The Canadian Journal of Statistics
Vol. 13, No. 3, 1985, Pages 211–216
La Revue Canadienne de Statistique

Tests based on equivariant estimators*

Louis-Paul RIVEST

Université Laval

Key words and phrases: Dispersive ordering, mixtures, power function, star-shaped ordering, spread.
AMS 1980 *subject classifications*: Primary 62F04; secondary 62J05.

ABSTRACT

Let $\hat{T}^{(1)}$ and $\hat{T}^{(2)}$ be location-equivariant estimators of an unknown location parameter μ. It is shown that the test for $H_0: \mu \leq \mu_0$ versus $H_A: \mu > \mu_0$ that rejects H_0 if $\hat{T}^{(1)}$ is large is uniformly more powerful than the one that rejects H_0 if $\hat{T}^{(2)}$ is large if and only if $\hat{T}^{(2)}$ is "more dispersed" than $\hat{T}^{(1)}$. A similar result is obtained for tests on scale using the star-shaped ordering. Examples are given.

RÉSUMÉ

Soient $\hat{T}^{(1)}$ et $\hat{T}^{(2)}$ des estimateurs équivariants d'un paramètre de localisation inconnu μ. On montre que le test pour $H_0: \mu \leq \mu_0$ versus $H_A: \mu > \mu_0$ qui rejette H_0 si $\hat{T}^{(1)}$ est grand est uniformément plus puissant que celui qui rejette H_0 si $\hat{T}^{(2)}$ est grand si et seulement si $\hat{T}^{(2)}$ est "plus dispersé" que $\hat{T}^{(1)}$. En utilisant l'ordre étoilé, on obtient un résultat semblable pour des tests portant sur un paramètre d'échelle. On présente ensuite des applications de ces résultats.

1. INTRODUCTION

The star-shaped ordering is a partial ordering on the set of distributions. It has found fruitful applications in reliability theory (Barlow and Proschan 1975) and mathematical statistics (see Rivest 1982 for references). This note shows that it also pertains to the theory of tests.

Section 2 recalls some properties of the star-shaped and the dispersive orderings. In Section 3 it is shown that if $\hat{T}^{(1)}$ and $\hat{T}^{(2)}$ are equivariant estimators of the location parameter μ, then the test for $H_0: \mu \leq \mu_0$ versus $H_A: \mu > \mu_0$ based on $\hat{T}^{(1)}$ is uniformly more powerful than the one based on $\hat{T}^{(2)}$ if and only if $\hat{T}^{(2)}$ is more dispersed than $\hat{T}^{(1)}$. For tests on the scale parameter a similar result holds with the star-shaped instead of the dispersive ordering. Examples are presented in Section 4. Tests on the slope of a regression model are also considered.

2. THE STAR-SHAPED AND THE DISPERSIVE ORDERING

Let A be a subset of R; define $\mathcal{F}_A$ as the set of absolutely continuous distribution with densities that have as their support an interval in A. Let F and G be in $\mathcal{F}_{(0,\infty)}$; F is said to be *star-shaped* with respect to G if $G^{-1}(F(x))/x$ increases in $(0, \infty)$. This is written $G \overset{*}{>} F$, or $Y \overset{*}{>} X$ if Y and X are distributed as G and F respectively.

*Research supported by NSERC grant A5244 and FCAC grant EQ1023.

The spread or dispersive ordering has been studied independently by many authors. Let F and G be in $\mathcal{F}_R$; G is said to be *more dispersed* than $F(G \overset{d}{>} F)$ if

$$G^{-1}(t) - F^{-1}(t)$$

is increasing in $(0, 1)$. Shaked (982) pointed out that it is equivalent to the star-shaped ordering: he showed that $G \overset{d}{>} F$ if and only if $G(\ln \cdot) \underset{*}{\geq} F(\ln \cdot)$.

PROPOSITION 1. *Let F and G be in $\mathcal{F}_R$. The following statements are equivalent:*

(i) $G \overset{d}{>} F$,
(ii) $G^{-1}(F(x)) - x$ *increases in* R,
(iii) $f(F^{-1}(t)) \geq g(G^{-1}(t))$ *for* $t \in (0, 1)$,

where F^{-1}, G^{-1}, f, and g are respectively the inverses and the densities of F and G.

Proof. It is clear that (i) $\Leftrightarrow$ (ii). The fact that

$$\frac{d}{dt} G^{-1}(t) - F^{-1}(t) = \frac{1}{g(G^{-1}(t))} - \frac{1}{f(F^{-1}(t))}$$

implies that (ii) and (iii) are equivalent. Q.E.D.

The following properties are used in Section 4.

PROPOSITION 2 (Barlow and Proschan 1975, p. 108). *Let F and G be distributions, and define F_{jn} and G_{jn} as the distributions of the jth order statistic for a sample of size n from F and G respectively; then*

(i) $F, G \in \mathcal{F}_{(0,\infty)}$ *and* $G \underset{*}{\geq} F \Rightarrow G_{jn} \underset{*}{\geq} F_{jn}$,
(ii) $F, G \in \mathcal{F}_R$ *and* $G \overset{d}{>} F \Rightarrow G_{jn} \overset{d}{>} F_{jn}$.

PROPOSITION 3 (Lewis and Thompson 1981; Rivest 1982). *Let X be a random variable whose distribution is in F_R and has a strongly unimodal density. If Y_1 and Y_2 are random variables independent of X, then*

$$Y_1 \overset{d}{>} Y_2 \Rightarrow X + Y_1 \overset{d}{>} X + Y_2.$$

If X, Y_1, and Y_2 are positive and if $\ln X$ has a strongly unimodal density, then

$$Y_1 \underset{*}{\geq} Y_2 \quad \Rightarrow \quad XY_1 \underset{*}{\geq} XY_2.$$

3. THE RESULTS

Let $X_1, X_2, \ldots, X_n$ be a sample from $F(x - \mu)$, where μ is an unknown location, and define $\hat{T}^{(j)} = \hat{T}^{(j)}(X_1, X_2, \ldots, X_n)$ for $j = 1, 2$ as two equivariant location estimators, i.e.

$$\hat{T}^{(j)}(X_1 + a, \ldots, X_n + a) = \hat{T}^{(j)}(X_1,, \ldots, X_n) + a.$$

If $\beta_j(\alpha, \mu)$ denotes the power function of the α level test for

$$H_0 : \mu \leq \mu_0 \quad \text{versus} \quad H_A : \mu > \mu_0 \tag{1}$$

that rejects if $\hat{T}^{(j)}$ is large, then

PROPOSITION 4. *If the distributions of $\hat{T}^{(1)}$ and $\hat{T}^{(2)}$ belong to $\mathcal{F}_R$, then*

$$\hat{T}^{(2)} \overset{d}{>} \hat{T}^{(1)} \quad \Leftrightarrow \quad \beta_1(\alpha, \mu) \geq \beta_2(\alpha, \mu) \quad \textit{for all } \alpha \in (0, 1), \textit{ for all } \mu > \mu_0. \tag{2}$$

Proof. Let G_j be the distribution of $\hat{T}^{(j)}$ when $\mu = \mu_0$, $j = 1, 2$. The right-hand side of (2) holds if and only if, for any α in $(0, 1)$ and $\mu > 0$,

$$1 - G_1(G_1^{-1}(1 - \alpha) - \mu) \geq 1 - G_2(G_2^{-1}(1 - \alpha) - \mu),$$

which if $1 - \alpha = G_1(x)$ is equivalent to

$$G_2^{-1}(G_1(x - \mu)) \leq G_2^{-1}(G_1(x)) - \mu \qquad \text{for all } \mu > 0, \quad x \in R.$$

By Proposition 1(ii), this holds if and only if the left-hand side of (2) is true. Q.E.D.

For scale parameters let $\beta_j(\alpha, \sigma)$ be the power function of the test for $H_0 : \sigma \leq \sigma_0$ versus $H_A : \sigma > \sigma_0$ that rejects if $\hat{S}^{(j)} = \hat{S}^{(j)}(X_1, \ldots, X_n)$ is large, where $\hat{S}^{(j)}$ is calculated using a sample from $F(x/\sigma)$ or $F((x - \mu)/\sigma)$ and satisfies $\hat{S}^{(j)}(aX_1, \ldots, aX_n) = a\hat{S}^{(j)}(X_1, \ldots, X_n)$ for $j = 1, 2$. Along the lines of Proposition 4 one proves:

PROPOSITION 5. *If the distributions of $\hat{S}^{(1)}$ and $\hat{S}^{(2)}$ belong to $\mathcal{F}_{(0,\infty)}$, then*

$$\hat{S}^{(2)} \underset{*}{\geq} \hat{S}^{(1)} \quad \Leftrightarrow \quad \beta_1(\alpha, \sigma) \geq \beta_2(\alpha, \sigma) \quad \textit{for all } \alpha \in (0, 1), \quad \textit{for all } \sigma > \sigma_0.$$

REMARK 1. Proposition 4 has been stated in what one might call the "classical setting": given a sample from a fixed distribution, different estimators are compared. In a "robust setting" (given an estimator, its behaviours for various underlying distributions are compared), Proposition 4 holds with G_j and β_j defined as the distribution of $\hat{T}$ and the power function of the test based on $\hat{T}$ when F_j is the underlying distribution for $j = 1, 2$. Proposition 5 can, in a similar way, be stated in a robust setting.

REMARK 2. Let $H_{n\mu}(x)$ be the null distribution of

$$\prod \frac{f(X_i - \mu)}{f(X_i - \mu_0)},$$

the Neymann-Pearson statistic for (1) with μ fixed. The power of the corresponding α level Neymann-Pearson test is easily seen to be $\int_{1-\alpha}^{1} H_{n\mu}^{-1}(t)\, dt$. Applying the Neymann-Pearson lemma yields a universal upper bound for $\beta(\alpha, \mu)$:

$$\int_{1-\alpha}^{1} H_{n\mu}^{-1}(t)\, dt \geq \beta(\alpha, \mu) \qquad \text{for all} \quad \alpha > 0.$$

Klaassen (1984) derived a similar bound for the local power.

REMARK 3. Proposition 4 and 5 are easily generalized to two-sample tests: if, for instance, $\beta_j(\alpha, d)$ denotes the power function of the α level test for $H_0 : \mu_1 - \mu_2 \leq d_0$ versus $H_A : \mu_1 - \mu_2 > d_0$ that rejects if $\hat{T}_1^{(j)} - \hat{T}_2^{(j)}$ is large, where $\hat{T}_i^{(j)}$ is an equivariant estimator of μ_i, the location of sample i, for $j = 1, 2$, then

$$\hat{T}_1^{(2)} - \hat{T}_2^{(2)} \overset{d}{>} \hat{T}_1^{(1)} - \hat{T}_2^{(1)}$$

$$\Leftrightarrow \quad \beta_1(\alpha, d) \geq \beta_2(\alpha, d) \quad \text{for all} \quad \alpha \in (0, 1), \quad \text{for all } d > d_0.$$

REMARK 4. For monotone two-sample rank tests Doksum (1969) obtained results similar to Propositions 4 and 5. If $\beta_j(\alpha, \mu)$ denotes the power function of an α level rank test for $H_0 : F_2(x) = F_1^{(j)}(x)$ versus $H_A : F_2(x) = F_1^{(j)}(x - \mu)$, $\mu > 0$ (under H_0, $F_1^{(j)}$ is the common distribution of the two samples while under H_A, $F_1^{(j)}$ and $F_1^{(j)}(x - \mu)$ are the distributions of samples 1 and 2 respectively) for $j = 1, 2$, then Doksum's Theorem 4.1 is

$$F_1^{(2)} \overset{d}{>} F_1^{(1)} \quad \Rightarrow \quad \beta_1(\alpha, \mu) \geq \beta_2(\alpha, \mu) \quad \text{for all } \alpha \in (0, 1), \quad \text{for all } \mu > 0.$$

For the alternatives $H_A: F_2(x) = F_1^{(j)}(x/\sigma)$, $\sigma > 1$, Doksum's Theorem 3.1 is

$$F_1^{(2)} \gtrsim_* F_1^{(1)} \quad \Rightarrow \quad \beta_1(\alpha, \sigma) \geq \beta_2(\alpha, \sigma) \quad \text{for all } \alpha \in (0, 1), \quad \text{for all } \sigma > 1.$$

4. EXAMPLES

First a new proof of the optimality of the location test based on $\bar{X}$ for a normal sample is given. The power of the classical F test on variances for samples from a contaminated normal is investigated. Tests based on an order statistic are considered. Finally an extension of Proposition 4 to regression estimators is presented.

EXAMPLE 1. *Let $X_1, \ldots, X_n$ be a sample from an $\mathsf{N}(\mu, 1)$. The test for* (1) *based on $\bar{X}$ is uniformly more powerful unbiased among all the tests based on equivariant estimators of μ.*

Proof. Let $\hat{T} = \hat{T}(X_1, \ldots, X_n)$ be an equivariant estimator of μ. Note that

$$\hat{T} = \bar{X} + \hat{T}(X_1 - \bar{X}, \ldots, X_n - \bar{X})$$

and $\{X_1 - \bar{X}, \ldots, X_n - \bar{X}\}$ is independent of $\bar{X}$; thus $\hat{T}$ is a location mixture of $\bar{X}$. Since $\bar{X}$ is distributed as $\mathsf{N}(\mu, 1/n)$, it has a strongly unimodal density; Proposition 4 implies, since $\hat{T}(X_1 - \bar{X}, \ldots, X_n - \bar{X}) \overset{d}{>} 0$, that $\hat{T} \overset{d}{>} \bar{X}$. Q.E.D.

EXAMPLE 2. *Let X_{ij}, $i = 1, 2$, $j = 1, \ldots, n_i$, be two independent samples from $F((x - \mu_i)/\sigma_i)$, where F is a normal scale mixture ($F(x) = \int_0^\infty \Phi(x/y)\, dH(y)$ for some distribution H on $(0, \infty)$, where $\Phi(x)$ is the standardized normal distribution). If $\beta(\alpha, r)$ is the power function of the α level test for $H_0: \sigma_1^2 \leq \sigma_2^2$ versus $H_A: \sigma_1^2 > \sigma_2^2$ that rejects if*

$$\hat{R} = \frac{\Sigma(X_{ij} - \bar{X}_{1\cdot})^2(n_2 - 1)}{\Sigma(X_{2j} - \bar{X}_{2\cdot})^2(n_1 - 1)} \tag{3}$$

is large, then $\beta_F(\alpha, r) \leq \beta_\Phi(\alpha, r)$ for all $\alpha \in (0, 1)$, for all $r > 1$.

Proof. By Proposition 3 of Rivest (1986), under F the distribution of (3) is an F_{n_1-1, n_2-1} scale mixture. Since the logarithm of an F_{n_1-1, n_2-1} random variable has a strongly unimodal density, Proposition 3 implies

$$P_F(\hat{R} \leq \cdot) \gtrsim_* F_{n_1-1}, n_2 - 1.$$

Q.E.D.

A similar reasoning shows that in a normal universe the power function of the F test is an increasing function of n_1 and n_2. Tan (1982) gave a survey of the literature studying the robustness of the standard normal procedures using analytical expansions.

EXAMPLE 3. *Let $X_{(i)}$ be the ith order statistic of a sample of size n from $F(x/\sigma)$, where $F \in \mathcal{F}_{(0,\infty)}$. Let $\beta_F(\alpha, \sigma)$ be the power function of the α level test for $H_0: \sigma \leq \sigma_0$ versus $H_A: \sigma > \sigma_0$ that rejects if $X_{(i)}$ is large. Then*

$$F_2 \gtrsim_* F_1 \quad \Rightarrow \quad \beta_{F_1}(\alpha, \sigma) \geq \beta_{F_2}(\alpha, \sigma) \quad \textit{for all } \alpha \in (0, 1), \quad \textit{for all } \sigma > \sigma_0.$$

The proof is a straightforward consequence of Proposition 2. Note that if F_1 belongs to the Gamma family and F_2 is a scale mixture of F_1, the result holds with $X_{(i)}$ replaced by $\bar{X}$. The argument follows that of Proposition 3 of Rivest (1986).

EXAMPLE 4. *Let*

$$Y_i = b_0 + b_1 x_i + \varepsilon_i,$$

where ε_i *is a sample from* F, *and* $\beta_F(\alpha, b_1)$ *be the power function of the* α *level test for* $H_0: b_1 \le b_{10}$ *versus* $H_A: b_1 \ge b_{10}$ *that rejects if*

$$\hat{b}_1 = \frac{\Sigma(x_i - \bar{x})Y_i}{\Sigma(x_i - \bar{x})^2}$$

is large. Then if F_1 *and* F_2 *are strongly unimodal distributions,*

$$F_2 \overset{d}{>} F_1 \quad \Rightarrow \quad \beta_{F_1}(\alpha, b_1) \ge \beta_{F_2}(\alpha, b_1) \quad \textit{for all } \alpha \in (0,1), \quad \textit{for all } b_1 > b_{10}.$$

Proof. It is easily seen that $\hat{b}_1$ is equivariant: if $G(x)$ is the distribution of $\hat{b}_1$ when b_{10} is the true slope, then $G(x - b_1 + b_{10})$ is its distribution when b_1 is. Thus Proposition 5 holds for equivariant regression estimators; in order to prove the result it suffices to show

$$F_2 \overset{d}{>} F_1 \quad \Rightarrow \quad P_{F_2}(\hat{b}_1 \le \cdot) \overset{d}{>} P_{F_1}(\hat{b}_1 \le \cdot).$$

Let $\varepsilon_i^{(j)}$, $i = 1, \ldots, n$, $j = 1, 2$, denote independent samples from F_1 and F_2 respectively. Since $F_2 \overset{d}{>} F_1$, successive applications of Proposition 3 yield

$$\begin{aligned}
\sum_1^n (x_i - \bar{x})\varepsilon_i^{(2)} &\overset{d}{>} \sum_1^{n-1} (x_i - \bar{x})\varepsilon_i^{(2)} + (x_n - \bar{x})\varepsilon_n^{(1)} \\
&\overset{d}{>} \sum_1^{j} (x_i - \bar{x})\varepsilon_i^{(2)} + \sum_{j+1}^{n} (x_i - \bar{x})\varepsilon_i^{(1)} \\
&\overset{d}{>} \sum_1^{n} (x_i - \bar{x})\varepsilon_i^{(1)}.
\end{aligned}$$

This implies

$$P_{F_2}(\hat{b}_1 \le \cdot) \overset{d}{>} P_{F_1}(\hat{b}_1 \le \cdot).$$

Q.E.D.

It can be shown that the condition $F_2 \overset{d}{>} F_1$ holds for the following pairs of distributions:

$$F_1(x) = \begin{cases} e^x/2, & x < 0 \\ 1 - e^{-x}/2, & x \ge 0 \end{cases} \quad \text{and} \quad F_2(x) = \frac{1}{1 + e^{-x}}, \quad x \in R;$$

$$F_1(x) = \Phi(x) \quad \text{and} \quad F_2(x) = \frac{1}{1 + e^{-x}}.$$

Another sufficient condition for the conclusion of Example 4 to hold is "F_1 is strongly unimodal and F_2 is a location mixture of F_1". Proposition 3 can be applied, since $P_{F_2}(\hat{b}_1 \le \cdot)$ is then a location mixture of $P_{F_1}(\hat{b}_1 \le \cdot)$, and $P_{F_1}(b_1 \le x)$ has a strongly unimodal density, being a convolution of strongly unimodal densities.

REFERENCES

Barlow, R.E., and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing. Probability Models.* Holt, Rinehart and Winston, New York.

Doksum, K.A. (1969). Starshaped transformations and the power of rank tests. *Ann. Math. Statist.*, 40, 1167–1176.

Klaassen, C.A. (1984). Location estimators and spread. *Ann. Statist.*, 12, 311–321.

Lewis, T., and Thompson, J.W. (1981). Dispersive distributions and the connection between dispersivity and strong unimodality. *J. Appl. Probab.*, 18, 76–90.

Rivest, L.P. (1986). Bartlett's, Cochran's and Hartley's tests on variances are liberal when the underlying distribution is long-tailed. To appear in the J. Amer. Statist. Assoc.

Rivest, L.P. (1982). Products of random variables and star-shaped ordering. *Canad. J. Statist.*, 10, 219–223.

Shaked, M. (1982). Dispersive ordering of distributions. *J. Appl. Probab.*, 19, 310–320.

Tan, W.Y. (1982). Sampling distributions and robustness of t, F and variance-ratio in two samples and ANOVA models with respect to departure from normality. *Comm. Statist. A—Theor. Methods*, 11, 2485–2511.

Received 5 August 1984
Revised 20 June 1985
Accepted 28 June 1985

Département de mathématiques statistique et actuariat
Université Laval
Cité Universitaire
Québec
Québec G1K 7P4