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

Lecture 17: December 1

Lecturer: Joseph Romano *Scribes: Kenneth Tay*

17.1 Invariant Tests

We have the following set-up:

- X lives on a sample space $\mathcal{S}$, $X \sim P_\theta$, $\theta \in \Omega$.
- There is a group G of 1-to-1 transformations such that for $g \in G$, $X \sim P_\theta \Rightarrow gX \sim P_{\theta'}$ for some θ'. We write $\theta' = \bar{g}\theta$.
- We are testing $H_0 : \theta \in \Omega_0$ vs $H_1 : \theta \in \Omega_1$. We assume that the parameter spaces are "preserved", i.e. $\bar{g}\Omega_i = \Omega_i$ for all $g \in G$.

In words, testing $\theta \in \Omega_0$ based on data X is identical to testing $\theta \in \Omega_0$ based on data gX.

Definition 17.1 *φ is **invariant w.r.t.** G if $\varphi(X) = \varphi(gX)$ for all X, g.*

Note that the transformations induce orbits, or equivalence classes, on the sample space $\mathcal{S}$. (X_1 and X_2 are in the same equivalence class if there is a transformation g such that $X_2 = g(X_1)$.) Thus, an invariant test must be constant on each orbit.

The goal is to find UMPI tests, i.e. a UMP test among all invariant tests.

17.1.1 Maximal Invariant Statistics

Definition 17.2 *A statistic $T = T(X)$ is **maximal invariant** if:*

1. *T is invariant (i.e. $T(x) = T(gx)$ for all x, g), and*
2. *If $T(x_1) = T(x_2)$, then there exists some $g \in G$ such that $x_2 = gx_1$.*

Theorem 17.3 *φ is invariant iff it is a function of a maximal invariant T.*

Proof: Assume $\varphi = h(T)$ for some function h. Since T is maximal invariant, it is also invariant, so for all x, g,

$$h(T)(x) = h(T(x)) = h(T(gx)) = h(T)(gx),$$

i.e. φ is invariant.

In the other direction, assume that φ is invariant. To show that φ is a function of T, it is enough to show that

$$T(x_1) = T(x_2) \quad \Rightarrow \quad \varphi(x_1) = \varphi(x_2).$$

Since T is maximal invariant, $T(x_1) = T(x_2)$ implies that there is a g such that $x_2 = g(x_1)$, so $\varphi(x_2) = \varphi(gx_1) = \varphi(x_1)$, as required. ■

This theorem characterizes invariant tests: **in the search for a UMPI test, we only have to look at functions of a maximal invariant T.**

17.1.1.1 Examples of maximal invariant statistics

In all the following examples, data is given by $X = (X_1, \dots, X_n)$.

1. Family of transformations $X_i' = X_i + a$, where a is a real constant.
   $T(X_1, \dots, X_n) = (X_1 - X_n, \dots, X_{n-1} - X_n)$ is maximal invariant.
2. Family of transformations $X_i' = cX_i$, where c is a non-zero real constant, and X_i's assumed to be non-zero.
   $T(X_1, \dots, X_n) = \left(\frac{X_1}{X_n}, \dots, \frac{X_{n-1}}{X_n}\right)$ is maximal invariant.
3. $G = n!$ permutations of $X_1, \dots, X_n$.
   $T(X_1, \dots, X_n) = (X_{(1)}, \dots, X_{(n)})$ order statistics is maximal invariant.
4. Family of transformations $X_i' = f(X_i)$, where f is a continuous, strictly increasing function.
   Here, the ranks of the data are maximal invariant.

17.1.2 Finding UMPI Tests

17.1.2.1 Example: One of 2 location models

$X = (X_1, \dots, X_n)$ real-valued with joint density $f_i(x_1 - \theta, \dots, x_n - \theta)$, $i = 0$ or 1 with θ unknown. Testing $H_0 : f_0$ is true with unknown θ vs. $H_1 : f_1$ is true with unknown θ.

In this set-up, the family of transformations $X_i' = X_i + a$ ($a \in \mathbb{R}$) applies, with $\bar{g}\theta = \theta + a$. Hence, if we let $Y_i = X_i - X_n$, the statistic $T = (Y_1, \dots, Y_{n-1})$ is maximal invariant. Any UMPI test must be a function of T.

Note that the distribution of T does not depend on θ, thus the problem is reduced to testing a simple null vs. a simple alternative!

The density of T under H_i is $\displaystyle\int f_i(y_1 + t, \dots, y_{n-1} + t, t)dt$, so by the Neyman-Pearson Lemma, the UMPI test rejects if

$$\frac{\displaystyle\int f_1(y_1 + t, \dots, y_{n-1} + t, t)dt}{\displaystyle\int f_0(y_1 + t, \dots, y_{n-1} + t, t)dt} > c$$

for some constant c, or equivalently, it rejects if

$$\frac{\displaystyle\int f_1(x_1 + u, \dots, x_n + u)du}{\displaystyle\int f_0(x_1 + u, \dots, x_n + u)du} > c.$$

17.1.2.2 Example: Testing variance in normal setting

$X_1, \ldots, X_n$ iid, $X_i \sim \mathcal{N}(\xi, \sigma^2)$ with both parameters unknown. Testing $H_0 : \sigma = \sigma_0$ vs. $H_1 : \sigma < \sigma_0$.

Recall that we can always reduce the search to tests which are functions of sufficient statistics. In this case, $T_1 = \bar{X}$, $T_2 = \sum(X_i - \bar{X})^2$.

Consider the family of transformations on the original data $X_i' = X_i + a$. These transformations should not affect the result of the test. This family of transformations induces transformations on the T_i: $T_1' = T_1 + a$, $T_2' = T_2$.

We claim that T_2 is maximal invariant: it is clearly invariant, and if $(T_1, T_2) = (\tilde{T_1}, T_2)$, the transformation with $a = \tilde{T_1} - T_1$ gives us $\tilde{T_1} = gT_1$.

Thus, we only have to look at functions of $T_2 = \sum(X_i - \bar{X})^2$. Note that the family of distributions for T_2 has monotone likelihood ratio, hence there is a UMP 1-sided test.

17.1.2.3 Example: Testing mean in normal setting

$X_1, \ldots, X_n$ iid, $X_i \sim \mathcal{N}(\xi, \sigma^2)$ with both parameters unknown. Testing $H_0 : \xi = 0$ vs. $H_1 : \xi > 0$.

Again, reduce to looking for functions of the sufficient statistic $(T_1, T_2) = \left(\bar{X}, \sqrt{\sum(X_i - \bar{X})^2}\right)$.

Consider the group of transformations $X_i' = cX_i$, where $c > 0$. This induces transformations on the sufficient statistic $T_1' = cT_1$, $T_2' = cT_2$. Hence, the ratio $\dfrac{T_1}{T_2}$ is maximal invariant.

Note also that the group of transformations induces transformations on the parameter space, with $\dfrac{\xi}{\sigma}$ being maximal invariant. Hence (by the next proposition), the family of distributions of $\dfrac{T_1}{T_2}$ depends only on $\theta = \dfrac{\xi}{\sigma}$. This is a 1-parameter family which has monotone likelihood ratio in $\dfrac{T_1}{T_2}$, and so a UMPI test exists. (Details in TSH.)

Proposition 17.4 *The group of transformations G on the data sample space induces a group of transformations $\bar{G}$ on the parameter space.*

Suppose $v(\theta)$ is maximal invariant based on $\bar{G}$. Then any invariant statistic $\varphi(X)$ has a distribution that depends on θ only through $v(\theta)$.

Proof: Take θ_1, θ_2 in the same orbit (i.e. $\bar{g}\theta_1 = \theta_2$ for some $\bar{g} \in \bar{G}$).

Note that if $X \sim P_\theta$, then $gX \sim P_{\bar{g}\theta}$. Hence, for any arbitrary event E,

$$\begin{aligned} P_{\theta_1}\{\varphi(X) \in E\} &= P_{\theta_1}\{\varphi(gX) \in E\} \\ &= P_{\bar{g}\theta_1}\{\varphi(X) \in E\} \\ &= P_{\theta_2}\{\varphi(X) \in E\}, \end{aligned}$$

which is what we need to prove. ■

17.1.3 Computing Maximal Invariant Statistics in Stages

There are some set-ups where 2 different subgroups of transformations leave the problem invariant (e.g. shifting and scaling). The following example shows that the order in which the reduction (or finding the maximal invariant statistic) is done matters.

Consider $X = (X_1, \dots, X_n)$ and the family of transformations $gx = (ax_1 + b, \dots, ax_n + b)$, $a \neq 0$.

- First let at the subgroup of transformations where $a = 1$ (shifting). We get $Y = (Y_1, \dots, Y_{n-1}) = (X_1 - X_n, \dots X_{n-1} - X_n)$ maximal invariant.

  Next, consider the subgroup $X_i'' = ax_i$ (scaling). This induces transformations on Y: $Y_i'' = aY_i$, which leads to maximal invariant $Z = (Z_1, \dots, Z_{n-2})$ with $Z_i = \frac{Y_i}{Y_{n-1}} = \frac{X_i - X_n}{X_{n-1} - X_n}$.

  Z is maximal invariant for the original group.

- Now let's consider the subgroups in the reverse order. For the scaling transformations, $U = (U_1, \dots, U_{n-1})$ with $Y_i = \frac{X_i}{X_n}$.

  Next, consider the shifting transformations. This induces transformations on U: $U_i \to \frac{X_i + a}{X_n + a}$, which is not a function of $\frac{X_i}{X_n}$. So we can't go any further.

In a nutshell, the order in which subgroups of transformations are considered is important. The following theorem makes precise the notion that if you can carry this process all the way through (as in the first method), you will get the "right answer" (something maximal invariant for the original setting).

Theorem 17.5 *Suppose that the group of transformations G is generated by D and E. Suppose that $y = s(x)$ is maximal invariant w.r.t. D, and for any $e \in E$, $s(x_1) = s(x_2) \Rightarrow s(ex_1) = s(ex_2)$.*

If $z = t(y)$ is maximal invariant under the induced group E^* of transformations e^* defined by $e^* y = s(ex)$ when $y = s(x)$, then $t(s(x))$ is maximal invariant w.r.t. G.

17.1.3.1 Example: Two-sample problem

$X_1, \dots, X_m$ iid, $X_i \sim \mathcal{N}(\xi, \sigma^2)$, $Y_1, \dots, Y_n$ iid, $Y_j \sim \mathcal{N}(\eta, \tau^2)$, all 4 parameters unknown. Testing $H_0 : \sigma^2 = \tau^2$ vs. $H_1 : \sigma^2 < \tau^2$.

Note that $X_i' = X_i + a$, $Y_j' = Y_j + b$ leaves the problem invariant, and $X_i' = cX_i$, $Y_j' = cY_j$ leaves the problem invariant.

First we reduce by sufficiency: look only at functions of $T_1 = \bar{X}$, $T_2 = \bar{Y}$, $T_3 = \sum (X_i - \bar{X})^2$, $T_4 = \sum (Y_j - \bar{Y})^2$.

Look now at the shifting transformations. They induce transformations on the sufficient statistics $T_1' = T_1 + a$, $T_2' = T_2 + b$, $T_3' = T_3$, $T_4' = T_4$. This leads to (T_3, T_4) being maximal invariant.

Next, look at the scaling transformations. This leads to $\dfrac{T_4}{T_3}$ being maximal invariant. Also, the distribution of $\dfrac{T_4}{T_3}$ only depends on $\Delta^2 := \dfrac{\tau^2}{\sigma^2}$.

The family of distributions of $\dfrac{T_4}{T_3}$ is a 1-parameter family. In fact, under H_0,

$$Z := \frac{\sum (Y_j - \bar{Y})^2 / (n-1)}{\sum (X_i - \bar{X})^2 / (m-1)} \sim F_{m-1, n-1}.$$

More generally, Z has non-central F distribution, with density

$$p_\Delta(z) \propto \frac{z^{\frac{1}{2}(n-3)}}{\left(\Delta + \frac{n-1}{m-1} z\right)^{\frac{1}{2}(m+n-2)}}.$$

It can be checked that this family has monotone likelihood ratio in Z as $\frac{p_{\Delta'}(z)}{p_\Delta(z)}$ is increasing in z. Hence, a UMPI test exists.

$$x$$

$$x$$

Definition 17.6 x

Theorem 17.7 x

Proof: x ■