18.650: Fundamentals of Statistics October 18, 2023

Lecture 15 — Hypothesis testing continued

Prof. Philippe Rigollet *Scribe: Anya Katsevich*

In this lecture we talk about how to

1. construct a test Ψ
2. compute the power function.

Recap of Lecture 14: Let Θ be the full parameter space, and let Θ_0, Θ_1 split Θ into two disjoint subsets. A hypothesis test takes the form

$$\begin{aligned} H_0 &: \theta \in \Theta_0 \quad \text{(null)} \\ &\text{vs} \\ H_1 &: \theta \in \Theta_1 \quad \text{(alternative).} \end{aligned}$$

Recall the size of a test Ψ is

$$\text{size}(\Psi) = \max_{\theta \in \Theta_0} \mathbb{P}_\theta(\Psi = 1).$$

The test Ψ is said to have *level* α (a number between 0 and 1) if $\text{size}(\Psi) \leq \alpha$.
We defined the *power* function as

$$\beta(\theta) = \mathbb{P}_\theta(\Psi = 1).$$

Remark.

If Ψ has level α then for all $\theta \in \Theta_0$, we have $\beta(\theta) \leq \alpha$. This is true by definition!

1 Constructing a test and computing the power

Recall our ER example: $X_1, \ldots, X_n$ i.i.d., $E[X_1] = \mu$, where X_i is the waiting time of a random patient. We test

$$H_0 : \mu \leq 30 \quad \text{vs} \quad H_1 : \mu > 30.$$

(By convention, the alternative always gets the strict inequality.) To estimate μ, we use $\hat{\mu} = \bar{X}_n$. For our test, we take

$$\Psi = \mathbb{1}\{\bar{X}_n - 30 > c_\alpha\}.$$

It remains to choose c_α to ensure the size of the test is at most α. Maximizing our budget for type I error, we find c_α so that

$$\max_{\mu \le 30} \mathbb{P}_\mu(\Psi = 1) = \alpha \tag{1}$$

exactly. Now, recall the maximum is achieved at $\mu = 30$ (the boundary), so we need to choose c_α so that

$$\mathbb{P}_{\mu=30}(\Psi = 1) = \mathbb{P}_{\mu=30}(\bar{X}_n - 30 > c_\alpha) = \alpha.$$

By the CLT, we know that if $\mu = 30$ then $\sqrt{n}(\bar{X}_n - 30) \rightsquigarrow \mathcal{N}(0, \sigma^2)$. Therefore,

$$\bar{X}_n - 30 \approx \mathcal{N}(0, \sigma^2/n).$$

Therefore, if we let Z denote a standard normal random variable, we have

$$\begin{aligned}
\mathbb{P}_{\mu=30}(\bar{X}_n - 30 > c_\alpha) &\approx \mathbb{P}_{\mu=30}\left(\frac{\sigma}{\sqrt{n}} Z > c_\alpha\right) \\
&= \mathbb{P}_{\mu=30}\left(Z > c_\alpha \frac{\sqrt{n}}{\sigma}\right) = 1 - \Phi\left(c_\alpha \frac{\sqrt{n}}{\sigma}\right)
\end{aligned}$$

Setting this last quantity equal to α, we get

$$c_\alpha = \frac{\sigma}{\sqrt{n}} \Phi^{-1}(1 - \alpha) = \frac{\sigma}{\sqrt{n}} z_\alpha. \tag{2}$$

If e.g. $\alpha = 0.05$ then $z_\alpha = z_{0.05} = 1.65$, and we take

$$c_\alpha = 1.65 \frac{\sigma}{\sqrt{n}} \approx 1.65 \frac{\hat{\sigma}}{\sqrt{n}}.$$

(As usual, we don't know σ so we replace it by $\hat{\sigma}$.)

Suppose $n = 164$, $\hat{\sigma} = 12$, and $\bar{X}_n = 33.4$. We compute $c_\alpha = 1.65 \cdot 12/\sqrt{164} = 1.54$. In other words, we reject the null if $\bar{X}_n - 30 \ge 1.54$. But indeed, we have observed $\bar{X}_n = 33.4$ and $33.4 - 30 = 3.4 \ge 1.54$. Therefore, we $\boxed{\text{REJECT } H_0}$. This means we found enough evidence to prove that $\mu > 30$.

1.1 Computing the power function

For our test of level α, we have

$$\beta(\mu) = \mathbb{P}_\mu(\bar{X}_n - 30 \ge c_\alpha).$$

We now use that when the true mean is μ, we have $\bar{X}_n \approx \mathcal{N}(\mu, \hat{\sigma}^2/n)$. Therefore,

$$\begin{aligned}\beta(\mu) &\approx \mathbb{P}\left(\mu + \frac{\hat{\sigma}}{\sqrt{n}} Z \geq 30 + c_\alpha\right) \\ &= \mathbb{P}\left(Z \geq \frac{\sqrt{n}}{\hat{\sigma}}(30 + c_\alpha - \mu)\right) = 1 - \Phi\left(\frac{\sqrt{n}}{\hat{\sigma}}(30 + c_\alpha - \mu)\right)\end{aligned} \tag{3}$$

In particular, using the numbers above we get

$$\beta(\mu) \approx 1 - \Phi\left(\frac{\sqrt{164}}{12}(31.54 - \mu)\right)$$

Note from (3) that $\beta(30) = \alpha$. Indeed, $\beta(30) \approx 1 - \Phi\left(\frac{\sqrt{n}}{\sigma} c_\alpha\right) = \alpha$, using the formula (2) for c_α. This should come as no surprise — indeed, we *chose* c_α *to ensure* $\beta(30) = \alpha$*!*.

Remark.

> Recall from Figure 2 in the Lecture 14 notes that in the ideal case, the power function transitions sharply from being zero for $\theta \in \Theta_0$ to being one for $\theta \in \Theta_1$. This isn't possible in reality, but the larger the slope of the power function at the transition point, the better. Let's compute $\beta'(\mu)$ in our example. Differentiating the last expression in (3) (and using the chain rule), we get
>
> $$\beta'(\mu) \approx \frac{\sqrt{n}}{\hat{\sigma}} \phi\left(30 + c_\alpha - \mu\right),$$
>
> where ϕ is the standard Gaussian pdf. At $\mu = 30$, we get
>
> $$\beta'(30) \approx \frac{\sqrt{n}}{\hat{\sigma}} \phi\left(c_\alpha\right).$$
>
> Note that if n increases or $\hat{\sigma}$ decreases, the slope at 30 increases.

1.2 Wald Test

The Wald test is based on asymptotic normality of $\hat{\theta}$. Consider any situation where

$$\frac{\hat{\theta} - \theta}{\widehat{\mathrm{se}}} \rightsquigarrow \mathcal{N}(0, 1), \quad n \to \infty. \tag{4}$$

Definition 1.1: Wald Test

Let $\hat{\theta}$ be asymptotically normal, i.e. suppose $\hat{\theta}$ satisfies (4). Then the Wald test for the hypothesis $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ is

$$\Psi = \mathbb{1}\left(\left|\frac{\hat{\theta} - \theta_0}{\hat{se}}\right| \geq z_{\alpha/2}\right),$$

i.e. reject H_0 if $\left|(\hat{\theta} - \theta_0)/\hat{se}\right| \geq z_{\alpha/2}$.
The Wald test for the hypothesis $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$ is

$$\Psi = \mathbb{1}\left(\frac{\hat{\theta} - \theta_0}{\hat{se}} \geq z_{\alpha}\right),$$

i.e. reject H_0 if $(\hat{\theta} - \theta_0)/\hat{se} \geq z_{\alpha}$.
The Wald test for the hypothesis $H_0 : \theta \geq \theta_0$ vs $H_1 : \theta < \theta_0$ is

$$\Psi = \mathbb{1}\left(\frac{\hat{\theta} - \theta_0}{\hat{se}} \leq -z_{\alpha}\right),$$

i.e. reject H_0 if $(\hat{\theta} - \theta_0)/\hat{se} \leq -z_{\alpha}$

Example.

> Recall that for the MLE, we have $\hat{\theta}^{\text{MLE}} \approx \mathcal{N}(\theta, \frac{1}{nI(\theta)})$. Therefore, $\hat{\text{se}} = 1/\sqrt{nI(\hat{\theta}^{\text{MLE}})}$. Then Wald's test for $H_0 : \theta = 0$ vs $H_1 : \theta \neq 0$ is given by
>
> $$\Psi = \mathbb{1}\left(\sqrt{nI(\hat{\theta}^{\text{MLE}})}|\hat{\theta}^{\text{MLE}} - 0| > z_{\alpha/2}\right).$$
>
> It is also valid to use $I(0)$ instead of $I(\hat{\theta}^{\text{MLE}})$, i.e. we can also use the test
>
> $$\Psi = \mathbb{1}\left(\sqrt{nI(0)}|\hat{\theta}^{\text{MLE}} - 0| > z_{\alpha/2}\right).$$

Example.

> Suppose we want to test $H_0 : \theta \leq 2$ vs $H_1 : \theta > 2$. Then
>
> $$\Psi = \mathbb{1}\left(\frac{\hat{\theta}^{\text{MLE}} - 2}{\hat{\text{se}}} > z_\alpha\right)$$
>
> Suppose we want to test $H_0 : \theta \geq 35$ vs $H_1 : \theta < 35$. Then
>
> $$\Psi = \mathbb{1}\left(\frac{\hat{\theta}^{\text{MLE}} - 35}{\hat{\text{se}}} < -z_\alpha\right)$$

Exercise. In all of the above examples, check that Wald's test has size α, and compute $\beta(\theta)$.