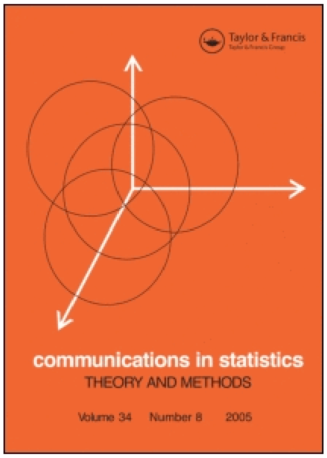

Communications in Statistics - Theory and Methods

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

The influence of the sample on the posterior distribution

Salwa Fahmy; A. Carlos [a]; B. Pereira [a]; Frank Proschan [b]; Moshe Shaked [c]
[a] Universidade De Sao Paulo IME-MAE-Cid Universitaria, Iguatemi, Sao Paulo, SP, Brazil
[b] Department of Statistics, Florida State University, Tallahassee, FL
[c] Department of Mathematics, University of Arizona, Tucson, Arizona

Online Publication Date: 01 January 1982
To cite this Article: Fahmy, Salwa, Carlos, A., Pereira, B., Proschan, Frank and Shaked, Moshe (1982) 'The influence of the sample on the posterior distribution', Communications in Statistics - Theory and Methods, 11:16, 1757 - 1768
To link to this article: DOI: 10.1080/03610928208828348
URL: http://dx.doi.org/10.1080/03610928208828348

COMMUN. STATIST.-THEOR. METH., 11(16), 1757-1768 (1982)

THE INFLUENCE OF THE SAMPLE ON THE POSTERIOR DISTRIBUTION

Salwa Fahmy

1 Road 208
Degla, Maad
Cairo, Egypt, A.R.E.

Carlos A. de B. Pereira

Universidade De Sao Paulo
IME-MAE-Cid Universitaria
CX. Postal 20570-Ag
Iguatemi, Sao Paulo, SP
Brazil

Frank Proschan

Department of Statistics
Florida State University
Tallahassee, FL. 32306

Moshe Shaked

Department of Mathematics
University of Arizona
Tucson, Arizona 85721

Key Words and Phrases: Monotone likelihood ratio property; multivariate totally positive functions; multivariate reverse rule functions; negative dependence; Bayesian operation (prior to posterior).

ABSTRACT

In this paper, we present conditions on the likelihood function and on the prior distribution which permit us to assess the effect of the sample on the posterior distribution. Our work is inspired by Whitt (1979) J. Amer. Statist. Assoc. 74, and is based on the notions of multivariate totally positive and (strongly) multivariate reverse rule functions introduced and studied by Karlin and Rinott (1980a, b).

1. INTRODUCTION

As usual, $\underline{\theta}$ is the parameter of interest and $\underline{x}$ is the data on which the inference about $\underline{\theta}$ is based. The Bayesian operation (prior to posterior) supplies the answer to the question of how to use the information (about $\underline{\theta}$) provided by the data, $\underline{x}$. Here, attention is shifted to another general question: What kind of information about $\underline{\theta}$ does the sample possess?

Whitt (1979) shows that in a Bayesian analysis, under certain general conditions, the larger the observations, the larger (or smaller in a reparametrization) stochastically will be the appropriate parameter of the posterior distribution. His interesting results are developed for the case of a univariate parameter, mainly in the hypergeometric distribution. In the present paper, it is shown that Whitt's key ideas may be extended to the case of multivariate distributions with multivariate parameters. The basic notions used here are multivariate total positivity of order 2, and multivariate reverse rule of order 2, introduced and studied by Karlin and Rinott (1980a, b). These concepts are briefly described in Section 2.

Section 3 presents sufficient conditions on the likelihood function and on the prior distribution under which θ_i, the i-th component of $\underline{\theta}$ in the posterior distribution, be increasing in x_i the i-th component of $\underline{x}$ and decreasing in x_j for $j \neq i$. In Section 4, these results are applied to some well known distributions.

2. PRELIMINARIES

In this section we present definitions, notation, and basic facts used throughout the paper.

Definition 1. A random vector $\underline{x}$ is said to be stochastically increasing in a random vector $\underline{y}$ if $E(\phi(\underline{x})|\underline{y})$ is increasing in $\underline{y}$ for every increasing bounded real function ϕ. (A function ϕ: $R^k \to R$ is said to be increasing if it is increasing in each of its arguments.)

The following concepts of total positivity of order 2 (TP_2), reverse rule of order 2 (RR_2), and Pólya frequency function of order 2 (PF_2) may be found in Karlin (1968).

Definition 2. (i) A nonnegative real function f: $R^2 \to R$ is TP_2 (RR_2) if

$$f(x_1, x_2)f(x_1', x_2') \geq (\leq) f(x_1, x_2')f(x_1', x_2)$$

whenever $x_1' \geq x_1$, and $x_2' \geq x_2$.

(ii) A nonnegative real function ζ: $R \to R$ is PF_2 if

$$f(x_1, x_2) = \zeta(x_1 - x_2) \text{ is } TP_2.$$

The definitions below appear in Karlin and Rinott (1980a, b). For every $\underline{x}$, $\underline{y} \in R^k$, denote:

$$\underline{x} \geq \underline{y} \text{ if } x_i \geq y_i \ \forall\, i = 1, \ldots, k,$$

$$\underline{x} \vee \underline{y} = (\max(x_1, y_1), \ldots, \max(x_k, y_k)),$$

$$\text{and } \underline{x} \wedge \underline{y} = (\min(x_1, y_1), \ldots, \min(x_k, y_k)).$$

The following is the natural generalization of Definition 2(i):

Definition 3. Consider a nonnegative real function f: $R^k \to R$. We say that $f(\underline{x})$ is multivariate totally positive of order 2 or MTP_2 (multivariate reverse rule of order 2 or MRR_2) if:

$$f(\underline{x} \vee \underline{y})\ f(\underline{x} \wedge \underline{y}) \geq (\leq)\ f(\underline{x})f(\underline{y})$$

for every $\underline{x}$, $\underline{y} \in R^k$.

Karlin and Rinott (1980a) show that MTP_2 is a concept of strong positive dependence. They show, however, in their second paper (1980b) that the MRR_2 property fails to be a "good" concept of negative dependence. In the same paper they solve the problem by introducing the following definition. For additional illustration of its usefulness we refer to Block, Savits, and Shaked (1982).

Let $(i_1, \ldots, i_k)$ be any permutation of $(1, 2, \ldots, k)$.

Definition 4. An MRR_2 function f: $R^k \to R$ is said to be strongly-MRR_2 (S-MRR_2) if for any set of k PF_2 functions $\{\zeta_1, \ldots, \zeta_k\}$, and for each $j \leq k$, the function

$$g_{k-j}(x_{i_{j+1}}, \ldots, x_{i_k}) = \int \cdots \int f(\underline{x}) \prod_{m=1}^{j} \zeta_m(x_{i_m})dx_{i_m}$$

is MRR_2 whenever the integral exists.

The multivariate generalization of the familiar monotone likelihood ratio property is given by:

Definition 5. Let f_1 and f_2 be two probability density or mass functions (p.d.f.) on R^k. If for every $\underline{x}$, $\underline{y} \in R^k$,

$$f_2(\underline{x} \vee \underline{y})\ f_1(\underline{x} \wedge \underline{y}) \geq f_1(\underline{x})\ f_2(\underline{y}),$$

then we say that f_2 is "larger than" f_1 in the TP_2 sense, and write $f_2 >_{TP_2} f_1$.

It is well known that the monotone likelihood ratio ordering implies stochastic ordering. The following result is a generalization of this fact.

Theorem 1. Let f_1 and f_2 be two p.d.f.'s on R^k such that $f_2 >_{TP_2} f_1$. If ϕ: $R^k \to R$ is an increasing function, then

$$\int\cdots\int\phi(\underline{x})\ f_1(\underline{x})d\underline{x} \leq \int\cdots\int\phi(\underline{x})\ f_2(\underline{x})d\underline{x}.$$

For a proof of this result, see Karlin and Rinott (1980a).

3. THEORETICAL RESULTS.

In the sequel, let $\underline{\theta}$ be a parameter taking values in a subset Θ (the parameter space) of R^k, and $\underline{x} \in R^n (n \geq k)$ be the data vector. The likelihood function or the sample p.d.f. is represented by $f(\underline{x}|\underline{\theta})$. The prior and the posterior p.d.f.'s are denoted respectively by $\xi(\underline{\theta})$ and $\xi^*(\underline{\theta}|\underline{x})$.

The following result is the TP_2 version of Theorem 4 of Whitt (1979). In the above notation, suppose that $k = n = 1$, θ denotes the parameter, and x denotes the data.

Theorem 2. Consider $f(x|\theta)$ and $\xi^*(\theta|x)$ as bivariate real functions of θ and x. Then $f(x|\theta)$ is $TP_2(RR_2)$ if and only if $\xi^*(\theta|x)$ is $TP_2(RR_2)$.

Proof. For $[H(x)]^{-1} = \int f(x|\theta)\xi(\theta)d\theta$, we notice that $\xi^*(\theta|x) = H(x)\xi(\theta)f(x|\theta)$. Then, $\frac{\xi^*(\theta|x)}{\xi^*(\theta'|x)} \geq(\leq) \frac{\xi^*(\theta|x')}{\xi^*(\theta'|x')}$ holds if and only if $\frac{f(x|\theta)}{f(x|\theta')} \geq(\leq) \frac{f(x'|\theta)}{f(x'|\theta')}$. □

Theorem 2 motivates the results of this section.

Let h: $R^n \to R$ and g: $R^{n+k} \to R$ be two nonnegative real functions.

Theorem 3. Suppose that

(a) $f(\underline{x}|\underline{\theta}) = h(\underline{x})\ g(\underline{x}, \underline{\theta})$, where g is MTP_2,

and (b) $\xi(\underline{\theta})$ is MTP_2. Then

$$\xi^*(\underline{\theta}|\underline{x}') >_{TP_2} \xi^*(\underline{\theta}|\underline{x})$$

for every $\underline{x}$, $\underline{x}' \in R^n$ such that $\underline{x}' \geq \underline{x}$.

Proof. The posterior p.d.f. may be factored as:

$$\xi^*(\underline{\theta}|\underline{x}) = H(\underline{x})\; g(\underline{x}, \underline{\theta})\xi(\underline{\theta}),$$

where

$$[H(\underline{x})]^{-1} = \int g(\underline{x}, \underline{\theta})\xi(\underline{\theta})d\underline{\theta}.$$

Consider two sample points $\underline{x}'$ and $\underline{x}$ such that $\underline{x}' \geq \underline{x}$, and denote

$$\xi_0(\underline{\theta}) = \xi^*(\underline{\theta}|\underline{x}) \text{ and } \xi_1(\underline{\theta}) = \xi^*(\underline{\theta}|\underline{x}').$$

Let $\underline{\theta}$ and $\underline{\theta}'$ be two points in the parameter space Θ. Now, we can write

$$\xi_0(\underline{\theta})\xi_1(\underline{\theta}') = H(\underline{x})H(\underline{x}')g(\underline{x}, \underline{\theta})g(\underline{x}', \underline{\theta}')\xi(\underline{\theta})\xi(\underline{\theta}')$$

$$\leq H(\underline{x})H(\underline{x}')g(\underline{x} \wedge \underline{x}', \underline{\theta} \wedge \underline{\theta}')g(\underline{x} \vee \underline{x}', \underline{\theta} \vee \underline{\theta}')\xi(\underline{\theta} \wedge \underline{\theta}')\xi(\underline{\theta} \vee \underline{\theta}')$$

since both g and ξ are MTP_2. Since $\underline{x}' \geq \underline{x}$, we finally have

$$\xi_0(\underline{\theta})\xi_1(\underline{\theta}') \leq [H(\underline{x})g(\underline{x}, \underline{\theta} \wedge \underline{\theta}')\xi(\underline{\theta} \wedge \underline{\theta}')] \times [H(\underline{x}')g(\underline{x}', \underline{\theta} \vee \underline{\theta}')\xi(\underline{\theta} \vee \underline{\theta}')]$$

$$= \xi_0(\underline{\theta} \wedge \underline{\theta}')\xi_1(\underline{\theta} \vee \underline{\theta}').$$

Thus, it is equivalent to say that $\xi_1 >_{TP_2} \xi_0$. □

Corollary. If conditions (a) and (b) of Theorem 3 hold, then $\underline{\theta}$ is stochastically increasing in $\underline{x}$.

Proof. The result follows immediately from Theorems 1 and 3. □

In many cases, conditions (a) and (b) of Theorem 3 are too strong. A more realistic result is presented below where $\underline{x}$ is considered to be the data reduced by sufficiency, and to have the same dimension of $\underline{\theta}$; that is, n = k.

Let c: $R^k \to R$, and g_i: $R^2 \to R (i = 1, \ldots, k)$ be nonnegative functions and represent the posterior marginal p.d.f. of θ_i $(i = 1, \ldots, k)$ by $\xi_i^*(\theta_i|\underline{x})$.

Theorem 4. Suppose that

$$f(\underline{x}|\underline{\theta}) = h(\underline{x}) \prod_{i=1}^{k} g_i(x_i, \theta_i)c(\underline{\theta})$$

where, for $i = 1, \ldots, k$, g_i is TP_2. Then, for every $i = 1, \ldots, k$, the following condition (for the posterior marginal density of θ_i) holds:

$$\xi_i^*(\theta_i|\underline{x}') >_{TP_2} \xi_i^*(\theta_i|\underline{x})$$

for $\underline{x}'$ equal to $\underline{x}$ except for the i-th coordinate, where $x_i'(\geq x_i)$ replaces x_i.

<u>Proof</u>. Without loss of generality we assume $i = 1$. The posterior marginal p.d.f. of θ_1 is

$$\xi_1^*(\theta_1|\underline{x}) = H(\underline{x})g_1(x_1, \theta_1)\int\cdots\int C(\underline{\theta}) \prod_{i=2}^{k} g_i(x_i, \theta_i)d\theta_i,$$

where $C(\underline{\theta}) = c(\underline{\theta})\xi(\underline{\theta})$, and $H(\underline{x})$ is defined as above. Now, define

$$G_1(\underline{x}, \theta_1) = \int\cdots\int C(\underline{\theta}) \prod_{i=2}^{k} g_i(x_i, \theta_i)d\theta_i,$$

which is constant in x_1. Thus,

$$\xi^*(\theta_1|\underline{x}) = H(\underline{x})g_1(x_1, \theta_1)G_1(\underline{x}, \theta_1).$$

Consider two sample points that differ only in the first coordinate, say

$$\underline{x} = (x_1, x_2, \ldots, x_k), \text{ and } \underline{x}' = (x_1', x_2, \ldots, x_k),$$

where $x_1' \geq x_1$. Define

$$\xi_{10}(\theta_1) = \xi_1^*(\theta_1|\underline{x})$$

and

$$\xi_{11}(\theta_1) = \xi_1^*(\theta_1|\underline{x}')$$

Since by definition

$$G_1(\underline{x}, \theta_1) = G_1(\underline{x}', \theta_1),$$

we thus have

$$\frac{\xi_{11}(\theta_1)}{\xi_{10}(\theta_1)} = \frac{H(\underline{x}')g_1(x_1', \theta_1)}{H(\underline{x})g_1(x_1, \theta_1)},$$

which is increasing in θ_1 by the TP_2 property of g_1. Hence,

$$\xi_{11} >_{TP_2} \xi_{10}. \quad \square$$

Remark 1. Note that the result in Theorem 4 pertains to the posterior marginal p.d.f. of θ_i. Also it holds irrespective of the choice of the prior distribution.

Remark 2. It follows from Theorems 1 and 4 that $E\{\phi(\theta_i)|\underline{x}\}$ is increasing in x_i for every increasing real function $\phi: R \to R$.

In many applications, we noted that the posterior distribution of θ_i stochastically decreases in x_j for every $j \neq i$. This fact is included in the following result.

Theorem 5. Suppose that

$$\xi^*(\underline{\theta}|\underline{x}) = H(\underline{x}) \prod_{i=1}^{k} g_i(x_i, \theta_i)C(\underline{\theta}),$$

where

(i) $g_i(x_i, \theta_i)$ is TP_2 $\forall$ $i = 1, \ldots, k$,

(ii) for fixed $x_i (i = 1, \ldots, k)$, $g_i(x_i, \theta_i)$ is PF_2(in θ_i).

and (iii) $C(\underline{\theta})$ is S-MRR_2.

Then, for every $i = 1, \ldots, k$,

$$\xi_i^*(\theta_i|\underline{x}) >_{TP_2} \xi_i^*(\theta_i|\underline{x}')$$

whenever $\underline{x}' \geq \underline{x}$ and the i-th coordinates of $\underline{x}'$ and $\underline{x}$ are equal; that is, $x_i = x_i'$ and $x_j \leq x_j'$ $\forall$ $j \neq i$. (For $k = 2$, condition (ii) is not required.)

Proof. Without loss of generality we assume $i = 1$,

$$\underline{x} = (x_1, x_2, x_3, \ldots, x_k), \text{ and}$$

$$\underline{x}' = (x_1, x_2', x_3, \ldots, x_k),$$

where $x_2' > x_2$; that is, $\underline{x}$ and $\underline{x}'$ differ only in the second coordinate.

(A) We consider first the case of $k > 2$. The posterior marginal p.d.f. of θ_1 is

$$\xi_1^*(\theta_1|\underline{x}) = H(\underline{x})g_1(x_1, \theta_1)\int g_2(x_2, \theta_2)G(\theta_1, \theta_2, x_3, \ldots, x_k)d\theta_2$$

where

$$G(\theta_1, \theta_2, x_3, \ldots, x_k) = \int\cdots\int C(\underline{\theta}) \prod_{i=3}^{k} g_i(x_i, \theta_i)d\theta_i.$$

Since $C(\underline{\theta})$ is S-MRR_2 and $g_i(x_i, \theta_i)$ is PF_2 in θ_i, it follows from Definition 4 that G is RR_2 in (θ_1, θ_2) for every fixed $(x_3, \ldots, x_k)$.

By the basic composition formula (Karlin [1968]) and the fact that $g_2(x_2, \theta_2)$ is TP_2, it follows that

$$G_1(\theta_1, \underline{x}) = \int g_2(x_2, \theta_2) G(\theta_1, \theta_2, x_3, \ldots, x_k) d\theta_2$$

is RR_2 in (θ_1, x_2) for every fixed $(x_3, \ldots, x_k)$. (Note that G_1 is constant in x_1.)

As before, let

$$\xi_{10}(\theta_1) = \xi_1^*(\theta_1 | \underline{x}), \text{ and } \xi_{11}(\theta_1) = \xi_1^*(\theta_1 | x')$$

and note that

$$\frac{\xi_{11}(\theta_1)}{\xi_{10}(\theta_1)} = \frac{H(\underline{x}')}{H(\underline{x})} \frac{G_1(\theta_1, \underline{x}')}{G_1(\theta_1, \underline{x})}$$

is decreasing in θ_1 since G_1 is RR_2 in (θ_1, x_2). Hence,

$$\xi_{10} >_{TP_2} \xi_{11}.$$

(B) When $k = 2$, from (iii), $C(\theta_1, \theta_2)$ is RR_2 and by the basic composition formula,

$$G_1(\theta_1, x_2) = \int g_2(\theta_2, x_2) C(\theta_1, \theta_2) d\theta_2$$

is RR_2. Thus, if $x_2' > x_2$

$$\frac{\xi_1^*(\theta_1 | (x_1, x_2'))}{\xi_1^*(\theta_1 | (x_1, x_2))} = \frac{H(x_1, x_2')}{H(x_1, x_2)} \frac{G_1(\theta_1, x_2')}{G_1(\theta_1, x_2)}$$

is decreasing in θ_1 and the result follows. □

Remark 3. Note that Theorem 5 involves conditions on the prior distribution.

Remark 4. It follows from Theorems 1 and 5 that $E\{\phi(\theta_i)|\underline{x}\}$ is decreasing in x_j whenever $j \neq i$ and ϕ: $R \to R$ is increasing.

4. APPLICATIONS

In this section, we show how the results of Section 3 apply to some important probability distributions.

Example 1. Multivariate Normal Distribution. Let $\underline{x}$ be a k-dimensional random vector whose coordinates, $x_i (i = 1, \ldots, k)$, are independent. Suppose that for i = 1, ..., k, $\theta_i = E\{x_i\}$ is unknown and $\sigma_i^2 = \mathrm{Var}\{x_i\}$ is known. Then, whatever appropriate prior we choose, Theorem 4 applies, and by Remark 2, $E\{\phi(\theta_i)|\underline{x}\}$ is increasing in x_i *for every increasing real function* ϕ.

Suppose that our prior opinion about $\underline{\theta}$ is represented by a non-singular k-dimensional normal distribution, with mean vector $\underline{\mu}$ and covariance matrix V^{-1}, which is a conjugate prior. Let v_{ij} (i, j = 1, ..., k) be the (i, j)-th element of V. If $v_{ij} \le 0$ for every $i \ne j$, then $\xi(\underline{\theta})$, the prior p.d.f., is MTP_2 (Barlow and Proschan [1981]). Thus, Theorem 3 and its corollary apply yielding the conclusion that $\underline{\theta}$ is stochastically increasing in $\underline{x}$.

If the prior $\xi(\underline{\theta})$ is negatively dependent in the S-MRR_2 sense, then Theorem 5 applies, and by Remark 4, for every i = 1, ..., k, $E\{\phi(\theta_i)|\underline{x}\}$ is decreasing in $x_j (\forall j \ne i)$ for every increasing real function ϕ. A normal prior is S-MRR_2 if

$$V^{-1} = D - \alpha'\alpha$$

where D is a positive definite diagonal matrix, say $D = \mathrm{diag}(d_1, \ldots, d_k)$ with $d_i > 0$, and $\alpha = (\alpha_1, \ldots, \alpha_k)$ with $\alpha_i \ge 0$ and $\sum_1^k \alpha_i^2 d_i^{-1} < 1$ (Karlin and Rinott (1980b] or Block, Savits, and Shaked [1982]). In particular, if the correlation matrix for the normal prior distribution is

$$\begin{pmatrix} 1 & \rho & \cdots & \rho \\ \vdots & & & \vdots \\ \rho & \rho & & 1 \end{pmatrix},$$

where $\rho \le 0$, then the prior p.d.f. is S-MRR_2. □

Example 2. Multivariate Bernoulli Trials. Let $\underline{y}_1, \underline{y}_2, \ldots$ be a sequence of i.i.d. k-dimensional vectors with common multinomial distribution with parameters n = 1 and $\underline{p} = (p_1, \ldots, p_k)$, where $\sum_1^k p_i = 1$. That is, $\underline{y}_1, \underline{y}_2, \ldots$ are independent and $\underline{y}_i \sim M(1, \underline{p})\ \forall\ i = 1, 2, \ldots$. Note that for any finite sequence $\underline{y}_1, \ldots, \underline{y}_m$, the vector $\underline{x} = \sum_1^m \underline{y}_i = (x_1, \ldots, x_k)$ is a sufficient statistic since the probability mass function of $\underline{y}_1, \ldots, \underline{y}_m$ is

$$L = \prod_{i=1}^{k} p_i^{x_i} I(\underline{p})$$

where $I(\underline{p})$ is the indicator function of $\sum_1^k p_i = 1$. Clearly, $f(\underline{x}|\underline{p}) = h(\underline{x})L$.

We notice now that Theorem 4 applies since $p_i^{x_i}$ is TP_2 in (x_i, p_i). Hence, by Remark 2, for $i = 1, \ldots, k$, $E\{\phi(p_i)|\underline{x}\}$ is increasing in x_i for every increasing real function ϕ.

Suppose that a Dirichlet distribution with parameters each no smaller than unity is chosen to represent our prior opinion about $\underline{p}$. With this choice, the prior p.d.f. for $\underline{p}$ is S-MRR_2 (see Karlin and Rinott [1980b] or Block, Savits, and Shaked [1982]). Thus, since $p_i^{x_i}$ is PF_2 in $p_i (i = 1, \ldots, k)$, Theorem 5 applies and by Remark 4, for $i = 1, \ldots, k$, $E\{\phi(p_i)|\underline{x}\}$ is decreasing in x_j whenever $j \neq i$ and ϕ increasing. □

<u>Remark 5</u>. Note that Example 2 includes both the multinomial and the negative multinomial models.

<u>Example 3</u>. <u>Multivariate Hypergeometric Distribution</u>. The probability mass function in this case may be expressed as

$$f(\underline{x}|\underline{\theta}) = h(\underline{x}) \prod_{i=1}^{k} \binom{\theta_j}{x_j} I(\underline{\theta}),$$

where $I(\underline{\theta})$ is the indicator function of $\sum_1^k \theta_i = N$. Again, Theorem 4 applies since $\binom{\theta_i}{x_i}$ is TP_2. Thus, for $i = 1, 2, \ldots, k$, $E\{\phi(\theta_i)|\underline{x}\}$ is increasing in x_i whenever ϕ is an increasing function.

Suppose that a Dirichlet-Multinomial distribution [denoted by $DM(n; \underline{\alpha})$] with shape parameters, α_i's, each no smaller than unity is chosen to represent our opinion about $\underline{\theta}$. In addition, note that $\binom{\theta_i}{x_i}$ is PF_2 in θ_i for every fixed x_i. Thus, Theorem 5 and Remark 4 apply. To conclude this example, recall that the posterior distribution of $\underline{\theta} - \underline{x}$ (the unsampled population) is $DM(N; \alpha_1 + x_1, \ldots, \alpha_k + x_k)$ whose component means are given by $(\sum_1^k(\alpha_j + x_j))^{-1}(\alpha_i + x_i)$ for $i = 1, \ldots, k$. Thus, for every $i = 1, \ldots, k$, $E\{\theta_i - x_i|\underline{x}\}$ is decreasing in x_j $\forall$ $j \neq i$. □

Remark 6. The above example may be viewed as a natural generalization of Theorems 2 and 3 of Whitt (1979).

Example 4. Uniform Distribution. Suppose that $t_1, \ldots, t_n$ is a random sample from a uniform distribution on the real interval (θ_1, θ_2). Let (x_1, x_2) be the usual sufficient statistic, that is, $x_1 = \min(t_1, \ldots, t_k)$ and $x_2 = \max(t_1, \ldots, t_k)$. Suppose that a bilateral Pareto distribution with parameter (r_1, r_2, α) represents our prior opinion. This distribution is a conjugate prior for the uniform distribution case (De Groot [1970], pp. 62-63 and pp. 172-174). The prior p.d.f. is given by

$$\xi(\theta_1, \theta_2) = \alpha(\alpha + 1)(r_2 - r_1)^{\alpha}(\theta_2 - \theta_1)^{-(\alpha+2)}I(\theta_1, \theta_2)$$

where $\alpha > 0$, $r_1 < r_2$, and $I(\theta_1, \theta_2)$ is the indicator function of $\theta_1 < r_1$ and $\theta_2 > r_2$. The likelihood function may be expressed as:

$$L = (\theta_2 - \theta_1)^{-n}I_1(x_1, \theta_1)I_2(x_2, \theta_2)$$

where I_1 and I_2 are the indicator functions of $x_1 > \theta_1$ and $x_2 < \theta_2$ respectively. Since I_1 and I_2 are TP_2 functions and $(\theta_2 - \theta_1)^{-b}$ is RR_2 for $b > 0$, Theorems 4 and 5 $(k = 2)$ apply. Then by Remarks 2 and 4, we conclude that $E\{\phi(\theta_1)|(x_1, x_2)\}$ increases in x_1 and decreases in x_2 for every increasing function ϕ. The analogous result for θ_2 is obvious. □

Example 5. The Exponential Family. This application is of a general nature. Consider that the distribution of $\underline{x}$ belongs to the exponential family. With a proper reparametrization we may consider $\underline{\theta}$ such that

$$f(\underline{x}|\underline{\theta}) = h(\underline{x}) \exp\{\underline{\theta}'\underline{x}\}c(\underline{\theta})$$

Since $e^{\theta_i x_i}$ is TP_2, Theorem 4 applies and for any increasing ϕ, $E\{\phi(\theta_i)|\underline{x}\}$ is increasing in x_i. With a suitable choice for the prior, Theorem 5 and Remark 4 apply.

5. ACKNOWLEDGEMENTS.

Originally, we stated Theorem 5 for the case $k = 2$. This fact greatly restricted our applications. We would like to thank Shelly Zacks and the referee for their suggestions which made possible the actual extended version.

BIBLIOGRAPHY

Barlow, R.E. and Proschan, F. (1981). Statistical Theory of Reliability and Life Testing. To Begin With, Silver Spring, MD.

Block, H.D., Savits, T.H., and Shaked, M. (1982). Some Concepts of Negative Dependence. Ann. Prob., to appear.

De Groot, M.H. (1970). Optimal Statistical Decisions. McGraw-Hill, New York.

Karlin, S. (1968). Total Positivity. Stanford University Press.

Karlin, S. and Rinott, Y. (1980a). Classes of Orderings of Measures and Related Correlation Inequalities. I. Multivariate Totally Positive Distributions. J. Mult. Analysis 10, no. 4; 467-498.

Karlin, S. and Rinott, Y. (1980b). Classes of Orderings of Measures and Related Correlation Inequalities. II. Multivariate Reverse Rule Distributions. J. Mult. Analysis 10, no. 4, 499-518.

Pereira, C.A.B. (1980). Bayesian Solutions to some Classical Problems of Statistics. Unpublished Ph.D. thesis submitted to Florida State University.

Whitt, W. (1979). A note on the Influence of the Sample on the Posterior Distribution. J. Amer. Statist. Assoc., 74 No. 366 424-426.

Received March, 1981; Revised February 1982.

Recommended by Shelemyahu Zacks, State University of New York at Binghamton, Binghamton, NY.

Refereed Anonymously.