GENERALIZED TOLERANCE LIMITS

By J. H. B. Kemperman

Purdue University

1. Summary. A method for constructing tolerance limits due to Fraser [8] is generalized by allowing that each step of the construction may depend not only on the blocks previously formed but also on all the known boundary observations and, moreover, on certain sets of indices. Furthermore, Tukey's [5] lexicographical ordering is replaced by a more general type of ordering.

2. Introduction. Let $\{\Omega, \mathfrak{A}, \mu(A)\}$ be a measure space with $\mu(\Omega) = 1$ and $\mathfrak{A}$ complete. Then the relation $P(X \varepsilon A) = \mu(A)$ $(A \varepsilon \mathfrak{A})$ defines a random variable X taking values in the space Ω. Let $W = (x_1, \cdots, x_n)$ be a set of n independent observations on X and let $D_j = D_j(W)$ $(j = 1, 2, \cdots)$ be disjoint measurable subsets of Ω depending on W. These D_j sets are called (nonparametric) tolerance limits when the joint distribution of the random "coverages" $\mu(D_j)$ does not depend on the true distribution $\mu(A)$ of X, given that the latter belongs to a certain rather wide class of probability measures. Such tolerance limits were first introduced by S. S. Wilks [1] whose method was generalized to a far extent by A. Wald, H. Scheffé, J. W. Tukey, R. Wormleighton, and D. A. S. Fraser ([2], [3], [4], [5], [6], [7], [8]).

3. Ordering. By a (generalized) ordering o in Ω we shall mean an assignment of exactly one of the relations $x_1 < x_2$, $x_1 \sim x_2$, or $x_1 > x_2$ to each pair x_1, x_2 of points in Ω, such that $x_1 \sim x_2$ is an equivalence relation and such that o induces an (ordinary) transitive ordering among the corresponding equivalence classes. Let $\Omega = A \cup B$ with $A < B$ in the obvious sense. We shall assume that always: (i) A is measurable. (ii) If A is non-empty, we have $A = \bigcup_k \{x \mid x \leqq a_k\}$ for some (at most denumerable) subsequence $\{a_k\}$ of A. Similarly, if B is non-empty, we have $A = \bigcap_k \{x \mid x < b_k\}$ for some subsequence $\{b_k\}$ of B.

One way of obtaining such a generalized ordering is as follows: Let M be a finite or denumerable well-ordered set and let, for each m in M, $g_m(x)$ be a real-valued measurable function on Ω. If $g_m(x_1) = g_m(x_2)$ for all m in M, we define $x_1 \sim x_2$. Otherwise, $x_1 < x_2$ if and only if $g_s(x_1) < g_s(x_2)$, where s is the smallest index such that $g_s(x_1) \neq g_s(x_2)$.

An ordering o is said to be *continuous* (with respect to the measure $\mu(A)$) when for each x_0 in Ω we have $\mu\{x \mid x \sim x_0\} = 0$.

Lemma 1. *Let o be a continuous ordering and let $q(x_0) = P(X < x_0) = \mu\{x \mid x < x_0\}$. Then $q(X)$ is a uniformly distributed random variable in* $[0, 1]$.

Proof. Let $0 \leqq q \leqq 1$, $A = \{x \mid q(x) \leqq q\}$, and $B = \{x \mid q(x) > q\}$. We have to show that

$$P(q(X) \leqq q) = P(X \varepsilon A) = \mu(A) = q.$$

Received September 11, 1954, revised January 29, 1955.

But for a_k in A, we have $\mu\{x \mid x \leqq a_k\} = \mu\{x \mid x < a_k\} = q(a_k) \leqq q$; hence, from (ii), $\mu(A) \leqq q$ whether or not A is empty. Moreover, for b_k in B, we have $\mu\{x \mid x < b_k\} = q(b_k) > q$; hence, from (ii), $\mu(A) \geqq q$ whether or not B is empty.

4. Partitioning. Let m, m_0, and m_1 be positive integers, $m = m_0 + m_1$. Let $x_1, \cdots, x_{m-1}$ be $m - 1$ points in a measurable subset D of Ω and let o be a given ordering. Denoting by x^* the m_0-th smallest ($= m_1$-th largest) point x_i with respect to o, the partition of D into the three disjoint subsets $D_0 = \{x \mid x < x^*, x \varepsilon D\}$, $D_1 = \{x \mid x > x^*, x \varepsilon D\}$, and $D^* = \{x \mid x \sim x^*, x \varepsilon D\}$ is called the (m_0, m_1)-partition of D with respect to o and to the $m - 1$ points x_i in D. Note that, when $x_i \sim x_j$ does not happen for $i \neq j$, the "boundary" element x^* is unique, while D_0, D_1, and D^* contain exactly $m_0 - 1$, $m_1 - 1$, and 1 elements x_i, respectively. If o is continuous, $\mu(D^*) = 0$, hence, $\mu(D) = \mu(D_0) + \mu(D_1)$.

LEMMA 2. *If $\mu(D) > 0$ we assume that o is continuous and that $x_1, \cdots, x_{m-1}$ are $m - 1$ independent observations on X restricted to $X \varepsilon D$. Then $\mu(D_0) = q\mu(D)$, where q is a random variable which has the incomplete Beta-function $I_q(m_0, m_1)$ as its cumulative distribution function.*

PROOF. We may assume that $\mu(D) > 0$. Let Y be the random variable whose distribution $\nu(A) = \mu(A)\mu(D)^{-1}(A \subset D)$ is that of X restricted to $X \varepsilon D$. Observing that o induces an ordering on D which is continuous with respect to $\nu(A)$, it follows from Lemma 1 (replacing Ω by D, and X by Y) that for $q(x_0) = \nu\{x \mid x < x_0, x \varepsilon D\}$ the variable $q(Y)$ is uniformly distributed in [0, 1]. Hence, $q(x^*)$ is the m_0-th smallest among $m - 1 = m_0 + m_1 - 1$ independent observations $q(x_i)$ on a uniformly distributed random variable in [0, 1]. This proves that $q(x^*) = \mu(D_0)\mu(D)^{-1}$ has the d.f. $I_q(m_0, m_1)$.

5. The construction. For the sake of brevity, we shall employ a somewhat colloquial language. In the construction two persons are involved: a statistician (S) and his assistant (A). A knows precisely the actual outcomes of the n independent observations $x_1, \cdots, x_n$ on X, while, at the very outset, S has no information at all about these outcomes. On the other hand, S has at his disposal a class H of orderings o in Ω known to be continuous with respect to the distribution $\mu(A)$ of X.

In the first step of the construction, S selects an ordering 0_1 from H and a positive integer m_0, $m_0 \leqq n$, and asks A to give him the m_0-th smallest observation $x^*(1)$ with respect to 0_1 (this element is unique with probability 1), together with the two sets of indices corresponding to the $m_0 - 1$ and $m_1 - 1 = n - m_0$ observations which are smaller or larger than $x^*(1)$, respectively. Now, S can draw the (m_0, m_1)-partition $\Omega = \Omega_0 \cup \Omega^* \cup \Omega_1$ of Ω with respect to 0_1 and the set of the n observations x_i in Ω. Let $D^0(0) = \Omega$, $D^1(j) = \Omega_j$ $(j = 0, 1)$, and $D_1^* = \Omega^*$.

After k steps, $0 \leqq k \leqq n - 1$, S has obtained a partition of Ω into $k + 1$ disjoint "blocks" $D^k(j)$ $(j = 0, 1, \cdots, k)$ and k boundary sets D_i^* $(i = 1, \cdots, k)$, each of μ-measure 0. Further, for each of these $2k + 1$ sets, S knows precisely the set of indices corresponding to the observations x_i within the set.

Finally, for each boundary set D_i^* $(i = 1, \cdots, k)$, S knows the actual value of the boundary observation $x^*(i)$ in D_i^* (with a probability 1 these boundary observations are unique).

Now, the $(k + 1)$-th step of the construction proceeds as follows: His choice depending, *in any way whatsoever*,[1] on the knowledge acquired, S chooses: (i) A distinguished block $D = D^k(j^*)$ among those of the $k + 1$ blocks $D^k(j)$ $(j = 0, \cdots, k)$ which contain at least one observation. (ii) A positive integer m_0 not larger than the number $m - 1$ of observations in D. (iii) An ordering 0_{k+1} from H.

He then asks A for the m_0-th smallest observation $x^*(k + 1)$ in D with respect to 0_{k+1}, together with the two sets of indices corresponding to the $m_0 - 1$ or $m_1 - 1 = m - m_0 - 1$ observations in D which are smaller or larger than $x^*(k + 1)$, respectively.

Using the acquired value $x^*(k + 1)$, S is now able to draw the (m_0, m_1)-partition $D = D_0 \cup D^* \cup D_1$ of D with respect to 0_{k+1} and the $m - 1$ observations in D. Afterwards, he renumbers the blocks $D^k(0), \cdots, D^k(j^* - 1), D_0, D_1, D^k(j^* + 1), \cdots, D^k(k)$ as $D^{k+1}(j)$ $(j = 0, \cdots, k + 1)$, in this order. Finally, let $D_{k+1}^* = D^*$.

After exactly n steps the construction stops. Then S has obtained a partition of Ω into $n + 1$ disjoint blocks $U_j = D^n(j)$ $(j = 0, \cdots, n)$ and, further, n boundary sets D_k^* $(k = 1, \cdots, n)$, each of μ-measure 0.

THEOREM 1. *The coverages* $c_j = \mu(U_j)$ $(j = 0, \cdots, n)$ *have the joint distribution* $dc_1\, dc_2 \cdots dc_n$, *where* $c_j \geqq 0$, $c_1 + \cdots + c_n = 1 - c_0 \leqq 1$. *Moreover, the union* U *of* m *distinct sets* U_j *has a coverage* $p = \mu(U)$ *with* d.f. $I_p(m, n + 1 - m)$.

Let $0 < \alpha < 1$, and let $p = p_m(\alpha)$ be such that $I_p(m, n + 1 - m) = \alpha$. Then with a probability $1 - \alpha$, the random set U contains at least a proportion $p_m(\alpha)$ of the total probability mass 1 in Ω (i.e., we have confidence limits on the distribution of X or its parameters). For $\alpha = .01$ or $.05$, the value $p_m(\alpha)$ may be determined by using F-tables. Let F_0 be the α-point of the F-distribution with $n_1 = 2(n + 1 - m)$ and $n_2 = 2m$ degrees of freedom. Then $p_m(\alpha) = (1 + F_0 n_1/n_2)^{-1}$.

Some warning seems desirable. If the construction stops after k steps, we have a partition of Ω into the blocks $D^k(j)$ $(j = 0, \cdots, k)$ and, further, k boundary sets of measure 0. Let the random variable $n_j - 1$ denote the number of observations in $D^k(j)$ $(j = 0, \cdots, k)$, and let $N_j = n_0 + \cdots + n_{j-1}$. One can easily see that $D^k(j)$ is the union of the "final" blocks $U_{N_j}, \cdots, U_{N_j+n_j-1}$ (which might be found by completing the construction) and, further, some set of measure 0. However (as certain counterexamples show), this does not imply that conditional to $n_j = m$ (m given) the coverage of $D^k(j)$ has the conditional dis-

[1] Chance decisions are also allowed. For example, instead of making each decision as the necessity for it arises, S could start with a complete plan which provides for all contingencies. Then we may as well assume that S has already determined beforehand the actual outcomes of the random decisions which might arise.

tribution $I_p(m, n + 1 - m)$. Generally, the latter conclusion is only justified when (with a probability 1) both N_j and n_j are constant.

6. Proof of Theorem 1. In order to keep the proof on an elementary level, we shall avoid an explicit use of the usual complicated measure preserving transformations (cf. Fraser [8], pp. 53–54). Let $t_j = t(j) = s_j\sqrt{-1}$ $(j = 0, \cdots, n)$ be complex parameters with s_j real. It suffices to show that the characteristic function

$$E[\exp (t_0 \log c_0 + \cdots + t_n \log c_n)] = E(c_0^{t_0} \cdots c_n^{t_n})$$

depends only on n and the t_j but not on the distribution of X or the actual mode of construction. For then the joint distribution of $\log c_0, \cdots, \log c_n$, and hence the joint distribution of $c_0, \cdots, c_n$, will not be affected when X is replaced by a real random variable, uniformly distributed in [0, 1], and when the ordering 0_k $(k = 1, \cdots, n)$ is replaced by the common ordering in [0, 1]. But then $c_0, \cdots, c_n$ become the differences between consecutive order statistics, and Theorem 1 now follows from a well-known result (cf. Wilks [1]).

After k steps, $0 \leqq k \leqq n$, the construction based on the sample $W = (x_1, \cdots, x_n)$, yields (with a probability 1) a partition of Ω into the $k + 1$ blocks $D^k(j)$ $(j = 0, \cdots, n)$ and, further, k boundary sets of measure 0. Let the random variable $n_j - 1$ equal the number of observations inside $D^k(j)$ and let $N_j = n_0 + n_1 + \cdots + n_{j-1}$. Now consider the quantity

$$\rho_k = \prod_{j=0}^{k} \frac{\Gamma(n_j)\, \mu(D^k(j))^{t(N_j)+\cdots+t(N_j+n_j-1)}}{\Gamma[n_j + t(N_j) + \cdots + t(N_j + n_j - 1)]},$$

depending on the parameters $t_0, \cdots, t_n$. Though for $1 < k < n$ the joint distribution of $\mu(D^k(0)), \cdots, \mu(D^k(k))$ depends strongly on S's method of construction, it turns out that, for any mode of construction and for each (arbitrary but fixed) set of values $t_0, \cdots, t_n$,

$$E(\rho_k) = n!\, \Gamma(n + 1 + t_0 + \cdots + t_n)^{-1} \qquad (k = 0, 1, \cdots, n). \tag{1}$$

For $k = n$, we have $n_j = 1$, $N_j = j$, $\mu(D^k(j)) = \mu(U_j) = c_j$ $(j = 0, \cdots, n)$, and (1) implies

$$E(c_0^{t_0} \cdots c_n^{t_n}) = n!\, \Gamma(n + 1 + t_0 + \cdots + t_n)^{-1} \prod_{j=0}^{n} \Gamma(t_j + 1),$$

where indeed the right-hand side depends only on n and the t_j.

Formula (1) is evident for $k = 0$. For, $\mu(D^0(0)) = \mu(\Omega) = 1$ and $n_0 = n + 1$, $N_0 = 0$ (when $k = 0$) imply that ρ_0 is always equal to the right-hand side of (1). Let k be a fixed integer, $0 \leqq k \leqq n - 1$; it suffices to prove that $E(\rho_k) = E(\rho_{k+1})$.

Let $D^k(j)$ $(j = 0, \cdots, k)$, $D = D^k(j^*)$, D_0, D_1, m, m_0, and m_1 be as defined in the $(k + 1)$-th step of the construction. Here, with probability 1, D_0 and D_1 contain precisely $m_0 - 1$ and $m_1 - 1$ observations, respectively; $(m_0 + m_1 = m)$. Moreover, $\mu(D) = \mu(D_0) + \mu(D_1)$. It now follows from the definitions of ρ_k,

ρ_{k+1}, and the blocks $D^{k+1}(j)$ $(j = 0, \cdots, k+1)$ that

$$\rho_{k+1} = \rho_k \frac{\Gamma(m_0)\,\Gamma(m_1)\,\Gamma(m + t' + t'')}{\Gamma(m)\,\Gamma(m_0 + t')\,\Gamma(m_1 + t'')} q^{t'}(1 - q)^{t''}, \tag{2}$$

where $\mu(D_0) = q\mu(D)$ and, with $N = N_{j^*}$,

$$t' = t(N) + \cdots + t(N + m_0 - 1), \qquad t'' = t(N + m_0) + \cdots + t(N + m - 1).$$

In view of the footnote to the construction, we may assume without loss of generality that S has a complete non-random plan of construction which provides for all contingencies. The following information Σ has been received by S from A during the first k steps of the construction: (i) For $i = 1, \cdots, k$, the value ξ_i and the index ν_i of the boundary observation $x^*(i)$. (ii) For $j = 0, \cdots, k$, the indices $\sigma(j, h)$ $(h = 1, \cdots, n_j - 1)$ of the observations in the block $D^k(j)$. Here, the n different integers ν_i and $\sigma(j, h)$ together constitute the full set of indices $(1, 2, \cdots, n)$.

Knowing only Σ, S can reconstruct the blocks $D^k(j)$ $(j = 0, \cdots, k)$ according to plan; hence, Σ is equivalent to the information known to S at the beginning of the $(k + 1)$-th step. Therefore, Σ completely determines the quantity ρ_k, the distinguished block $D = D^k(j^*)$, together with the ordering 0_{k+1}, and the positive integers m_0 and m_1 $(m_0 + m_1 = m = n_{j^*})$ mentioned in the $(k + 1)$-th step of the construction.

To almost all samples W there corresponds a set of information Σ of the above type. Among these corresponding Σ's, let Σ_0 be a specific set of information (i) and (ii). Denoting the ith observation by $x(i)$, it is evident that in an actual construction Σ_0 will arise if and only if: (i) $x(\nu_i) = \xi_i$ $(i = 1, \cdots, k)$. (ii) For $j = 0, \cdots, k$, we have $x(\sigma(j, h)) \in D^k(j)$ $(j = 1, \cdots, n_j - 1)$, where the set $D^k(j)$ is uniquely determined by Σ_0. Hence, we have for $0 \leqq j \leqq k$ that, given $\Sigma = \Sigma_0$, the observations $x(\sigma(j, h))$ $(h = 1, \cdots, n_j - 1)$ behave as $n_j - 1$ independent observations on the random variable X restricted to $X \in D^k(j)$ (provided $\mu(D^k(j)) > 0$).

Further, D_0 is obtained as the "lower" set in the (m_0, m_1)-partition of $D = D^k(j^*)$ with respect to the $n_{j^*} - 1 = m - 1$ observations $x(\sigma(j^*, h))$ in D and the continuous ordering 0_{k+1}. It follows from Lemma 2 that, given $\Sigma = \Sigma_0$, we have $\mu(D_0) = q\mu(D)$, where q has the conditional d.f. $I_q(m_0, m_1)$.

Moreover, given $\Sigma = \Sigma_0$, the quantities ρ_k, m, m_0, m_1, t', and t'' are constants. It now follows from (2) that

$$E(\rho_{k+1} \mid \Sigma = \Sigma_0) = \rho_k = E(\rho_k \mid \Sigma = \Sigma_0),$$

implying that $E(\rho_{k+1}) = E(\rho_k)$.

7. A remark. The above proof is not completely rigorous because the very last step ("implying that") is still open to doubt for lack of a precise definition of the expected values $E(\rho_{k+1} \mid \Sigma = \Sigma_0)$, $E(\rho_{k+1})$, etc. The latter omission is also the root of the following difficulty.

If, in the construction, S's decisions depend too wildly (that is, in a non-measurable way) on the available information, it may easily happen that the coverage c_j of the final block U_j is a non-measurable function (with respect to the Borel field $\mathfrak{A}^n$ in Ω^n) of the sample point $W = (x_1, \cdots, x_n)$. Then the question arises as to what (in the assertion of Theorem 1) is meant by the probability $\Pr(c_j \leqq a_j)$ $(j = 0, \cdots, n)$. The following approach to this question, which avoids additional measurability assumptions, was indicated to me by Prof. D. A. S. Fraser.

For simplicity, let us assume that S starts with a complete *non-random* plan which provides for all contingencies. Let Q stand for a specific (a priori possible) outcome of the indices of the observations inside the blocks $D^k(j)$ $(k = 1, \cdots, n;$ $j = 0, \cdots, k)$ and the indices i_k of the boundary observations $x^*(k)$ $(k = 1, \cdots, n)$. Let the (finitely many) different possible outcomes Q be denoted by $Q_1, \cdots, Q_p$. Let $f_r(W) = 1$ $(r = 1, \cdots, p)$ when the construction based on W yields the outcome Q_r; otherwise, $f_r(W) = 0$. Thus, $\sum_r f_r(W) = 1$ for almost all W.

Let F_n be the class of all the subsets B of Ω^n such that, for $r = 1, \cdots, p$, the integral

$$P_r(B) = \int_{\Omega^n} [[f_r(W)\, \chi_B(W)\, d\mu(x_{i_n})] \cdots d\mu(x_{i_1})]$$

has a meaning and exists as a *repeated* Lebesgue-Stieltjes integral; here $(i_1, \cdots, i_n)$ corresponds to Q_r, while $\chi_B(W)$ denotes the characteristic function of B. Let F_k $(0 \leqq k < n)$ be the class of F_n-sets B such that $W_1 \subset B$ implies $W_2 \subset B$ whenever the two constructions based on W_1 and W_2 yield, at the end of the kth step, exactly the same information Σ (cf. the above proof).

One can show that: (1) F_k is a Borel field $(k = 0, \cdots, n)$ and $F_0 \subset F_1 \subset \cdots \subset F_n$. (ii) $P(B) = \sum_r P_r(B)$ defines a probability measure on F_n. (iii) The function $\rho_k(W)$, employed in the above proof, is F_k-measurable $(k = 0, \cdots, n)$; hence, $c_j = c_j(W)$ is F_n-measurable. (iv) The above proof becomes exact by defining (at the $(k + 1)$-th step) $E(y \mid \Sigma = \Sigma_0)$ as the conditional expectation of y relative to F_k with $\{F_n, P(B)\}$ as the underlying measure space. (v) Consequently, interpreting the assertion of Theorem 1 in terms of this same measure space, we have a meaningful and true result.

8. The discontinuous case. The above procedure imposes one restriction on the distribution $\mu(A)$ of X; namely, that each ordering (which might be used in the construction) of the given class H is a continuous ordering with respect to $\mu(A)$. In the so-called discontinuous case, the distribution $\mu(A)$ of X is completely unrestricted. However, in this case the above construction might break down with a positive probability in the sense that some boundary set will contain more than one observation. This defect will be repaired as follows (cf. Fraser [8], p. 50).

Let Y be a real random variable, uniformly distributed in the unit interval

$L = [0, 1]$, which is independent of X and let $X' = (X, Y)$, taking values in $\Omega' = \Omega \times L$. To each ordering o in Ω we associate the following ordering o' in Ω':

$$(x_1, y_1) < (x_2, y_2) \qquad \text{if } x_1 < x_2 \text{ or } x_1 \sim x_2 \text{ and } y_1 < y_2 .$$

Let H' consist of all orderings in Ω' which are associated to some ordering in Ω. Then, even in the discontinuous case, each ordering o' in H' is continuous with respect to the distribution $\mu'(B)$ of X'.

Let $x_1, \cdots, x_n$ and $y_1, \cdots, y_n$ be independent observations on X and Y, respectively. Then $x_i' = (x_i, y_i)$ $(i = 1, \cdots, n)$ are n independent observations on X'. Replacing in the above construction Ω, H, and x_i by Ω', H', and x_i', respectively, we obtain a partition of Ω' into the final blocks U_j' $(j = 0, \cdots, n)$ and the set of measure 0 consisting of the n observations x_i'. Clearly, the coverages $c_j = \mu'(U_j')$ satisfy the assertions of Theorem 1·. Thus we are able to set precise tolerance limits on the distribution of $X' = (X, Y)$ which will yield some information on the distribution of X.

As a simple illustration: Let o be any ordering in Ω and let $x'(1) \leqq x'(2) \leqq \cdots \leqq x'(n)$ be the ordered set (with respect to o') of the n observations x_i' on X'. Then $U = \{x' \mid x' < x'(m)\}$ has a coverage $p = \mu'(U)$ with a cumulative d.f. $I_p(m, n + 1 - m)$. But, for $x'(m) = (x(m), y(m))$,

$$\begin{aligned} \mu'(U) &= \mu\{x \mid x < x(m)\} + y(m)\mu\{x \mid x \sim x(m)\} \\ &\geqq \mu\{x \mid x < x(m)\} = P(X < x(m)) = c \qquad \text{(say).} \end{aligned}$$

Hence,

$$P(c \leqq p) \geqq P(\mu'(U) \leqq p) = I_p(m, n + 1 - m),$$

a well-known result due to Scheffé and Tukey ([3], p. 191).

REFERENCES

[1] S. S. Wilks, "Determination of sample sizes for setting tolerance limits," *Ann. Math. Stat.*, Vol. 12 (1941), pp. 91–96.

[2] A. Wald, "An extension of Wilks' method for setting tolerance limits," *Ann. Math. Stat.*, Vol. 14 (1943), pp. 45–55.

[3] H. Scheffé and J. W. Tukey, "Non-parametric estimation: I. Validation of order statistics," *Ann. Math. Stat.*, Vol. 16 (1945), pp. 187–192.

[4] J. W. Tukey, "Nonparametric estimation: II. Statistically equivalent blocks and tolerance regions—the continuous case," *Ann. Math. Stat.*, Vol. 18 (1947), pp. 529–539.

[5] J. W. Tukey, "Nonparametric estimation: III. Statistically equivalent blocks and multivariate tolerance regions—the discontinuous case," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 30–39.

[6] D. A. S. Fraser and R. Wormleighton, Nonparametric estimation: IV.," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 294–298.

[7] D. A. S. Fraser, "Sequentially determined statistically equivalent blocks," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 372–381.

[8] D. A. S. Fraser, "Nonparametric tolerance regions," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 44–55.