

Lecture 15 – Most powerful tests and Neyman-Pearson lemma

Last Modified: October 21, 2026

Outline:

15.1 Optimal tests	1
15.2 Examples using Neyman-Pearson	3

15.1 Optimal tests

Recall that a test $\psi(\underline{x})$ has level α if

$$\sup_{\theta \in \Theta_0} \beta_{\psi}(\theta) \leq \alpha$$

where $\beta_{\psi}(\theta)$ is the power function of ψ . We used the subscript ψ to emphasize that the power function depends on the test being used. Recall that $\sup = \max$ when taking max over an infinite set.

Definition 1 (Uniformly most powerful (UMP) test). Consider the class of all the tests of level α for

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta \in \Theta_1.$$

A test ψ^* in this class is called **uniformly most powerful** (UMP) of level α if for all ψ in the class,

$$\beta_{\psi^*}(\theta) \geq \beta_{\psi}(\theta), \quad \forall \theta \in \Theta_1.$$

Remark 1. We restrict to the tests with level α (control of Type I errors) and we select the one with the highest power when $\theta \in \Theta_1$ (lowest Type II error).

Remark 2. If H_1 is *simple*, i.e., $\Theta_1 = \{\theta_1\}$, the definition becomes

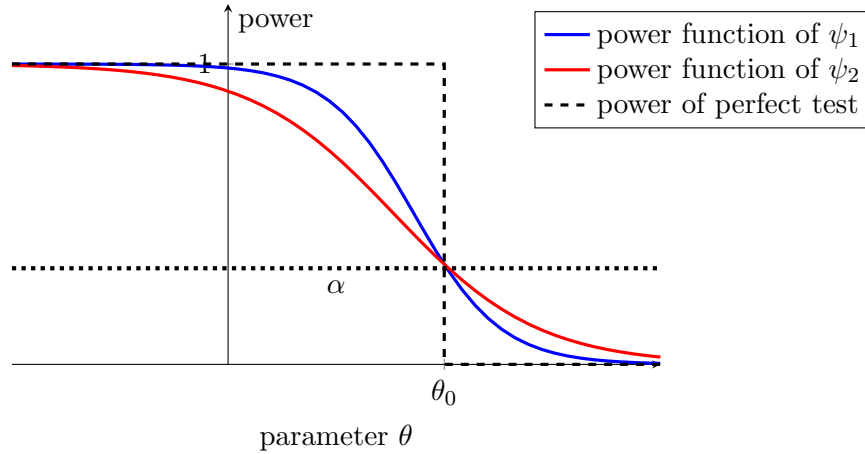
$$\beta_{\psi^*}(\theta_1) \geq \beta_{\psi}(\theta_1), \quad \forall \text{level-}\alpha \psi.$$

and ψ is called the **most powerful** (MP) test (rather than uniformly most powerful, since we only have to compare the power at the one point θ_1).

Example 1. Consider the test for

$$H_0 : \theta \geq \theta_0 \quad \text{versus} \quad H_1 : \theta < \theta_0.$$

In the figure below, the test ψ_1 is uniformly more powerful than ψ_2 , since it always has a higher probability of rejecting when $\theta \in \Theta_1$.



Theorem 1 (Neyman-Pearson Lemma). Consider a statistical model $\{f_\theta : \theta \in \Theta\}$ and a testing problem with simple hypotheses $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$. If a test ψ^* is of the form

$$\psi^*(\underline{x}) = \mathbb{1}_{\{L_{\underline{x}}(\theta_1) > k L_{\underline{x}}(\theta_0)\}}$$

for some $k \geq 0$, and

$$\mathbb{P}_{\theta_0}(\psi^*(\underline{X}) = 1) = \alpha,$$

then ψ^* is the level- α most powerful test.

Remark 3. The most powerful test rejects H_0 if the likelihood ratio $L_{\underline{x}}(\theta_1)/L_{\underline{x}}(\theta_0)$ is large.

Proof in discrete case. Suppose the space $\mathcal{X}^n$ is discrete. Let the pmfs for the two hypotheses be denoted by

$$f_0(\underline{x}) := f_{\underline{X}, \theta_0}(\underline{x}) = L_{\underline{x}}(\theta_0), \quad f_1(\underline{x}) := f_{\underline{X}, \theta_1}(\underline{x}) = L_{\underline{x}}(\theta_1).$$

Define the function

$$g(\underline{x}) = f_1(\underline{x}) - k f_0(\underline{x}).$$

Let $\mathcal{R}_*$ be the rejection region for ψ^* . By definition, $\mathcal{R}_*$ is precisely all points $\underline{x}$ in the sample space for which $g(\underline{x}) \geq 0$. Next, let ψ be any other level α test, and suppose it has rejection region $\mathcal{R}$. We claim that

$$\sum_{\underline{x} \in \mathcal{R}_*} g(\underline{x}) \geq \sum_{\underline{x} \in \mathcal{R}} g(\underline{x}). \quad (15.1)$$

The proof is by Figure 15.1. The sum of $g(\underline{x})$ over $\underline{x} \in \mathcal{R}_*$ is a sum of all +’s. The sum of $g(\underline{x})$ over $\underline{x} \in \mathcal{R}$ includes some of these +’s, as well as some −’s. Therefore the first sum is bigger. Now let’s work out what (15.1) gives us. Note that

$$\sum_{\underline{x} \in \mathcal{R}_*} g(\underline{x}) = \sum_{\underline{x} \in \mathcal{R}_*} f_1(\underline{x}) - k \sum_{\underline{x} \in \mathcal{R}_*} f_0(\underline{x}) = \mathbb{P}_{\theta_1}(\underline{X} \in \mathcal{R}_*) - k \mathbb{P}_{\theta_0}(\underline{X} \in \mathcal{R}_*)$$

Similarly,

$$\sum_{\underline{x} \in \mathcal{R}} g(\underline{x}) = \mathbb{P}_{\theta_1}(\underline{X} \in \mathcal{R}) - k \mathbb{P}_{\theta_0}(\underline{X} \in \mathcal{R}).$$

This allows us to rewrite (15.1) as follows:

$$\underbrace{\mathbb{P}_{\theta_1}(\underline{X} \in \mathcal{R}_*)}_{\beta_{\psi^*}(\theta_1)} - \underbrace{\mathbb{P}_{\theta_1}(\underline{X} \in \mathcal{R})}_{\beta_{\psi}(\theta_1)} \geq k \left(\underbrace{\mathbb{P}_{\theta_0}(\underline{X} \in \mathcal{R}_*)}_{=\alpha} - \underbrace{\mathbb{P}_{\theta_0}(\underline{X} \in \mathcal{R})}_{\leq \alpha} \right) \geq 0.$$

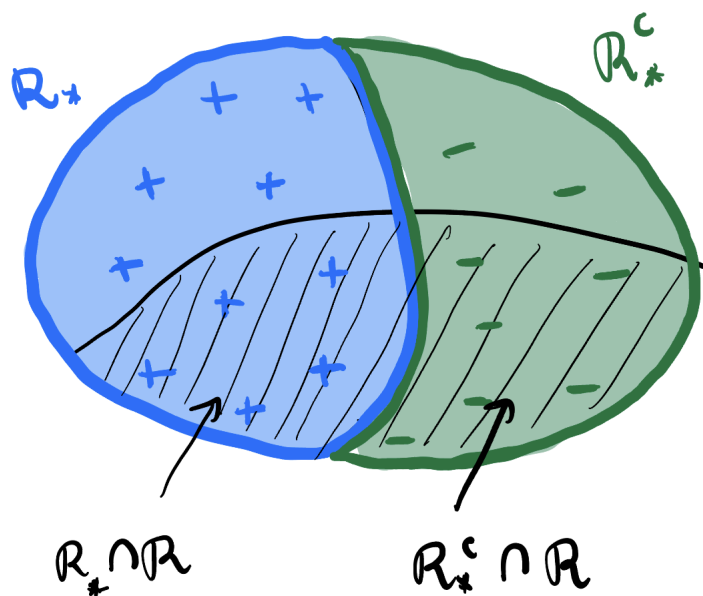


Figure 15.1: Diagram for Neyman-Pearson proof

Therefore,

$$\beta_{\psi^*}(\theta_1) = \mathbb{P}_{\theta_1}(\underline{X} \in \mathcal{R}^*) \geq \mathbb{P}_{\theta_1}(\underline{X} \in \mathcal{R}) = \beta_{\psi}(\theta_1).$$

□

15.2 Examples using Neyman-Pearson

Recall the Neyman–Fisher factorization theorem from Lectures 3–4: if $T(\underline{X})$ is sufficient for θ , then

$$L_{\underline{x}}(\theta) = g(T(\underline{x}), \theta)h(\underline{x}),$$

where h does not depend on θ . Consequently, wherever the denominator is positive, the likelihood ratio simplifies to

$$\frac{L_{\underline{x}}(\theta_1)}{L_{\underline{x}}(\theta_0)} = \frac{g(T(\underline{x}), \theta_1)h(\underline{x})}{g(T(\underline{x}), \theta_0)h(\underline{x})} = \frac{g(T(\underline{x}), \theta_1)}{g(T(\underline{x}), \theta_0)}.$$

Thus the likelihood ratio depends on the data only through the sufficient statistic. Equivalently, writing the Neyman–Pearson rejection rule without division, we can take

$$\mathcal{R}_* = \{\underline{x} : g(T(\underline{x}), \theta_1) > k g(T(\underline{x}), \theta_0)\}.$$

This rule depends only on the value of $T(\underline{x})$: two samples with the same sufficient statistic lead to the same decision. Points where $h(\underline{x}) = 0$ have probability zero under both hypotheses and do not affect the size or power. In the examples below, this reduces the likelihood-ratio calculation to a function of the sample mean or sum.

Example 2 (Gaussian mean with known variance). Suppose $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ with $\sigma^2 > 0$ known. Find the most powerful level α test for

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu = \mu_1,$$

where $\mu_1 < \mu_0$ and $0 < \alpha < 1$.

The likelihood function is

$$L_{\underline{x}}(\mu) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right\}.$$

By the Neyman-Pearson lemma, we should reject when $L_{\underline{x}}(\mu_1)/L_{\underline{x}}(\mu_0) > k$, choosing k so that the test has size α . The likelihood ratio is

$$\begin{aligned} \frac{L_{\underline{x}}(\mu_1)}{L_{\underline{x}}(\mu_0)} &= \frac{\exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_1)^2\}}{\exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu_0)^2\}} \\ &= \exp\left\{\frac{\mu_1 - \mu_0}{\sigma^2} \sum_{i=1}^n x_i + \frac{n(\mu_0^2 - \mu_1^2)}{2\sigma^2}\right\}. \end{aligned}$$

Since $\mu_1 - \mu_0 < 0$, this ratio is decreasing in $\bar{x}_n$. Thus the rejection region has the form $\mathcal{R} = \{\underline{x} : \bar{x}_n < c\}$.

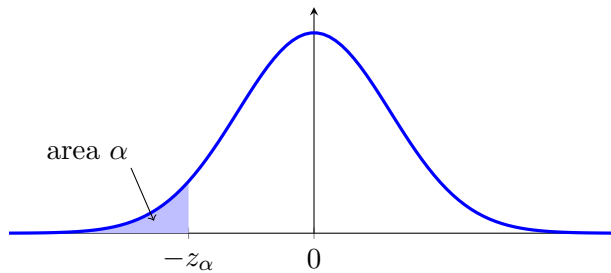
To choose c , note that under H_0 ,

$$\bar{X}_n \sim \mathcal{N}\left(\mu_0, \frac{\sigma^2}{n}\right), \quad Z = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu_0) \sim \mathcal{N}(0, 1).$$

Recall the *upper-tail* convention from Lecture 14:

$$z_\alpha = \Phi^{-1}(1 - \alpha), \quad \mathbb{P}(Z > z_\alpha) = \alpha.$$

By symmetry, $\mathbb{P}(Z < -z_\alpha) = \alpha$.



Therefore, we find c to ensure

$$\alpha = \mathbb{P}_{\mu_0}(\bar{X}_n < c) = \mathbb{P}\left(Z < \frac{\sqrt{n}}{\sigma}(c - \mu_0)\right).$$

So we should set

$$\frac{\sqrt{n}}{\sigma}(c - \mu_0) = -z_\alpha \implies c = \mu_0 - \frac{\sigma}{\sqrt{n}}z_\alpha.$$

The most powerful level α test consequently rejects when

$$\boxed{\bar{X}_n < \mu_0 - \frac{\sigma}{\sqrt{n}}z_\alpha.}$$

In other words, $\bar{X}_n$ being sufficiently smaller than μ_0 provides evidence for $\mu_1 < \mu_0$. As α decreases, $-z_\alpha$ becomes more negative and the cutoff decreases, making rejection more difficult. To give an example, $z_{0.05} \approx 1.645$, so the level-0.05 MP test has cutoff $\mu_0 - 1.645\sigma/\sqrt{n}$.

Example 3 (Exponential rate). **Note: we didn't get to this in class - please read on your own!**

Suppose $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Exp}(\theta)$, where $\theta > 0$ is the rate. Find the most powerful level α test for

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta = \theta_1,$$

where $\theta_1 > \theta_0$ and $0 < \alpha < 1$.

The likelihood function is

$$L_{\underline{x}}(\theta) = \theta^n \exp \left\{ -\theta \sum_{i=1}^n x_i \right\} \mathbb{1}_{\{x_1 > 0, \dots, x_n > 0\}}.$$

On the common support $(0, \infty)^n$, the likelihood ratio is

$$\frac{L_{\underline{x}}(\theta_1)}{L_{\underline{x}}(\theta_0)} = \left(\frac{\theta_1}{\theta_0} \right)^n \exp \left\{ -(\theta_1 - \theta_0) \sum_{i=1}^n x_i \right\}.$$

Since $\theta_1 > \theta_0$, this ratio is decreasing in $\sum_{i=1}^n x_i$. By the Neyman-Pearson lemma, the most powerful test therefore rejects when $\sum_{i=1}^n X_i < c$, with c chosen to give size α . This makes sense because θ_1 being larger than θ_0 means the mean of $\text{Exp}(\theta_1)$ is *smaller* than the mean of $\text{Exp}(\theta_0)$.

Let us now choose c . Under H_0 , we have that $\theta_0 X_i \stackrel{\text{iid}}{\sim} \text{Exp}(1)$, so

$$\theta_0 \sum_{i=1}^n X_i \sim \text{Gamma}(n, 1).$$

Let $F_{n,1}$ be the cdf of $\text{Gamma}(n, 1)$ and let $q_\alpha = F_{n,1}^{-1}(\alpha)$ be its *lower-tail* quantile: $\mathbb{P}(Y < q_\alpha) = \alpha$ for $Y \sim \text{Gamma}(n, 1)$. Then we should choose c so that

$$\alpha = \mathbb{P}_{\theta_0} \left(\sum_{i=1}^n X_i < c \right) = \mathbb{P}_{\theta_0} \left(\theta_0 \sum_{i=1}^n X_i < \theta_0 c \right) = \mathbb{P}(\text{Gamma}(n, 1) \leq \theta_0 c).$$

Thus we should set $\theta_0 c = q_\alpha$, or $c = q_\alpha / \theta_0$. We conclude that the most powerful level α test rejects when

$$\boxed{\sum_{i=1}^n X_i < \frac{q_\alpha}{\theta_0}}.$$