CHAPTER 11

Type III Problems: Global Slice

In Type III problems there is, as in Type I problems, a group H such that $K = GH$ is a group transitive over $\mathcal{X}$ but not all items of Assumption 8.11 are satisfied. All assumptions except 8.11(ii) are rather mild and can be expected to hold in applications. On the other hand, Assumption 8.11(ii) turns out to be fairly strong and can easily fail. Inspection of the examples in Chapter 9 reveals that in Section 9.2, Case 1, 8.11(ii) holds because G and H commute (and therefore G_0 and H commute), whereas in all other examples the validity of 8.11(ii) is a consequence of $G_0 = \{e\}$. If neither G_0 and H commute nor $G_0 = \{e\}$, then 8.11(ii) is likely to fail. In that case $\mathcal{Z} = Hx_0$ need no longer be a cross section and it seems as though the additional structure provided by the group H is useless. However, it turns out that H and $\mathcal{Z}$ can still be useful provided a different kind of group structure exists. Consider:

11.1. ASSUMPTION. *Let Assumption 8.11 be satisfied except that 8.11(ii) is changed to*

(ii)′ $gH_0g^{-1} = H_0$ *for every* $g \in G$,

and in addition assume

(iii) $gHg^{-1} = H$ *for every* $g \in G$.

Thus, Assumption 11.1 is Assumption 8.11 with G and H interchanged (note that in 8.11 all assumptions except (ii) are symmetric in G and H), and in addition H is assumed to be normal in K.

The normality of H in K is actually quite common, as the examples in Chapter 9 show. We shall see that the consequences of Assumption 11.1 are as follows: the set $\mathcal{Z} = Hx_0$, although not a cross section, is transformed into itself by G_0 and G_0 acts on $\mathcal{Z}$; the intersections of the G-orbits in $\mathcal{X}$ with $\mathcal{Z}$ (which for a cross section would be exactly one point per orbit) are precisely the G_0-orbits in $\mathcal{Z}$; there is a 1-1 correspondence between the open G-invariant subsets of $\mathcal{X}$ and the open G_0-invariant subsets of $\mathcal{Z}$ so that the structure of the integrals of G-invariant functions on $\mathcal{X}$ according to the Bourbaki theory is the same as that of G_0-invariant functions on $\mathcal{Z}$; given a probability distribution on $\mathcal{X}$ it is possible to write down (in integral form) a probability distribution on $\mathcal{Z}$ such that the probabilities of corresponding G-invariant subsets of $\mathcal{X}$ and G_0-invariant subsets of $\mathcal{Z}$ coincide. Then a maximal invariant on $\mathcal{Z}$ under G_0, together with its distribution, is a solution of the original problem. In this way the original problem, with space $\mathcal{X}$ and group G, has been reduced to a (presumably) simpler problem with the smaller space $\mathcal{Z}$ and smaller group G_0. As in the case of Type I and II problems we shall usually represent $\mathcal{Z}$ by $\mathcal{T} = H/H_0$. It should be emphasized that under Assumption 11.1 there is not an obvious function $\mathcal{X} \to \mathcal{Z}$ that preserves probabilities of invariant sets and that can be used to induce a distribution on $\mathcal{Z}$ from one on $\mathcal{X}$. (This point was overlooked in Theorem 8.1 of Wijsman (1986); also the present Assumption 11.1(ii)′ was erroneously omitted.) The usefulness of the structure that Assumption 11.1 provides was first shown by Woteki and Mayer (1976) in several examples.

11.2. LEMMA. *Let Assumption 11.1 be satisfied and define* $\mathcal{Z} = Hx_0$, $\mathcal{T} = H/H_0$. *Then (i)* G_0 *acts on the left of* $\mathcal{Z}$ *and of* $\mathcal{T}$; *(ii) for any* $z \in \mathcal{Z}$, $\mathcal{Z} \cap Gz = G_0 z$. *Thus, there is a 1-1 correspondence between the* G*-orbits in* $\mathcal{X}$ *and the* G_0*-orbits in* $\mathcal{Z}$.

PROOF. (i) Let $z = hx_0$ $(h \in H)$ be an arbitrary point of $\mathcal{Z}$ and g_0 an arbitrary element of G_0. Then $g_0 z = h' x_0$, where $h' = g_0 h g_0^{-1} \in H$. This shows that G_0 transforms $\mathcal{Z}$ into itself, and since G_0 acts on the left of $\mathcal{X}$ it follows that G_0 acts on the left of $\mathcal{Z}$.

(ii) Obviously, $G_0 z \subset \mathcal{Z} \cap Gz$. In order to show $\mathcal{Z} \cap Gz \subset G_0 z$ let

$z_1 \in \mathcal{Z} \cap Gz$ so that $z_1 \in \mathcal{Z}$ and there exists $g \in G$ such that $z_1 = gz$. For some $h, h_1 \in H$, $z = hx_0$, $z_1 = h_1 x_0$. Then $h_1 x_0 = ghx_0$. Since H is normal in K, $gh = h_2 g$ for some $h_2 \in H$. Therefore, $h_2^{-1} h_1 x_0 = gx_0$. By Assumption 8.11(i) both members of the last equation must equal x_0. This implies $g \in G_0$ so that $z_1 = g_0 z \in G_0 z$. □

The 1-1 correspondence between the G-orbits in $\mathcal{X}$ and the G_0-orbits in $\mathcal{Z}$ provides a 1-1 correspondence between the G-invariant subsets of $\mathcal{X}$ and the G_0-invariant subsets of $\mathcal{Z}$ since an invariant set is a union of orbits. We shall show now that this correspondence preserves open sets.

11.3. LEMMA. *Under Assumption 11.1 there is a 1-1 correspondence between the G-invariant open subsets of $\mathcal{X}$ and the G_0-invariant open subsets of $\mathcal{Z}$.*

PROOF. Since Assumption 11.1 implies Assumption 8.11 with G and H interchanged, Theorem 8.12 applies with G and H interchanged. Therefore, define $\varphi^* : G/G_0 \times H/H_0 \to \mathcal{X}$ by

$$\varphi^*(gG_0, hH_0) = hgx_0, \tag{11.1}$$

then φ^* is a homeomorphism. Put $\mathcal{Y} = G/G_0$, $\mathcal{T} = H/H_0$, then φ^* is a homeomorphism of $\mathcal{Y} \times \mathcal{T}$ and $\mathcal{X}$. Equivalently,

$$\varphi^{**}(gG_0, hH_0) = hgK_0 \tag{11.2}$$

is a homeomorphism of $\mathcal{Y} \times \mathcal{T}$ and K/K_0. This homeomorphism is even analytic, by Theorem 5.9.9 with G and H interchanged. By Lemma 11.2 there is a 1-1 correspondence between the G-invariant subsets of $\mathcal{X}$ and the G_0-invariant subsets of $\mathcal{Z}$, where to $A \subset \mathcal{X}$ corresponds $A \cap \mathcal{Z} \subset \mathcal{Z}$, and to $B \in \mathcal{Z}$ corresponds $GB \subset \mathcal{X}$. We have to show (i) if $A \subset \mathcal{X}$ is G-invariant and open, then $A \cap \mathcal{Z}$ is open in $\mathcal{Z}$; and (ii) if $B \subset \mathcal{Z}$ is G_0invariant and open in $\mathcal{Z}$, then GB is open in $\mathcal{X}$. Part (i) follows from the homeomorphism between $\mathcal{Y} \times \mathcal{T}$ and $\mathcal{X}$, and between $\mathcal{T}$ and $\mathcal{Z}$ (but note that the action of G on $\mathcal{T}$ is not trivial so that $\varphi^{*-1}(A)$ is not a product set). Then (i) reduces to the statement that if A is an open subset of the product space $\mathcal{Y} \times \mathcal{T}$ endowed with the product topology, then $A \cap \mathcal{T}$ is open in $\mathcal{T}$. This is an immediate consequence of the definition of product topology, for A can be written as a union of open product sets. In order to show (ii) it is sufficient to show that if B_1 is an open subset of $\mathcal{T}$, then GB_1 is open in $\mathcal{Y} \times \mathcal{T}$, or, equivalently, open in K/K_0 which is homeomorphic to $\mathcal{Y} \times \mathcal{T}$ (note that the action of G on $\mathcal{Y} \times \mathcal{T}$ is derived from the action of G on $\mathcal{X}$). Let $G \times H$ be endowed with the product topology and define $f_1 : G \times H \to K/K_0$ by $f_1(g,h) = hgK_0$. Similarly, f_2 by $f_2(g,h) = ghK_0$. Since $B_1 = \{hH_0 : h \in B_2\}$ for some open $B_2 \subset H$, it suffices to show that f_2 is an open mapping. Now f_1 is the composition of the open orbit projection $G \times H \to \mathcal{Y} \times \mathcal{T}$ and the homeomorphism φ^{**} given by (11.2). Therefore, f_1 is open. Consider the function $f_3 : G \times H \to G \times H$ defined by $f_3(g,h) = (g, ghg^{-1})$, where we have used Assumption 11.1(iii). Obviously, f_3 is continuous, and its inverse $f_3^{-1}(g,h') = (g, g^{-1}h'g)$ is also continuous so that f_3 is a homeomorphism. Then observe that $f_2 = f_1 \circ f_3$ and conclude that f_2 is open. □

11.4. THEOREM. *Let Assumption 11.1 be satisfied and let $P^X(dx) = p(x)\lambda(dx)$ be a probability distribution on $\mathcal{X}$, with λ relatively invariant with respect to K with multiplier χ. Define the following probability distribution on $\mathcal{T} = H/H_0$:*

$$(11.16) \qquad P^T(dt) = c\chi(t)\mu_{\mathcal{T}}(dt) \int p(ghx_0)\chi(g)\mu_G(dg), \quad [h] = t,$$

with suitable $c > 0$, in which $\mu_{\mathcal{T}}$ and μ_G have the same meaning as in Theorem 8.14. Then the distribution on $\mathcal{X}/G$ induced by P^X is the same as the distribution on $\mathcal{T}/G_0$ induced by P^T.

PROOF. Since there is a homeomorphism between $\mathcal{T}$ and $\mathcal{Z} = Hx_0$, we may and shall switch back and forth between $(\mathcal{T}, \mathcal{T}/G_0)$ on one hand and $(\mathcal{Z}, \mathcal{Z}/G_0)$ on the other. Since by Lemma 11.2 there is a 1-1 correspondence between the G-orbits in $\mathcal{X}$ and the G_0-orbits in $\mathcal{Z}$ we may identify $\mathcal{X}/G$ and $\mathcal{Z}/G_0$ as point sets. By Lemma 11.3 this correspondence is a homeomorphism, so that we may identify $\mathcal{X}/G$

and $\mathcal{Z}/G_0$ as l.c. spaces. By Lemma 11.2 there is also a 1-1 correspondence between G-invariant functions $\mathcal{X} \to R$ and G_0-invariant functions $\mathcal{Z} \to R$, obtained by equating the functions on corresponding orbits. Write this correspondence as $f_0 = \alpha(f)$, $f = \alpha^{-1}(f_0)$, for f on $\mathcal{X}$, f_0 on $\mathcal{Z}$. Let $\mathcal{F}[\mathcal{F}_0]$ be the family of G-invariant functions $\mathcal{X} \to R$ [G_0-invariant functions $\mathcal{Z} \to R$] that are continuous and bounded. Then the 1-1 correspondence between the invariant open sets shown by Lemma 11.3 guarantees that α is a 1-1 correspondence between $\mathcal{F}$ and $\mathcal{F}_0$. Now the distribution on $\mathcal{X}/G$ induced by P^X is determined by the values of the expectations $\int f dP^X$, $f \in \mathcal{F}$. Similarly, for any probability distribution P^Z on $\mathcal{Z}$, the distribution on $\mathcal{Z}/G_0$ induced by P^Z is determined by the values of $\int f_0 dP^Z$, $f_0 \in \mathcal{F}_0$. Therefore, if $\int f dP^X = \int \alpha(f) dP^Z$ for every $f \in \mathcal{F}$, then the distribution on $\mathcal{X}/G$ induced by P^X equals the distribution on $\mathcal{Z}/G_0$ induced by P^Z. Now switch to $(\mathcal{T}, P^T)$ and consider the functions f_0 of $\mathcal{F}_0$ to be on $\mathcal{T}$ rather than on $\mathcal{Z}$. Then with P^X and P^T of the hypotheses of the theorem it is to be shown that

$$\int f dP^X = \int \alpha(f) dP^T, \quad f \in \mathcal{F}. \tag{11.17}$$

Here $\alpha(f)$ is obtained from f by equating the two functions on $\mathcal{Z}$. Since $z \in \mathcal{Z}$ is of the form $z = hz_0$, $h \in H$, we have

$$\alpha(f)(hx_0) = f(hx_0), \quad h \in H. \tag{11.18}$$

Write down (8.24) for λ-integrable f; this only uses the first part of Assumption 8.11. Combine this with (7.6.6), valid for H normal in K, and get

$$\int f(x)\lambda(dx) = c \int f(ghx_0)\chi(g)\chi(h)\mu_G(dg)\mu_H(dh). \tag{11.19}$$

Now replace f by fp, with $f \in \mathcal{F}$ and p the density with respect to λ of P^X. Then

$$\begin{aligned} &\int f(x)p(x)\lambda(dx) \\ &\qquad = c \int f(ghx_0)p(ghx_0)\chi(g)\chi(h)\mu_G(dg)\mu_H(dh). \end{aligned} \tag{11.20}$$

Use the invariance of $f : f(ghx_0) = f(hx_0)$, and write the right-hand side of (11.20) as an iterated integral:

$$\int f(x)p(x)\lambda(dx) = c\int f(hx_0)\chi(h)\mu_H(dh)\int p(ghx_0)\chi(g)\mu_G(dg). \tag{11.21}$$

The left-hand sides of (11.17) and (11.21) agree, and so do the right-hand sides, using (11.16) and (11.18), when the integration over $\mathcal{T}$ in (11.17) is carried back to an integration over H. □

11.5. REMARK. Under Assumption 11.1, the set $\mathcal{Z}$ is in general not a cross section but it is a so-called **global slice** for the group G. As defined by Palais (1961), a global slice at $x_0 \in \mathcal{X}$ is a set $\mathcal{Z} \subset \mathcal{X}$ containing x_0 such that (i) $G\mathcal{Z} = \mathcal{X}$, and (ii) there is an equivariant continuous function $f : \mathcal{X} \to G/G_0$ such that $f^{-1}(G_0) = \mathcal{Z}$. A (global) cross section is a global slice with the additional property that it has exactly one point in common with each orbit. In our case the function f can be defined by $f(hgx_0) = gG_0$, which is well-defined because of the 1-1 function φ^* of (11.1). This also shows that $f(x) = G_0$ if and only if x is of the form hx_0, i.e., $x \in \mathcal{Z}$. Therefore, $f^{-1}(G_0) = \mathcal{Z}$. The equivariance follows from the following computation: if $x = hgx_0$ and $g_1 \in G$ then by the normality of H, $g_1hg_1^{-1} = h_1 \in H$ so that $f(g_1x) = f(g_1hgx_0) = f(h_1g_1gx_0) = g_1gG_0 = g_1f(x)$. The continuity of f follows from the homeomorphism φ^* of (11.1) by writing $f = \mathrm{pr}_1 \circ \varphi^{*-1}$ where pr_1 is the projection of $\mathcal{Y} \times \mathcal{T}$ on $\mathcal{Y}$; then observe that both pr_1 and φ^{*-1} are continuous. □

Theorem 11.4 shows that if we start with a space $\mathcal{X}$ and probability distribution of the form $P^X(dx) = p(x)\lambda(dx)$, then the problem of giving an explicit expression for the distribution of a maximal invariant under the action of G may be replaced by the analogous problem where the space is $\mathcal{T}$, the group is G_0, and the distribution is P^T given by (11.16). Note that in contrast with Type I or II problems we don't have a factorization result such as (8.10) or (8.22). Consequently, for obtaining the constant c in (11.16) the method that was used in Chapters 9 and 10, and which consisted in obtaining, by differentiation, an explicit factorization at a special point or points (as in Example 8.7), is not available here. Below we shall give an example of the use of Theorem 11.4.

11.6. EXAMPLE. Suppose we have a sample from a p-variate population that is partitioned into three p_i-variate subpopulations, where $p_1 + p_2 + p_3 = p$. At first we shall assume that the population is multivariate normal. Suppose it is given that the first subpopulation is independent of the third and we want to test that the first is also independent of the second (as in Das Gupta, 1977, Problem(ii) (with 2 and 3 interchanged) and in Marden, 1981, Problem P_2). Let inference depend only on the sample covariance matrix S. Thus, $\mathcal{X} = PD(p)$. We take λ to be Lebesgue measure on $\mathcal{X}$. Partition S into 3×3 blocks according to the three subpopulations. The group G of invariance transformations may be chosen to consist of all matrices C of the form

$$C = \begin{bmatrix} A_1 & & \\ & A_2 & B \\ & & A_2 \end{bmatrix} \tag{11.22}$$

in which $A_i \in GL(p_i)$, $i = 1,2,3$, and $B \in M(p_2, p_3)$. For H we take the group of all matrices

$$C = \begin{bmatrix} I_{p_1} & D & E \\ & I_{p_2} & \\ & & I_{p_3} \end{bmatrix} \tag{11.23}$$

in which $D \in M(p_1, p_2)$, $E \in M(p_1, p_3)$. The action of both G and H on $\mathcal{X}$ is defined by $S \to CSC'$. Then $K = GH$ is a transitive group over $\mathcal{X}$ and H is normal in K. Take $x_0 = \text{diag}(I_{p_1}, I_{p_2}, I_{p_3})$, then it is seen that H_0 is trivial and G_0 consists of all block-diagonal matrices $\text{diag}(\Gamma_1, \Gamma_2, \Gamma_3)$, $\Gamma_i \in O(p_i)$, $i = 1,2,3$. Here Assumption 8.11(ii) is violated but Assumption 11.1 is satisfied so that Theorem 11.4 applies. In (11.16), $\mathcal{T} = H$ since $H_0 = \{e\}$. For $\mu_{\mathcal{T}}(dt) = \mu_H(dh)$ we may take

$(dD)(dE)$. The multiplier χ for both G and H is $\chi(C) = |C|^{p+1}$, by (9.1.4). But for $C \in H$, $|C| = 1$ so that $\chi(h) = 1$. On the other hand, for $C \in G$ we have $\chi(g) = \prod_{i=1}^{3} |A_i|^{p+1}$. To get an explicit expression for $\mu_G(dg)$ it is convenient to write G as G_1G_2, where G_1 consist of all $\operatorname{diag}(A_1, A_2, A_3)$, $A_i \in GL(p_i)$, and G_2 of all matrices of the form (11.22) with $A_i = I_{p_i}$. Then G_2 is normal in G, and we can use Corollary 7.6.2 (replace there K, G, H by G, G_1, G_2 here) with the result $\int f(g)\mu_G(dg) = \int f(g_1g_2)\mu_{G_1}(dg_1)\mu_{G_2}(dg_2)$. Here $\mu_{G_2}(dg_2)$ can be taken as (dB) and $\mu_{G_1}(dg_1)$ as $\prod_{i=1}^{3} |A_i|^{-p_i}(dA_i)$, by (7.7.1). Substitution of all this into (11.16) yields a formula for P^T, which we shall rename P_1^T for later use:

$$P_1^T(dt) = p_1(D, E)(dD)(dE), \tag{11.24}$$

in which

$$p_1(D, E) = c \int p(S) \prod_{i=1}^{3} |A_i|^{p-p_i+1}(dA_i)(dB), \tag{11.25}$$

where S, partitioned into the submatrices S_{ij}, $i, j = 1, 2, 3$, with $S_{ji} = S'_{ij}$, depends on the matrices A_i, etc., as follows:

$$\begin{aligned} &S_{11} = A_1(I + DD' + EE')A_1', \quad S_{12} = A_1(D + EB')A_2', \\ &S_{13} = A_1EA_3', \quad S_{22} = A_2(I + BB')A_2', \\ &S_{23} = A_2BA_3', \quad S_{33} = A_3A_3'. \end{aligned} \tag{11.26}$$

We may now drop insistence that S be based on a sample from a multivariate normal distribution and let $p(S)$ be any probability density with respect to Lebesgue measure λ on $PD(p)$.

The space $\mathcal{T}$ consists of the pairs of matrices (D, E), $D \in M(p_1, p_2)$, $E \in M(p_1, p_3)$. The action of G_0 on $\mathcal{T}$ is given by $(D, E) \to (\Gamma_1 D\Gamma_2', \Gamma_1 E\Gamma_3')$. A maximal invariant and its distribution can be obtained in two steps. First consider $(D, E) \to (D\Gamma_2', E\Gamma_3')$ with maximal invariant, say, $(S_1, S_2) = (DD', EE')$. This can be handled by Section 9.2, using density p_1 given by (11.25). In the second step G_0 acts on (S_1, S_2) by $S_i \to \Gamma_1 S_i \Gamma_1'$, $i = 1, 2$. A maximal invariant and its distribution can be handled by Section 10.6. □

11.7. REMARK. Although Example 11.6 is a good illustration of the use of Theorem 11.4 it is less fortunate in that the problem of Example 11.6 can also be solved by a succession of Type I and II problems, thereby avoiding the use of Theorem 11.4 altogether. The group G_1 of matrices $\text{diag}(A_1, A_2, A_3)$ can be further factored into $G_1 = G_0 G_3$, where G_0 is as before and G_3 consists of matrices of the form $\text{diag}(T_1, T_2, T_3)$, with $T_i \in UT(p_i)$. This factorization follows, for each $i = 1, 2, 3$ separately, from Section 7.7.4. Then a maximal invariant may be obtained in two steps: first under the group $G_3 G_2 = G'$, say (where G_2 was defined in Example 11.6), and then under G_0. In the first step $G'H$ is a transitive group over $\mathfrak{X}$ and $G'_0 = \{e\}$. As a result, Assumption 8.11 is satisfied and therefore Theorem 8.14 can be used. A maximal invariant is again (D, E), as in the first step of Example 11.6, but its distribution, say P_2^T, is in general different from the distribution P_1^T defined by (11.24) and (11.25). The second step, reduction by G_0, is the same as in the second step of Example 11.6 and leads to the same final result even though P_1^T and P_2^T are in general different. The relation between these latter two distributions can be described by saying that P_2^T is P_1^T averaged with help of G_0 (which is compact) so that—loosely speaking—the conditional distribution on each G_0-orbit becomes uniform. It is not known whether every Type III problem can be reduced to a succession of Type I and II problems. But even in cases where it can be done this may not be obvious by inspection. And, finally, even if one knows how to make this reduction (as in the problem of Example 11.6), it is not a priori clear which of the integrals in (8.23) and (11.16) is the easier of the two to work out. □