STATS 300A: Theory of Statistics I **Autumn 2016/17**

Lecture 18: December 6

Lecturer: Joseph Romano *Scribes: Kenneth Tay*

18.1 UMPI Tests

We have our usual set-up $X \sim P_\theta$, testing $H_0 : \theta \in \Omega_0$ vs. $H_1 : \theta \in \Omega_1$.

Definition 18.1 *A test φ is **almost invariant** w.r.t. group G if it satisfies*

$$\varphi(x) = \varphi(g(x)) \text{ a.e.} \tag{18.1}$$

for all $g \in G$. Here, the null set N_g of Equation 18.1 can depend on g.

Theorem 18.2 *Assume that there exists a unique (a.e.) UMPU test φ^* and also a UMPaI test w.r.t. group G.*

Then the latter test is unique (a.e.) and equal to φ^* (a.e.).

Proof: Let $U(\alpha)$ be the set of all unbiased level α tests. For any test ϕ, define ϕg by $\phi g(x) := \phi(g(x))$.

Claim: $\phi \in U(\alpha)$ if and only if $\phi g \in U(\alpha)$.

If $\phi \in U(\alpha)$, then for any θ,

$$\begin{aligned} \mathbb{E}_\theta \phi g(X) &= \mathbb{E}_\theta \phi(g(X)) \\ &= \mathbb{E}_{\bar{g}\theta} \phi(X). \end{aligned}$$

Since Ω_0 and Ω_1 are preserved under G, it follows that $\phi g \in U(\alpha)$.

In the other direction, if $\phi g \in U(\alpha)$, then $\phi g h \in U(\alpha)$ for any $h \in G$. We obtain the desired conclusion by setting $h = g^{-1}$.

Now, let $\beta_\phi(\theta) := \mathbb{E}_\theta \phi(X)$ be the power function of test ϕ. Note that

$$\beta_{\phi g}(\theta) = \mathbb{E}_\theta \phi(g(X)) = \mathbb{E}_{\bar{g}\theta} \phi(X) = \beta_\phi(\bar{g}\theta).$$

Hence, for $\theta \in \Omega_1$,

$$\begin{aligned} \beta_{\phi^* g}(\theta) &= \beta_{\phi^*}(\bar{g}\theta) \\ &= \sup_{\phi \in U(\alpha)} \beta_\phi(\bar{g}\theta) \\ &= \sup_{\phi \in U(\alpha)} \beta_{\phi g}(\theta) \\ &= \sup_{\phi \in U(\alpha)} \beta_\phi(\theta) \\ &= \beta_{\phi^*}(\theta), \end{aligned}$$

which means that ϕ^* and $\phi^* g$ have the same power function. By uniqueness (a.e.) of the UMPU test, we must have $\phi^*(x) = \phi^*(g(x))$ a.e. for each $g \in G$, i.e. ϕ^* is almost invariant. ■

18.1.1 Example: UMPI tests need not be admissible (Stein)

Let $\begin{pmatrix} X_{11} \\ X_{12} \end{pmatrix}$ and $\begin{pmatrix} X_{21} \\ X_{22} \end{pmatrix}$ be independent bivariate normals with all means 0 and unknown covariance matrices $\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ and $\Delta \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ respectively. Testing $H_0 : \Delta = 1$ vs. $H_1 : \Delta > 1$.

Let $Z = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$. Without loss of generality, we can assume that Z is non-singular. Consider the group of transformations $Z \mapsto AZ$, where A is a non-singular 2×2 matrix. These transformations leave the problem invariant.

However, for any 2 datasets Z and Z', there exists an A such that $Z' = AZ$. This means that there is only 1 orbit, and so the only invariant level α test is the constant test $\varphi = \alpha$. Thus, $\varphi = \alpha$ is UMPI.

To produce a "better" test, ignore the second components and test $H_0 : \Delta = 1$ on data X_{11} and X_{21}. In this setting, $T = \dfrac{X_{21}^2}{X_{11}^2}$ is maximal invariant. The test based on T is unbiased and has a non-trivial power function, and so it makes the constant test inadmissible.

18.2 Univariate Linear Hypotheses

Consider the following general set-up:

- Data $X = (X_1, \ldots, X_n)$ with the X_i's independent, $X_i \sim \mathcal{N}(\xi_i, \sigma^2)$ with σ^2 unknown. (Note that the following analysis does not change much if σ is known.)
- $\xi = (\xi_1, \ldots, \xi_n) \in \Pi_\Omega$, where Π_Ω is some s-dimensional subspace of $\mathbb{R}^n$. ($s < n$, can have $s = n$ if σ is known.)
- H_0 imposes r linear constraints on ξ, i.e. $\xi \in \Pi_\omega \in (s - r)$-dimensional subspace.

Here are 3 examples of this set-up:

- X_i iid, $X_i \sim \mathcal{N}(\xi, \sigma^2)$. Here, $\Pi_\Omega = \xi(1, \ldots, 1)$, $s = 1$. For $H_0 : \xi = 0$, $s = r = 1$.
- (Two-sample problem) $X_1, \ldots, X_{n_1} \sim \mathcal{N}(\xi, \sigma^2)$, $X_{n_1+1}, \ldots, X_{n_1+n_2} \sim \mathcal{N}(\eta, \sigma^2)$. Here, Π_Ω is the span of $\xi(\underbrace{1, \ldots, 1}_{n_1}, \underbrace{0, \ldots, 0}_{n_2})$ and $\eta(\underbrace{0, \ldots, 0}_{n_1}, \underbrace{1, \ldots, 1}_{n_2})$, $s = 2$.

  If $H_0 : \xi = \eta$, $r = 1$. If $H_0 : \xi = \eta = 0$, $r = 2$.
- $\xi_i = \alpha = \beta t_i$, where the t_i's are fixed and known. Here, Π_Ω is the span of $(1, \ldots, 1)$ and $(t_1, \ldots, t_n)$, and $s = 2$.

18.2.1 Reduction to Canonical Form

First, we reduce the problem to a simpler canonical form. Let $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = C \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$, where $C = \begin{pmatrix} \underset{\sim}{c}_1 \\ \vdots \\ \underset{\sim}{c}_n \end{pmatrix}$ is an $n \times n$ orthogonal matrix constructed so that $\underset{\sim}{c}_1, \ldots, \underset{\sim}{c}_s$ span Π_Ω and $\underset{\sim}{c}_{r+1}, \ldots, \underset{\sim}{c}_s$ span Π_ω.

Let $\eta := \mathbb{E}Y = C\xi$. With this construction, we have

$$\begin{aligned} (\xi_1, \ldots, \xi_n) \in \Pi_\Omega \quad &\Leftrightarrow \quad \eta_{s+1} = \cdots = \eta_n = 0, \\ (\xi_1, \ldots, \xi_n) \in \Pi_\omega \quad &\Leftrightarrow \quad \eta_1 = \cdots = \eta_r = \eta_{s+1} = \cdots = \eta_n = 0. \end{aligned}$$

Thus, we have a new parameter space $(\eta_1, \ldots, \eta_s, 0 \ldots, 0)^T$, and the null hypothesis specifies $\eta_1 = \cdots = \eta_r = 0$.

Let us now restate the testing problem: Data $Y = (Y_1, \ldots, Y_n)$, Y_i's independent with $Y_i \sim \mathcal{N}(\eta_i, \sigma^2)$, with $\eta_i = 0$ for $i > s$. Testing $H_0 : \eta_1 = \cdots = \eta_r = 0$.

To find the UMPI test, we consider a series of groups which leave the problem invariant:

1. $$Y_i' = \begin{cases} Y_i + c_i & i = r+1, \ldots, s, \\ Y_i & \text{otherwise.} \end{cases}$$

   The maximal invariant is $(Y_1, \ldots, Y_r, Y_{s+1}, \ldots, Y_n)$.

2. The group of orthogonal $r \times r$ transformations on the first r components of the vector above. The maximal invariant is $\left(\sum_{i=1}^{r} Y_i^2, Y_{s+1}, \ldots, Y_n \right)$. Using sufficiency, we can reduce this further to $\left(\sum_{i=1}^{r} Y_i^2, \sum_{j=s+1}^{n} Y_i^2 \right) =: (T_1, T_2)$.

3. $Y_i' = cY_i$. Here, the maximal invariant is $W := \sum_{i=1}^{r} Y_i^2 \Big/ \sum_{j=s+1}^{n} Y_i^2$.

Consider how the series of transformations transform the parameter space:

$$\begin{aligned} \left((\eta_1, \ldots, \eta_s), \sigma^2\right) \quad &\to \quad \left((\eta_1, \ldots, \eta_r), \sigma^2\right), \\ \left((\eta_1, \ldots, \eta_r), \sigma^2\right) \quad &\to \quad \left(\sum_{i=1}^{r} \eta_i^2, \sigma^2 \right), \\ \left(\sum_{i=1}^{r} \eta_i^2, \sigma^2 \right) \quad &\to \quad \frac{\sum_{i=1}^{r} \eta_i^2}{\sigma^2} =: \psi^2. \end{aligned}$$

Thus, the distribution of W depends only on ψ^2, i.e. it is a 1-parameter family. The density of W is given by

$$p_{\psi^2}(w) = e^{-\psi^2/2} \sum_{k=0}^{\infty} \frac{(\psi^2/2)^k}{k!} \frac{w^{\frac{r}{2}-1+k}}{(1+w)^{\frac{r+n-s}{2}+k}} \cdot c_k,$$

where

$$c_k = \frac{\Gamma\left(\frac{r+n-s}{2}+k\right)}{\Gamma\left(\frac{r}{2}+k\right)\Gamma\left(\frac{n-s}{2}\right)}.$$

Since we are testing $\psi^2 = 0$, for a UMPI test to exist, we just need to show that $\dfrac{p_{\psi^2}(w)}{p_0(w)}$ is an increasing function of w. We have

$$\frac{p_{\psi^2}(w)}{p_0(w)} = e^{-\psi^2/2}\sum_{k=0}^{\infty}\frac{(\psi^2/2)^k}{k!}\left(\frac{w}{1+w}\right)^k \cdot c_k.$$

Since $\frac{w}{1+w}$ is increasing in w, each term in the sum above is increasing in w. Thus, we have monotone likelihood ratio in W, implying that there is a UMPI test.

To obtain the critical value above which we reject the null hypothesis, let

$$W^* = \frac{\sum_{i=1}^{r} Y_i^2 \Big/ r}{\sum_{j=s+1}^{n} Y_j^2 \Big/ (n-s)}.$$

Under the null hypothesis, $W^* \sim F_{r,n-s}$.

18.2.2 Returning to the X_i's

We seek to express W^* in terms of the X_i's. Note that

$$\begin{aligned}
\sum_{j=s+1}^{n} Y_j^2 &= \min_{(\eta_1,\ldots,\eta_n)\in\Pi_\Omega^Y}\sum_{i=1}^{n}(Y_i-\eta_i)^2 \\
&= \min_{(\xi_1,\ldots,\xi_n)\in\Pi_\Omega}\sum_{i=1}^{n}(X_i-\xi_i)^2 \\
&= \sum_{i=1}^{n}(X_i-\hat{\xi}_i)^2,
\end{aligned}$$

where $(\hat{\xi}_1,\ldots,\hat{\xi}_n)$ is the least squares estimate of $(\xi_1,\ldots,\xi_n)$ subject to π_Ω.

Similarly, we have

$$\sum_{i=1}^{r} Y_i^2 + \sum_{j=s+1}^{n} Y_j^2 = \sum_{i=1}^{n}(X_i-\hat{\hat{\xi}}_i)^2,$$

where $(\hat{\hat{\xi}}_1, \ldots, \hat{\hat{\xi}}_n)$ is the least squares estimate of $(\xi_1, \ldots, \xi_n)$ subject to π_ω. Thus,

$$W^* = \frac{\left[\sum_{i=i}^{n}(X_i - \hat{\hat{\xi}}_i)^2 - \sum_{i=i}^{n}(X_i - \hat{\xi}_i)^2\right] \Big/ r}{\sum_{i=i}^{n}(X_i - \hat{\xi}_i)^2 \Big/ (n-s)}$$

$$= \frac{\sum_{i=i}^{n}(\hat{\xi}_i - \hat{\hat{\xi}}_i)^2 \Big/ r}{\sum_{i=i}^{n}(X_i - \hat{\xi}_i)^2 \Big/ (n-s)}.$$

18.2.3 Example: Two-sample problem

$X_1, \ldots, X_{n_1} \sim \mathcal{N}(\xi, \sigma^2)$, $X_{n_1+1}, \ldots, X_{n_1+n_2} \sim \mathcal{N}(\eta, \sigma^2)$.

With no other constraints (i.e. subject to π_Ω), we have the least squares estimate $(\hat{\xi}_1, \ldots, \hat{\xi}_n) = (\underbrace{\hat{\xi}, \ldots, \hat{\xi}}_{n_1}, \underbrace{\hat{\eta}, \ldots, \hat{\eta}}_{n_2})$, where

$$\hat{\xi} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i, \quad \hat{\eta} = \frac{1}{n_2}\sum_{i=n_1+1}^{n_2} X_i.$$

Under H_0 (i.e. subject to π_ω), the least squares estimate is simply $\bar{X}$.

Plugging these values into the formula for W^*, we get the classical two-sample t-statistic.