The Annals of Statistics
1976, Vol. 4, No. 4, 791–795

HYPOTHESES GENERATING GROUPS FOR TESTING MULTIVARIATE SYMMETRY

BY LUDGER RÜSCHENDORF

University of Hamburg

Hypotheses generating groups are constructed for the class of multivariate symmetric distributions and for the corresponding multisample problems. Defining ranks as maximal invariants under these groups we get distribution free procedures. Applications to further multivariate test problems are indicated.

1. Introduction. An open problem stated in a paper by Bell and Haller (cf. [1], page 263, cf. also [2], [4]) is to find a group of monotone, invertible and bimeasurable transformations of $\mathbb{R}^p$, which generates the class $\tilde{\Omega}_s{}^p$ of p-dimensional symmetric and continuous distributions or a "dense" subset of it. In this context a group G is called a hypothesis generating (HG) group for the class of distributions Ω_0, if for each $P \in \Omega_0$ we have $\Omega_0 = \{gP; g \in G\}$, where gP is the image of P under g (cf. [1], [2]), and a p-dimensional distribution P is called symmetric, if $\pi P = P$ for all π out of the symmetric group γ_p of all $p!$ permutations, with $\pi(x_1, \cdots, x_p) = (x_{\pi(1)}, \cdots, x_{\pi(p)})$.

The solution of this problem is of great importance for the construction of multivariate symmetry rank tests, since it is useful to define ranks as maximal invariants under G for two reasons. Firstly, these rank statistics yield distribution free (DF) procedures for Ω_0, and secondly, the theorem of Berk and Bickel (cf. [5]) is applicable if an additional completeness result is known.

For multivariate hypotheses, constructions of HG-groups are known primarily for hypotheses which reduce to one-dimensional problems, as e.g. for the hypothesis of total independence (cf. [6], page 50, [2]). An exception is a construction of a HG-group G_1 by Smith [6] for the class Ω^p of all continuous distributions on $(\mathbb{R}^p, \mathfrak{B}^p)$ with strictly monotone conditional distribution functions (df's). But for this class of distributions—being too large for the application of DF-techniques—this method leads to tests, which are essentially univariate and, therefore, of low power.

In Section 2 we provide a HG-group G_2 for the "dense" subset $\Omega_s{}^p$ of $\tilde{\Omega}_s{}^p$, containing all elements of $\tilde{\Omega}_s{}^p$ with strictly monotone conditional df's. The resulting maximal invariant is intuitively quite appealing. For our construction we mainly use some results derived for the multivariate Rosenblatt transformation τ_F, which is defined for $F \in \Omega^p$ (here and for the rest of the paper we identify

Received October 1974; revised June 1975.

AMS 1970 *subject classifications.* Primary 62A05, 62G10, 62H15.

Key words and phrases. Hypotheses generating groups, multivariate symmetry, Rosenblatt's transformation, ranks, bounded completeness.

distributions and df's) by:

$$\tau_F(x_1, \cdots, x_p) = (F_1(x_1), F_{2|1}(x_2 \mid x_1), \cdots, F_{p|1,\cdots,p-1}(x_p \mid x_1, \cdots, x_{p-1}))\,,$$

where $F_{i|1,\cdots,i-1}$ are the conditional df's of F. The simple idea for the construction of G_2 is then applied to the multi-sample case. Further applications are indicated in Section 3.

2. The multivariate symmetry case. For $F \in \Omega^p$, $\tau_F : \mathbb{R}^p \to (0, 1)^p$ is a bijective, bimeasurable transformation. Furthermore, generalizing a well-known fact for the one-dimensional case, Smith [6] has proven that for $F_0 \in \Omega^p$ the set $G_1 = \{\tau_F^{-1} \circ \tau_{F_0}; F \in \Omega^p\}$ is a HG-group for Ω^p.

Let $E = \{x \in \mathbb{R}^p; x_1 < x_2 < \cdots < x_p\}$, and let $\Omega^p(E)$ be the class of continuous distributions on $(E, \mathfrak{B}^p \cap E)$ which have strictly conditional df's on E. For different reasons one is led to the conjecture that E should be (as in the independence case) the "natural" rank set. If one tries to prove this by taking Ω_s^p instead of Ω^p in the definition of Smith's group G_1 one fails, since $x \in E$ does not imply $\tau_F^{-1} \circ \tau_{F_0}(x) \in E$ for $F \in \Omega_s^p$, which can be seen for $p = 2$ by taking F as two-dimensional normal df.

For $F \in \Omega^p(E)$ we define the modified Rosenblatt transformation τ_F by the restriction of τ_F to E, $\tau_F : E \to (0, 1)^p$, and by means of the result obtained by Smith for Ω^p we get the following analog for $\Omega^p(E)$.

Lemma 1.

a) *For $F \in \Omega^p(E)$, τ_F is a $1 - 1$ transformation of E onto $(0, 1)^p$.*

b) *For $F_0 \in \Omega^p(E)$, $G_2' = \{\tau_F^{-1} \circ \tau_{F_0}; F \in \Omega^p(E)\}$ is a HG-group for $\Omega^p(E)$.*

For $F, F_0 \in \Omega^p(E)$ we now define $g_F : \mathbb{R}^p \to \mathbb{R}^p$ by $g_F(x) = \pi \circ \tau_F^{-1} \circ \tau_{F_0} \circ \pi^{-1}(x)$ for $x \in \pi E$, $\pi \in \gamma_p$. (g_F is defined almost everywhere, cf. [7]).

Theorem 1.

a) *$G_2 = \{g_F; F \in \Omega^p(E)\}$ is a HG-group for Ω_s^p.*

b) *A maximal invariant under G_2 is the p-tuple of ranks $(r_1, \cdots, r_p)$, where $r_i(x_1, \cdots, x_p)$ denotes the rank of x_i among $\{x_1, \cdots, x_p\}$.*

Proof.

a) The group properties of G_2 are easily derived from the corresponding properties of G_2'. We have to show that

(1) $g_F L \in \Omega_s^p$ for all $L \in \Omega_s^p$, $F \in \Omega^p(E)$ and

(2) the G_2-orbit of L, $G_2(L)$, equals Ω_s^p for all $L \in \Omega_s^p$.

To (1): By definition of g_F and from Lemma 1 we obtain that g_F maps πE into πE for all $\pi \in \gamma_p$ and $F \in \Omega^p(E)$ and that $g_F L$ has strictly monotone marginals for $L \in \Omega_s^p$. For $L \in \Omega_s^p$ and $\sigma \in \gamma_p$ we have

$$\begin{aligned} g_F L(\sigma A) &= \textstyle\sum_{\pi \in \gamma_p} L(\pi \circ \tau_{F_0}^{-1} \circ \tau_F(\pi^{-1} \circ \sigma A \cap E)) \\ &= \textstyle\sum_{\pi \in \gamma_p} L(\tau_{F_0}^{-1} \circ \tau_F(\pi^{-1} A \cap E)) \end{aligned}$$

which is independent of $\sigma \in \gamma_p$. Therefore, $g_F L \in \Omega_s^p$.

To (2): For $L \in \Omega_s^p$ we define $h: \Omega_s^p \to \Omega^p(E)$ by $h(L) = L(\cdot \mid E)$, where $L(\cdot \mid E)$ is the conditional distribution of L under $E(L(E) = 1/p! > 0)$; $h(L)$ has strictly increasing marginals if L has. Furthermore h is bijective, and for $P \in \Omega^p(E)$ the inverse of h is given by

$$h^{-1}(P)(A) = 1/p! \sum_{\pi \in \gamma_p} P(\pi^{-1}A \cap E) .$$

For $A \in \mathfrak{B}^p$,

$$\begin{aligned} h^{-1}(\tau_F^{-1} \circ \tau_{F_0} h(L))(A) &= 1/p! \sum_{\pi \in \gamma_p} \tau_F^{-1} \circ \tau_{F_0} L(\pi A \cap E \mid E) \\ &= \sum_{\pi \in \gamma_p} L(\tau_F^{-1} \circ \tau_F(\pi A \cap E)) = g_F(L)(A) . \end{aligned}$$

Lemma 1 now implies:

$$\begin{aligned} G_2(L) &= \{g_F(L); F \in \Omega^p(E)\} \\ &= \{h^{-1}(\tau_F^{-1} \circ \tau_{F_0} h(L)); F \in \Omega^p(E)\} \\ &= h^{-1}\Omega^p(E) = \Omega_s^p . \end{aligned}$$

b) From Lemma 5.3 in [6] it follows that the orbit $G_2(x)$ equals πE for all $x \in \pi E$ and all $\pi \in \gamma_p$. Therefore, a maximal invariant for G_2 is given by the rank statistic $(r_1, \cdots, r_p)$.

In the multisample case of testing whether the df's of n independent p-dimensional symmetric random variables are identical, we have the hypothesis $\Omega_0 = \{F^{(n)}; F \in \Omega_s^p\}$, where $F^{(n)}$ is the n-fold produce measure. Let

$$\begin{aligned} E' = \{(x_1, \cdots, x_n); x_i \in \mathbb{R}^p, x_{1i} < x_{2i} < \cdots < x_{pi}, 1 \leqq i \leqq n, \\ x_{11} < x_{12} < \cdots < x_{1n}\} \subset \mathbb{R}^{pn} , \qquad x_i = (x_{1i}, \cdots, x_{pi}) , \end{aligned}$$

and let $\Omega^{pn}(E')$ be defined analogously to $\Omega^p(E)$.

The following theorem shows that E' is the "natural" rank set for the multisample problem. To prove this, we first observe that each $F^{(n)} \in \Omega_0$ is invariant under the elements of the wreath product $S^I = \gamma_p \sim \gamma_n = \{(\pi_1, \cdots, \pi_n, v); \pi_i \in \gamma_p, 1 \leqq i \leqq n, v \in \gamma_n\}$, with $(\pi_1, \cdots, \pi_n, v)(x_1, \cdots, x_n) = (\pi_1(x_{v(1)}), \cdots, \pi_n(x_{v(n)}))$, $x_i \in \mathbb{R}^p$. S^I is called the symmetry group of Ω_0, $|S^I| = (p!)^n n!$.

For $F \in \Omega^{pn}(E')$ we define the modified Rosenblatt-transformation τ_F by restriction to E' and further define functions $h_F: \mathbb{R}^{pn} \to \mathbb{R}^{pn}$ for $F \in \Omega^{pn}(E')$ by $h_F(x) = \pi \circ \tau_F^{-1} \circ \tau_{F_0}(\pi^{-1}x)$ for $x \in \pi E'$, $\pi \in \gamma_p \sim \gamma_n$, $F_0 \in \Omega^{pn}(E')$, and obtain:

THEOREM 2.

a) $G_3 = \{h_F; F \in \Omega^{pn}(E')\}$ *is a* HG-*group for the multisample problem* Ω_0.

b) *A maximal invariant for* G_3 *is the pn-tuple* $(r_{11}, \cdots, r_{1n}, g_{11}, \cdots, g_{p1}, \cdots, g_{1n}, \cdots, g_{pn})$, *where* $g_{ij}(x_1, \cdots, x_n)$ *is the rank of* x_{ij} *among* $\{x_{1j}, \cdots, x_{pj}\}$, $1 \leqq i \leqq p$, $i \leqq j \leqq n$, *and where* r_{1k} *is the rank of* $\min\{x_{1k}, \cdots, x_{pk}\}$ *among* $\{\min\{x_{1l}, \cdots, x_{pl}\}; 1 \leqq l \leqq n\}$, $1 \leqq k \leqq n$.

PROOF. The proof of Theorem 2 is analogous to the one of Theorem 1. Analogously to Lemma 1, $G_3' = \{\tau_F^{-1} \circ \tau_{F_0}; F \in \Omega^{pn}(E')\}$ is a HG-group for $\Omega^{pn}(E')$. Furthermore, the fact that $\gamma_p \sim \gamma_n$ is a symmetry group for Ω_0 implies that $h: \Omega_0 \to \Omega^{pn}(E')$, defined by $h(P) = P(\cdot \mid E)$, is bijective. This together with the proof of Theorem 1 implies Theorem 2.

REMARK 1. a) The basic rank sets for the multisample problem are E' and the $(p!)^n n!$ images under $\gamma_p \sim \gamma_n$. These rank sets are essentially different from those in the independence case; compare the different types of rank sets proposed by Bell and Haller [1] in the case $p = 2$.

b) Using the map $h: \Omega_0 \to \Omega^{pn}(\pi E')$ defined by $h(P) = P(\cdot \mid \pi E')$, $\pi \in \gamma_p \sim \gamma_n$ fixed, and the group G_3'' consisting of transformations h_F' defined by

$$h_F'(x) = \pi' \circ \tau_F^{-1} \circ \tau_{F_0}(\pi'^{-1}x)$$
$$\text{for} \quad x \in \pi' \circ \pi E', \quad \pi' \in \gamma_p \sim \gamma_n, \quad F \in \Omega^{pn}(\pi E'),$$

we get different HG-groups with maximal invariants $(r_{k1}, \cdots, r_{kn}, g_{11}, \cdots, g_{pn})$, where r_{ki} is the rank of the kth order statistics of the components of x_i under the set of kth order statistics of the components of $x_1, \cdots, x_n$.

c) A different HG-group for Ω_0 is

$$G_2^{(n)} = \{g_F^{(n)}; g_F \in G_2\}, \qquad \text{with} \quad g_F^{(n)}(x_1, \cdots, x_n) = (g_F(x_1), \cdots, g_F(x_n)).$$

But it is difficult to determine the maximal invariant for $G_2^{(n)}$. For the relation between different HG-groups cf. [3].

In the nonstationary case of testing whether the df's F_i of the p_i-dimensional random variables X_i, $1 \leqq i \leqq n$, are symmetric, we get $\Omega_0 = \{\prod_{i=1}^n F_i; F_i \in \Omega_s^{p_i}, 1 \leqq i \leqq n\}$, where $\prod_{i=1}^n F_i$ is the product measure of $F_1, \cdots, F_n$. The symmetry group for Ω_0 is $S^I = \prod_{i=1}^n \gamma_{p_i}$, and we obtain the following result.

THEOREM 3.

a) *A* HG-*group for the nonstationary multisample problem is* $G_4 = \prod_{i=1}^n G_2(p_i)$, *where* $G_2(p_i)$ *is the group constructed in Theorem* 1 *for the* p_i*-dimensional case.*

b) *A maximal invariant for* G_4 *is the N-tuple*

$$(g_{11}, \cdots, g_{p_1 1}, \cdots, g_{1n}, \cdots, g_{p_n n}), \quad N = \sum_{i=1}^n p_i.$$

3. Further applications and an open problem. The results of Section 2 are easily applied to further multivariate problems, which are characterized by symmetry groups in a similar way as Ω_s^p, e.g. to the hypothesis of coordinatewise symmetry about 0. To give a concrete example, let $s: R^p \to \mathbb{R}^p$ be the shift, defined by $s(x_1, \cdots, x_p) = (x_2, \cdots, x_p, x_1)$, and let $\Omega(s, p) = \{F \in \Omega^p; sF = F\}$, be the so-called hypothesis of stationarity.

The symmetry group for $\Omega(s, p)$ is the cyclic group S^I generated by s, $S^I = \{1, s, s^2, \cdots, s^{p-1}\}$; a corresponding minimal set E'' is given by $E'' = \bigcup_{\pi \in \gamma_{p-1}} \pi E$, with E as in Section 2 and $\pi(x_1, \cdots, x_p) = (x_{\pi(1)}, \cdots, x_{\pi(p-1)}, x_p)$ for $\pi \in \gamma_{p-1}$. A completeness theorem for $\Omega(s, p)$ follows easily from the theorem of Bell, Blackwell and Breiman (cf. [4]).

Let $\Omega^p(E'')$ be the class of all continuous distributions concentrated on E'' which have strictly monotone conditional df's on E'', then $g: \Omega(s, p) \to \Omega^p(E'')$ defined by $g(P) = P(\cdot \mid E'')$ is bijective. Defining now $k_F: R^p \to \mathbb{R}^p$ by $k_F(x) = s^k \circ \tau_F^{-1} \circ \tau_{F_0}(s^{-k}x)$ for $x \in s^k E''$, $1 \leqq k \leqq n$, $F \in \Omega^p(E'')$, we get the following analog of Theorem 1.

THEOREM 4.

a) $G_5 = \{k_F; F \in \Omega^p(E'')\}$ *is a* HG-*group for* $\Omega(s, p)$.
b) *A maximal invariant for* G_5 *is* $R(x) = \sum_{i=1}^{p} i I_{s^i E''}(x)$, $x \in \mathbb{R}^p$.

EXAMPLE 1. In the case $p = 3$ we have the rank sets

$$E'' = \{(x_1, x_2, x_3); x_3 > \max\{x_1, x_2\}\}$$
$$sE'' = \{(x_1, x_2, x_3); x_2 > \max\{x_1, x_3\}\}$$
$$s^2E'' = \{(x_1, x_2, x_3); x_1 > \max\{x_2, x_3\}\}$$

OPEN QUESTION. An alternative method for constructing DF rank tests is given in a paper by Witting [7]. In this paper ranks are defined as maximal invariants under a group G, so that Ω_0 is invariant under G, i.e. $G\Omega_0 \subset \Omega_0$ with the further requirement that the (finite) symmetry group S^I for Ω_0 has the following cross-sectional property:

a) for all x, y there exists a $\pi \in S^I$ so that $\pi G(x) = G(y)$ and
b) for all x and $\pi \in S^I$ there exists a y so that $\pi G(x) = G(y)$.

It is noteworthy that all constructions of HG-groups I have found in literature—including the one in this paper—have this cross-sectional property and conversely, that all constructions of groups in this second sense lead to HG-groups. The question remains open whether these two concepts are equivalent is some sense.

Acknowledgment. I thank the referees for several helpful suggestions.

REFERENCES

[1] BELL, C. B. and HALLER, H. S. (1969). Bivariate symmetry tests: parametric and nonparametric. *Ann. Math. Statist.* **40** 259–269.

[2] BELL, C. B., WOODROOFE, M. and AVADHANI, V. (1970). Some nonparametric tests for stochastic processes. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.). Cambridge Univ. Press.

[3] BELL, C. B. and KUROTSCHKA, V. (1971). Einige Prinzipien zur Behandlung nichtparametrischer Hypothesen. In *Studi di probability*, statistica e ricerca operativa in onore di Giusseppe Pompili Odensi-Gubbio.

[4] BELL, C. B. and SMITH, P. I. (1972). Completeness theorems for characterizing distribution—free statistics. *Ann. Inst. Statist. Math.* **24** 435–453.

[5] BERK, R. R. and BICKEL, P. J. (1968). On invariance and almost invariance. *Ann. Math. Statist.* **39** 1573–1576.

[6] SMITH, P. I. (1969). Structure of nonparametric tests of some multivariate hypothesis. Ph.D. thesis, Case Western Reserve Univ.

[7] WITTING, H. (1970). On the theory of nonparametric tests. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.). Cambridge Univ. Press.

INSTITUT FÜR MATH. STOCHASTIK
2 HAMBURG 13
ROTHENBAUMCHAUSSEE 45
GERMANY