The Annals of Statistics
1991, Vol. 19, No. 2, 889–898

ANOMALIES OF THE LIKELIHOOD RATIO TEST FOR TESTING RESTRICTED HYPOTHESES[1]

BY J. A. MENÉNDEZ AND B. SALVADOR

Universidad de Valladolid

The first anomaly in the L.R.T. for testing restricted hypotheses was observed by Warrack and Robertson. They found the L.R.T. for testing an order restriction in a normal model to be dominated by a different test.

In this paper we deal with a more general situation in which the L.R.T. for testing a face of an acute cone is dominated by a different test that does not take into account some of the information in the model.

1. Introduction. In a normal model, the likelihood ratio test (L.R.T.) provides a frequently used method for testing means when the hypotheses define order restrictions on the parameters. The L.R.T. performs well in some testing problems, as can be seen in the book of Barlow, Bartholomew, Bremner and Brunk (1972). The first anomaly of the L.R.T. was observed by Warrack and Robertson (1984). They showed a problem with some order restrictions about means in a normal model where the L.R.T. is dominated by another test and they asked for the cause of such an anomaly. We examine this issue in a general context, where answers can be given.

Consider a k-dimensional random normal vector $N_k(\theta, \Gamma)$, with unknown mean vector $\theta = (\theta_1, \dots, \theta_k)'$ and a known covariance. We deal with the L.R.T. for testing the hypotheses:

$$(1.1)\qquad \begin{aligned} &H_0\colon a_j'\theta = 0, \quad j = 1, \dots, r; \quad a_j'\theta \geq 0, \quad j = r+1, \dots, n.\\ &H_A\colon a_j'\theta \geq 0, \quad j = 1, \dots, n. \end{aligned}$$

given by k-dimensional fixed vectors $a_1, \dots, a_n$ in such a way that H_A defines a polyhedric closed convex cone in R^k.

The statistic $T(x) = -2 \ln l(x)$ that defines the L.R.T. $\{T > t\}$ for testing H_0 against $H_A - H_0$ is given by

$$(1.2)\qquad T(x) = \|x - x^0\|^2 - \|x - x^A\|^2,$$

where $\|x\|^2 = x'\Gamma^{-1}x$ and x^0, x^A are the projections of x on H_0, H_A, respectively.

Let us consider the hypotheses:

$$(1.3)\qquad \begin{aligned} &H_0^*\colon a_j'\theta = 0, \quad j = 1, \dots, r,\\ &H_A^*\colon a_j'\theta \geq 0, \quad j = 1, \dots, r. \end{aligned}$$

Received October 1987; revised February 1990.
[1]Research partially supported by the DGICYT under grant PB-87-0905-C02-01.
AMS 1980 *subject classifications*. Primary 62F03; secondary 62H15.
Key words and phrases. Restricted inference, likelihood ratio test, acute cone.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics*. www.jstor.org

The L.R.T. $\{T^* > t\}$ for testing H_0^* against $H_A^* - H_0^*$ is given by the statistic

$$T^*(x) = \|x - x^{*0}\|^2 - \|x - x^{*A}\|^2 = \|x^{*A} - x^{*0}\|^2, \tag{1.4}$$

the last equality is true since H_0^* is the least dimension face in the cone H_A^* and x^{*0} and x^{*A} are the projections of x on H_0^* and H_A^*, respectively.

We are concerned with the dominance of the L.R.T. $\{T > t\}$ when H_A is an acute cone. The concept of an acute cone was first introduced by Martín and Salvador (1988) who studied the relation between acute cones and the usefulness of the pool adjacent violator algorithm (PAV).

We now give some notation, definitions and a result, which will be useful later. Let C be a cone, $C = \{x \in R^k\colon a_j'x \geq 0,\ j = 1, \dots, n\}$. Any face of C may be denoted by

$$K_B = \{x\colon a_j'x = 0,\ j \in B;\ a_j'x \geq 0,\ j \in B^c\} \tag{1.5}$$

for some subset B of $\{1, \dots, n\}$. The subspace associated with K_B is $L_B = \{x\colon a_j'x = 0,\ j \in B\}$. Denote by $p(x|C)$ the Γ^{-1}-orthogonal projection onto the cone C, so that

$$\|x - p(x|C)\|^2 = \inf_{y \in C} \|x - y\|^2.$$

We will consider the following two definitions which are equivalent to Definition 2.2 in Martín and Salvador (1988).

Definition 1.1. The cone $C_{ij} = \{x\colon a_i'x \geq 0,\ a_j'x \geq 0\}$ is said to be acute (strictly acute) if $x'\Gamma^{-1}y \geq 0$ (> 0) whenever $x \in L_i \cap C_{ij}$ and $y \in L_j \cap C_{ij}$ with x in the Γ^{-1}-orthogonal subspace to L_{ij}.

Definition 1.2. The cone C is said to be acute (strictly acute) if C_{ij} is acute (strictly acute) for any couple i, j in $\{1, \dots, n\}$.

Note that the acuteness of a cone is preserved by linear transformations of the entire statistical problem and therefore we could use the identity matrix for the covariance of the normal model and the unit metric on R^k, after performing a linear transformation.

Proposition 1.1. *C is an acute cone if and only if $a_j'p(x|C) = 0$ for any x such that $a_j'x \leq 0$.*

Proof. See Martín and Salvador (1988). □

In an obvious reference to the PAV algorithm, Martín and Salvador (1988) say that the cone C is PAV when the sufficient condition in the proposition holds.

We now explain briefly the PAV process for obtaining $p(x|C)$ when C is an acute cone [cf. Martín and Salvador (1988), Theorem 2.2].

At the first step, we project x onto the subspace S_1 defined by the constraints in C which are violated by x. If $p(x|S_1) \in C$, then $p(x|S_1) = p(x|C)$. If $p(x|S_1) \notin C$, at the second step we project x or $p(x|S_1)$ onto $S_1 \cap S_2$, where S_2 is the subspace associated with the constraints in C which are violated by $p(x|S_1)$. If $p(x|S_1 \cap S_2) \notin C$, we begin a new step and so on. In a finite number of steps we reach $p(x|C)$. At any rate the solution $p(x|C)$ is in a subspace of S_1, as shown by Proposition 1.1.

In Section 2, we find the test $\{T > t\}$ to be dominated by $\{T^* > t\}$ when testing H_0 against $H_A - H_0$, so that T and T^* being equally sized, T^* becomes more powerful than T. In order to prove that, we give three lemmas, also useful in their own right. The proofs of the lemmas are given in Section 3.

2. Dominance of the L.R.T. The next lemma is a very useful property of an acute cone.

LEMMA 2.1. *Let C be an acute cone and let B be a subset of $\{1, \dots, n\}$ such that K_B is not empty. Then, for any x in C, $p(x|K_B) = p(x|L_B)$. Moreover, $p(x|K_B) \neq 0$, whenever $x \neq 0$ and C is strictly acute.*

The next two results generalize to arbitrary acute cones Lemmas 2.1 and 2.2 in Warrack and Robertson (1984).

LEMMA 2.2. *If C is an acute cone, then for any $x \in R^k$ and $\delta \in C$: $a'_j p(x + \delta|C) \geq a'_j p(x|C)$, $j = 1, \dots, n$.*

LEMMA 2.3. *Let x and y be two elements in a cone C, such that $a'_j x \leq a'_j y$, $j = 1, \dots, n$. If C is acute, then $\|x - p(x|K_{1,\dots,n})\|^2 \leq \|y - p(y|K_{1,\dots,n})\|^2$.*

As noted earlier we give proofs of Lemmas 2.1–2.3 in Section 3.

Before presenting the main result (Theorem 2.2), we need to prove the following theorem about the statistic T defined in (1.2).

THEOREM 2.1. *Let θ be an element of H_0. Then, for any $x \in R^k$, $T(x + \theta) \geq T(x)$.*

PROOF. Through this proof we shall denote by U_m, $m = 1, \dots, n$, the subspace $L_{1 \cdots m}$ and we write x^0 and x^A instead of $p(x|H_0)$ and $p(x|H_A)$, respectively.

Since $x^0 = p(x|U_m)$ for some $m \geq r$ and x^0 is reached projecting x onto a subspace S defined by all such constraints which are satisfied with equality by x^0, the three cases considered below cover all possible situations.

Note that H_0 is an acute cone in U_r.

The restrictions that define U_r are always verified with equality by x^0, so that S can be defined by L_B, with $\{1, 2, \dots, r\} \subset B$. Without loss of generality we can take $B = \{1, \dots, m\}$, $m \geq r$, for each x under consideration.

CASE 1. Let $x \in R^k$ be such that $x^0 = p(x|U_r)$. If $\theta \in H_0$, then $(x + \theta)^0 = x^0 + \theta$, so that

$$\|x - x^0\|^2 = \left\|(x + \theta) - (x + \theta)^0\right\|^2.$$

By Theorem 2.1 in Robertson and Wegman (1978), $\|x - x^A\|^2 \geq \|(x + \theta) - (x + \theta)^A\|^2$, and therefore $T(x + \theta) \geq T(x)$.

CASE 2. Let $x \in R^k$ and $\theta \in H_0$ be such that $x^0 = p(x|U_m)$ and $(x + \theta)^0 = p(x + \theta|U_m)$ for some $m > r$. Consider T', the L.R.T. for testing H_0': $a_j'\theta = 0$, $j = 1, \dots, m$; $a_j'\theta \geq 0$, $j = m + 1, \dots, n$ against $H_A - H_0'$.

If $\theta \in U_m$, then by Case 1, $T'(x + \theta) \geq T'(x)$ and it is easy to prove that $T'(x) = T(x)$ and $T'(x + \theta) = T(x + \theta)$.

If $\theta \notin U_m$, consider $\theta^m = p(\theta|U_m)$, then as before, $T(x + \theta^m) \geq T(x)$.

Let us consider the cone $C^{(m)} = \{x\colon a_j'x \geq 0,\ j = 1, \dots, m\}$ and $y^{(m)} = p(x + \theta|C^{(m)})$ and $z^{(m)} = p(x + \theta^m|C^{(m)})$.

It is obvious that $y^0 = p(x + \theta|U_m) = p(x + \theta^m|U_m) = z^0$. Note that $y^0 = y^{(m)0}$ and $z^0 = z^{(m)0}$.

Decomposing $x + \theta = x + \theta^m + \theta - \theta^m$ and since U_m is the least dimension face in $C^{(m)}$, the Lemma 2.2 guarantees that $a_j'y^{(m)} \geq a_j'z^{(m)} \geq 0$, $j = 1, \dots, m$, and by Lemma 2.3, $\|y^{(m)} - y^0\|^2 \geq \|z^{(m)} - z^0\|^2$ and therefore $\|y^{(m)}\|^2 \geq \|z^{(m)}\|^2$.

Also we note that $y^{(m)} - z^{(m)} \in C^{(m)} \cap U_m^{\perp}$, therefore $p(y^{(m)} - z^{(m)}| C^{(m+1)} \cap U_m) = 0$, being $C^{(m+1)} = \{x\colon a_j'x \geq 0,\ j = 1, \dots, m + 1\}$, which implies, by Lemma 2.1, that $y^{(m)} - z^{(m)} \notin C^{(m+1)}$ or $y^{(m)} = z^{(m)}$. In any case, $a'_{m+1}y^{(m)} \leq a'_{m+1}z^{(m)}$. In the same way, $a_j'y^{(m)} \leq a_j'z^{(m)}$, $j = m + 1, \dots, n$. Now suppose $y^{(m)} \in H_A$, then $z^{(m)} \in H_A$, and both of them coincide respectively with $(x + \theta)^A$ and $(x + \theta^m)^A$ and the result follows. In the other case, $a_j'y^{(m)} < 0$ for some j. Without loss of generality we can assume $j = m + 1$. Consider $y^{(m+1)} = p(x + \theta|C^{(m+1)})$ and $z^{(m+1)} = p(x + \theta^m|C^{(m+1)})$. We deal with two situations:

(a) $$a'_{m+1}z^{(m)} \leq 0.$$

Let us consider the affine hyperplanes $H_j^z = \{x\colon a_j'x = a_j'z^{(m)}\}$, $j = 1, \dots, m + 1$.

Let A be the set of the indices j for which H_j^z separates $y^{(m)}$ and $y^{(m+1)}$. A is not empty, since $m + 1 \in A$. For each $j \in A$, there exists λ_j such that $\lambda_j y^{(m)} + (1 - \lambda_j)y^{(m+1)} \in H_j^z$.

Consider $\lambda_0 = \max_{j \in A} \lambda_j$ and $y' = \lambda_0 y^{(m)} + (1 - \lambda_0)y^{(m+1)}$.

Then, we have $y' - z^{(m)} \in C^{(m+1)}$ and $y'^{(m+1)} = y^{(m+1)}$, so that $p(y'^{(m+1)}|U_{m+1}) = p(y^{(m+1)}|U_{m+1}) = p(z^{(m+1)}|U_{m+1})$, the last equality because U_{m+1} is the least dimension face of $C^{(m+1)}$ and $y^{(m)0} = z^{(m)0}$.

If we decompose $y' = z^{(m)} + (y' - z^{(m)})$, then by Lemma 2.2, $a_j'y^{(m+1)} \geq a_j'z^{(m+1)} \geq 0$, $j = 1, \dots, m + 1$, and by Lemma 2.3, $\|y'^{(m+1)} - p(y'^{(m+1)}|U_{m+1})\|^2 \geq \|z^{(m+1)} - p(z^{(m+1)}|U_{m+1})\|^2$. Therefore $\|y^{(m+1)}\|^2 \geq \|z^{(m+1)}\|^2$.

(b) $$a'_{m+1}z^{(m)} > 0.$$

The hyperplane $\{a'_{m+1}x = 0\}$ separates $y^{(m)}$ and $z^{(m)}$, so there is $\lambda \in (0,1)$ such that $y' = \lambda y^{(m)} + (1-\lambda)z^{(m)} \in C^{(m+1)}$ and $\|y'\|^2 \geq \|z^{(m)}\|^2 = \|z^{(m+1)}\|^2$. The points $y^{(m+1)}$ and y' are in case (a) and therefore $\|y^{(m+1)}\|^2 \geq \|y'\|^2$. In both (a) and (b) situations, the points $y^{(m+1)}$ and $z^{(m+1)}$ verify the conditions that $y^{(m)}$ and $z^{(m)}$ verified at the previous step, so that we could repeat the same procedure with the cone $C^{(m+2)}$ and so on.

The PAV algorithm for acute cones guarantees [Martín and Salvador (1988)], that in a finite number of steps, $(x+\theta)^A$ and $(x+\theta^m)^A$ are reached. For these points, $\|(x+\theta)^A\|^2 \geq \|(x+\theta^m)^A\|^2$. Therefore $T(x+\theta) \geq T(x+\theta^m)$, since $(x+\theta)^0 = (x+\theta^m)^0$.

Case 3. Consider x such that $x^0 = p(x|U_m)$ and let θ be in H_0 such that $(x+\theta)^0 = p(x+\theta|U_s)$, $r \leq s < m \leq n$. Denote $x^{(s)} = p(x|U_s)$ and $\theta^{(s)} = p(\theta|U_s)$. Note that $x^{(s)} \notin H_A$, whenever $s < m$. Without loss of generality, we can suppose $a'_{s+1}x^{(s)} < 0$.

$(x+\theta)^0 \in H_0 \subset H_A$, so that $a'_{s+1}(x+\theta)^0 = a'_{s+1}(x^{(s)} + \theta^{(s)}) \geq 0$.

Therefore $a'_{s+1}\theta^{(s)} \geq -a'_{s+1}x^{(s)} > 0$.

Consider $\lambda_1 = -(a'_{s+1}x^{(s)})/(a'_{s+1}\theta^{(s)})$, $0 < \lambda_1 \leq 1$.

$a'_{s+1}(x^{(s)} + \lambda_1\theta^{(s)}) = 0$ and we can write $p(x+\lambda_1\theta|U_s) = p(x+\lambda_1\theta|U_{s+1})$.

On decomposing $x+\theta = x+\lambda_1\theta+\theta-\lambda_1\theta$, we have $\theta-\lambda_1\theta \in H_0$ and $x+\theta$ and $x+\lambda_1\theta$ are in Case 2 hence $T(x+\theta) \geq T(x+\lambda_1\theta)$. Repeating the procedure, we obtain $(x+\lambda_1\theta)^0 = p(x+\lambda_1\theta|U_{s+1})$.

Consider $\lambda_2\theta$ in such a way that $\lambda_1\theta - \lambda_2\theta \in H_0$ and $p(x+\lambda_2\theta|U_{s+1}) = p(x+\lambda_2\theta|U_{s+2})$ and therefore $T(x+\lambda_1\theta) \geq T(x+\lambda_2\theta)$.

In this way, after $m-s-1$ steps we obtain $\lambda_{m-s-1}\theta$ such that $(x+\lambda_{m-s-1}\theta)^0 = p(x+\lambda_{m-s-1}\theta|U_m)$. Therefore $T(x+\lambda_{m-s-1}\theta) \geq T(x)$. The chain of inequalities obtained proves that $T(x+\theta) \geq T(x)$. □

Theorem 2.2. *The L.R.T., $\{T > t\}$ for testing H_0 against $H_A - H_0$ is dominated by $\{T^* > t\}$.*

Proof. (a) Fix a point θ_0 in $RI(H_0) = \{a'_j\theta = 0,\ j = 1,\dots,r;\ a'_j\theta > 0,\ j = r+1,\dots,n\}$. For each x in R^k, there is a λ, depending on x, such that

$$T(x+\lambda\theta_0) = T^*(x+\lambda\theta_0).$$

Let x be a point of R^k and $z = p(x|H_A^*)$. Consider

$$\delta'_j = a'_j z, \qquad j = 1,\dots,r,$$

$$\delta'_j = -\frac{a'_j z}{a'_j\theta_0}, \qquad j = r+1,\dots,n \quad \text{and} \quad \lambda = \max\{\delta_1,\dots,\delta_n\}.$$

Then $z + \lambda\theta_0 \in H_A$.

$\theta_0 \in H_0 \subset H_0^*$, with H_0^* the least dimension face in H_A^*, so that

$$p(x+\lambda\theta_0|H_A^*) = z + \lambda\theta_0 = p(x+\lambda\theta_0|H_A)$$

and

$$p(x + \lambda\theta_0 | H_0^*) = p(x + \lambda\theta_0 | H_0).$$

As a consequence $T(x + \lambda\theta_0) = T^*(x + \lambda\theta_0)$.

(b) $$P_\theta(T > t) \leq P_\theta(T^* > t) \qquad \forall\, \theta, \qquad \forall\, t.$$

We shall prove that $T(x) \leq T^*(x)\ \forall\ x$, so that (b) will become an obvious consequence. Let x be a point in R^k and $z = p(x|H_A^*)$. H_A is an acute cone, so that z^A and x^A can be reached by projecting x on the same subspace. The same is true for z^0 and x^0. Therefore $z^A = x^A$ and $z^0 = x^0$ and $T(x) = T(z)$. By Theorem 2.1, $T(z) \leq T(z + \theta)\ \forall\ \theta \in H_0$. According to (a), we can choose $\theta \in H_0$ in such a way that $z + \theta \in H_A$. Using then the Lemma 2.1, we conclude that $T(z + \theta) = T^*(z + \theta)$.

$T^*(z + \theta) = T^*(z)$ follows, since $\theta \in H_0 \subset H_0^*$ and H_0^* is the least dimension face in H_A^*.

Finally, from the definition of T^*, $T^*(z) = T^*(x)$ and we can assure that, $\forall\ x$, $T(x) \leq T^*(x)$.

Inequality (b) proves the L.R.T. to be less powerful than the test $\{T^* > t\}$. Now, we only need to prove that the same significance level is reached by both tests.

(c) $$\forall\, t, \ \sup_{\theta \in H_0} P_\theta(T > t) = P_0(T^* > t).$$

Let t be a real number with $P_0(T^*(X) > t) = \alpha$, where $X \sim N_k(0, \Gamma)$.

Consider $\delta > 0$ and E a sphere centered at the origin, such that

$$P_0(\{T^*(X) > t\} \cap E) \geq \alpha - \delta.$$

This is always possible by considering E with $P_0(X \in E) \geq 1 - \delta$. For all λ and $\theta_0 \in RI(H_0)$, we have

$$P_0(\{T^*(X) > t\} \cap E) = P_{\lambda\theta_0}(\{T^*(X + \lambda\theta_0) > t\} \cap \{E + \lambda\theta_0\}).$$

From (a) and the boundedness of E, there exists λ_0 such that $T^*(x + \lambda_0\theta_0) = T(x + \lambda_0\theta_0)\ \forall\ x \in E$.

Therefore $P_{\lambda_0\theta_0}(\{T(X + \lambda_0\theta_0) > t\} \cap \{E + \lambda_0\theta_0\}) \geq \alpha - \delta$.

This inequality beside (b) proves (c) and the theorem follows. □

Figure 1 sketches the results in the proof of Theorem 2.2. H_A is given by $a_1 = (0, 1)$ and $a_2 = (\frac{1}{2}, -1)$ and we can see $\{T > t\}$, the striped region, to be contained in $\{T^* > t\}$, the dotted and striped region which shows the critical region for testing H_0^* (defined by a_1) against $H_A^* - H_0^*$.

REMARK 1. If we consider part (c) in the proof of Theorem 2.2 and Figure 1, we may obtain an intuitive idea for getting the significance level of the test $\{T > t\}$.

The significance level is reached as we consider $\lambda_0 \to \infty$ since for $\theta_0 \in RI(H_0)$, $P_{\lambda\theta_0}(T(X) > t)$ is an increasing function of λ. This also implies the test $\{T > t\}$ is biased. When $\theta_0 \in RI(H_0)$ and $\lambda_0 \to \infty$, the only sensible

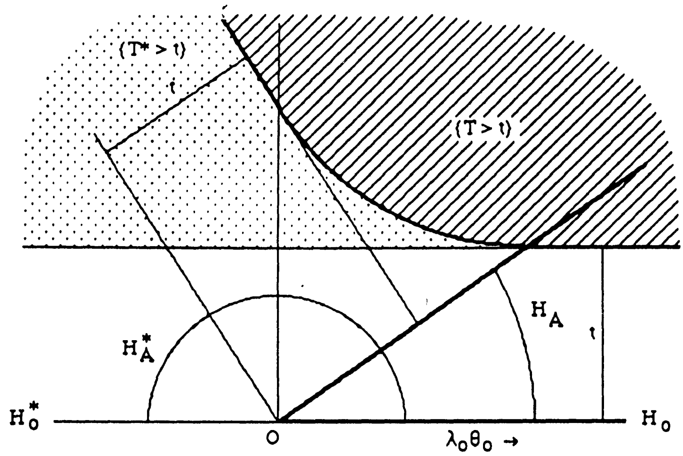

FIG. 1. *Critical regions given by T and T*.*

constraints defining H_A are those defining H_A^*, so that at infinity T^* and T become equivalents.

REMARK 2. The region $\{T^* > t\}$ in Figure 1 yields the uniformly most powerful level α test of H_0 against $H_A - H_0$. Although in general that is not true, possibly $\{T^* > t\}$ will always be admissible. (We are in debt to a referee for this remark).

Figure 2 shows how the Theorem 2.2 fails when H_A is not acute. H_A is defined by $a_1 = (0,1)$ and $a_2 = (1,1)$. By considering t in such a way that $P_\theta(T^* > t) = \alpha \;\forall\; \theta \in H_0^*$, it can be seen $\{T^* > t\} \subset \{T > t\}$, so that $P_\theta(T > t) > \alpha \;\forall\; \theta \in H_0$, and the L.R.T. for testing H_0 against $H_A - H_0$ is not

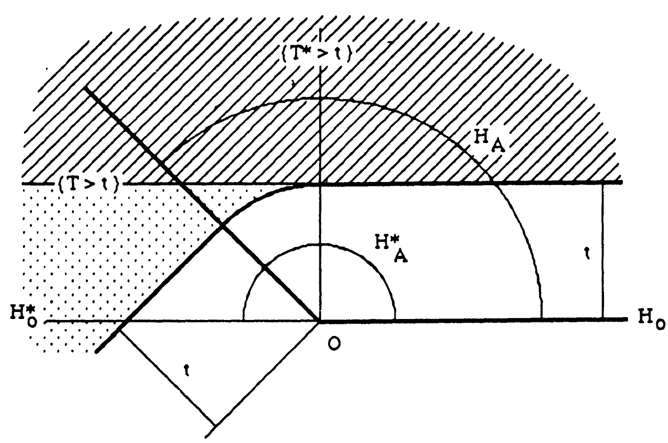

FIG. 2. *Critical regions given by T and T* for H_A nonacute.*

dominated by $\{T^* > t\}$ ($\{T > t\}$ is the dotted region added to the striped region $\{T^* > t\}$).

3. Proof of the lemmas. We give proofs of the Lemmas used in Section 2.

PROOF OF LEMMA 2.1. Let x be in C. We use induction on the number of elements in B. Consider $B = \{i\}$ and $x^i = p(x|L_i)$. Suppose $x^i \notin C$. Then there is a $j \neq i$ such that $a'_j x^i < 0$.

Consider $x^{ij} = p(x|L_{ij})$. We have $x^{ij} - x^i \in L_i \cap C_{ij}$, since $a'_i(x^{ij} - x^i) = 0$ and $a'_j(x^{ij} - x^i) = -a'_j x^i > 0$. Also, $x^{ij} - x^i$ is orthogonal to L_{ij} because $x^{ij} = p(x^i|L_{ij})$.

Let $z = x - x^{ij}$ and $y = x^i - x^{ij}$. They satisfy $a'_i z \geq 0$, $a'_j z \geq 0$, $a'_i y = 0$ and $a'_j y < 0$, so that L_j separates y and z. There is an $\alpha \in (0, 1]$ such that $t = \alpha z + (1 - \alpha) y \in L_j$. Consequently $a'_j(t - x^{ij}) = 0$ and $a'_i(t - x^{ij}) \geq 0$, that is, $t - x^{ij} \in L_j \cap C_{ij}$.

On the other hand $(x^{ij} - x^i)'(t - x^{ij}) = (x^{ij} - x^i)'(x^i - x^{ij} + \alpha(x - x^i) - \alpha x^i) = -\|x^i - x^{ij}\|^2 < 0$, which is in contradiction with the acuteness of C_{ij}. Now, let us consider $x \neq 0$ and C strictly acute and suppose $x^i = 0$. Then we can write $x = x - x^i \in L_i^{\perp}$ so that $x = \lambda a_i$, where $\lambda > 0$ since $x \neq 0$ and $x \in C$. C strictly acute implies $a'_j a_i < 0$, $j \neq i$, therefore $a'_j x = \lambda a'_j a_i < 0$, in contradiction with $x \in C$, so that we conclude $x^i = p(x|L_i) \neq 0$.

Now, we suppose the result is right when B contains r elements. Let B be some subset with $r + 1$ elements in $\{1, \dots, n\}$. Consider $x^{B-i} = p(x|L_{B-i})$ for $i \in B$.

By the induction hypothesis, $x^{B-i} \in L_{B-i} \cap C$.

Now $L_B = L_i \cap L_{B-i}$, so that $x^B = p(x|L_B) = p(x^{B-i}|L_B)$.

In the subspace L_{B-i}, the cone $L_{B-i} \cap C$ is acute and we can use the preceding arguments in order to obtain $x^B \in C$ and also $x^B \neq 0$, whenever $x \neq 0$ and C strictly acute. □

PROOF OF LEMMA 2.2. If $x \in C$ the result is obvious.

Given x in R^k, consider $B = \{i\colon a'_i x^c = 0\}$ and $B^\delta = \{i\colon a'_i(x + \delta)^c = 0\}$, where $x^c = p(x|C)$ and $(x + \delta)^c = p(x + \delta|C)$.

Let δ be in C. We begin by showing $B^\delta \subset B$.

Consider $B_r = \{i\colon a'_i x^r \leq 0\}$ with $x^{r+1} = p(x|L_{B_r})$ and

$$B_r^\delta = \left\{i\colon a'_i(x + \delta)^r \leq 0\right\} \quad \text{with } (x + \delta)^{r+1} = p\left(x + \delta \middle| L_{B_r^\delta}\right)$$

for $r = 0, 1, \dots,$ $[x^0 = x$ and $(x + \delta)^0 = x + \delta]$. It is obvious that $B_0 \subset B_1 \subset \cdots \subset B$ and $B_0^\delta \subset B_1^\delta \subset \cdots \subset B^\delta$.

$$B_0^\delta \subset B_0, \quad \text{since } a'_i x \leq a'_i x + a'_i \delta \leq 0 \quad \text{for } i \in B_0^\delta.$$

We shall now prove that $B_1^\delta \subset B_1$.

$$(x + \delta)^1 = p\left(x + \delta \middle| L_{B_0^\delta}\right) = p\left(x \middle| L_{B_0^\delta}\right) + p\left(\delta \middle| L_{B_0^\delta}\right).$$

Let $i \in B_1^\delta$, then $a'_i(x + \delta)^1 \leq 0$.

By Lemma 2.1, $p(\delta|L_{B_0^\delta}) \in C$, so that $a_i' p(\delta|L_{B_0^\delta}) \geq 0$, therefore $a_i' p(x|L_{B_0^\delta}) \leq 0$.

But $L_{B_0} \subset L_{B_0^\delta}$, since $B_0^\delta \subset B_0$, so that $x^1 = p(x|L_{B_0}) = p(p(x|L_{B_0^\delta})|L_{B_0})$ and we can assume that x^1 has been obtained by projecting x on L_{B_0} after projecting x on $L_{B_0^\delta}$ in a step of the PAV process applied to x. In this way, either x^1 has the restriction given by a_i as an active constraint or does not, as it happens with $p(x|L_{B_0^\delta})$. Therefore $a_i' x^1 \leq 0$ and $i \in B_1$.

In the same way, it can be proved that $\forall\ r$, $B_r^\delta \subset B_r$.

If $B_r^\delta = B^\delta$ and $B_s = B$, then $r \leq s$, because if $p(x|L_{B_s}) \in C$, then $p(x + \delta|L_{B_s}) \in C$ by Lemma 2.1. Therefore, $B^\delta \subset B$.

Let $x \notin C$, then for any $j \in B$, $a_j'(x + \delta)^c \geq a_j' x^c = 0$ and

$$(x + \delta)^c = p(x + \delta|L_{B^\delta}) = p(x|L_{B^\delta}) + p(\delta|L_{B^\delta}).$$

Denote by y, z the first and second terms on the right-hand side. By Lemma 2.1, $z \in C$. Moreover $L_B \subset L_{B^\delta}$, since $B^\delta \subset B$ and so $x^c = p(x|L_B) = p(y|L_B)$.

Now, we prove, for $y \notin C$,

$$a_j' y \geq a_j' x^c \qquad \forall\, j \notin B. \tag{3.1}$$

Assume that there is $j \notin B$ such that $a_j' y < a_j' x^c$. Consider $x_j^c = p(x|L_B \cap L_j)$ and $y_j = x_j^c + (y - x^c)$. Then $a_j' y_j < 0$ and $a_j' x_j^c = 0$.

Consider $t = \lambda(y - y_j)$. There is a t such that $a_j' t = 0$. For this t, set $t_B = t - (y - x^c)$. Then $t_B = p(t|L_B)$ and $a_j' t_B > 0$ which is not possible since C is acute. Therefore (3.1) holds and we can say $\forall\ j \notin B$, $a_j' x^c \leq a_j' y + a_j' z = a_j'(x + \delta)^c$.

Now, when $y \in C$, we have $B = B^\delta$ and $x^c = y$ and therefore,

$$\forall\, j,\ a_j' x^c = a_j' y \leq a_j' y + a_j' z = a_j'(x + \delta)^c. \qquad \square$$

Proof of Lemma 2.3. Let x, y be elements in C with $a_j' x \leq a_j' y$, $j = 1, \ldots, n$. If $a_j' x = a_j' y$ for all j, then both x and y are in $K_{1,\ldots,n} + x$ and the result follows.

Consider $D = \{j\colon\ a_j' x < a_j' y\}$, $x^D = p(x|L_D)$ and $y^D = p(y|L_D)$ and denote $x^\phi = p(x|K_{1,\ldots,n})$ and $y^\phi = p(y|K_{1,\ldots,n})$. We have

$$\|x - x^\phi\|^2 = \|x - x^D\|^2 + \|x^D - x^\phi\|^2,$$
$$\|y - y^\phi\|^2 = \|y - y^D\|^2 + \|y^D - y^\phi\|^2,$$
$$\|y - y^D\|^2 \geq \|y^{Dx} - y^D\|^2 = \|x - x^D\|^2,$$

where $y^{Dx} = p(y|L_D + x)$ and $L_D + x = \{z\colon\ z = y + x,\ y \in L_D\}$ so that the result will be proved if we prove that $\|y^D - y^\phi\|^2 \geq \|x^D - x^\phi\|^2$.

C is an acute cone and $y - x \in C$, therefore $y^D - x^D \in C$ by Lemma 2.1, so that $a_j' y^D \geq a_j' x^D$, $j = 1, \ldots, n$.

Also y^D, x^D are in $C \cap L_D$, which is an acute cone in the subspace L_D. In this way, x^D and y^D are, with respect to $C \cap L_D$, in the same situation as x and y were respect to C. If $\forall\ j$, $a_j' x^D = a_j' y^D$, then $\|y^D - y^\phi\|^2 = \|x^D - x^\phi\|^2$.

In the other case, we can apply to x^D and y^D the procedure applied to x and y and so on. This iterative procedure gives pairs x^F, y^F satisfying for all j, $a'_j x^F \leq a'_j y^F$. If we have at least one strict inequality for $F \subset \{1, \ldots, n\}$ in every step, then we shall obtain x^ϕ and y^ϕ that verify the result. □

Acknowledgments. The authors are grateful to the Associate Editor, Professor R. Dykstra and the referees for their careful reading of the manuscript and for their valuable comments and suggestions that resulted in a much improved version of the paper.

REFERENCES

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

MARTÍN, M. and SALVADOR, B. (1988). Validity of the pool-adjacent-violator algorithm. *Statist. Probab. Lett.* **6** 143–145.

ROBERTSON, T. and WEGMAN, E. J. (1978). Likelihood ratio tests for order restriction in exponential families. *Ann. Statist.* **6** 485–505.

WARRACK, G. and ROBERTSON, T. (1984). A likelihood ratio test regarding two nested but oblique order restricted hypotheses. *J. Amer. Statist. Assoc.* **79** 881–886.

DEPARTAMENTO DE ESTADISTICA E I.O.
FACULTAD DE CIENCIAS
UNIVERSIDAD DE VALLADOLID
47071 VALLADOLID
SPAIN