A note on nonparametric trend conformity

Friedrich Pukelsheim

Angaben zur Veröffentlichung / Publication details:

Pukelsheim, Friedrich. 1984. "A note on nonparametric trend conformity." *The Annals of Statistics* 12 (2): 775–77. https://doi.org/10.1214/aos/1176346524.

The Annals of Statistics
1984, Vol. 12, No. 2, 775–777

A NOTE ON NONPARAMETRIC TREND CONFORMITY

BY FRIEDRICH PUKELSHEIM

Universität Freiburg im Breisgau

The question is discussed when of two ordered sets of continuous distribution functions one shows less positive trend than the other. The resulting ordering has applications to power function monotonicity of isotonic rank tests, and is related to certain orderings in parametric models.

In a parametric context, Robertson and Wright (1982) recently investigated various notions specifying that, of two sets of parameters, one more closely conforms to a positive trend than the other. Their discussion does not readily generalize to the nonparametric trend problem, as discussed in Lehmann (1959, page 233), and Barlow, Bartholomew, Bremner and Brunk (1972, page 200). It is the purpose of this note to show that conformity to a nonparametric trend may be quantified by an ordering which was employed by Yanagimoto and Okamoto (1969, page 495), and by Yanagimoto and Sibuya (1972, page 432) in the nonparametric problems of testing against positive regression dependence, and positive asymmetry, respectively; see also Rüschendorf (1983).

Let $\mathscr{F}_c^n$ be the family of all n-tuples $\mathbf{F} = (F_1, \cdots, F_n)$ where each component F_i is a continuous distribution function on the real line, and write $F_i \le F_j$ whenever $F_i(x) \le F_j(x)$ for all $x \in \mathbb{R}$. Define the *negative trend* subfamily by

$$\mathscr{F}_\downarrow^n = \{\mathbf{F} \in \mathscr{F}_c^n \mid \forall\ i < j\colon F_i \le F_j\}, \tag{1}$$

and the *positive trend* subfamily $\mathscr{F}_\uparrow^n$ by replacing $F_i \le F_j$ by $F_i \ge F_j$ in (1). Their common intersection is the *trendless* subfamily $\mathscr{F}_=^n = \{(F, \cdots, F) \mid F \in \mathscr{F}_c^1\}$. For $\mathbf{F}, \mathbf{G} \in \mathscr{F}_c^n$ we shall say that $\mathbf{F}$ *shows less positive trend than* $\mathbf{G}$, denoted by $\mathbf{F} \le \mathbf{G}$, if

$$\forall\ i < j\colon\quad F_i \circ F_j^- \le G_i \circ G_j^-, \tag{2}$$

where as usual $F_j^-(u) = \inf\{x \in \mathbb{R} \mid F(x) \ge u\}$. This ordering is reflexive, and transitive. It is antisymmetric in a nonparametric sense: if the function g from $\mathbb{R}$ onto $\mathbb{R}$ is strictly isotonic and $F_i = G_i \circ g$, for all i, then $\mathbf{F} \le \mathbf{G} \le \mathbf{F}$. Conversely suppose that all F_i and G_i are strictly isotonic: if $\mathbf{F} \le \mathbf{G} \le \mathbf{F}$ then $G_i^{-1} \circ F_i = G_j^{-1} \circ F_j = g$, say, and g is strictly isotonic and onto, and satisfies $F_i = G_i \circ g$, for all i.

The ordering $\le$ was introduced by Yanagimoto and Okamoto (1969), and Yanagimoto and Sibuya (1972) in order to quantify larger positive regression dependence, and more positive biasedness. It also characterizes a positive or a negative trend relative to the trendless boundary cases:

$$\mathbf{F} \in \mathscr{F}_\downarrow^n \Leftrightarrow \forall\ \mathbf{G} \in \mathscr{F}_=^n\colon \mathbf{F} \le \mathbf{G}, \tag{3}$$

Received March 1983; revised July 1983.
AMS 1980 *subject classification.* Primary 62G10, secondary 62E10.
Key words and phrases. Isotonic inference, isotonic rank statistics.

(4) $$\mathbf{F} \in \mathscr{F}^n_{\uparrow} \Leftrightarrow \forall\, \mathbf{G} \in \mathscr{F}^n_{=}: \mathbf{G} \le \mathbf{F}.$$

This follows from the facts, due to continuity, that $F_i \circ F_i^-(u) = u$, and that $F_i \circ F_j^- \circ F_j \le F_j$ forces $F_i \le F_j$. Furthermore, (3) and (4) show that $\mathbf{F}$ lies in $\mathscr{F}^n_=$ if and only if $\mathbf{G} \le \mathbf{F} \le \mathbf{G}$, for all $\mathbf{G} \in \mathscr{F}^n_=$, which is in line with the nonparametric antisymmetry mentioned earlier.

The ordering $\le$ of the full family $\mathscr{F}^n_c$ may of course be restricted to subfamilies of interest. For instance, in the problem of whether the first of two samples is stochastically smaller than the second, the antisymmetric ordering of Lehmann (1959, page 187) places (F_I, F_{II}) before (G_I, G_{II}) if $F_I = G_I$ and $F_{II} \ge G_{II}$. This entails that (F_I, F_{II}) shows less positive trend than (G_I, G_{II}). But ordering (2) is larger in that it orders more members of $\mathscr{F}^2_c$, and hence preferable.

Next consider a parametric situation $F_i = H_{\alpha_i}$ and $G_i = H_{\beta_i}$ where $\{H_\theta \mid \theta \in \Theta \subset \mathbb{R}\}$ is a given one-parameter family. Then $\mathbf{F} \le \mathbf{G}$ induces between the vector parameters $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)$ and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_n)$ an ordering $\boldsymbol{\alpha} \le \boldsymbol{\beta}$. The location family $H_\theta(x) = H(x - \theta)$ leads to

(5) $$\boldsymbol{\alpha} \le \boldsymbol{\beta} \Leftrightarrow \alpha_1 - \beta_1 \ge \cdots \ge \alpha_n - \beta_n,$$

which is the $\gtrsim$ ordering discussed by Robertson and Wright (1982, page 1235). The exceptional compatibility of this ordering with the Euclidean structure of the parameter space cannot be expected to hold in general. The scale family $H_\theta(x) = H(x/\theta)$, $H(0) = 0$, $\theta > 0$, as well as the Lehmann family $H_\theta(x) = \{H(x)\}^\theta$, $\theta > 0$, both result in $\boldsymbol{\alpha} \le \boldsymbol{\beta} \Leftrightarrow \alpha_1/\beta_1 \ge \cdots \ge \alpha_n/\beta_n$.

The counterpart of the ordering $\le$ on the parameter space $\mathscr{F}^n_c$ is any ordering on the sample space $\mathbb{R}^n$ such that any isotonic real statistic $T(\mathbf{x})$ is stochastically smaller under $\mathbf{F}$ than under $\mathbf{G}$ whenever $\mathbf{F} \le \mathbf{G}$. Due to the nonparametric antisymmetry of $\le$ only statistics of the form $T(\mathbf{x}) = f \circ \mathbf{R}(\mathbf{x})$ need be considered where $\mathbf{R}(\mathbf{x}) = (R_1(\mathbf{x}), \cdots, R_n(\mathbf{x}))$ is the rank vector of $\mathbf{x} \in \mathbb{R}^n$, and f is a real function on the permutation group $\mathscr{S}_n$. Thus from the very beginning the sample space ordering may be confined to comparing two permutations $\mathbf{r} = (r_1, \cdots, r_n)$ and $\mathbf{s} = (s_1, \cdots, s_n)$. We shall call $\mathbf{r}$ *less significant for a positive trend than* $\mathbf{s}$, denoted by $\mathbf{r} \le \mathbf{s}$, if

(6) $$\forall\, i < j: \quad r_i < r_j \Rightarrow s_i < s_j.$$

This is the ordering of Lehmann (1966, equation (7.2)). That our terminology is justified results from the following.

THEOREM. *If $T(\mathbf{x})$ is an isotonic real rank statistic, i.e., $T(\mathbf{x}) = f \circ \mathbf{R}(\mathbf{x})$ with $f(\mathbf{r}) \le f(\mathbf{s})$ whenever $\mathbf{r} \le \mathbf{s}$, then $\mathbf{F} \le \mathbf{G}$ implies that T is stochastically smaller under $\mathbf{F}$ than under $\mathbf{G}$.*

PROOF. For $\mathbf{u} = (u_1, \cdots, u_n) \in (0, 1)^n$, define $\mathbf{F}^-(\mathbf{u}) = (F_1^-(u_1), \cdots, F_n^-(u_n))$. The claim is that

$$\lambda^n\{\mathbf{u} \in (0, 1)^n \mid T \circ \mathbf{F}^-(\mathbf{u}) > t\} \le \lambda^n\{\mathbf{u} \in (0, 1)^n \mid T \circ \mathbf{G}^-(\mathbf{u}) > t\},$$

for all $t \in \mathbb{R}$, where λ^n denotes n-dimensional Lebesgue measure. Fix $\mathbf{u}$, and put

$\mathbf{x} = \mathbf{F}^{-}(\mathbf{u})$ and $\mathbf{y} = \mathbf{G}^{-}(\mathbf{u})$. If $i < j$ and $x_i \le x_j$, then $\mathbf{F} \le \mathbf{G}$ entails $y_i = G_i^{-} \circ F_i(x_i) \le G_i^{-} \circ F_i \circ F_j^{-}(u_j) \le G_i^{-} \circ G_i(y_j) \le y_j$. Therefore $\mathbf{R}(\mathbf{x}) \le \mathbf{R}(\mathbf{y})$, and monotonicity of f gives $T(\mathbf{x}) \le T(\mathbf{y})$. The theorem is proved. □

As an application we note that if $T(\mathbf{x})$ is an isotonic real rank statistic then the one-sided rank test with large values of T significant has an isotonic power function with respect to the ordering (2). Therefore (3) and (4) entail that in the nonparametric problem of testing against a positive trend, i.e., $\mathbf{F} \in \mathscr{F}^n_{\downarrow}$ vs. $\mathbf{F} \in \mathscr{F}^n_{\uparrow}$, $\mathbf{F} \notin \mathscr{F}^n_{=}$, any such rank test has size α on $\mathscr{F}^n_{\downarrow}$ and is unbiased on $\mathscr{F}^n_{\uparrow}$ provided it is α-similar on $\mathscr{F}^n_{=}$. Also notice that this trend problem analysis makes no use of wedge sets and thus drastically shortens the arguments for the case of regression dependence, as laid out by Yanagimoto and Okamoto (1969, page 498).

The permutation ordering (6) compares favorably with the (positive trend versions of the) orderings of Robertson and Wright (1982): $\gtrsim$ was in (5) identified as a parameter rather than a sample space ordering. It induces the dual cone ordering $\gg^*$ which, when restricted to permutations, coincides with $\gg$. A smaller ordering, in that it orders fewer members of $\mathscr{S}_n$, is that of Savage (1957, page 968). But ordering (6) is yet smaller and hence, admitting more isotonic statistics, preferable. Notice, however, that $\gg^*$ and (6) do not distinguish between isotonic functions which are also linear: $\sum c_i r_i$ is isotonic on $\mathscr{S}_n$ if and only if $c_1 \le \cdots \le c_n$, and the same holds for the $\gg^*$ ordering, see Robertson and Wright (1982, page 1240).

REFERENCES

BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference under Order Restrictions.* Wiley, New York.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.

ROBERTSON, T. and WRIGHT, F. T. (1982). On measuring the conformity of a parameter set to a trend, with applications. *Ann. Statist.* **10** 1234–1245.

RÜSCHENDORF, L. (1983). Unbiasedness of tests via a.s. constructions. Preprint.

SAVAGE, I. R. (1957). Contributions to the theory of rank order statistics—the "trend" case. *Ann. Math. Statist.* **28** 968–977.

YANAGIMOTO, T. and OKAMOTO, M. (1969). Partial orderings of permutations and monotonicity of a rank correlation statistic. *Ann. Inst. Statist. Math.* **21** 489–506.

YANAGIMOTO, T. and SIBUYA, M. (1972). Test of symmetry of a one-dimensional distribution against positive biasedness. *Ann. Inst. Statist. Math.* **24** 423–434.

LEHRSTUHL FÜR STOCHASTIK
UND IHRE ANWENDUNGEN
UNIVERSITÄT AUGSBURG
MEMMINGER STRASSE 6
D-8900 AUGSBURG
FEDERAL REPUBLIC OF GERMANY