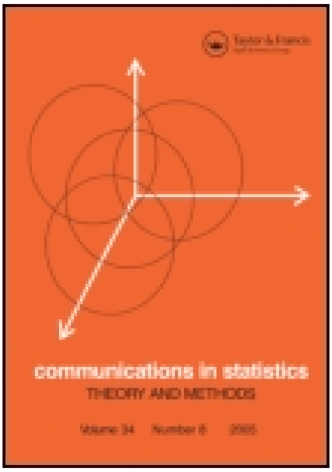

Communications in Statistics - Theory and Methods

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Maximal Invariant and Weakly Equivariant Estimators

M. Shams[a], M. Emadi[a] & N. R. Arghami[a]

[a] Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

Published online: 17 Mar 2015.

Communications in Statistics—Theory and Methods, 44: 1293–1317, 2015

ISSN: 0361-0926 print / 1532-415X online
DOI: 10.1080/03610926.2013.781637

Maximal Invariant and Weakly Equivariant Estimators

M. SHAMS, M. EMADI, AND N. R. ARGHAMI

Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

Equivariant functions can be useful for constructing of maximal invariant statistic. In this article, we discuss construction of maximal invariants based on a given weakly equivariant function under some additional conditions. The theory easily extends to the case of two or more weakly equivariant functions. Also, we derive a maximal invariant statistic when the group contains a sharply transitive and a characteristic subgroup. Finally, we consider the independence of invariant and weakly equivariant functions under some special conditions.

Keywords Topological group; G-space; Sharply transitive group; Maximal invariant statistic; Weakly equivariant function; Weakly isovariant function; Basu's Theorem.

Mathematics Subject Classification Primary 62F10; Secondary 54H11.

1. Introduction

Statistical decisions should not be affected by transformations on the data, so we study invariance. For some additional discussion on the development of invariance arguments, see Hall et al. (1965). It should be mentioned that this article is based on topological groups and since statisticians commonly examine the invariance theory from the point of view of transformation groups, its literature may seems unfamiliar to them. It is strongly recommended to refer to Eaton (1989) to get a better view of the invariance theory from the point of view of topological groups. At first, we list some symbols and well known results about the topological groups and related arguments.

Definition 1.1 (Deitmar and Echterhoff, 2009). A map $f : X \to Y$ between topological spaces X and Y is called continuous if $f^{-1}(U)$ is open in X for every open set $U \subset Y$. This is equivalent to the condition that $f^{-1}(C)$ is closed in X for every closed $C \subset Y$.

Received June 12, 2012; Accepted February 14, 2013.
Address correspondece to Mahdi Emadi, Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran; E-mail: emadi_mahdi@yahoo.com

Definition 1.2. (Folland, 1995). A topological group is a group G, together with a topology on the set G such that the group multiplication and inversion,

$$\begin{array}{ll} G \times G \to G & G \to G \\ (g_1, g_2) \mapsto g_1 g_2 & g \mapsto g^{-1}, \end{array}$$

are both continuous maps.

It suffices to assert that the map $\zeta : (g_1, g_2) \mapsto g_1^{-1} g_2$ is continuous. To see this, assume that ζ is continuous. Since $x \mapsto (x, e)$ is continuous (where e is the unit element of the group G), $g \mapsto (g, e) \mapsto g^{-1}e = g^{-1}$ and $(g_1, g_2) \mapsto (g_1^{-1}, g_2) \mapsto g_1 g_2$ are continuous (Deitmar and Echterhoff, 2009).

Definition 1.3 (Deitmar and Echterhoff, 2009). A topological space X is called a Hausdorff space, if any two different points can be separated by disjoint neighborhoods, i.e., if for any two $x \neq y$ in X there are open sets $U, V \subset X$ with $x \in U, y \in V, U \cap V = \phi$.

A topological group is called a locally compact group if it is Hausdorff and every point possesses a compact neighborhood.

Definition 1.4 (Deitmar and Echterhoff, 2009). A bijective map $f : X \to Y$ is called a homeomorphism if $f$and f^{-1} are contiuous.

Homeomorphisms are the mappings which preserve all the topological properties of a given space. Two spaces with a homeomorphism between them are called homeomorphic, and from a topological viewpoint they are the same. For example, every nonempty open interval $(a, b) \subset R$ is homeomorphic to the real line R when both are equipped with the usual topology.

The useful groups in statistics are transformation groups acting on some set or space. Transformation groups were introduced for the first time by Fraser (1961, 1968). Group actions are first induced on the sample space, which induce group actions on the parameter space.

Definition 1.5 (Deitmar and Echterhoff, 2009). The set X is called a G-space iff a map $G \times X \to X$, given by $(g, x) \mapsto gx$, satisfies the conditions $g_1(g_2 x) = (g_1 g_2)x$ and $ex = x$ for all $g_1, g_2 \in G, x \in X$. In this case, we can also say G acts on X.

A subgroup H of G, written $H \leq G$, is a subset that is again a group under the same composition. An important class of subgroups are the normal subgroups, written $H \lhd G$, which are subgroups H such that $gH = Hg$ for all $g \in G$. This means that the subgroup H is closed under conjugation, i.e., $gHg^{-1} = \{ghg^{-1} : h \in H\} = H$ for all $g \in G$. Let G and H be groups. A function $\rho : G \to H$ is a homomorphism if $\rho(g_1 g_2) = \rho(g_1)\rho(g_2)$, for all $g_1, g_2 \in G$. When there is a homomorphism ρ, H is called a homomorphic image of G and we write $\bar{G} = H$. In this case, $\overline{g_1 g_2} = \overline{g_1 g_2}$ and $\overline{g^{-1}} = \bar{g}^{-1}$. Also, if e is the identity in G, then $\bar{e}$ is the identity in $\bar{G}$. A bijective homomorphism from G on itself is called an automorphism. $Aut(G)$ denotes the group of all continuous automorphisms of G. If X is a G-space and $x \in X$, then $Gx = \{gx : g \in G\}$ is called the orbit of G (through x) and $G_x = \{g : gx = x\}$ is called the stabilizer or stability subgroup of G at x (Deitmar and Echterhoff, 2009).

Example 1.1. Let $G = Z_2 = \{\pm 1\}$ acts on $X = \{x = (x_1, x_2) \in R^2 : x_1^2 + x_2^2 = 1\}$ defined by $g(x_1, x_2) = (gx_1, x_2)$. For a given $x \in X$,

$$Gx = \begin{cases} \{x\} & x = (x_1, x_2) = (0, \pm 1) \\ \{(\pm x_1, x_2)\} & x = (x_1, x_2) \neq (0, \pm 1) \end{cases},$$

$$G_x = \begin{cases} Z_2 & x = (x_1, x_2) = (0, \pm 1) \\ \{1\} & x = (x_1, x_2) \neq (0, \pm 1) \end{cases}.$$

Definition 1.6. (Bredon, 1972). Let G be a group and X a G-space. The action of G on X is said to be:

(i) trivial, if $G_x = G$ for every $x \in X$;
(ii) free, if $G_x = \{e\}$ for every $x \in X$;
(iii) transitive, if for every pair $x, x' \in X$, there is a $g \in G$ with $x' = gx$; and
(iv) sharply transitive, if for all $x, x' \in X$there is a unique$g \in G$ such that $x' = gx$.

Remark 1.1. If G acts transitively on X, then this action is free iff it is sharply transitive.

Example 1.2. Consider $G = GL_n$ of all real invertible $n \times n$ matrices and $X = R^n - \{0\}$. Thus, G acts on X by gx, where gx means the matrix g times the vector x. Also, GL_n acts transitively on $R^n - \{0\}$ and $G = GL_1 = R - \{0\}$ is free and sharply transitive on $X = R - \{0\}$.

Definition 1.7. Suppose that X and Y are two G- spaces. Then we have the following.

(i) A measurable function $f : X \to Y$ is called G-equivariant map if $f(gx) = gf(x)$, for all $g \in G$ and $x \in X$ (Lehmann and Romano, 2005).
(ii) A measurable function $f : X \to Y$ is called weakly G-equivariant map if there exists a continuous automorphism α_f of G such that $f(gx) = \alpha_f(g)f(x)$, for all $g \in G$ and $x \in X$ (Bredon, 1972).
(iii) A measurable function $f : X \to Y$ is said to be maximal invariant if it is G-invariant i.e., $f(gx) = f(x)$, for all $g \in G$ and $x \in X$ and also, if $f(x_1) = f(x_2)$ implies $x_2 = gx_1$ for some$g \in G$ (Lehmann and Romano, 2005).
(iv) A measurable function $f : X \to Y$ is called G-isovariant map if $G_x = G_{f(x)}$ for all $x \in X$ (Palais, 1960).
(v) A measurable function $f : X \to Y$ is said to be weakly G-isovariant if $\alpha_f(G_x) = G_{f(x)}$ for all $x \in X$ and some $\alpha_f \in Aut(G)$.

Isovariant maps were introduced by Palais (1960) in order to study a classification problem for orbit maps of *G*-spaces. It should be clear by now that the concept of isovariance can be easily extended to the case of weakly isovariance when we study weakly G-equivariant maps (see Definition 1.7(v)).

Example 1.3. Assume that the G-spaces X and Y are Hausdorff for which the topology has a countable base. Consider the invariant integral $J(f) = \int_X f(x)\mu(dx)$, $f \in K(X)$ (the real vector space K(X) is the set of all continuous real valued functions f defined on X which have compact support) which clearly satisfies the following.

(i) $J(a_1 f_1 + a_2 f_2) = a_1 J(f_1) + a_2 J(f_2)$ for all $f_1, f_2 \in K(X), a_1, a_2 \in R$;

(ii) $J(f) \geq 0$ whenever $f \in K(X)$ and f is non negative;
(iii) $J(f) > 0$ for some $f \in K(X)$; and
(iv) $J(L_g f) = \int_X f(g^{-1}x)\mu(dx) = \int_X f(x)\mu(dx) = J(f)$ for all $f \in K(X)$, $g \in G$,

where the transformation $L_g : K(X) \to K(X)$, given by $(L_g f)(x) = f(g^{-1}x)$ for all $x \in X$, $g \in G$, $f \in K(X)$ (Nachbin and Bechtolsheim, 1965) and μ is the corresponding Radon measure i.e., $\mu \neq 0$ and $\mu(\mathrm{C}) < +\infty$ for all compact sets $C \subseteq X$ (Folland, 1995). Given a Borel measurable weakly equivariant function φ from X onto Y, there is a natural measure ν induced on Y s.t. $\nu(B) \equiv \mu(\varphi^{-1}(B))$ for measurable subset $B \subseteq Y$. Suppose that $\mu(\varphi^{-1}(K)) < +\infty$ for all compact sets $K \subseteq Y$, then the integral $J_1(f) = \int_Y f(y)\nu(dy)$, $f \in K(Y)$ is well defined. Since φ is weakly equivariant with onto $\alpha_f \in Aut(G)$ and J_1 is invariant, we have

$$\begin{aligned} J_1(L_g f) &= \int_Y f\left((\alpha_f(g_0))^{-1}y\right)\nu(dy) = \int_X f\left((\alpha_f(g_0))^{-1}\phi(x)\right)\mu(dx) \\ &= \int_X f\left(\phi(g_0^{-1}x)\right)\mu(dx) = \int_X f\left(\phi(x)\right)\mu(dx) = J_1(f) \end{aligned}$$

for all $f \in K(X)$, $g \in G$ and some $g_0 = \alpha_f^{-1}(g) \in G$, so the integral J_1 is invariant under G.

The following lemma is a useful example for bijective weakly G-equivariant functions.

Lemma 1.1. *Assume that G is a compact group which acts on X and $x \in X$, $\alpha \in Aut(G)$. A natural map $\tau_x : G/\alpha^{-1}(G_x) \to Gx$, given by $g\alpha^{-1}(G_x) \mapsto \alpha(g)x$, is a bijective weakly G-equivariant mapping.*

Proof. Suppose that $\tau_x(g_1\alpha^{-1}(G_x)) = \alpha(g_1)x = \alpha(g_2)x = \tau_x(g_2\alpha^{-1}(G_x))$, then $\alpha(g_2^{-1}g_1)x = x$ and so $g_2^{-1}g_1 \in \alpha^{-1}(G_x)$. It follows that $g_1\alpha^{-1}(G_x) = g_2\alpha^{-1}(G_x)$. This means that τ_x is injective. For any $y \in Gx$, there exists a $g\prime \in G$ such that $y = g\prime x$, but there is a $g \in G$ that $g\prime = \alpha(g)$, hence $\tau_x(g\alpha^{-1}(G_x)) = \alpha(g)x = g\prime x = y$, and thus τ_x is surjective. Also,

$$\tau_x((h(g\alpha^{-1}(G_x))) = \tau_x((hg\alpha^{-1}(G_x))) = \alpha(h)(\alpha(g)x) = \alpha(h)\tau_x(g\alpha^{-1}(G_x))$$

for all $g, h \in G$ and therefore τ_x is weakly G-equivariant. □

Now, if $G = O_n$ acts on $X = R^n - \{0\}$ and we take $H = \{diag(1, \Gamma_2) : \Gamma_2 \in O_{n-1}\}$, then the left coset ΓH of any $\Gamma \in G$ consists of all orthogonal matrices with the same first column that Γ has. Also, if $x = (1, 0, \ldots, 0)^t$, then $G_x = H$. Since G is compact, Lemma 1.1 implies that each orbit Gx of G acting on X, is homogeneous for G and can be regarded as G/G_x, provided $\alpha = 1_G$. For all $y \in Gx$, the equation $y = \Gamma x$, $\Gamma \in G$ is equivalent to the first column of Γ which is y. But all such Γ's constitute a coset of $G/H = G/G_x$. Thus, the function $\tau_x : G/G_x \to Gx$ assigns to an $n \times n$ orthogonal matrix its first column (Eaton, 1983).

Remark 1.2. If G acts freely on X, then, by Lemma 1.1, there is a one-to-one correspondence between G and Gx. Furthermore, if G acts transitively on X, then by Remark 1.1, G is sharply transitive on X and in this case, there is a one-to-one correspondence between G and X.

Weakly isovariant G-equivariant maps have nice homeomorphism properties as follows.

Proposition 1.1. *If $f : X \to Y$ is a weakly G- equivariant map between G-spaces, then:*

(i) $\alpha_f(G_x) \subseteq G_{f(x)}$ for all $x \in X$ and
(ii) f is one-to-one on Gx if and only if equality holds.

Proof.

(i) If $g \in G_x$, then $f(x) = f(gx) = \alpha_f(g)f(x)$. Thus, $\alpha_f(g) \in G_{f(x)}$.
(ii) Assume that $\alpha_f(G_x) = G_{f(x)}$ for all $x \in X$. Let $x \in X$ and suppose $g_1, g_2 \in G$. Then

$$f(g_1x) = f(g_2x) \Rightarrow \alpha_f(g_1)f(x) = \alpha_f(g_2)f(x) \Rightarrow f(x) = \alpha_f(g_1^{-1})\alpha_f(g_2)f(x)$$
$$\Rightarrow \alpha_f(g_1^{-1}g_2) \in G_{f(x)} = \alpha_f(G_x) \Rightarrow g_1^{-1}g_2x = x \Rightarrow g_1x = g_2x.$$

It follows that f is injective on Gx.
Conversely, suppose $f|Gx$ is injective for each $x \in X$, then

$$\alpha_f(g) \in \alpha_f(G_x) \Leftrightarrow g \in G_x \Leftrightarrow f(gx) = f(x) \Leftrightarrow \alpha_f(g) \in G_{f(x)}.$$

Thus, we get $\alpha_f(G_x) = G_{f(x)}$, that is f is weakly isovariant.
The following is a useful consequence of Proposition 1.1. □

Corollary 1.1. *Let G be a compact topological group and let X and Y be G-spaces. A weakly G-equivariant function $f : X \to Y$ is weakly isovariant if and only if its restriction to each orbit in X is a homeomorphism onto its image in Y.*

Proof. First assume that f is weakly isovariant, then since $f|Gx : Gx \to Gf(x)$ is a closed, surjective and continuous map for each $x \in X$, then $f|Gx$ is a homeomorphism if and only if it is injective. Thus by using Proposition 1.1(ii), we will complete the proof. □

Remark 1.3. In general, if f is one-to-one, then it is one-to-one on *Gx*. Converse is true if G acts transitively on X, i.e., $Gx = X$ for all $x \in X$. Consequently, if G acts transitively on X, then f is one-to-one iff its restriction to each orbit in X is one-to-one.

Eaton introduced a method of construct a maximal invariant in terms of a given equivariant function (Eaton, 1989, pp. 28–40). In this article, we will introduce a method for finding maximal invariant functions by using weakly equivariant functions, which is a generalization of Eaton's method. To apply this method we need a condition, which is only satisfied by equivariant functions when the group of transformations is free. Thus, in this case, our method is the same as Eaton's. In this case, we limit ourselves to sharply transitive groups. Thus we can consider the parameter space as a group with a new binary action. Based on this concept, we will find maximal invariant functions by using weakly equivariant estimators, with the difference that, in this case, the functions are estimators. Notice that when the group is sharply transitive, it is possible to change weakly G-equivariant functions into weakly G-equivariant estimators and vice versa. In some cases, the transformation group acting on the parameter space is not sharply transitive but it contains a subgroup with this property. In a special case, we will find a maximal invariant statistic when the group contains a sharply transitive and a characteristic subgroup. Instead of using the above methods, one can use two (or more) weakly G-equivariant functions to obtain a maximal invariant function. This method immediately extends to the case where we have two (or more) weakly G-equivariant estimators. Finally, we deal with the independence of an invariant function and a weakly equivariant function under some special conditions.

2. Maximal Invariant and Weakly Equivariant Estimators

In this section, we investigate the connection between maximal invariant statistics and weakly equivariant estimators. At first, we show that a weakly equivariant function can be used for constructing a maximal invariant and then we give a general form for the maximal invariant. Then, under some special conditions, we construct a general maximal invariant from two (or more) given weakly equivariant functions? Toward the end of Sec. 3, we improve this method for weakly equivariant estimators when the group is sharply transitive.

The elementary relationship between maximal invariant functions and weakly equivariant functions is given by the following proposition.

Proposition 2.1. *If X, Y, and Z are G-spaces, then:*

(i) if $f : X \to Y$ is weakly equivariant and $h : Y \to Z$ is invariant, then $k_1 = h \circ f : X \to Z$ is invariant;
(ii) if $f : X \to Y$ is one-to-one and $h : Z \to X$ is maximal invariant, then $k_2 = f \circ h : Z \to Y$ is maximal invariant; and
(iii) if $f : X \to Y$ is weakly equivariant and weakly isovariant and $h : Y \to Z$ is maximal invariant, then $k_1 = h \circ f : X \to Z$ is maximal invariant.

Proof.

(i) For all $x \in X$, $g \in G$: $k_1(gx) = h(f(gx)) = h(\alpha_f(g)f(x)) = h(f(x)) = k_1(x)$.
(ii) $k_2 = f \circ h$ is invariant because $k_2(gx) = f(h(gx)) = f(h(x)) = k_2(x)$ for all $x \in X$, $g \in G$. Now, if $k_2(x_1) = k_2(x_2)$, then $f(h(x_1)) = f(h(x_2))$ and since f is one-to-one, we have $h(x_1) = h(x_2)$ and by maximal invariance of h, there exists $g \in G$ such that $x_1 = gx_2$. It follows that $k_2 = f \circ h$ is maximal invariant.
(iii) By part (i), $k_1 = h \circ f$ is invariant. If $k_1(x_1) = k_1(x_2)$, then $h(f(x_1)) = h(f(x_2))$ and since h is maximal invariant, there exists $g\prime \in G$ such that $f(x_1) = g\prime f(x_2)$. Since α_f is onto, there is a $g \in G$ s.t. $g\prime = \alpha_f(g)$, and therefore $f(x_1) = \alpha_f(g)f(x_2) = f(gx_2)$. But Proposition 1.1(ii) implies that f is one-to-one on Gx and thus $x_1 = gx_2$ for some $g \in G$. Thus, $k_1 = h \circ f$ is maximal invariant. w

A maximal invariant function is constant on the orbits but for each orbit takes on a different value. If G acts transitively on X, for all invariant functions like f, if $f(x_1) = y_1$ and $f(x_2) = y_2$, then $y_2 = f(x_2) = f(gx_1) = f(x_1) = y_1$ for some $g \in G$ and in this way, only (maximal) invariant functions are constant and so Proposition 2.1 is trivial. The following lemma states how are we can construst a general maximal invariant from a given equivariant function Afterward, in Proposition 2.2, we improve it for weakly equivariant function. □

Lemma 2.1. *(Eaton, 1989). If $\delta_0 : X \to G$ is any G-equivariant function, then $f(x) = (\delta_0(x))^{-1}x$ is maximal invariant.*

Proof. f is invariant because $f(gx) = (g\delta_0(x))^{-1}gx = (\delta_0(x))^{-1}g^{-1}gx = f(x)$. Suppose $f(x_1) = f(x_2)$, then $x_2 = (\delta_0(x_2)(\delta_0(x_1))^{-1})x_1 = gx_1$ and it follows that f is maximal invariant. □

To see how Lemma 2.1 together with Proposition 2.1 can be used for constructing a maximal invariant statistic, suppose that $G = R$ acts on $\mathbf{X} = R^n$ by $gx = x + ge_n$, $x \in R^n$, $g \in R$ where $e_n = (1, \ldots, 1) \in R^n$. Since $\delta_0(x) = \bar{x}$ is G-equivariant, then by Lemma 2.1, $h(x) = (\delta_0(x))^{-1}x = x - \bar{x}e_n = (x_1 - \bar{x}, \ldots, x_n - \bar{x})$ is maximal invariant. Hence, because of Proposition 2.1(iii), for G-isovariant and weakly G-equivariant function $f : R^n \to R^n$, given by $f(x) = (f_1(x), \ldots, f_n(x))$, we conclude $k_1(x) = h \circ f(x) = (f_1(x) - \overline{f(x)}, \ldots, f_n(x) - \overline{f(x)})$ is a maximal invariant statistic where $\overline{f(x)} = \frac{1}{n}\sum_{i=1}^{n} f_i(x)$. In a special case, take the order statistic $f(x) = (x_{(1)}, \ldots, x_{(n)})$ and so $k_1(x) = (x_{(1)} - \bar{x}, \ldots, x_{(n)} - \bar{x})$ is a maximal invariant statistic. Furthermore, for another G-equivariant estimator as $\delta\prime_0(x) = x_n$, Lemma 2.1 implies that $(\delta\prime_0(x))^{-1}x = x - x_n e_n = (x_1 - x_n, \ldots, x_{n-1} - x_n, 0)$ and so $h\prime(x) = (x_1 - x_n, \ldots, x_{n-1} - x_n)$ is maximal invariant. Similarly, $k_1'(x) = h' \circ f(x) = (x_{(1)} - x_{(n)}, \ldots, x_{(n-1)} - x_{(n)})$ is maximal invariant.

In Lemma 2.1, and so on G acts on X and $Y = G$. Thus, for an equivariant function $\delta_0 : X \to G$ where $\delta_0(gx) = g\delta_0(x)$ for all $g \in G$, $x \in X$, we have $\delta_0(x) \in G$ and in this way, $g\delta_0(x) \in G$ means composition of g and $\delta_0(x)$, while for $\delta_0(x) \in Y$, $g\delta_0(x) \in Y$ means g acts on $\delta_0(x)$.

Example 2.1. Suppose that the action of G on G is given by conjugation. For a map $\eta : X \to G$ we have $\eta(gx) = \alpha(g)\eta(x)\alpha(g^{-1})$ for all $g \in G, x \in X$ and some $\alpha \in Aut(G)$. Hence, if $(\alpha(g))^{-1}g \in G_x$, then $f(x) = (\eta(x))^{-1}x$ is weakly equivariant because:

$$f(gx) = (\eta(gx))^{-1}gx = \alpha(g)(\eta(x))^{-1}\alpha(g^{-1})gx = \alpha(g)f(x).$$

To improve Lemma 2.1, for weakly G-equivariant function, first we give the following exa mples.

Example 2.2. Assume that G is an abelian group and there exists $\tau : X \to G$ with $\tau(gx) = \alpha(g)\tau(x)$, for all $g \in G, x \in X$. Define $f : X \to X$ by $f(x) = (\tau(x))^{-1}x$. Then

$$f(gx) = (\tau(gx))^{-1}gx = (\tau(x))^{-1}(\alpha(g))^{-1}gx = \beta(g)f(x)$$

for all $g \in G, x \in X$, where $\beta(g) = (\alpha(g))^{-1}g$, but f is not weakly equivariant because in general $\beta \notin Aut(G)$. In a special case, if τ is G- equivariant (i.e., $\alpha(g) = g$), then $\beta(g) = e$, and by Lemma 2.1, f is maximal invariant. Also, if τ is G-invariant (i.e., $\alpha(g) = e$), then $\beta = 1_G \in Aut(G)$ and in this way, f is G-equivariant.

Example 2.3. If the action of $H \lhd G$ on X is trivial and there exists $\tau : X \to G$ with $\tau(gx) = h^{-1}gh\tau(x)$, for all $g \in G$, $x \in X$ and some $h \in H$, then $f(x) = (\tau(x))^{-1}x$ is maximal invariant because since $H_x = H \lhd G$ for all $x \in X$, we have:

$$\begin{aligned} f(gx) &= (\tau(gx))^{-1}gx = (\tau(x))^{-1}h^{-1}g^{-1}hgx = (\tau(x))^{-1}h^{-1}g^{-1}gh'x \\ &= (\tau(x))^{-1}h^{-1}h'x = f(x) \end{aligned}$$

for all $g \in G, x \in X$ and some $h\prime \in H$.

We can extend Examples 2.2 and 2.3, and add some conditions for constructing maximal invariant in terms of a given weakly equivariant function.

Proposition 2.2. *Assume that there exists a weakly G-equivariant mapping $\tau : X \to G$ with $\tau(gx) = \alpha(g)\tau(x)$, $\beta(g) = (\alpha(g))^{-1}g \in G_x$ for all $g \in G$, $x \in X$ and some $\alpha \in Aut(G)$. Then $f(x) = (\tau(x))^{-1}x$ is maximal invariant.*

Proof. Since $\beta(g) = (\alpha(g))^{-1}g \in G_x$, we obtain:

$$f(gx) = (\tau(gx))^{-1}gx = (\tau(x))^{-1}(\alpha(g))^{-1}gx = f(x)$$

for all $g \in G$, $x \in X$, so f is invariant. Furthermore, f is maximal invariant (see Lemma 2.1). □

Remark 2.1. In parts of statistical inference theory it is important that group actions can be defined both on the sample space and on the parameter space, and these two types of group actions are connected. When the results are of interest for maximal invariants in the parameter space, similar to Proposition 2.2, if there exists a weakly G-equivariant mapping $\tau : \Theta \to G$ with $\tau(g\theta) = \alpha(g)\tau(\theta)$, $\beta(g) = (\alpha(g))^{-1}g \in G_\theta$ for all $g \in G$, $\theta \in \Theta$ and some $\alpha \in Aut(G)$, then $f(x) = (\tau(\theta))^{-1}\theta$ is maximal invariant. In other words, by using Proposition 2.2, we can construct a general ?maximal invariant on the sample space X based on a given weakly equivariant function from X to G. In a similar manner, Proposition 2.2 gives a method to derive a maximal invariant statistic on the parameter space Θ by using weakly equivariant functions from Θ to G.

Note that if G acts freely on X, $\beta(g) = (\alpha(g))^{-1}g \in G_x = \{e\}$ implies that $\alpha = 1_G$ and in this way, Proposition 2.2 and Lemma 2.1 ought to be coinciding because there is no weakly G-equivariant mapping $\tau : X \to G$ with $\alpha \neq 1_G$ such that $\beta(g) = (\alpha(g))^{-1}g \in G_x$ for all $g \in G, x \in X$ and some $\alpha \in Aut(G)$. Hence, when $\tau : X \to G$ is a weakly equvariant function and $G_x \neq \{e\}$, we need a general condition for invariance of $f(x) = (\tau(x))^{-1}x$ provided in Proposition 2.2, while f is maximal and hence we don't need any condition for maximality of f. Also, in a special case for G-equivariant mapping $\tau : X \to G$, we have $\tau(gx) = \alpha(g)\tau(x)$ for all $g \in G$, $x \in X$ where $\alpha = 1_G$ and so $\beta(g) = (\alpha(g))^{-1}g = e \in G_x$. Thus, by Proposition 2.2, $f(x) = (\tau(x))^{-1}x$ is maximal invariant and in this case, Proposition 2.2 and Lemma 2.1 ought to be coinciding. Hence, in Proposition 2.2, we improve Lemma 2.1 (Eaton's method) for weakly equivariant function. Now we can say in Example 2.3,

$$\beta(g) = (\alpha(g))^{-1}g = h^{-1}g^{-1}hg = h^{-1}g^{-1}gh' = h^{-1}h' \in G_x$$

for all $g \in G, x \in X$ and some $h, h\prime \in H = H_x$, thus by Proposition 2.2, $f(x) = (\tau(x))^{-1}x$ is maximal invariant. Similarly, in Example 2.2, if $\alpha = 1_G$, since $\beta(g) = (\alpha(g))^{-1}g = e \in G_x$, Proposition 2.2 implies that $f(x) = (\tau(x))^{-1}x$ is maximal invariant.

Example 2.4. Consider iid p-dimensional random vectors $x_1, \ldots, x_n$, which have a multivariate normal distribution $N_p(0, \Sigma)$. The problem considered here is the estimation of the $p \times p$ covariance matrix Σ which is assumed to be symmetric and nonsingular and unknown. Further, it is assumed that $n > p$, so that the sufficient statistic$s = \sum_{i=1}^{n} x_i x_i'$ is positive definite with probability 1. Without loss of generality, estimators of Σ are functions of s. Obviously, s has a Wishart distribution $W(\Sigma, p, n)$. It is supposed that $G = GL_n \cap S_n$

acts on $S = GL_n \cap S_n$ by $gs = \beta(g)s(\beta(g))^t$, for all $g \in G$, $s \in S$, and some $\beta \in Aut(G)$, respectively, where S_n is the set of $n \times n$ real symmetric matrices and GL_n is the set of all real invertible $n \times n$ matrices. It is easy to show that $f(s) = c\gamma(s)$ is weakly G-equivariant with $\alpha \in Aut(G)$ for some $c > 0$ where $\gamma = \beta \circ \alpha \circ \beta^{-1} \in Aut(G)$ because

$$\begin{aligned} f(gs) &= c\gamma[\beta(g)s(\beta(g))^t] = c(\beta \circ \alpha \circ \beta^{-1}[(\beta(g)])\gamma(s)(\beta \circ \alpha \circ \beta^{-1}[\beta(g^t)]) \\ &= \beta[\alpha(g)]f(s)\beta[\alpha(g^t)] = \alpha(g)f(s) \end{aligned}$$

for all $s \in S$, $g \in G$. By Proposition 2.2, if $(\alpha(g))^{-1}g \in G_s$ for all $g \in G$, $s \in S$, then $(f(s))^{-1}s = k\gamma(s^{-1})s$ is maximal invariant where $k > 0$.

The following proposition states how we can construct a general maximal invariant two given ?weakly equivariant functions. Notice that the result in Proposition 2.2 holds even if the condition $(\alpha(g))^{-1}g \in G_x$ for all $g \in G$, $x \in X$ is omitted. Instead of using this condition, one could use two weakly G-equivariant functions to obtain a maximal invariant function.

Proposition 2.3. *Assume that there exist two weakly* G-*equivariant functions* τ_i, $i = 1, 2$ *with the same* $\alpha \in Aut(G)$. *Then:*

(i) the function $f : X \to G$ *defined by* $f(x) = (\tau_1(x))^{-1}\tau_2(x)$ *is invariant and*
(ii) f *is maximal invariant if at least one of the* $\tau_i : X \to G$ *is weakly* G-*isovariant.*

Proof.

(i) f is invariant because

$$f(gx) = (\tau_1(gx))^{-1}\tau_2(gx) = (\tau_1(x))^{-1}(\alpha(g))^{-1}\alpha(g)\tau_2(x) = f(x).$$

(ii) Assume that one of the two, say τ_2 weakly G-isovariant. Let $x \in X$ and suppose that $g_1, g_2 \in G$. Since τ_2 is weakly G-isovariant and weakly G-equivariant, Proposition 1.1 (ii) implies that τ_2 is one-to-one on the orbit Gx. By part (i), it is enough to show that f is maximal. To see this, suppose that

$$f(x_1) = (\tau_1(x_1))^{-1}\tau_2(x_1) = (\tau_1(x_2))^{-1}\tau_2(x_2) = f(x_2),$$

since $\alpha \in Aut(G)$ is onto and $\tau_1(x_1), \tau_1(x_2) \in G$, there exist $g_1, g_2 \in G$ such that $(\tau_1(x_1))^{-1} = \alpha(g_1)$, $(\tau_1(x_2))^{-1} = \alpha(g_2)$, and so $\alpha(g_1)\tau_2(x_1) = \alpha(g_2)\tau_2(x_2)$. Using the fact that τ_2 is weakly G-equivariant, thus $\tau_2(g_1x_1) = \tau_2(g_2x_2)$. On the other hand, τ_2 is one-to-one on each orbit Gx, and hence $x_1 = gx_2$ where $g = g_1^{-1}g_2 \in G$, and the result follows.

□

Example 2.5. Let (x_1, x_2) be a single random variable with density $(\sigma_1 e^{-\sigma_1 x_1})$ $(\sigma_2 e^{-\sigma_2 x_2})$ where $\sigma_1, \sigma_2 > 0$ and $x_1, x_2 > 0$. Suppose that $G = R^+$ acts on $X = \{(x_1, x_2) : x, y \in R^+\}$, $Y = R^+$ by $g \times (x_1, x_2) = (gx_1, gx_2)$, $g \otimes y = g^n y$ for all $g \in G$, $(x_1, x_2) \in X$, $y \in Y$ and some integer n, respectively. (Since $G(x_1, x_2) = X$ and $G_{(x_1,x_2)} = \{1\}$ for all $(x_1, x_2) \in X$, we can say the action of G on X is sharply transitive.)

The functions $\tau_i : X \to G, i = 1, 2$, given by $\tau_1(x_1, x_2) = x_1^m x_2^m$, $\tau_2(x_1, x_2) = x_1^{2m-1}x_2$ are two weakly G-equivariant functions with the same $\alpha(g) = \sqrt[n]{g^{2m}}$ where m is an even integer, respectively. For example, we can see that τ_1 is weakly G-equivariant because

$$\tau_1(g(x_1, x_2)) = \tau_1((gx_1, gx_2)) = g^{2m} x_1^m x_2^m = (\alpha(g))^n \tau_1(x_1, x_2) = \alpha(g) \otimes \tau_1(x_1, x_2)$$

for all $g \in G$, $(x_1, x_2) \in X$. Proposition 2.3(i) implies that

$$f(x_1, x_2) = (\tau_1(x_1, x_2))^{-1} \tau_2(x_1, x_2) = x_1^{2m-1} x_2 / x_1^m x_2^m = (x_1/x_2)^{m-1}$$

is invariant. By the way $\alpha(G_{(x_1,x_2)}) = G_{\tau_1(x_1,x_2)} = \{1\}$ and so τ_1 is weakly G-isovariant. Hence, by Proposition 2.3(ii), $f(x_1, x_2) = (x_1/x_2)^{m-1}$ is maximal invariant.

Notice that $(\alpha(g))^{-1} g = \sqrt[n]{g^{n-2m}} \notin G_{(x_1,x_2)} = \{1\}$ for all $g \in G$, $(x_1, x_2) \in X$, hence the condition in Proposition 2.2 does not hold. In this example, since G acts freely on X, we thus can not use this proposition for constructing a maximal invariant statistic.

Proposition 2.3 can be improved by using more weakly G-equivariant functions for constructing a maximal invariant statistics as follows.

Corollary 2.1. *Assume that $\tau_i : X \to G$, $i = 1, \ldots, 2n$ are weakly G-equivariant functions with the same $\alpha \in Aut(G)$. Then the function $f : X \to G$ defined by $f(x) = \prod_{i=1}^n (\tau_{2i-1}(x))^{-1} \tau_{2i}(x)$ is maximal invariant if at least one of the τ_1 or τ_{2n} is weakly G-isovariant.*

Proof. Proposition 2.3(i) implies that $f_i(x) = (\tau_{2i-1}(x))^{-1} \tau_{2i}(x)$, $i = 1, \ldots, n$ are invariant, thus $f(x) = \prod_{i=1}^n f_i(x)$ is invariant too. Assume that one of the two, say or $\tau_{2n}$is weakly G-isovariant.

By Proposition 1.1(ii), τ_{2n} is one-to-one on the orbit Gx. Now, if $f(x_1) = f(x_2)$, then

$$\prod_{i=1}^{n-1} f_i(x_1)(\tau_{2n-1}(x_1))^{-1} \tau_{2n}(x_1) = \prod_{i=1}^{n-1} f_i(x_2)(\tau_{2n-1}(x_2))^{-1} \tau_{2n}(x_2).$$

But $\alpha \in Aut(G)$ is onto and so there exist $g_1, g_2 \in G$ such that

$$\prod_{i=1}^{n-1} f_i(x_j)(\tau_{2n-1}(x_j))^{-1} = \alpha(g_j), j = 1, 2$$

and then $\alpha(g_1)\tau_{2n}(x_1) = \alpha(g_2)\tau_{2n}(x_2)$. Using the fact that τ_{2n} is weakly G-equivariant and one-to-one on each orbit Gx, the result follows, similar to the proof of Proposition 2.3(ii). □

We illustrate above corollary by the following examples.

Example 2.6. Let $x_1, \ldots, x_n$ be iid normal distribution $N(\theta, \sigma^2)$, where $\sigma^2 > 0$ is known and θ is unknown. Suppose that $G = R$ acts on $X = R^n$, by $g \times x = (x_1 + g, \ldots, x_n + g)$, $g \otimes y = y + g$ for all $g \in G$, $x = (x_1, \ldots, x_n) \in X$, $y \in Y$, respectively. The functions $\tau_i : X \to G$, given by $\tau_i(x_1, \ldots, x_n) = \sum_{j=1}^n m_{i,j} x_j$, $i = 1, \ldots, 2n$, are weakly G-equivariant functions with the same $\alpha(g) = (\sum_{j=1}^n m_{i,j})g$, where $m_{i,j}$'s, $i = 1, \ldots, 2n$, $j = 1, \ldots, n$ are chosen to satisfy $\sum_{j=1}^n m_{i,j} = c$, for all $i = 1, \ldots, 2n$ and some integer c. Corollary 2.1 implies that

$$f(x_1, \ldots, x_n) = \sum_{i=1}^n (\tau_{2i}(x_1, \ldots, x_n) - \tau_{2i-1}(x_1, \ldots, x_n))$$

$$= \sum_{i=1}^{n}\left(\sum_{j=1}^{n}[m_{2i,j} - m_{2i-1,j}]x_j\right) = \sum_{j=1}^{n}\omega_j x_j$$

is invariant where $\omega_j = \sum_{i=1}^{n}(m_{2i,j} - m_{2i-1,j})$ for all $j = 1, \ldots, n$. Notice that $(\alpha(g))^{-1}g = g - (\sum_{j=1}^{n} m_{i,j})g \in G_{(x_1,\ldots,x_2)} = \{0\}$ for all $g \in G$, $(x_1, \ldots, x_2) \in X$ iff $\sum_{j=1}^{n} m_{i,j} = 1$, and then in this case, the condition in Proposition 2.2 hold the result follows. Thus, Proposition 2.2 implies that

$$h(x_1, \ldots, x_n) = \left(\sum_{j=1}^{n} m_{i,j}x_j\right)^{-1} \times (x_1, \ldots, x_n)$$

$$= \left(x_1 - \sum_{j=1}^{n} m_{i,j}x_j, \ldots, x_n - \sum_{j=1}^{n} m_{i,j}x_j\right)$$

is a maximal invariant statistic provided $\sum_{j=1}^{n} m_{i,j} = 1$. In the general case of example 2.6 suppose that

$$G = \left\{(g_1, \ldots, g_n) \in (R^+)^n : \prod_{i=1}^{n/2} g_{2i-1}/g_{2i} = 1\right\}$$

acts on $X = \{(x_1, \ldots, x_n) \in (R^+)^n : \prod_{i=1}^{n/2} x_{2i-1}/x_{2i} = 1\}$, $Y = R^+$ by $g \times x = (g_1x_1, \ldots, g_nx_n), g \otimes y = \prod_{j=1}^{n} g_j^{m_{i,j}} y$ for all $g = (g_1, \ldots, g_n) \in G, x = (x_1, \ldots, x_n) \in X$, $y \in Y$ and some integers $m_{i,j}$ where $i = 1, \ldots, 2n$, $j = 1, \ldots, n$, respectively, s.t. n is an even integer and $\pi_i = \prod_{j=1}^{n} g_j^{m_{i,j}} = \pi$ for all $i = 1, \ldots, 2n$ and some positive numberπ. The functions $\tau_i : X \to G$, given by $\tau_i(x_1, \ldots, x_n) = \prod_{j=1}^{n} x_j^{m_{i,j}}, i = 1, \ldots, 2n$, are G-equivariant functions. Corollary 2.1 implies that $f(x_1, \ldots, x_n) = \prod_{j=1}^{n} x_j^{\omega_j} = \prod_{j=1}^{n/2}(x_{2j-1}/x_{2j})^k$ is invariant where $\omega_j = \sum_{i=1}^{n}(m_{2i,j} - m_{2i-1,j})$ for all $j = 1, \ldots, n$ and $\omega_{2j} = -k$, $\omega_{2j-1} = k$ for all $j = 1, \ldots, n/2$ and some odd integer k. In this case, $f(x_1, \ldots, x_n) = \prod_{j=1}^{n/2}(x_{2j-1}/x_{2j})^k$ is maximal invariant. This can also be seen directly from the fact that if $f(x_1, \ldots, x_n) = f(x'_1, \ldots, x'_n)$, there exist $g_j = x_j/x\prime_j$, for all $j = 1, \ldots, n$ where $(x_1, \ldots, x_n) = g(x'_1, \ldots, x'_n)$ s.t. $g = (g_1^{m_{i,1}}, \ldots, g_n^{m_{i,n}}) \in G$ and

$$m_{i,j} = \begin{cases} 1 & j = 1, 3, \ldots, n-1 \\ -1 & j = 2, 4, \ldots, n \end{cases}$$

for all $i = 1, \ldots, 2n$.

Example 2.7. Let x_i's be iid according to the exponential distribution with density $\sigma e^{-\sigma x_i}$, $i = 1, \ldots, n$, where $\sigma > 0$ and $x_i > 0$. Suppose that $G = R^+$ acts on $X = (R^+)^n$, $Y = R^+$ by $g \times x = g(x_1, \ldots, x_n) = (gx_1, \ldots, gx_n)$, $g \otimes y = gy$ for all $g \in G$, $x = (x_1, \ldots, x_n) \in X$, $y \in Y$, respectively. The functions $\tau_i : X \to G$, given by $\tau_i(x_1, \ldots, x_n) = \prod_{j=1}^{n} x_j^{m_{i,j}}$, $i = 1, \ldots, 2n$, are weakly G-equivariant functions with the same $\alpha(g) = g^{\sum_{j=1}^{n} m_{i,j}}$ where $m_{i,j}$'s, $i = 1, \ldots, 2n$, $j = 1, \ldots, n$ are chosen to satisfy

$\sum_{j=1}^{n} m_{i,j} = t$, for all $i = 1, \ldots, 2n$ and some odd integer t. Corollary 2.1 implies that

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} \tau_{2i}(x_1, \ldots, x_n)/\tau_{2i-1}(x_1, \ldots, x_n) = \prod_{j=1}^{n} x_j^{\sum_{i=1}^{n}(m_{2i,j}-m_{2i-1,j})}$$

is invariant. Assume that n is an even integer. Let $\omega_j = \sum_{i=1}^{n}(m_{2i,j} - m_{2i-1,j})$ for all $j = 1, \ldots, n$. If $\omega_{2j} = -k$, $\omega_{2j-1} = k$ for all $j = 1, \ldots, n/2$ and some odd integer k, we conclude that $f(x_1, \ldots, x_n) = \prod_{j=1}^{n} x_j^{\omega_j} = \prod_{j=1}^{n/2}(x_{2j-1}/x_{2j})^k$ is invariant. Also, $f(x_1, \ldots, x_n) = \prod_{j=1}^{n/2}(x_{2j-1}/x_{2j})^k$ is maximal invariant if $n = 2$. But $(\alpha(g))^{-1}g = g/g^{\sum_{j=1}^{n} m_{i,j}} \in G_{(x_1,\ldots,x_2)} = \{1\}$ for all $g \in G$, $(x_1, \ldots, x_2) \in X$ iff $\sum_{j=1}^{n} m_{i,j} = 1$. Therefore, Proposition 2.2 implies that

$$h(x_1, \ldots, x_n) = \left(\prod_{j=1}^{n} x_j^{m_{i,j}}\right)^{-1} \times (x_1, \ldots, x_n) = \left(x_1 \Big/ \prod_{j=1}^{n} x_j^{m_{i,j}}, \ldots, x_n \Big/ \prod_{j=1}^{n} x_j^{m_{i,j}}\right)$$

is another maximal invariant statistic provided $\sum_{j=1}^{n} m_{i,j} = 1$.

3. Maximal Invariant Under Sharply Transitive Groups

In this section, we limit ourselves to sharply transitive groups. In this class, we can consider the parameter space as a group with a new binary action. Based on this concept, it is possible to change weakly G-equivariant functions into weakly G-equivariant estimators and so we can find maximal invariant functions by using weakly equivariant estimators. In a special case, we derive maximal invariant statistic when the group contains a sharply transitive and a characteristic subgroup. Furthermore, Proposition 2.2 immediately extends to the case of weakly equivariant estimators. Afterwards ?we offer a simple way for constructing a maximal invariant function based on two given weakly G-equivariant estimators and improve Proposition 2.3 for weakly G-equivariant estimators.

Let G be a group and Θ a G-space. When G is sharply transitive on Θ, we may index G by Θ. (see Remark 1.2). We can pick one arbitrary basis point $\theta_0 \in \Theta$ and $\alpha \in Aut(G)$, then write every element $\theta \in \Theta$ in a unique way as $\alpha(g_\theta)\theta_0 = \theta$. Clearly, e corresponds to θ_0. Since G is sharply transitive on Θ, by Remark 1.1, G acts freely on Θ and in this way, the identity $\alpha(g_{h\theta})\theta_0 = h\theta = h\alpha(g_\theta)\theta_0$ implies that $\alpha(g_{h\theta}) = h\alpha(g_\theta)$ for all $\theta \in \Theta, h \in G$. Also, we can say $\eta_\alpha : \Theta \to G$, given by $\eta_\alpha(\theta) = \alpha(g_\theta)$, is a bijective G-equivariant mapping because $\eta_\alpha(h\theta) = \alpha(g_{h\theta}) = h\alpha(g_\theta) = h\eta_\alpha(\theta)$ for all $\theta \in \Theta, h \in G$. Similarly, $\eta(\theta) = g_\theta = \alpha^{-1}(\eta_\alpha(\theta))$ is a bijective and weakly G-equivariant mapping with $\alpha^{-1} \in Aut(G)$. Since G acts transitively on Θ, the only (maximal) invariant functions are constant which, by Lemma 2.1, are equal to $f(\theta) = (\eta(\theta))^{-1}\theta = (\alpha(g_\theta))^{-1}\theta = \theta_0$.

Example 3.1. Let Θ be the class of all increasing and continuous cdf's and G be the class of all strictly increasing bijective continuous maps acting coordinatewise on X. G acts on Θ by $g.F = F \circ g^{-1}$. Now, G is sharply transitive on Θ, because for $F_1, F_2 \in \Theta$, the transformation g given by $g(x) = F_1^{-1}F_2(x)$ is the unique member of g satisfying $g.F_1 = F_2$. Fix an arbitrary point $\theta_0 = F_0 \in \Theta$ and define g_F to be the unique $g \in G$ satisfying $g.F_o = F$, for all $F \in \Theta$. Thus, $g_F = F^{-1}F_0$.

Let **P** be the family of all continuous distributions on the real line having unique medians. The group G of real-valued bijective increasing functions on R acts coordinatewise on $X = R^n$. The group action on X induces a group action on $\Theta = R$. For $P \in \mathbf{P}$, let $\tau(P) = F_P^{-1}(1/2)$ where F_P is the cdf associated with P. $\theta(P)$ is thus the median associated with P. The function $\tau : \mathbf{P} \to R$ is G-equivariant because:

$$\tau(gP) = F_{gP}^{-1}(1/2) = (g.F_P)^{-1}(1/2) = (F_P \circ g^{-1})^{-1}(1/2) = gF_P^{-1}(1/2) = g\tau(P),$$

for all $g \in G$, $P \in \mathbf{P}$. In fact, Berk (1967) proved that equivariant estimators of medians (or any other fractile) must be order statistics.

Definition 3.1 (Robinson, 1995). A subgroup H of a group G is said to be characteristic in G if $\alpha(H) \leq H$ for all $\alpha \in Aut(G)$. Notice that if H is characteristic in G and $\alpha \in Aut(G)$, then $\alpha(H) = H$ since $\alpha(H) \leq H$ and $\alpha^{-1}(H) \leq H$.

Obviously, characteristic subgroupbs are normal. To see this, take $\alpha \in Aut(G)$ such that $\alpha(g) = g_0 g g_0^{-1}$ for all $g_0 \in G$, hence $H = \alpha(H) = \{\alpha(h) : h \in H\} = \{g_0 h g_0^{-1} : h \in H\} = g_0 H g_0^{-1}$ for all $g_0 \in G$. It follows that $H \lhd G$.

Lemma 3.1. *Assume that G contains a characteristic subgroup H which is sharply transitive on Θ. Fix an arbitrary point $\theta_0 \in \Theta$ and define h_θ to be the unique $h \in H$ satisfying $\alpha(h)\theta_0 = \theta$ where $\alpha \in Aut(G)$. Then:*

(i) $h_{\alpha(g)\theta} = g h_\theta g^{-1}$ for all $g \in \alpha^{-1}(G_{\theta_\circ})$, $\theta \in \Theta$ and some $h_\theta \in H$,
(ii) $h_{g\theta} = \alpha^{-1}(g) h_\theta \alpha^{-1}(g^{-1})$ for all $g \in G_{\theta_\circ}$, $\theta \in \Theta$ and some $h_\theta \in H$, and
(iii) $\eta : \Theta \to H$, given by $\eta(\theta) = \alpha(h_\theta)$, is weakly $G_{\theta_\circ}-$equivariant.

Proof.

(i) For all $g \in G$, let $\theta = \alpha(g)\theta_0$. Since H is sharply transitive on Θ, there is a unique $h_\theta \in H$ s.t. $\theta = \alpha(h_\theta)\theta_0$. Thus, we have $g_\circ = \alpha(h_\theta^{-1} g) \in G_{\theta_\circ}$ and hence, since H is characteristic in G, we conclude that $g = h_\theta \alpha^{-1}(g_\circ) \in H\alpha^{-1}(G_{\theta_\circ}) = \alpha^{-1}(HG_{\theta_\circ})$. Therefore, $G = HG_{\theta_\circ}$, that is H and G_{θ_0} generate G. For any $g \in \alpha^{-1}(G_{\theta_\circ})$, the identity $\alpha(h_{\alpha(g)\theta})\theta_\circ = \alpha(g)\theta = \alpha(g)\alpha(h_\theta)\theta_\circ$ shows that $g_\circ = h_{\alpha(g)\theta}^{-1} g h_\theta \in \alpha^{-1}(G_{\theta_\circ})$. Since $H \lhd G$, we can find $h' \in H$ such that $gh_\theta = h_{\alpha(g)\theta} g_\circ = g_\circ h'$. Thus, $g_\circ^{-1} g = h' h_\theta^{-1} \in H \cap \alpha^{-1}(G_{\theta_\circ})$, i.e., $\alpha(g_\circ^{-1} g)\theta_\circ = \theta_\circ$ such that $g_\circ^{-1} g \in H$. But H is sharply transitive and characteristic and also α is one-to-one. It follows that $g = g_\circ = h_{\alpha(g)\theta}^{-1} g h_\theta$ and then $h_{\alpha(g)\theta} = g h_\theta g^{-1}$ for all $g \in \alpha^{-1}(G_{\theta_\circ})$, $\theta \in \Theta$ and some $h_\theta \in H$.

(ii) Substituting $g_o = \alpha(g) \in G_{\theta_0}$ into (i), yields (ii).

(iii) By part (ii), we have $\eta(g\theta) = \alpha(h_{g\theta}) = g\eta(\theta)g^{-1}$ for all $g \in G_{\theta_0}$, $\theta \in \Theta$, where the action of $G_{\theta_\circ}$ on H is given by conjugation. □

An immediate consequence of Lemma 3.1 is the following theorem. Here, we find a maximal invariant statistic when the group contains a sharply transitive and a characteristic subgroup.

Theorem 3.1. *Assume that G contains a characteristic subgroup H which is sharply transitive on Θ. Furthermore, if*

(i) $\alpha(h_0) = h_0$ for all $h_0 \in H$,
(ii) there exists a weakly $G_{\theta_\circ}-$equivariant mapping $\tau_0 : \Theta \to G_{\theta_0}$ with $\alpha \in Aut(G)$, s.t. $(\alpha(g_o))^{-1} g_o \in G_\theta$ for all $g_o \in G_{\theta_0}$, $\theta \in \Theta$, and

(iii) $\tau : \Theta \to \Theta$ *is an* $G_{\theta_\circ}$*−equivariant function where* τ*is one-to-one on each orbit* $H\theta$*, then* $f'(\theta) = (\tau_0 \circ \tau(h_\theta^{-1}\theta)))^{-1}\tau(h_\theta^{-1}\theta)$ *is maximal invariant on* Θ*, where* $h_\theta \in H$ *is a unique element of* H *such that* $\alpha(h_\theta)\theta_\circ = \theta$*, for fixed* $\theta_0 \in \Theta$*,* $\alpha \in Aut(G)$*.*

Proof. Let $\theta_0 \in \Theta$, $\alpha \in Aut(G)$. By (i) and Remark 2.1, $f_0 : \Theta \to \Theta$, given by $f_0(\theta) = (\tau_0(\theta))^{-1}\theta$, is maximal $G_{\theta_\circ}$−invariant. Define $f_H : \Theta \to \Theta$ by $f_H(\theta) = \tau(\alpha(h_\theta^{-1})\theta)$ where $h_\theta \in H$ is a unique element of H such that $\alpha(h_\theta)\theta_\circ = \theta$. Use Lemma 3.1(ii) to verify

$$f_H(g\theta) = \tau\big(\alpha(h_{g\theta}^{-1})g\theta\big) = \tau(g\alpha(h_\theta^{-1})g^{-1}g\theta) = gf_H(\theta)$$

for all $g \in G_{\theta_0}$, $\theta \in \Theta$. It follows that f_H is $G_{\theta_\circ}$−equivariant. Clearly,

$$f_H(h'\theta) = \tau\big(\alpha(h_{h'\theta}^{-1})h'\theta\big) = \tau\big(\alpha(h_\theta^{-1})(h')^{-1}h'\theta\big) = f_H(\theta)$$

for all $h' \in H$, $\theta \in \Theta$. This proves is H-invariant. It's easy to show that f_H is maximal H-invariant provided τ is one-to-one on each orbit $H\theta$. Now, define $f' : \Theta \to \Theta$ by $f' = f_0 \circ f_H$. Since Lemma 3.1 implies that $G = HG_{\theta_\circ}$, for all $g \in G$, we can choose $g_o \in G_{\theta_0}$ and $h_\theta \in H$ that $g = h_\theta g_0$ and in this way

$$f'(g\theta) = f_0 \circ f_H(h_\theta g_0\theta) = f_0 \circ f_H(g_0\theta) = f_0(g_0 f_H(\theta)) = f_0(f_H(\theta)) = f'(\theta)$$

for all $g \in G$, $\theta \in \Theta$. This proves f' is G-invariant. Now, if

$$f'(\theta_1) = f_0(f_H(\theta_1)) = f_0(f_H(\theta_2)) = f'(\theta_2),$$

it follows from f_0 being maximal $G_{\theta_\circ}$−invariant and f_H being $G_{\theta_\circ}$−equivariant, that

$$f_H(\theta_1) = g_0 f_H(\theta_2) = f_H(g_0\theta_2),$$

for some $g_0 \in G_{\theta_0}$. But f_H is maximal H-invariant if τ is one-to-one on $H\theta$ for all $\theta \in \Theta$. Thus, there exists $h_\theta \in H$ s.t. $\theta_1 = h_\theta g_0\theta_2$. Consequently, $\theta_1 = g\theta_2$ for some $g \in G$, and in this way,

$$\begin{aligned} f'(\theta) &= f_0\big[\tau\big(\alpha\big(h_\theta^{-1}\big)\theta\big)\big] = \big[\tau_0 \circ \tau\left(\alpha(h_\theta^{-1})\theta\right)\big]^{-1} \tau\left(\alpha(h_\theta^{-1})\theta\right) \\ &= \big[\gamma\left(\alpha(h_\theta^{-1})\theta\right)\big]^{-1} \tau\left(\alpha(h_\theta^{-1})\theta\right) \end{aligned}$$

is maximal G-invariant where $\gamma : \Theta \to G_{\theta_0}$, given by $\gamma = \tau_0 \circ \tau$, is weakly $G_{\theta_\circ}$−equivariant, such that $\gamma(g_o x) = \alpha(g_o)\gamma(x)$ for all $g_o \in G_{\theta_0}$, $x \in X$ that $\tau : \Theta \to \Theta$ is one-to-one on each orbit $H\theta$ and $G_{\theta_\circ}$−equivariant where $(\alpha(g_o))^{-1}g_o \in G_\theta$ for all $g_o \in G_{\theta_0}$, $\theta \in \Theta$.

Similar to the proof of Theorem 3.1, suppose that G contains a characteristic subgroup H which is sharply transitive on Θ, for fixed $\theta_0 \in \Theta$. If there exist $f_0 : \Theta \to \Theta$ which is maximal $G_{\theta_\circ}$−invariant and $f_H : \Theta \to \Theta$ which is $G_{\theta_\circ}$−equivariant and maximal H-invariant, then $f' : \Theta \to \Theta$, given by $f' = f_0 \circ f_H$, is maximal G-invariant. Theorem 3.1, as in the case $\alpha = 1_G \in Aut(G)$ is trivial, which is illustrated by the following example. □

Example 3.2. Let $G = \{diag(a_1, \ldots, a_n) : a_1, \ldots, a_n \neq 0\} \subset GL_n$ be the group of $n \times n$ diagonal matrices where$diag(a_1, ..., a_n)$ is an $n \times n$ diagonal matrix with positive diagonal elements $(a_1, ..., a_n)$. Assume that G acts on $\Theta = \{\theta = (\theta, ..., \theta)^t : \theta \in R\}$ by $(g, \theta) \mapsto g\theta$

for all $g \in G$, $\theta \in \Theta$. (This is usefull when we have i.i.d. p-dimensional random vectors $x_1, \ldots, x_n$ distributed as $N_p(\theta, \Sigma)$ with unknown $\theta = (\theta, \ldots, \theta)^t \in \Theta$ and known Σ.) Clearly, G contains a characteristic subgroup $H = \{diag(a, \ldots, a) : a \neq 0\} \leq G$ which is sharply transitive on Θ. Fix a reference point $\theta_0 = (\theta_0, \ldots, \theta_0)^t \in \Theta$ and index G by Θ and for all $\theta \in \Theta$, define h_θ to be the unique $h \in H$ satisfying $\alpha(h_\theta)\theta_0 = \theta$ with $\alpha = 1_G \in Aut(G)$. Thus, we have $\alpha(h_0) = h_0$ for all $h_0 \in H$. The stabilizer subgroup G_{θ_0} is given by $G_{\theta_0} = \{I\}$, where I is the $n \times n$ identity matrix, which is the identity element of the group. It then follows that G is free and hence, by Remark 1.1, G is sharply transitive on Θ. Suppose that $\tau_0 : \Theta \to G_{\theta_0}$is a weakly $G_{\theta_0}-$equivariant function with $\alpha \in Aut(G)$ and $\tau : \Theta \to \Theta$ is an $G_{\theta_0}-$equivariant function which is one-to-one on each orbit $H\theta = \{a\theta = (a\theta, \ldots, a\theta)^t : a \neq 0\} = \Theta$, for all $\theta \in \Theta$. Theorem 3.1 implies that $f'(\theta) = (\tau_0 \circ \tau(\theta_0)))^{-1}\tau(\theta_0) = \tau(\theta_0)$ is maximal invariant.

Definition 3.2 (Robinson, 1995). In a group G, for $G_1, G_2 \leq G$ the subgroup

$$[G_1, G_2] = \{[g_1, g_2] = g_1 g_2 g_1^{-1} g_2^{-1} : g_1 \in G_1, g_2 \in G_2\} \leq G$$

is called commutator G_1 and G_2. The subgroup of G generated by all the commutators in G is called the derived subgroup of G. It is commonly denoted by $G' = [G, G]$. Alternatively, one may define G' as the smallest subgroup that contains all the commutators.

Remark 3.1. It is easy to see that the derived subgroup G' is characteristic in G. Because for all $\alpha \in Aut(G)$:

$$\begin{aligned}\alpha(G') &= \alpha([G, G]) = \{\alpha[g_1, g_2] : g_1, g_2 \in G\} \\ &= \{\alpha(g_1)\alpha(g_2)(\alpha(g_1))^{-1}(\alpha(g_2)))^{-1} : g_1, g_2 \in G\} \\ &= [\alpha(G), \alpha(G)] = [G, G] = G'.\end{aligned}$$

The following example provides a few illustrations and additional comments.

Example 3.3. It is supposed that

$$G = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} : a \neq 0 \right\} \subset GL_2$$

acts on $\Theta = \{\theta = (\theta', c)^t : \theta' \in R\}$ by $(g, \theta) \mapsto g\theta$ for all $g \in G, \theta \in \Theta$ where $c \neq 0$. (This is usefull when we have i.i.d. p-dimensional random vectors $x_1, \ldots, x_n$ with multivariate normal distribution $N_p(\theta, \Sigma)$, where the mean vector is $\theta \in \Theta$, s.t. $\theta'$is unknown and c is known and the $p \times p$ covariance matrix Σ is non singular and known.) Clearly, the derived subgroup of G is

$$G' = \left\{ \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} : b \in R \right\}$$

which is characteristic in G (see Remark 3.1). Also, G contains a sharply transitive and a characteristic subgroup $H = G'$. Let $\theta_0 = (\theta_0', c)^t \in \Theta$ be a fixed point where $\theta_0' = 0$ and suppose that

$$\tau_0 : \Theta \to G_{\theta_0},$$
$$(\theta', c)^t \mapsto \begin{bmatrix} k_o\theta' & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$\tau : \Theta \to \Theta,$$
$$(\theta', c)^t \mapsto (k\theta', c)^t$$

are $G_{\theta_0}-$equivariant functions for some $k_0, k \neq 0$ s.t. the stabilizer subgroup G_{θ_0} is given by

$$G_{\theta_0} = \left\{ \begin{bmatrix} a & (1-a)\theta_0'/c \\ 0 & 1 \end{bmatrix} : a \neq 0 \right\} = \left\{ \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} : a \neq 0 \right\}.$$

Notice that τ is one-to-one on each orbit

$$H\theta = \left\{ \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta' \\ c \end{bmatrix} = \begin{bmatrix} \theta' + bc \\ c \end{bmatrix} : b \in R \right\} = \Theta$$

for all $\theta = (\theta', c)^t \in \Theta$. Now, define $f_H : \Theta \to \Theta$ by $f_H(\theta) = \tau(\alpha(h_\theta^{-1})\theta) = \tau(\theta_0)$, where

$$h_\theta = \begin{bmatrix} 1 & (\theta_0' - \theta')/c \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\theta'/c \\ 0 & 1 \end{bmatrix} \in H$$

is a unique element of H such that $\alpha(h_\theta)\theta_\circ = \theta$, for all $\theta = (\theta', c)^t \in \Theta$ and some $\alpha \in Aut(G)$ given by

$$\alpha\left(\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \right) = \begin{bmatrix} 1 & -b \\ 0 & 1 \end{bmatrix}.$$

Clearly, f_H is $G_{\theta_\circ}-$equivariant and maximal H-invariant and thus Theorem 3.1 implies that $f'(\theta) = f_0(f_H(\theta)) = (\tau_0(\tau(\theta_0)))^{-1}\tau(\theta_0)$ is maximal invariant on Θ.

When G is sharply transitive on Θ, by Lemma 1.1, since $G_\theta = \{e\}$, $G\theta = \Theta$, we conclude that $\lambda : G \to \Theta$, given by $g \mapsto \alpha(g)\theta_0$, is a homeomorphism where $\alpha \in Aut(G)$ and $\theta_0 \in \Theta$ is a fixed point. Thus the group elements correspond to elements of a parameter space Θ. The action of G on X requires that $(\Theta, *)$ is a group, with binary operation $*$ which $g_\theta g_\omega = g_{\theta * \omega}$. Since $\lambda(e) = \alpha(e)\theta_0 = \theta_0$, we have e which corresponds to θ_0. Hence, if we define $g_\theta \in G$ as a unique element of G such that $\alpha(g_\theta)\theta_0 = \theta$, for some $\alpha \in Aut(G)$, we can rewrite

$$\lambda(g_\theta) = \theta, \theta * \omega = \alpha(g_\theta)\alpha(g_\omega)\theta_0 = \alpha(\lambda^{-1}(\theta)\lambda^{-1}(\omega))\lambda(e).$$

Thus, $(\Theta, *)$ is a group with identity element $\theta_0 = \lambda(e)$ and inverse element $\theta^{-1} = \alpha(g_\theta^{-1})\theta_0$. The group action on X induces a group action on Θ, such that for each $g_\theta \in G$, there is a $\alpha(g_\theta) \in G$ for some $\alpha \in Aut(G)$, satisfying $\alpha(g_\theta)\omega = \theta * \omega$.

Example 3.4. Let $x_1, \ldots, x_n$ be iid from $N(\theta, \theta^2)$, where $\theta > 0$ and $x, y > 0$. Clearly, $(z_1, z_2) = (\bar{x}, \sum_{i=1}^n (x_i - \bar{x})^2)$ is sufficient for $\theta > 0$. Suppose $G = R^+$ acts on $Z = R \times R^+$ by $g \times (z_1, z_2) = (gz_1, g^2 z_2)$. Define the G-equivariant estimator $\delta_0 : Z \to G$ by $\delta_0(z_1, z_2) = \sqrt{z_2}$, then by Lemma 2.1, $f(z_1, z_2) = (\delta_0(z_1, z_2))^{-1} \times (z_1, z_2) = (z_1/\sqrt{z_2}, 1)$

and so $h(z_1, z_2) = z_1/\sqrt{z_2}$ is maximal invariant. Now, $\delta_1 : Z \to G$, given by $\delta_1(z_1, z_2) = z_2^{m_1}$, is weakly G-equivariant with $\alpha_1(g) = g^{2m_1}$, for some $m_1 \in Q$ because:

$$\delta_1(g \times (z_1, z_2)) = \delta_1(gz_1, g^2 z_2) = g^{2m_1} z_2^{m_1} = \alpha_1(g)\delta_1(z_1, z_2).$$

But G is sharply transitive on $\Theta = \{(\theta_1, \theta_2) : \theta_1 = \theta, \theta_2 = \theta^2\}$ and the action is given by $(g, (\theta_1, \theta_2)) \mapsto (g\theta_1, g^2\theta_2)$ and in this way, we can define $g_{(\theta_1,\theta_2)}$ as the unique $g \in G$ satisfying $(\alpha(g), (1, 1)) \mapsto (\theta_1, \theta_2)$ for some $\alpha \in Aut(G)$. Thus,

$$\alpha(g_{(\theta_1,\theta_2)}) = \theta_1 = \sqrt{\theta_2} = \theta,\ \lambda(g_{(\theta_1,\theta_2)}) = (\theta_1, \theta_1^2).$$

Let $\alpha(g) = g^{2m}$, for some $m \in Q$, so $g_{(\theta_1,\theta_2)} = \sqrt[2m]{\theta_1}$. Also, let $(\Theta, *)$ be a group such that

$$(\theta_1, \theta_2) * (\omega_1, \omega_2) = \lambda(g_{(\theta_1,\theta_2)} g_{(\omega_1,\omega_2)}) = \lambda\left(\sqrt[2m]{\theta_2}\sqrt[2m]{\omega_2}\right) = (\theta_2\omega_2, (\theta_2\omega_2)^2).$$

An immediate concequence is Corollary 3.1. In this corollary, we construct maximal invariant based on a given weakly equivariant estimator when the group is sharply transitive on the parameter space.

Corollary 3.1. *Assume that G is sharply transitive on Θ. Let $\theta_0 \in \Theta$ be a fixed point and write every element $\theta \in \Theta$ in a unique way as $\alpha(g)\theta_0 = \theta$, where $\alpha \in Aut(G)$. The function $\lambda : G \to \Theta$, given by $\lambda(g_\theta) = \theta$, is weakly G-equivariant. Also, if there exists a weakly equivariant estimator $\tau : X \to \Theta$ with $\tau(g_\theta x) = \alpha(g_\theta)\tau(x)$, for all $g_\theta \in G$ and $x \in X$, then $f(x) = g^{-1}_{\tau(x)}x$ is maximal invariant.*

Proof. λ is weakly G-equivariant because

$$\lambda(gg_\theta) = \alpha(gg_\theta)\theta_0 = \alpha(g)\alpha(g_\theta)\theta_0 = \alpha(g)\theta = \alpha(g)\lambda(g_\theta),$$

for all $g \in G$, $\theta \in \Theta$. Hence, λ is a bijective weakly G-equivariant mapping.

f is invariant because

$$\begin{aligned} f(g_\theta x) &= g^{-1}_{\tau(g_\theta x)} g_\theta x = g^{-1}_{\alpha(g_\theta)\tau(x)} g_\theta x = g^{-1}_{\theta * \tau(x)} g_\theta x = (g_\theta g_{\tau(x)})^{-1} g_\theta x \\ &= g^{-1}_{\tau(x)} g_\theta^{-1} g_\theta x = f(x) \end{aligned}$$

for all $g_\theta \in G$ and $x \in X$. If $f(x) = f(x')$, then $g^{-1}_{\tau(x)}x = g^{-1}_{\tau(x')}x'$, and so $x = g_\omega x'$ for some $g_\omega = g_{\tau(x)} g^{-1}_{\tau(x')} \in G$, that is $f(x) = g^{-1}_{\tau(x)}x$ is maximal invariant. $\square$

Remark 3.2. In the general situation similar to Corollary 3.1, $f(x) = \rho(g^{-1}_{\tau(x)})x$ is maximal invariant for some $\rho \in Aut(G)$, provided $(\rho(g))^{-1}g \in G_x$ is true for all $g \in G$ and $x \in X$ (see Proposition 2.2).

Example 3.5. In location-scale model $x = \mu + \sigma z$, where z has a known density f_0, and the parameter space is $\Theta = \{(\mu, \sigma) : \mu \in R, \sigma \in R^+\}$, define a group action by $g_\theta x = g_{(\mu,\sigma)}x = \mu + \sigma x$, so that the group operation is

$$g_{(\mu_1,\sigma_1)} g_{(\mu_2,\sigma_2)} = \mu_1 + \sigma_1\mu_2 + \sigma_1\sigma_2 = g_{(\mu_1+\sigma_1\mu_2,\sigma_1\sigma_2)} = g_{(\mu_1,\sigma_1)*(\mu_2,\sigma_2)}.$$

The set of such transformations is closed with identity $g_{(0,1)}$. It is easy to check that $g^{-1}_{(\mu,\sigma)} = g_{(-\frac{\mu}{\sigma},\frac{1}{\sigma})}$. It follows that $G = \{g_{(\mu,\sigma)} : (\mu,\sigma) \in \Theta\}$ constitutes a group under the composition of functions operation $\circ$, as defined above. The induced group action on Θ is given by

$$\alpha(g_{(\mu_1,\sigma_1)})(\mu_2,\sigma_2) = (\mu_1,\sigma_1) * (\mu_2,\sigma_2) = (\mu_1 + \sigma_1\mu_2, \sigma_1\sigma_2).$$

Identity and inverse element of the group Θ are $\theta_0 = \lambda(g_{(0,1)}) = (0,1)$, $\theta^{-1} = \alpha(g^{-1}_{(\mu,\sigma)})\theta_0 = \alpha(g_{(-\frac{\mu}{\sigma},\frac{1}{\sigma})})\theta_0 = (-\frac{\mu}{\sigma},\frac{1}{\sigma})$, respectively. Take a random sample $\mathbf{x} = (x_1,\ldots,x_n)$ of this model, so $\tau(\mathbf{x}) = (\bar{x}, s)$ is a weakly equivariant estimator, where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$, because:

$$\begin{aligned}\tau(g_{(\mu,\sigma)}\mathbf{x}) &= \left(\mu + \sigma\bar{x}, \sqrt{(n-1)^{-1}\sum_{i=1}^{n}(\mu + \sigma X_i - \mu - \sigma\bar{x})^2}\right) = (\mu + \sigma\bar{x}, \sigma s)\\ &= \alpha(g_{(\mu,\sigma)})\tau(\mathbf{x})\end{aligned}$$

It follows that

$$f(\mathbf{x}) = g^{-1}_{\tau(\mathbf{x})}\mathbf{x} = g^{-1}_{(\bar{x},s)}\mathbf{x} = g_{(-\frac{\bar{x}}{s},\frac{1}{s})}\mathbf{x} = -\frac{\bar{x}}{s} + \frac{1}{s}\mathbf{x} = \left(\frac{x_1-\bar{x}}{s},\ldots,\frac{x_n-\bar{x}}{s}\right)$$

is maximal invariant.

It is possible to change weakly G-equivariant functions into weakly G-equivariant estimators and vice versa. This is summarized in the following proposition.

Proposition 3.1. *Assume that G is sharply transitive on Θ. Let $\theta_0 \in \Theta$ be a fixed point, and write every element $\theta \in \Theta$ in a unique way as $\alpha(g)\theta_0 = \theta$, where $\alpha \in Aut(G)$. As previously mentioned, define $\lambda : G \to \Theta$ by $\lambda(g_\theta) = \theta$. For a given weakly G-equivariant function $\tau : X \to G$, with $\alpha_\tau \in Aut(G)$, there exists a weakly G-equivariant estimator $\delta : X \to \Theta$ with $\alpha_\delta = \alpha\circ\alpha_\tau \in Aut(G)$, given by $\delta = \lambda\circ\tau$. Conversely, there exists a weakly G-equivariant function $\tau : X \to G$ with $\alpha_\tau = \alpha^{-1}\circ\alpha_\delta \in Aut(G)$, given by $\tau = \lambda^{-1}\circ\delta$, such that $\delta : X \to \Theta$ is a weakly G-equivariant estimator with $\alpha_\delta \in Aut(G)$.*

Proof. Suppose that τ is a weakly G-equivariant function. It follows from Corollary 3.1 that λ is weakly G-equivariant with $\alpha \in Aut(G)$, thus

$$\delta(gx) = \lambda(\tau(gx)) = \lambda(\alpha_\tau(g)\tau(x)) = \alpha(\alpha_\tau(g))\lambda(\tau(x)) = \alpha_\delta(g)\delta(x),$$

for all $g \in G$, $x \in X$. Thus, δ is a weakly G-equivariant estimator with $\alpha_\delta = \alpha \circ \alpha_\tau \in Aut(G)$.

On the contrary, if δ is a weakly G-equivariant estimator with $\alpha_\delta \in Aut(G)$, as previously mentioned, $\alpha(g_{h\theta}) = h\alpha(g_\theta)$ for all $\theta \in \Theta$, $h \in G$, we can conclude that:

$$\begin{aligned}\tau(gx) &= \lambda^{-1}(\delta(gx)) = \lambda^{-1}(\alpha_\delta(g)\delta(x)) = g_{\alpha_\delta(g)\delta(x)} = \alpha^{-1}(\alpha(g_{\alpha_\delta(g)\delta(x)}))\\ &= \alpha^{-1}(\alpha_\delta(g)\alpha(g_{\delta(x)}))\\ &= \alpha^{-1}(\alpha_\delta(g))\alpha^{-1}(\alpha(g_{\delta(x)})) = \alpha^{-1}\circ\alpha_\delta(g)g_{\delta(x)} = \alpha_\tau(g)\lambda^{-1}(\delta(x)) = \alpha_\tau(g)\tau(x)\end{aligned}$$

for all $g \in G$, $x \in X$. Hence, τ is a weakly G-equivariant function with $\alpha_\tau = \alpha^{-1}\circ\alpha_\delta \in Aut(G)$. $\square$

Example 3.6. It is easy to see that in Example 3.4, G is sharply transitive on Θ and in this way, we can define $g_{(\theta_1,\theta_2)}$ as the unique $g \in G$ satisfying $\alpha(g_{(\theta_1,\theta_2)}) = \theta_1 = \sqrt{\theta_2} = \theta$ for some $\alpha \in Aut(G)$. Thus, $\lambda(g_{(\theta_1,\theta_2)}) = (\theta_1, \theta_1^2)$. Let $\alpha(g) = g^{1/m}$, for some $m \in Q$, so $g_{(\theta_1,\theta_2)} = \theta_1^m$ and in this way, $\lambda(\theta_1^m) = (\theta_1, \theta_1^2)$. By the way $\tau : Z \to G$, given by $\tau(z_1, z_2) = z_2^m$, is weakly G-equivariant with $\alpha_\tau(g) = g^{2m}$. Hence, by Proposition 3.1, there exists a weakly G-equivariant estimator $\delta : Z \to \Theta$ with $\alpha_\delta(g) = \alpha(\alpha_\tau(g)) = \alpha(g^{2m}) = g^2$ given by

$$\delta((z_1, z_2)) = \lambda(\tau(z_1, z_2)) = \lambda(z_2^m) = \left(z_2, z_2^2\right) = \left(\sum_{i=1}^n (x_i - \bar{x})^2, \left(\sum_{i=1}^n (x_i - \bar{x})^2\right)^2\right).$$

Now we will be interested in improving Proposition 2.2 for weakly equivariant estimators. Proposition 2.2 together with Proposition 3.1 yields the following result.

Corollary 3.2. *Assume that G is sharply transitive on Θ. Let $\theta_0 \in \Theta$ be a fixed point, and write every element $\theta \in \Theta$ in a unique way as $\alpha(g)\theta_0 = \theta$, where $\alpha \in Aut(G)$. For a given weakly G-equivariant estimator $\delta : X \to \Theta$ with $\alpha_\delta \in Aut(G)$ such that $\beta(g) = \alpha^{-1}(\alpha_\delta(g^{-1}))g \in G_x$ for all $g \in G$, $x \in X$, $f(x) = [\lambda^{-1}(\delta(x))]^{-1}x$ is maximal invariant.*

Notice that the condition in Proposition 2.2 leads to the following condition in Corollary 3.2:

$$\beta(g) = (\alpha_\tau(g))^{-1}g = (\alpha^{-1}(\alpha_\delta(g)))^{-1}g = \alpha^{-1}(\alpha_\delta(g^{-1}))g \in G_x,$$

for all $g \in G$, $x \in X$. In the special case when $\alpha = \alpha_\delta$, we can omit this condition.

Example 3.7. In Example 3.4, $\tau_0 : Z \to G$, given by $\tau(z_1, z_2) = \sqrt{z_2}$, is G-equivariant. Proposition 3.1 implies that there exists a weakly G-equivariant estimator $\delta : Z \to \Theta$ with $\alpha = \alpha_\delta \in Aut(G)$ given by

$$\delta((z_1, z_2)) = \lambda(\sqrt{z_2}) = (\sqrt[2m]{z_2}, \sqrt[m]{z_2}) = \left(\sqrt[2m]{\sum_{i=1}^n (x_i - \bar{x})^2}, \sqrt[m]{\sum_{i=1}^n (x_i - \bar{x})^2}\right).$$

On the other hand, condition $\beta(g) = \alpha^{-1}(\alpha_\delta(g^{-1}))g = 1 \in G_x$ for all $g \in G, x \in X$ is satisfied. Corollary 3.2 implies that:

$$\begin{aligned} f(z_1, z_2) &= [\lambda^{-1}(\delta((z_1, z_2)))]^{-1} \times (z_1, z_2) \\ &= \left[\lambda^{-1}\left(\sqrt[2m]{z_2}, \sqrt[m]{z_2}\right)\right]^{-1} \times (z_1, z_2) \\ &= \left(\sqrt{z_2}\right)^{-1} \times (z_1, z_2) = (z_1/\sqrt{z_2}, 1) \end{aligned}$$

and hence $z_1/\sqrt{z_2} = \bar{x}/\sum_{i=1}^n (x_i - \bar{x})^2$ is maximal invariant.

To illustrate the difference between Proposition 2.2 and Corollary 3.2, note that in Proposition 2.2, we will find maximal invariant by using weakly equivariant functions, while in Corollary 3.2, we will want to introduce a way for finding maximal invariant by using weakly equivariant estimators. Notice that Proposition 2.2 usually has applications in mathematics, because it suggests that, we should use weakly equivariant functions to find maximal invariant, but to make it useful in statistics, we combine weakly equivariant functions with a suitable function to make it a weakly equivariant estimator and then use Corollary 3.2.

Proposition 2.3 states how are can construct a general ?maximal invariant two given weakly equivariant functions? Instead of using two ?weakly equivariant functions, one could use two weakly G-equivariant estimators to obtain a maximal invariant function. At first, we change weakly G-equivariant estimators into weakly G-equivariant functions (Proposition 3.1), then we can construct a general ?maximal invariant (Proposition 2.3). The details are given in the following proposition.

Proposition 3.2. *Suppose that G is sharply transitive on Θ. Let $\theta_0 \in \Theta$ be a fixed point, and write every element $\theta \in \Theta$ in a unique way as $\alpha(g)\theta_0 = \theta$, where $\alpha \in Aut(G)$. Assume that there exist two weakly G-equivariant estimators $\delta_i : X \to \Theta$, $i = 1, 2$ with the same $\alpha_\delta \in Aut(G)$ such that at least one of the δ_i is weakly G-isovariant. Then $f(x) = (\lambda^{-1}(\delta_1(x)))^{-1}\lambda^{-1}(\delta_2(x))$ is maximal invariant.*

Proof. By Proposition 3.1, $\tau_i : X \to G$, given by $\tau_i = \lambda^{-1} \circ \delta_i$, $i = 1, 2$, are weakly G-equivariant functions with the same $\alpha_\tau = \alpha^{-1} \circ \alpha_\delta \in Aut(G)$. Assume that one of the two, say δ_2 is weakly G-isovariant, i.e., $G_{\delta_2(x)} = \alpha_\delta(G_x)$ for all $x \in X$. Since λ^{-1} is one-to-one, we conclude that τ_2 is weakly G-isovariant because:

$$\begin{aligned} G_{\tau_2(x)} &= \{g : g\tau_2(x) = \tau_2(x)\} \\ &= \{g : g\lambda^{-1}(\delta_2(x)) = \lambda^{-1}(\delta_2(x))\} \\ &= \{g : \lambda^{-1}(\alpha(g)\delta_2(x)) = \lambda^{-1}(\delta_2(x))\} \\ &= \{\alpha^{-1}(\alpha(g)) : \alpha(g)\delta_2(x) = \delta_2(x)\} \\ &= \alpha^{-1}(G_{\delta_2(x)}) = \alpha^{-1} \circ \alpha_\delta(G_x) = \alpha_\tau(G_x) \end{aligned}$$

for all $x \in X$. Hence, Proposition 2.3 implies that

$$f(x) = (\tau_1(x))^{-1}\tau_2(x) = [\lambda^{-1}(\delta_1(x))]^{-1}\lambda^{-1}(\delta_2(x))$$

is maximal invariant. □

We illustrate this proposition with the following example.

Example 3.8. Let (x_i, y_i)‘s be iid from the pdf $(\sigma e^{-\sigma x})(\sigma^{-1}e^{-y/\sigma})$, $i = 1, \ldots, n$, where $\sigma > 0$ and $x, y > 0$. This is the model in the Nile problem considered by Fisher (1973).

Clearly, $(z_1, z_2) = (\bar{x}, \bar{y})$ is a minimal sufficient statistic. If $G = R^+$ acts on $Z = R^+ \times R^+$ by

$$g \times (z_1, Z_2) = (g^{-1}z_1, gz_2)$$

then G acts on $\Theta = \{(\theta_1, \theta_2) : \theta_2 = \theta_1^{-1} = \sigma > 0\}$ by

$$(g, (\theta_1, \theta_2)) \mapsto (g^{-1}\theta_1, g\theta_2)$$

Furthermore, since $G_{(z_1,z_2)} = \{1\}$ for all $(z_1, z_2) \in Z$, G is sharply transitive on Θ and in this way, we can define $g_{(\theta_1,\theta_2)}$ as the unique $g \in G$ satisfying $(\alpha(g), (1, 1)) \mapsto (\theta_1, \theta_2)$ for some $\alpha \in Aut(G)$. Thus, $\alpha(g_{(\theta_1,\theta_2)}) = \theta_2 = \theta_1^{-1} = \sigma$, $\lambda(g_{(\theta_1,\theta_2)}) = (\theta_2^{-1}, \theta_2)$. Let $\alpha(g) =$

g^{2k}, for some $k \in Q$, so $g_{(\theta_1,\theta_2)} = \sqrt[2k]{\theta_2}$ and then $\lambda(\theta_2) = (\theta_2^{-2k}, \theta_2^{2k})$. The functions $\tau_i : Z \to G$, $i = 1, 2$, given by $\tau_1(z_1, z_2) = z_2^m / z_1^n$, $\tau_2(z_1, z_2) = z_2^{2m-n} / z_1^{2n-m}$ for some $m, n \in Q$, are weakly G-equivariant with the same $\alpha_\tau(g) = g^{m+n}$, respectively. Using Proposition 3.1, these functions can be improved upon by using the weakly G-equivariant estimators $\delta_i : Z \to \Theta$, $i = 1, 2$ as follows

$$\begin{aligned}
\delta_1\left((z_1, z_2)\right) &= \lambda\left(\tau_1(z_1, z_2)\right) = \lambda\left(z_2^m / z_1^n\right) = \left(z_2^{-2mk} / z_1^{-2nk}, z_2^{2mk} / z_1^{2nk}\right) \\
\delta_2\left((z_1, z_2)\right) &= \lambda\left(\tau_2(z_1, z_2)\right) = \lambda\left(z_2^{2m-n} / z_1^{2n-m}\right) \\
&= \left(z_2^{4nk-2mk} / z_1^{4mk-2nk}, z_2^{4mk-2nk} / z_1^{4nk-2mk}\right),
\end{aligned}$$

with the same $\alpha_\delta(g) = \alpha(\alpha_\tau(g)) = \alpha(g^{m+n}) = g^{2mk+2nk}$.

On the other hand, since G is sharply transitive on Θ, we have $\alpha(G_{(z_1,z_2)}) = G_{\delta_1(z_1,z_2)} = \{1\}$ and so δ_1 is weakly G- isovariant. It follows from Proposition 3.2 that

$$f(x) = (\lambda^{-1}(\delta_1(x)))^{-1}\lambda^{-1}(\delta_2(x)) = \left(z_2^m / z_1^n\right)^{-1} z_2^{2m-n} / z_1^{2n-m} = (z_1 z_2)^{m-n}$$

is maximal invariant.

4. Independence, Invariance and Weakly Equivariance

In the rest of this article, we deal with the independence of an invariant function and a weakly equivariant function under some special conditions, which is a generalization of Bondesson's method. If a weakly equivariant estimator $\tau(x)$ is sufficient then it contains all the information in the sample about the parameters. Thus if, $h(x)$ is an ancillary, there is some reason to believe that $\tau(x)$ and $h(x)$ might be independent. We know if the group G acts transitively on Θ, then any invariant function will be ancillary (Lehmann and Romano, 2005, pp. 395–396). Also, h is invariant iff h is a function of a maximal invariant statistic. This suggests that a maximal invariant statistic $h(x)$ will be independent of weakly equivariant estimator $\tau(x)$ as explained as follows (Eaton, 1983). Bondesson (1997) proved the independence of invariant and equivariant functions under special conditions. This property also holds for weakly equivariant functions.

Proposition 4.1. *Suppose (X, β_1) and (Y, β_2) are measurable spaces acted on by a locally compact and σ-compact topological group G. The mapping $G \times Y \to Y$, given by $(\alpha(g), x) \mapsto \alpha(g)x$, is jointly measurable where $\alpha \in Aut(G)$ and G acts transitively on Y. Assume that $\tau : X \to Y$ is a measurable weakly G-equivariant function. Also, let (Z, β_3) be a measurable space and let $h : X \to Z$ be a measurable G-invariant function. For a random variable $x \in X$, with distribution P_o, set $y = \tau(x)$, $z = h(x)$ and assume that $\tau(x)$ is a sufficient statistic for the family $\{gP_o : g \in G\}$ of distributions on (X, β_1). Under these assumptions, y and z are independent when $x \in X$ has distribution gP_o for $g \in G$.*

Proof. First, assume that $x \in X$ has distribution P_o and let Q_o be the induced distribution of $\tau(x)$. Fix a bounded measurable function f on Z and let $H_g(y) = E_{gP_0}(f(h(x))|\tau(x) = y)$ (the conditional expectations are well defined since f is bounded). Since $\tau(x)$ is a sufficient statistic for the family $\{gP_o : g \in G\}$, there is a measurable function H on Y s.t. $H_g(y) = H(y)$ for $y \notin N_g$, where N_g is a set of $\alpha(g)Q_0$−measure zero (see Eaton, 1983, Proposition 7.17, p. 288). Hence, $\alpha(g)Q_0(N_g) = Q_0(\alpha(g^{-1})N_g) = 0$. Also, $H_e(y)$ is the unique a.e.

(Q_0) function that satisfies the equation

$$\begin{aligned}\int_Y k(y)H_e(y)Q_0(dy) &= E_{Q_0}(k(y)H_e(y))\\ &= E_{Q_0}E_{P_0}(k(y)f(h(x))|\tau(x)=y)\\ &= E_{P_0}(k(\tau(x))f(h(x)))\\ &= \int_X k(\tau(x))f(h(x))P_0(dx)\end{aligned} \tag{4.1}$$

for all bounded measurable k. The probability measure gP_0, satisfies the equation

$$\int_X l(x)gP_0(dx) = \int_X l(gx)P_0(dx) \tag{4.2}$$

for all bounded *l*. Since τ is weakly G-equivariant, this implies that $\tau(x)$ has distribution $\alpha(g)Q_0$ when x has distribution gP_o for $g \in G$. Using this and the invariance of $f \circ h$ (see Proposition 2.1(ii)), we have, for all bounded k,

$$\begin{aligned}\int_Y k(y)H_e(y)Q_0(dy) &\overset{(4.1)}{=} \int_X k(\tau(x))f(h(x))P_0(dx)\\ &= \int_X k(\alpha(g^{-1})\tau(gx))f(h(gx))P_0(dx)\\ &\overset{(4.2)}{=} \int_X k(\alpha(g^{-1})\tau(x))f(h(x))gP_0(dx)\\ &\overset{(4.1)}{=} \int_X k(\alpha(g^{-1})y)H_g(y)(\alpha(g)Q_0)(dy)\\ &\overset{(4.2)}{=} \int_X k(y)H_g(\alpha(g)y)Q_0(dy).\end{aligned}$$

Therefore, $H_e(y) = H_g(\alpha(g)y)$ a.e. (Q_0). Since $H(y) = H_e(y)$ a.e. (Q_0) and $H_g(\alpha(g)y) = H(\alpha(g)y)$ for $\alpha(g)y \notin N_g$, where $Q_0(\alpha(g^{-1})N_g) = 0$, we have $H_g(\alpha(g)y) = H(\alpha(g)y)$ a.e. (Q_0), and this implies that $H(y) = H(\alpha(g)y)$ a.e. (Q_0). Hence, there exists a G-invariant measurable function, say $\tilde{H}$, s.t. $H = \tilde{H}$ a.e. (Q_0). Since G acts transitively on Y, $\tilde{H}$ must be a constant, so H is a constant a.e. (Q_0). Therefore, $H_e(y) = E_{P_0}(f(h(x))|\tau(x) = y)$ is a constant a.e. (Q_0). Now, if k is a bounded function on (Y, β_2) and $H(y) = H(y_0)$ for $y \in Y$, then

$$\begin{aligned}E_{P_0}[k(\tau(x))f(h(x))] &= E_{P_0}E_{Q_0}(k(\tau(x))f(h(x))|\tau(x)=y)\\ &= \int_Y E_{P_0}[k(\tau(x))f(h(x))|\tau(x)=y]Q_0(dy)\\ &= \int_Y k(y)E_{P_0}[f(h(x))|\tau(x)=y]Q_0(dy)\\ &= \int_Y k(y)H(y)Q_0(dy) = H(y_0)\int_Y k(y)Q_0(dy)\\ &= \int_Y H(y)Q_0(dy)E_{P_0}k(\tau(x)) \overset{(4.1)}{=} E_{P_0}f(h(x))E_{P_0}k(\tau(x))\end{aligned}$$

and this implies that $y = \tau(x)$ and $z = h(x)$ are independent when $x \in X$ has distribution P_o. When x has distribution $\tilde{P}_o = g_1 P_o$, we note that $\{gP_o : g \in G\} = \{g\tilde{P}_o : g \in G\}$, so $\tau(x)$ is sufficient for $\{g\tilde{P}_o : g \in G\}$. The argument given for P_o now applies for $\tilde{P}_o$. Thus, $y = \tau(x)$ and $z = h(x)$ are independent when $x \in X$ has distribution $g_1 P_o$. □

Since G acts transitively on $\{gP_o : g \in G\}$ and $z = h(x)$ is G-invariant, the distribution of z is the same under each gP_o, $g \in G$, and so z is an ancillary statistic. Basu's Theorem (Basu, 1955), asserts that a complete sufficient statistic is independent of an ancillary statistic. Although no assumptions concerning invariance are made in the statement of Basu's Theorem, most applications are to problems where invariance is used to show that a statistic is ancillary. In Proposition 4.1, the completeness assumption of Basu's Theorem has been replaced by the invariance assumptions and, most particularly, by the assumption that the group G acts transitively on the space Y (Eaton, 1983). Let (X, β) be a measurable space acted on by locally compact and σ-compact topological group G and assume β_S is a sufficient subfield of β, β_I is the σ-field of invariant members of β, and β_{SI} is the intersection of β_I and β_S. Hall, Wijsman and Ghosh (1965) proved that under certain conditions β_I and β_S are conditionally independent given β_{SI}. Hence, Proposition 4.1 is similar to their result. We illustrate this proposition with the following example.

In a special case, suppose that in Proposition 4.1, $\tau : X \to Y$ is weakly G-isovariant, then by Proposition 1.1(ii), τ is one-to-one on the orbit Gx. We can take $y_1 = \tau(x_1)$, $y_2 = \tau(x_2)$ for all $x_1, x_2 \in X$, and since G acts transitively on Y and $\alpha \in Aut(G)$ is onto, we can say there exists $g \in G$ such that $y_2 = \tau(x_2) = \alpha(g)\tau(x_1) = \alpha(g)y_1$. But τ is weakly G-equivariant and so $\tau(x_2) = \tau(gx_1)$ and then $x_2 = gx_1$, for some $g \in G$. Thus in Proposition 4.1 we conclude that G acts transitively on X, provided τ is weakly G-isovariant. In this case, any invariant function on X will be ancillary. Therefore, Proposition 4.1 asserts that a weakly G-equivariant sufficient statistic is independent of an invariant (ancillary) statistic. Hence, in Proposition 4.1, the completeness assumption of Basu's Theorem has been replaced by the assumptions that the group G acts transitively on the space Y and τ is weakly G-isovariant and G-equivariant sufficient statistic.

Example 4.1. Let $x_1, \ldots, x_n$ be iid from Gamma distribution $\Gamma(a, b)$ with density

$$f_{a,b}(x) = \frac{x^{a-1}e^{-\frac{x}{b}}}{\Gamma(a)b^{a-1}}, \quad x > 0,$$

where $a > 0$ is known and $b > 0$ is unknown. If $G = R^+$ acts on $X = (R^+)^n$ by $g \times x = gx$, then the complete sufficient statistics for b, that is $\delta_0(x) = \sum_{i=1}^n x_i$ is equivariant too. By Lemma 2.1,

$$f(x) = (\delta_0(x))^{-1} \times x = \left(x_1 \Big/ \sum_{i=1}^n x_i, \ldots, x_n \Big/ \sum_{i=1}^n x_i\right)$$

is maximal invariant. But $k_0 : X \to X$ given by

$$k_0(a_1, \ldots, a_n) = \left(a_1 \sum_{i=1}^n a_i \Big/ a_n, \ldots, a_{n-1} \sum_{i=1}^n a_i \Big/ a_n, a_n \sum_{i=1}^n a_i \Big/ |a_n|\right)$$

is one-to-one and in this way, by Proposition 2.1(ii),

$$h(x) = k_0 \circ f(x) = (x_1/x_n, \ldots, x_1/x_{n-1}, x_n/|x_n|)$$

is maximal invariant. For a continuous automorphism $\alpha_1(g) = g^{m_1}$, for some $m_1 \in Q$, $\tau(x) = \alpha_1(\delta_0(x)) = (\sum_{i=1}^n x_i)^{m_1}$ is weakly G-equivariant with $\alpha_1 \in Aut(G)$, because

$$\tau(gx) = g^{m_1}\left(\sum_{i=1}^n x_i\right)^{m_1} = \alpha_1(g)\tau(x).$$

(Notice that since G is sharply transitive on $\Theta = R^+$, we can define g_b as the unique $g \in G$ satisfying $\alpha(g) = b$ for some $\alpha \in Aut(G)$.)

Finally, Proposition 4.1 implies that $\tau(x) = (\sum_{i=1}^n x_i)^{m_1}$ and $h(x) = (x_1/x_n, \ldots, x_1/x_{n-1}, x_n/|x_n|)$ are independent.

The following example shows that Proposition 4.1 provides an easy way to prove the independence of weakly G-equivariant sufficient statistic and G-invariant statistic when weakly G-equivariant sufficient statistic is not complete and Basu's Theorem is not applicable.

Example 4.2. Let X be distributed as

$$P_\theta(X = x) = \begin{cases} \theta^2(1-\theta)/2 & x = \pm 5 \\ \theta(1-\theta)^2/2 & x = \pm 4 \\ \theta^3/2 & x = \pm 3 \\ (1-\theta)^3/2 & x = \pm 2 \\ \theta(1-\theta) & x = \pm 1 \end{cases}$$

with range $D_X = D = \{\pm 1, \pm 2, \pm 3, \pm 4, \pm 5\}$. The totality of ancillary statistics is then obtained by

$$V(a) = \begin{cases} v_1 & a \in D' \\ v_2 & a \in D - D' \end{cases}$$

s.t. $D' = \{d_1, d_2, d_3, d_4, d_5\} \subset D$ satisfying $d_i \neq \pm d_j$ for all $d_i, d_j \in D'$, $i \neq j$ where $i, j = 1, 2, 3, 4, 5$. This can be described as follows. All statistics are given by $V(a) = v_i$ for all $a \in D_i$, s.t. $\{D_i\}$ is a countable partition of D. Let D' be a non empty subset of D, where $V(a) = v$ for all $a \in D\prime$. Thus,

$$\begin{aligned} P_\theta(V = v) &= \sum_{i \in D'} P_\theta(X = i) = \sum_{i \in D} \varepsilon_i P_\theta(X = i) \\ &= \frac{1}{2}[\sigma_4 - \sigma_5 + \sigma_3 - \sigma_2]\theta^3 + \left[\frac{1}{2}\sigma_5 - \sigma_4 + \tfrac{3}{2}\sigma_2 - \sigma_1\right]\theta^2 \\ &\quad + \left[\frac{1}{2}\sigma_4 - \frac{3}{2}\sigma_2 + \sigma_1\right]\theta + \frac{1}{2}\sigma_2, \end{aligned}$$

where $\varepsilon_i = I_{i \in D\prime}, \sigma_i = \varepsilon_{-i} + \varepsilon_i$ for all $i = 1, \ldots, 5$. It is easy to check that $P_\theta(V = v)$ does not depend on θ iff $\sigma_i = \sigma$ for all $i = 1, \ldots, 5$ where $\sigma \in \{0, 1, 2\}$.

If $\sigma = 0$, then $\varepsilon_i = 0$ for all $i = 1, \ldots, 5$ and hence$D\prime = \phi$, which is clearly impossible.

If $\sigma = 2$, then $\varepsilon_i = 1$ for all $i = 1, \ldots, 5$ and so in this case, the ancillary is trivially satisfying $V(a) = v$ for all $a \in D$.

Finally, if $\sigma = 1$, then we have $\varepsilon_i = 0, \varepsilon_{-i} = 1$ or $\varepsilon_i = 1, \varepsilon_{-i} = 0$ for all $i = 1, \ldots, 5$. Therefore, $D' = \{d_1, \ldots, d_5\}$ satisfying $d_i \neq \pm d_j$ for all $d_i, d_j \in D'$, $i \neq j$, where $i, j = 1, \ldots, 5$ and then the totality of ancillary statistics is obtained by

$$V(a) = \begin{cases} v_1 & a \in D' \\ v_2 & a \in D - D' \end{cases}.$$

Now, let $G = \{g_1, g_2\}$ acts on $D = \{\pm 1, \ldots, \pm 5\}$, where $g_1(x) = x$, $g_2(x) = -x$, which induce a group action on the parameter space $\Theta = (0, 1)$ as $g\theta = g(\theta) = \theta$ for all $\theta \in \Theta$, $g \in G$. Obviously, the model is invariant under the group G acting on X and Θ. Ancillary statistics are G-invariant if they are trivial. Thus, here only trivially ancillary statistics are invariant. Let $V_0(a) = v$ for all $a \in D$ which is G-invariant. Clearly, $T : D_X = \{\pm 1, \ldots, \pm 5\} \to D_Y = \{1, \ldots, 5\}$, given by $T = |X|$, is a minimal sufficient statistic for θ, which is not complete and so in this case, Basu's Theorem is not applicable. But T is G-invariant or equivalently weakly $\bar{G}$-equivariant with $\alpha = 1_{\bar{G}} \in Aut(\bar{G})$ where $\bar{G} = \{g_1\} \leq G$ acts trivially on $D_Y = \{1, \ldots, 5\}$ (see Definition 1.6(i)). It is easy to show that $P(V_0 = v_i | T = t) = P(V_0 = v_i) = \frac{1}{2}$ for all $t = 1, \ldots, 5, i = 1, 2$ and this implies that V_0 is independent of T where we can conclude that the same result by using Proposition 4.1. Similarly, $U : D_X \to D_X$, given by $U = U(X) = X$, is weakly G-equivariant sufficient statistic with $\alpha = 1_G \in Aut(G)$ and so by Proposition 4.1, U and V_0 are independent.

Acknowledgments

The authors are deeply grateful to the referees for constructive criticisms which greatly facilitated an improved presentation of this article.

References

Basu, D. (1955). On statistics independent of a complete sufficient statistic. *Sankhya* 15:377–380.

Bondesson, L. (1977). *A note on sufficiency and independence*. Preprint, University of Lund, Lund, Sweden.

Bredon, G. H. (1972). *Introduction to Compact Transformation Groups*. New York: Academic Press.

Berk, R. H. (1967). A special group structure and equivariant estimation. *Ann. Math. Statist.* 38(5):1436–1445.

Deitmar, A., Echterhoff, S. (2009). *Principles of Harmonic Analysis*. New York: Springer.

Eaton, M. L. (1983). *Multivariate Statistics: A Vector Space Approach*. New York: Wiley.

Eaton, M. L. (1989). Group Invariance Applications in Statistics. Hayward, CA: *Institute of Mathematical Statistics and American Statistical Association*.

Fisher, R. A. (1973). *Statistical Methods and Scientific Inference*. New York: Hafner.

Folland, G. B. (1995). *A Course in Abstract Harmonic Analysis*. Boca Raton, FL: CRC Press.

Fraser, D. A. S. (1961). The fiducial method and invariance. *Biometrika* 48:261–280.

Fraser, D. A. S. (1968). *The Structure of Inference*. New York: Wiley.

Hall, W. J., Wijsman, R. A., Ghosh, J. K. (1965). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Stat.* 36:575–614.

Lehmann, E. L., Romano, J. P. (2005). *Testing Statistical Hypotheses. 3rd ed.* New York: Springer.

Nachbin, L., Bechtolsheim, L. (1965). *The Haar Integral*. Princeton N.J, Toronto.:D. Van Nostrand.

Palais, R. S. (1960). Classification of G-spaces, *Mem. Am. Math. Soc.* 36.

Robinson, D. J. S. (1995). *A Course in the Theory of groups. 2nd ed*. New York: Springer-Verlag.