) + P\left(z < \frac{-c + \mu_0 - \mu}{\sigma}\sqrt{n}\right)$$

For $\mu \in \Omega_0$:

$$\sup_{\pi \in \Omega_0} \pi(\mu \mid \delta) = \pi\left(\mu_0 \mid \sigma\right) = P\left(z > \frac{c\sqrt{n}}{\sigma}\right) + P\left(z < \frac{-c}{\sigma}\sqrt{n}\right) =$$
$$= P\left(|z| > \frac{c}{\sigma}\sqrt{n}\right)$$

- Large $c \rightarrow$ reject very few nulls $\rightarrow$ Many Type II errors
- Small $c \rightarrow$ reject many nulls $\rightarrow$ Many Type I errors

Size/Level of a test

Size of a test

$a(\delta) = \sup_{\theta \in \Omega_0}(\theta \mid \delta)$

Level of a test

If $\sup_{\theta \in \Omega_0}(\theta \mid \delta) \leq a_0$ then the test has a level of significance a_0.

If the size of the test is at most α_0, then δ is an α_0-level test

Example

Find the size and level of a test for testing the mean of a normal distribution with known variance

- $H_0 : \mu = 4 \quad \text{vs} \quad H_1 : \mu \neq 4.$
- $\sigma = 3, n = 26, c = 1$
- $\sigma = 3, n = 26, c = 3$

Another Example

- $X_1 \ldots X_n \sim$ Uniform $[0, \theta]$.
- $H_0 : \theta \in [3, 4]$ vs $H_1 : \theta < 3$ or $\theta > 4$.

Another Example

- $X_1 \ldots X_n \sim$ Uniform $[0, \theta]$.
- $H_0 : \theta \in [3, 4]$ vs $H_1 : \theta < 3$ or $\theta > 4$.
- Statistic: $\max\{X_1 \ldots X_n\}$: If $\max\{X_1 \ldots X_n\} < 2.9$ or $y \geqslant 4$ reject the null hypothesis

Another Example

- $X_1 \dots X_n \sim$ Uniform $[0, \theta]$.
- $H_0 : \theta \in [3, 4]$ vs $H_1 : \theta < 3$ or $\theta > 4$.
- Statistic: $\max\{X_1 \dots X_n\}$: If $\max\{X_1 \dots X_n\} < 2.9$ or $y \geqslant 4$ reject the null hypothesis

$$P(T \in R \mid \theta) = P(y \leq 2.9 \mid \theta) + P(y \geqslant 2 \mid \theta)$$

$$\pi(\theta/\delta) = \begin{cases} 1, & \theta \in (0, 2.9] \\ (2.9/\theta)^n, & \theta \in (2.9, 4) \\ (2.9/\theta)^n + (4/\theta)^n, & \theta \in [4, \infty) \end{cases}$$

Level of the test:

$$sup_{\theta \in \Omega_0} \pi(\theta|\delta) = \pi(3) = (2.9/3)^n$$

Selecting a test

- If $\theta \in \Omega_0 : \pi(\theta \mid \delta) =$ probability of type I error.
- If $\theta \in \Omega_1 : 1 - \pi(\theta \mid \delta) =$ probability of type II error.

Strategy for picking δ

Pick the most powerful test that has at most size α_0.

Example

- $H_0 : \mu = 4 \quad \text{vs} \quad H_1 : \mu \neq 4.$
- Reject the null if $|\bar{x}_n - 4| > c$.
- Find c if you want the most powerful test with level 0.05.

Selecting a test

- If $\theta \in \Omega_0 : \pi(\theta \mid \delta) =$ probability of type I error.
- If $\theta \in \Omega_1 : 1 - \pi(\theta \mid \delta) =$ probability of type II error.

Strategy for picking δ

Pick the most powerful test that has at most size α_0.

Example

- $H_0 : \mu = 4 \quad \text{vs} \quad H_1 : \mu \neq 4.$
- Reject the null if $|\bar{x}_n - 4| > c$.
- Find c if you want the most powerful test with level 0.05. $c = 1.96\sigma/\sqrt{n}$.
- Equivalently $|Z| = |\frac{\bar{x}_n - \mu_0}{\sigma}\sqrt{n}| > 1.96$

The p-value

- If $|Z| = 1.97$ you reject the null.
- If $|Z| = 2.78$, you reject the null.
- These two are not exactly the same.

p-value

The smallest significance level at which the null hypothesis would be rejected is called the $p - value$.

- Alternatively, the p-value is the probability of seeing data at least as unfavorable to H_0 as ours, if the null hypothesis is true.
- Strength of the Neyman-Pearson paradigm: You only need to consider the distribution of the statistic under the null.

Recap

- The power function gives us the probability of rejecting the null hypothesis for every parameter value.
- The size of the test is the supremum of the power function for the null values of the parameter.
- Strategy for selecting a test: Pick the most powerful test that has the a size at most α_0 (typically 0.05 or close).
- You can construct a test using the likelihood ratio.