

Topic 6: Hypothesis Testing

Ethan P. Marzban University of California, Santa Barbara PSTAT 120B



Outline

1. Power of a Test

Power of a Test



Power

- Recall that α (the significance level) denotes the probability of committing a Type I error, and β denotes the probability of committing a Type II error.
- We can analogously define a quantity that represents the probability that a given test will lead to rejection of the null:



Power

Definition (Power)

Suppose that W is the test statistic and \mathcal{R} is the rejection region for a test of a hypothesis involving the value of a parameter θ . Then the power of the test, denoted by $\text{power}(\theta)$, is the probability that the test will lead to rejection of H_0 when the actual parameter value is θ . That is,

$$\text{power}(\theta) = \mathbb{P}(W \in \mathcal{R} \text{ when the parameter value is } \theta)$$



Power

Theorem (Relationship between Power and β)

If θ_A is a value of θ in the alternative hypothesis H_A , then

$$\text{power}(\theta_A) = 1 - \beta(\theta_A)$$

where $\beta(\theta_A)$ denotes the probability of committing a Type II error when the true value of θ is θ_A .



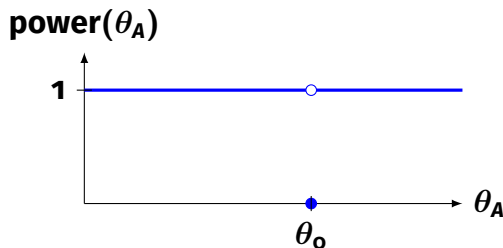
Power

- As the notation suggests, we typically view power as a function of the true value of θ_A .
- Plotting the power of a given test at a series of specified values in the alternative space yields a so-called **power curve**.
- Let's think through what the “ideal” power curve looks like.
- What would we like $\text{power}(\theta_o)$ to be?
- Well, since $\text{power}(\theta_A)$ is, by definition and for any point θ_A , the probability of rejecting $H_o : \theta = \theta_o$ when the true value of θ is θ_A , we'd like $\text{power}(\theta_o) = 0$.



Power

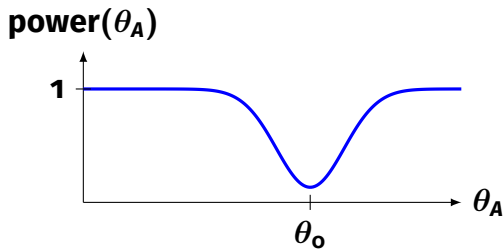
- Similarly, for any $\theta_A \neq \theta_o$, we'd like $\text{power}(\theta_A) = 1$.
- So, the ideal power curve for a test would look like





Power

- Now, keep in mind that all tests are performed at a fixed α level of significance.
- As we discussed before, it's impossible to simultaneously minimize α and β - hence, it's impossible to get a power of exactly zero.
- A more realistic power curve for a test of $H_0 : \theta = \theta_0$ vs $H_A : \theta \neq \theta_0$ might look like





Example

Example

Let $Y_1, \dots, Y_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, 1)$ for some unknown $\mu \in \mathbb{R}$, and suppose we wish to conduct a test of $H_0 : \mu = \mu_0$ vs $H_A : \mu > \mu_0$ at an $\alpha = 0.05$ level of significance. We propose two tests:

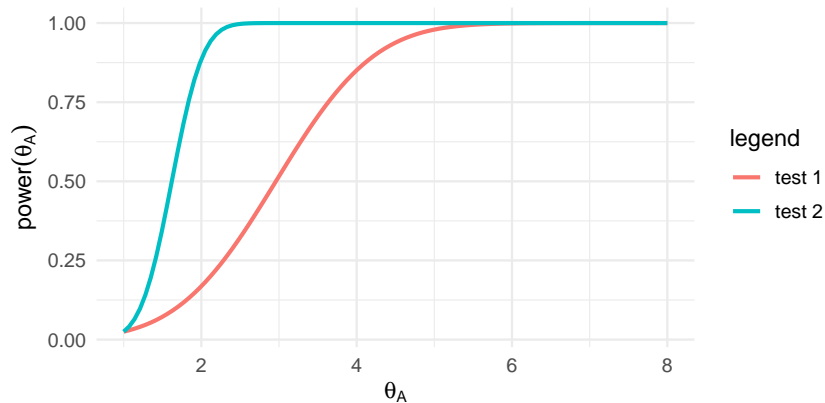
Test 1: Reject H_0 when $Y_1 - \mu_0 > \Phi^{-1}(0.975)$

Test 2: Reject H_0 when $\frac{\bar{Y}_n - \mu_0}{1/\sqrt{n}} > \Phi^{-1}(0.975)$

Derive expressions for the power functions for these two tests, and use this to determine if one test outperforms the other in terms of power for *all* values of θ in the alternative.

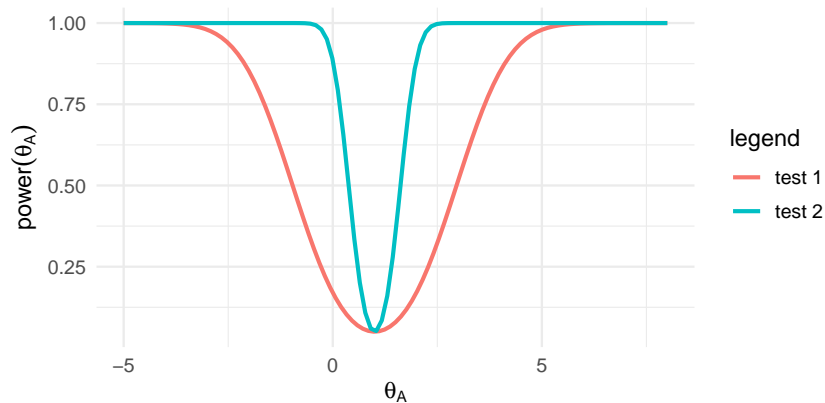


Power





Power





Power

- Since we want the power of our test to be 1 nearly everywhere, we often seek **uniformly most powerful tests**.
- In general, finding such tests is very challenging (and, indeed, such tests don't always exist).
- However, if we restrict ourselves to a *simple-vs-simple* test, we actually *can* construct a most powerful test at a level α , using what is known as the **Neyman-Pearson Lemma**.



Neyman-Pearson Lemma

- Since we want the power of our test to be 1 nearly everywhere, we often seek **uniformly most powerful tests**.
- In general, finding such tests is very challenging (and, indeed, such tests don't always exist).
- However, if we restrict ourselves to a *simple-vs-simple* test, we actually *can* construct a most powerful test at a level α , using what is known as the **Neyman-Pearson Lemma**.



Neyman-Pearson Lemma

THEOREM 10.1

The Neyman–Pearson Lemma Suppose that we wish to test the simple null hypothesis $H_0 : \theta = \theta_0$ versus the simple alternative hypothesis $H_a : \theta = \theta_a$, based on a random sample Y_1, Y_2, \dots, Y_n from a distribution with parameter θ . Let $L(\theta)$ denote the likelihood of the sample when the value of the parameter is θ . Then, for a given α , the test that maximizes the power at θ_a has a rejection region, RR , determined by

$$\frac{L(\theta_0)}{L(\theta_a)} < k.$$

The value of k is chosen so that the test has the desired value for α . Such a test is a most powerful α -level test for H_0 versus H_a .



Likelihood Ratio Test

Definition (Likelihood Ratio Test)

Consider hypotheses $H_0 : \Theta \in \Omega_0$ and $H_A : \Theta \in \Omega_A$. Define

$$\Lambda := \frac{\mathcal{L}(\hat{\Omega}_0)}{\mathcal{L}(\hat{\Omega})} = \frac{\max_{\Theta \in \Omega_0} \mathcal{L}_{\vec{Y}}(\Theta)}{\max_{\Theta \in \Omega_0 \cup \Omega_A} \mathcal{L}_{\vec{Y}}(\Theta)}$$

A **likelihood ratio test** (named as such because we call Λ a **likelihood ratio**) rejects H_0 whenever $\{\Lambda < k\}$.



Example

Example

Let $Y_1, \dots, Y_n \stackrel{\text{i.i.d.}}{\sim} \text{Exp}(\theta)$. Construct the likelihood ratio test for $H_0 : \theta = \theta_0$ vs $H_A : \theta \neq \theta_0$, using an α level of significance. You do not need to explicitly solve for constants; just derive the general form for the LRT.