STATS300A - Lecture 14

Dominik Rothenhaeusler
Scribed by Michael Howes

11/08/21

Contents

1 Recap

We have been doing hypothesis testing of $H_0 : \theta \in \Omega_0$ against $H_1 : \theta \in \Omega_1$ where $\Omega_0 \cap \Omega_1 = \emptyset$ and $\Omega_0 \cup \Omega_1 = \Omega$. Our goal has been to find uniformly most powerful (UMP) tests that have level α. That is we wise to find tests ϕ that maximize

$$\beta(\theta) = \mathbb{E}_\theta \phi, \quad \text{for all } \theta \in \Omega_1,$$

subject to

$$\mathbb{E}_\theta \phi \leq \alpha, \quad \text{for all } \theta \in \Omega_0.$$

We have seen several stratergies for certain special cases such as

(a) For simple against simple we can use Neyman Pearson to construct most powerful (MP) tests via likelihood ratios.

(b) For a simple null against a composite alternative we have the stratergy:

- Fix a $\theta_1 \in \Omega_1$ and use Neyman Pearson to construct a MP test.
- If this MP test does not depend on θ_1, then it is a UMP test.

(c) For the null $H_0 : \theta \leq \theta_0$ against the alternative $H_1 : \theta > \theta_0$ we have a result for the special cases of one-dimensional exponential families and monotone likelihood ratio families.

There are several things that we would like to test that we haven't yet. Our "to-do" list is

(a) Two sided tests $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$.

(b) Testing with nuisance parameters: $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ where there are additional unknown parameters such as the variance σ^2.

Today we will use the following stratergy for testing a composite null against a composite alternative:

- Fix a simple alternative within the full alternative.
- Put a prior on Ω_1 to collapse the null to a simple hypothesis.
- Use Neyman Pearson to find an MP test.
- Argue that the NP test is optimal for the full null and full alternative.

2 Collapsing the null

Consider testing a composite null against a simple alternative. That is, we are testing

$$H_0 : X \sim f_\theta, \theta \in \Omega_1 \text{ against } H_1 : X \sim g.$$

If we let Λ be a probability distribution on Ω_0 and then we can consider the collapsed null of testing the marginal of Λ against g. That is, we introduce a new null

$$H_\Lambda : X \sim f_\Lambda = \int_{\Omega_0} f_\theta(x) d\Lambda(\theta).$$

Testing H_Λ against H_1 is a simple null against a simple alternative.

Definition 1. Let β_Λ be the power of a MP test at level α of H_Λ against H_1.

Definition 2. A probability distribution Λ is called *least favourable* if β_Λ is minimized. That is, for all other probaility distributions Λ', $\beta_\Lambda \leq \beta_{\Lambda'}$.

Theorem 1 (TSH 3.8.1)**.** *Suppose ϕ_Λ is a MP level α test for testing H_Λ against H_1. If ϕ_Λ is level α for the original null hypothesis H_0, then*

(a) The test ϕ_Λ is MP for H_0 against H_1.

(b) The distribution Λ is least favourable.

Proof. Both statements can be proved using

$$\mathbb{E}_{\Lambda'}[\phi] \leq \sup_{\theta \in \Omega_0} \mathbb{E}_\theta[\phi],$$

where Λ is any probability distribution on Ω_0 and ϕ is any test function. This holds because

$$\begin{aligned} \mathbb{E}_{\Lambda'}[\phi] &= \int_{\Omega_0} \mathbb{E}_\theta[\phi] d\Lambda'(\theta) \\ &\leq \sup_{\theta \in \Omega_0} \mathbb{E}_\theta[\phi]. \end{aligned}$$

We will now show that ϕ_Λ is most powerful. Let ϕ^* be a level α test of H_0 against H_1. Then

$$\mathbb{E}_\Lambda[\phi] \leq \sup_{\alpha \in \Omega_0} \mathbb{E}_\theta[\phi] \leq \alpha.$$

So ϕ^* is a level α test of H_Λ against H_1. Thus

$$\mathbb{E}_1[\phi^*] \leq \mathbb{E}_1[\phi_\Lambda],$$

since ϕ_Λ is MP for H_Λ against H_1. Thus we have proved (a). To see that (b) also holds suppose that Λ' is a probability distribution on Ω_0. Then

$$\mathbb{E}_{\Lambda'}[\phi_\Lambda] \leq \sup_{\theta \in \Omega_0} \mathbb{E}_\theta[\phi_\Lambda] \leq \alpha.$$

Thus ϕ_Λ is a level α test of $H_{\Lambda'}$ against H_1. It follows that

$$\beta_\Lambda = \mathbb{E}_1[\phi_\Lambda] \leq \beta_{\Lambda'},$$

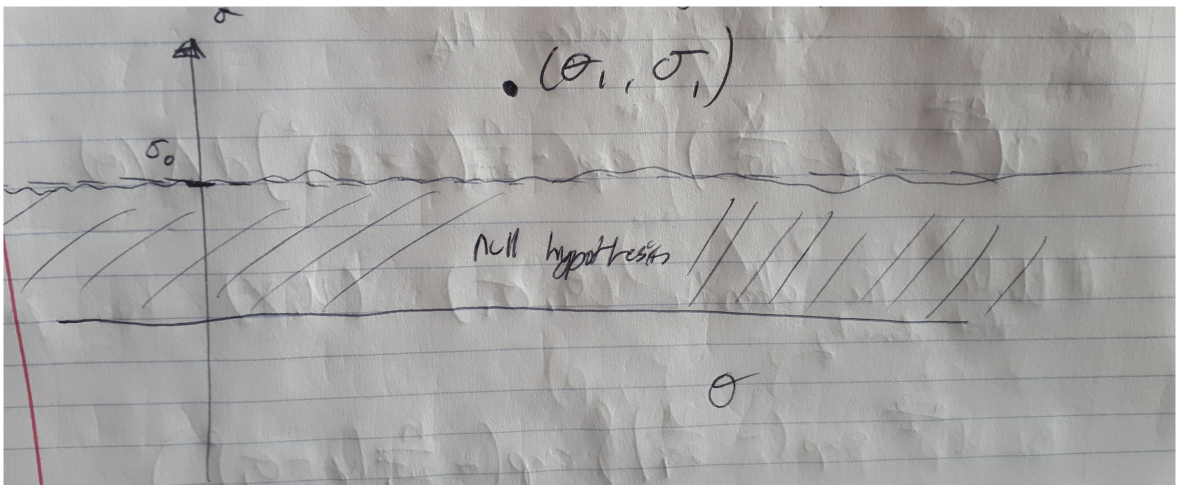
showing that Λ is least favourable. □

Remark 1. Note that the fact that Λ is least favourable is a consequence of the above theorem not an assumption. Thus when testing H_0 against H_1 we know that we are going to have to have a least favourable probability distribution to use the above theorem. This helps us narrow our search space. Inuitively we want a distribution Λ such that the distribution uder H_Λ is as close as possible to the distribution under H_1.

3 Examples

3.1 Testing the variance

Suppose $X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\theta, \sigma^2)$ where both θ and σ are unknown. We wish to test $H_0 : \sigma \leq \sigma_0$ against $H_1 : \sigma > \sigma_0$. Thus we are testing a composite test against a composite test and the parameter θ is nuisance parameter. Our goal is to find the UMP level α test. To start, fix a simple alternative (θ_1, σ_1) where $\sigma_1 > \sigma_0$. Our parameter space looks like this:



We have the idea that testing is hard when $\sigma = \sigma_0$ since the σ is as close as possible to σ_1. Thus we guess that our probability distribution Λ should be supported on

$$\{(\theta, \sigma_0) : \theta \in \mathbb{R}\}.$$

We can further simplify things by working with sufficient statistics. This is because if we are given a test function ϕ and a sufficient statistic T, then we can define a test function η given by

$$\eta(t) = \mathbb{E}[\phi(X)|T = t].$$

The test function η has power and level equal to that of ϕ and η is a function of only the sufficient statistic T. Furthermore η is well-defined by sufficiency.

In our example, a sufficient statistic is

$$(Y, U) = \left(\bar{X}, \sum_{i=1}^{n} (X_i - \bar{X})^2\right).$$

By Basu's theorem we know that Y and U are independent under any choice of (θ, σ). We also know that $Y \sim \mathcal{N}(\theta, \sigma^2/n)$ and $U/\sigma^2 \sim \chi^2_{n-1}$. For a fixed simple null (θ, σ_0) we know that the joint distribution of (U, Y) is given by

$$f(u, y; \theta, \sigma_0) \propto u^{\frac{n-3}{2}} \exp\left\{-\frac{u}{2\sigma_0^2}\right\} \exp\left\{-\frac{n}{2\sigma_0^2}(y - \theta)^2\right\}.$$

Thus if we have a distribution Λ on $\{(\theta, \sigma_0) : \theta \in \mathbb{R}\}$, then the joint distribution of (Y, U) under H_Λ is

$$f(u, y; \theta, \Lambda) \propto u^{\frac{n-3}{2}} \exp\left\{-\frac{u}{2\sigma_0^2}\right\} \int_{-\infty}^{\infty} \exp\left\{-\frac{n}{2\sigma_0^2}(y-\theta)^2\right\} d\Lambda(\theta).$$

Under the simple alternative (θ_1, σ_1), (U, Y) has joint density

$$f(u, y; \theta_1, \sigma_1) \propto u^{\frac{n-3}{2}} \exp\left\{-\frac{u}{2\sigma_1^2}\right\} \exp\left\{-\frac{n}{2\sigma_1^2}(y-\theta_1)^2\right\}.$$

We wish to choose a distribution Λ so that the above distributions are as close as possible. Under the alternative $(\theta, \sigma) = (\theta_1, \sigma_1)$, we have $Y \sim \mathcal{N}(\theta, \sigma_1^2/n)$. Under H_Λ, Y has distribution $Z + \Theta$ where Z and Θ are independent, $\Theta \sim \Lambda$ and $Z \sim \mathcal{N}(0, \sigma_0^2/n)$. This last claim can be seen by observing that the distribution of Y under H_Λ is given by a convolution. If we let $\Lambda = \mathcal{N}(\theta_1, \sigma_1^2/n - \sigma_0^2/n)$, then $Y \sim \mathcal{N}(\theta_1, \sigma_1^2/n)$ under H_Λ. The likelihood ratio thus simplifies to

$$\frac{\exp\left\{-\frac{u}{2\sigma_1^2}\right\}}{\exp\left\{-\frac{u}{2\sigma_0^2}\right\}} = \exp\left\{u\left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}\right)\right\},$$

which is an increasing function of u. Thus the MP test is one which rejects when U is large. Under H_Λ, $U/\sigma_0^2 \sim \chi^2_{n-1}$. Thus the MP level α test is given by

$$\phi(x) = \begin{cases} 1 & \text{if } \frac{1}{\sigma_0^2}\sum_{i=1}^n (x_i - \bar{x})^2 > k, \\ 0 & \text{else.} \end{cases}$$

where k is the $1-\alpha$ quartile of χ^2_{n-1}. We know have to ask if ϕ is MP for H_0 against $(\theta, \sigma) = (\theta_1, \sigma_1)$. To do this we have to show that ϕ has level α at (θ, σ) for all $\sigma \leq \sigma_0$. This is true since if X_i has variance $\sigma^2 < \sigma_0^2$, then

$$\frac{1}{\sigma_0^2}\sum_{i=1}^n (x_i - \bar{x})^2 \leq \frac{1}{\sigma^2}\sum_{i=1}^n (x_i - \bar{x})^2 \sim \chi^2_{n-1}.$$

Thus the probability of rejection decreases as σ decreases. Thus ϕ is MP for H_0 against the simple alternative $(\theta, \sigma) = (\theta_1, \sigma_1)$. Since ϕ does not depend on (θ_1, σ_1) we can conclude that θ is in fact UMP for $H_0 : \sigma \leq \sigma_0$ against $H_1 : \sigma > \sigma_0$.

Remark 2. One can ask: Why doesn't taking $(\theta, \sigma) = (\theta_1, \sigma_0)$ work? One could argue intuitively that this is the distribution "closest" to (θ_1, σ_1). This fails because the MP test of $(\theta, \sigma) = (\theta_1, \sigma_0)$ against $(\theta, \sigma) = (\theta_1, \sigma_1)$ is given by

$$\phi_{\theta_1}(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^n (x_i - \theta)^2 > k, \\ 0 & \text{else.} \end{cases}$$

This test is problematic since it depends on θ_1 but also it is not level α for the null $H_0 : \sigma \leq \sigma_0$. If we let $\theta_1 \to \infty$, the level approaches 1.

3.2 Non-parametric quantile test

Suppose $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathbb{P} \in \mathcal{P}$ where $\mathcal{P}$ is the set of all distributions on $\mathbb{R}$. For fixed $\mu \in \mathbb{R}$ and $p_0 \in [0, 1]$ we wish to test

$$H_0 : \mathbb{P}(X \leq \mu) \geq p_0 \quad \text{against} \quad H_1 : \mathbb{P}(X \leq \mu) < p_0.$$

Inuitively, a test which counts the number of i's such that $X_i \leq \mu$ might be UMP. To check this consider the following reparametrization:

$$\mathbb{P} \longleftrightarrow (\mathbb{P}^-, \mathbb{P}^+, p),$$

where

$$\begin{aligned}\mathbb{P}^{+} &= \text{the conditional distribution of } X \mid X \le \mu,\\ \mathbb{P}^{-} &= \text{the conditional distribution of } X \mid X > \mu,\\ p &= \mathbb{P}(X \le \mu).\end{aligned}$$

There is a correspondence between the triples $(\mathbb{P}^{+}, \mathbb{P}^{-}, p)$ and the distributions $\mathbb{P}$. Let p_{-} and p_{+} be the densities for $\mathbb{P}^{-}$ and $\mathbb{P}^{+}$, the density of $\mathbb{P}$ is thus

$$p(x) = p \cdot p_{-}(x)\mathbf{1}_{x\le\mu} + (1-p)\cdot p_{+}(x)\mathbf{1}_{x>\mu}.$$

If we sort $X_1, \ldots, X_n$ such that

$$X_{i_1}, \ldots, X_{i_m} \le \mu \text{ and } X_{j_1}, \ldots, X_{j_{n-m}} \ge \mu,$$

then under a simple alternative $(\mathbb{P}^{-}, \mathbb{P}^{+}, p_1)$ where $p_1 < p_0$, the joint density of $(X_1, \ldots, X_n)$ is

$$p(x) = p_1^m \prod_{k=1}^{m} p_{-}(x_{i_k})(1-p_1)^{n-m} \prod_{k=1}^{n-m} p_{+}(x_{j+k}).$$

If we wish to pick distribution which gives us a null which is close to the above distribution we should take $(\mathbb{P}_{-}, \mathbb{P}_{+}, p_0)$. This formalizes the idea that there is “no information” in the tails of $\mathbb{P}$ and that “all information” is contained in $\mathbb{P}(X \le \mu)$ which is the quantity we are interested in. When testing the simple null $(\mathbb{P}_{-}, \mathbb{P}_{+}, p_0)$ against the simple alternative $(\mathbb{P}_{-}, \mathbb{P}_{+}, p_1)$, the likelihood ratio is

$$\frac{p_1^m(1-p_1)^{n-m}}{p_0^m(1-p_0)^{n-m}},$$

since $p_1 < p_0$ we reject when the quantity

$$m = |\{i : X_i \le \mu\}|,$$

is small. Under our simple null, m has a binomial (n, p_0) distribution thus the MP test is

$$\phi = \begin{cases} 1 & \text{if } m < k,\\ \gamma & \text{if } m = k,\\ 0 & \text{if} m > k. \end{cases}$$

where k and γ are determined by $\mathbb{E}_{p_0}\phi = \alpha$.

Note that this test is level α for the original composite H_0. This is because the power function $\beta(\theta)$ depends on $(\mathbb{P}^{-}, \mathbb{P}^{+}, p)$ only through p and the power function is a decreasing funcion of p. Thus we can conclude that ϕ is MP for the oringal H_0. Furthermore ϕ does not depend on $(\mathbb{P}^{-}, \mathbb{P}^{+}, p_1)$ and thus ϕ is UMP for

$$H_0 : \mathbb{P}(X \le \mu) \ge p_0 \quad \text{against} \quad H_1 : \mathbb{P}(X \le \mu) > p_0.$$