The Influence of the Sample on the Posterior Distribution

by

Salwa Fahmy¹, Carlos A. de B. Pereira², Frank Proschan³, and Moshe Shaked⁴

Department of Statistics
Florida State University
November, 1981

FSU Report M-545R
AFOSR Technical Report 78-107R

Key Words and Phrases: Monotone Likelihood Ratio Property, Multivariate Totally Positive Functions, Multivariate Reverse Rule Functions, Negative Dependence, Bayesian Operation (Prior to Posterior).

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

² Universidade De Sao Paulo, IME-UAE-Cid Universitaria, CX. Postal 20570-Ag Iguatemi, Sao Paulo, SP, BRAZIL


⁴ University of Arizona, Tucson.
The Influence of the Sample on the Posterior Distribution

by

Salwa Fahmy¹, Carlos A. de B. Pereira², Frank Proschan³, and Moshe Shaked⁴

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).
The Influence of the Sample on the Posterior Distribution

by

Salwa Fahmy\textsuperscript{1}, Carlos A. de B. Pereira\textsuperscript{2}, Frank Proschan\textsuperscript{3}, and Moshe Shaked\textsuperscript{4}

1. Introduction

As usual, \( \theta \) is the parameter of interest and \( x \) is the data on which the inference about \( \theta \) is based. The Bayesian operation (prior to posterior) supplies the answer to the question of how to use the information (about \( \theta \)) provided by the data, \( x \). Here, attention is shifted to another general question: What kind of information about \( \theta \) 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 $\theta_i$, the $i$-th component of $\theta$ in the posterior distribution, by increasing in $x_i$ and the $i$-th component of $x$ 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 $x$ is said to be stochastically increasing in a random vector $y$ if $E[f(x) | y]$ is increasing in $y$ for every increasing bounded real function $f$. (A function $f: \mathbb{R}^k \to \mathbb{R}$ 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 $f: \mathbb{R}^2 \to \mathbb{R}$ 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 $\xi: \mathbb{R} \to \mathbb{R}$ is $PF_2$ if

$$f(x_1, x_2) = \xi(x_1 - x_2)$$

is $TP_2$.

The definitions below appear in Karlin and Rinott (1980a, b).

For every $x, y \in \mathbb{R}^k$, denote:

$$x \succeq y \text{ if } x_i \geq y_i \forall i = 1, \ldots, k,$$
\[ x \vee y = (\max(x_1, y_1), \ldots, \max(x_k, y_k)). \]
and \[ x \wedge y = (\min(x_1, y_1), \ldots, \min(x_k, y_k)). \]

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

**Definition 3**

Consider a nonnegative real function \( f: \mathbb{R}^k \to \mathbb{R} \). We say that \( f(x) \) is multivariate totally positive of order 2 or MTP\(_2\) (multivariate reverse rule of order 2 or MRR\(_2\)) if:

\[ f(x \vee y) f(x \wedge y) \geq (\leq) f(x) f(y) \]

for every \( x, y \in \mathbb{R}^k \).

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 (1980).

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

**Definition 4**

An MRR\(_2\) function \( f: \mathbb{R}^k \to \mathbb{R} \) is said to be **strongly-MRR\(_2\)** (S-MRR\(_2\)) if for any set of \( k \) PF\(_2\) functions \( \{r_1, \ldots, r_k\} \), and for each \( j \leq k \), the function
\[ g_{k-j}(x_{i_1}, \ldots, x_{i_k}) = \int \cdots \int f(x) \prod_{m=1}^{j} \xi_m(x_{i_m}) dx_{i_m} \]

is \( \text{MRR}_2 \) 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 \( \mathbb{R}^k \). If for every \( x, y \in \mathbb{R}^k \),

\[ f_2(x \vee y) f_1(x \wedge y) \geq f_1(x) f_2(y), \]

then we say that \( f_2 \) is "larger than" \( f_1 \) in the \( \text{TP}_2 \) sense, and write \( f_2 > \text{TP}_2 f_1 \).

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 \( \mathbb{R}^k \) such that \( f_2 > \text{TP}_2 f_1 \).

If \( \phi: \mathbb{R}^k \to \mathbb{R} \) is an increasing function, then

\[ \int \cdots \int \phi(x) f_1(x) dx \leq \int \cdots \int \phi(x) f_2(x) dx. \]

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

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

The following result is the TP\(_2\) version of Theorem 4 of Whitt (1979). In the above notation, suppose that \( k = n = 1 \), \( \theta \) denotes the parameter, and \( x \) denotes the data.

**Theorem 2**

Consider \( f(x|\theta) \) and \( \xi^*(\theta|x) \) as bivariate real functions of \( \theta \) 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, \( \frac{\xi^*(\theta|x)}{\xi^*(\theta'|x')} \geq (s) \frac{\xi^*(\theta|\theta')}{\xi^*(\theta'|\theta')} \) holds

if and only if \( \frac{f(x|\theta)}{f(x|\theta')} \geq (s) \frac{f(x'|\theta)}{f(x'|\theta')} \). \( \Box \)

Theorem 2 motivates the results of this section.

Let \( h: \mathbb{R}^n \rightarrow \mathbb{R} \) and \( g: \mathbb{R}^{n+k} \rightarrow \mathbb{R} \) be two nonnegative real functions.
Theorem 3

Suppose that

(a) $f(x|\theta) = h(x) g(x, \theta)$, where $g$ is MTP$_2$,

and (b) $\xi(\theta)$ is MTP$_2$.

Then

$$\xi^*(\theta|x') > \text{TP}_2 \xi^*(\theta|x)$$

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

Proof

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

$$\xi^*(\theta|x) = H(x) g(x, \theta) \xi(\theta),$$

where

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

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

$$\xi_0(\theta) = \xi^*(\theta|x)$$

and

$$\xi_1(\theta) = \xi^*(\theta|x').$$

Let $\theta$ and $\theta'$ be two points in the parameter space $\Theta$. Now, we can write

$$\xi_0(\theta) \xi_1(\theta') \leq H(x)H(x')g(x, \theta)g(x', \theta')\xi(\theta)\xi(\theta')$$

$$\leq H(x)H(x')g(x \land x', \theta \land \theta')g(x \lor x', \theta \lor \theta')\xi(\theta \land \theta')\xi(\theta \lor \theta')$$
since both $g$ and $\xi$ are MTP$_2$. Since $x' \geq x$, we finally have

$$\xi_0(\theta)\xi_1(\theta') \leq [H(x)g(x, \theta \land \theta')\xi(\theta \land \theta')] \times [H(x')\xi(x', \theta \lor \theta')\xi(\theta \lor \theta')]$$

$$= \xi_0(\theta \land \theta')\xi_1(\theta \lor \theta').$$

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

**Corollary**

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

**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 $x$ is considered to be the data reduced by sufficiency, and to have the same dimension of $\theta$; that is, $n = k$.

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

**Theorem 4**

Suppose that

$$f(x|\theta) = h(x) \prod_{i=1}^{k} g_i(x_i, \theta_i)c(\theta)$$
where, for $i = 1, \ldots, k$, $g_i$ is TP. Then, for every $i = 1, \ldots, k$, the following condition (for the posterior marginal density of $\theta_i$) holds:

$$
\xi_i^*(\theta_i | x^*) > TP_2 \xi_i^*(\theta_i | x)
$$

for $x^*$ equal to $x$ except for the $i$-th coordinate, where $x_i^* (\geq x_i)$ replaces $x_i$.

**Proof**

Without loss of generality we assume $i = 1$. The posterior marginal p.d.f. of $\theta_1$ is

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

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

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

which is constant in $x_1$. Thus,

$$
\xi_i^*(\theta_1 | x) = H(x)g_1(x_1, \theta_1)G_1(x, \theta_1)
$$

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

$$
x = (x_1, x_2, \ldots, x_k), \text{ and } x^* = (x_1^*, x_2, \ldots, x_k),
$$

where $x_1^* \geq x_1$.
Define

\[ \xi_{10}(\theta_1) = \xi_1^*(\theta_1|x) \]

and

\[ \xi_{11}(\theta_1) = \xi_1^*(\theta_1|x^-). \]

Since by definition

\[ g_1(x, \theta_1) = g_1(x^-, \theta_1), \]

we thus have

\[ \frac{\xi_{11}(\theta_1)}{\xi_{10}(\theta_1)} = \frac{H(x^-)g_1(x^-, \theta_1)}{H(x)g_1(x_1, \theta_1)}, \]

which is increasing in \( \theta_1 \) by the TP₂ property of \( g_1 \). Hence,

\[ \xi_{11} > \text{TP}_2 \xi_{10}. \]

Remarks

1 - Note that the result in Theorem 4 pertains to the posterior marginal p.d.f. of \( \theta_i \). Also it holds irrespective of the choice of the prior distribution.

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

In many applications, we noted that the posterior distribution of \( \theta_i \) stochastically decreases in \( x_j \) for every \( j \neq i \). This fact is included in the following result.
Theorem 5

Suppose that

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

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 \( \theta_i), \)

and (iii) \( C(\theta) \) is S-MRR_2.

Then, for every \( i = 1, \ldots, k, \)

$$\xi_i^*(\theta_i | x) > TP_2 \xi_i^*(\theta_i | x^-)$$

whenever \( x^- \geq x \) and the i-th coordinates of \( x^- \) and \( x \) 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 \( i = 1, \)

\[ x = (x_1, x_2, x_3, \ldots, x_k), \]

and

\[ x^- = (x_1, x_2^-, x_3, \ldots, x_k), \]

where \( x_2^- > x_2; \) that is, \( x \) and \( x^- \) differ only in the second coordinate.

(A) We consider first the case of \( k > 2. \)

The posterior marginal p.d.f. of \( \theta_1 \) is
\[ \xi^*_1(\theta_1 \mid x) = H(x) g_1(x_1, \theta_1) g_2(x_2, \theta_2) G(\theta_1, \theta_2, x_3, \ldots, x_k) d\theta_2, \]

where

\[ G(\theta_1, \theta_2, x_3, \ldots, x_k) = \int \cdots \int C(\theta) \Pi_{i=1}^k g_1(x_i, \theta_1) d\theta_i. \]

Since \( C(\theta) \) is S-MRR, and \( g_1(x_i, \theta_1) \) is PF in \( \theta_1 \), it follows from Definition 4 that \( G \) is RR in \((\theta_1, \theta_2)\) 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, it follows that

\[ G_1(\theta_1, x) = \int g_2(x_2, \theta_2) G(\theta_1, \theta_2, x_3, \ldots, x_k) d\theta_2 \]

is RR in \((\theta_1, x_2)\) for every fixed \((x_3, \ldots, x_k)\). (Note that \( G_1 \) is constant in \( x_1 \).)

As before, let

\[ \xi_{10}(\theta_1) = \xi^*_1(\theta_1 \mid x), \quad \text{and} \quad \xi_{11}(\theta_1) = \xi^*_1(\theta_1 \mid x^-) \]

and note that

\[ \frac{\xi_{11}(\theta_1)}{\xi_{10}(\theta_1)} = \frac{H(x^-) G_1(\theta_1, x^-)}{H(x) G_1(\theta_1, x)} \]

is decreasing in \( \theta_1 \) since \( G_1 \) is RR in \((\theta_1, x_2)\). Hence,

\[ \xi_{10} > \text{TP}_2 \xi_{11}. \]
When }k = 2\text{, from (iii), }C(\theta_1, \theta_2) \text{ is } RR_2 \text{ 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 \text{. Thus, if } x_2^- > x_2 \text{,

\[ \frac{\xi_1(\theta_1^* | (x_1, x_2^-))}{\xi_1(\theta_1^* | (x_1, x_2))} = \frac{H(x_1, x_2^-) G_1(\theta_1, x_2^-)}{H(x_1, x_2) G_1(\theta_1, x_1)} \]

is decreasing in } \theta_1 \text{ and the result follows. } \square

Remarks

\[ 3 \text{- Note that Theorem 5 involves conditions on the prior distribution.} \]

\[ 4 \text{- It follows from Theorems 1 and 5 that } E(\phi(\theta_i | x) \mid x) \text{ is decreasing in } x_j \text{ whenever } j \neq i \text{ and } \phi : \mathbb{R} \rightarrow \mathbb{R} \text{ is increasing.} \]

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 } x \text{ be a } k \text{-dimensional random vector whose coordinates, } x_i (i = 1, \ldots, k), \text{ are independent. Suppose that for } i = 1, \ldots, k, \theta_i = E(x_i) \text{ is unknown and } \sigma_i^2 = \text{Var}(x_i) \text{ is known.}

Then, whatever appropriate prior we choose, Theorem 4 applies, and
by Remark 2, \( E(\phi(0, x)| x) \) is increasing in \( x \) for every increasing real function \( \phi \).

Suppose that our prior opinion about \( \theta \) is represented by a nonsingular \( k \)-dimensional normal distribution, with mean vector \( \mu \) and covariance matrix \( V^{-1} \), which is a conjugate prior. Let \( v_{ij} \) (\( i, j = 1, \ldots, k \)) be the \((i, j)\)-th element of \( V \). If \( v_{ij} \leq 0 \) for every \( i \neq j \), then \( \xi(\theta) \), the prior p.d.f., is \( \text{MTP}_2 \) (Barlow and Proschan [1975]). Thus, Theorem 3 and its corollary apply yielding the conclusion that \( \theta \) is stochastically increasing in \( x \).

If the prior \( \xi(\theta) \) is negatively dependent in the \( S-\text{MRR}_2 \) sense, then Theorem 5 applies, and by Remark 4, for every \( i = 1, \ldots, k \), \( E(\phi(0, x)| x) \) is decreasing in \( x_j \) (\( v_j \neq i \)) for every increasing real function \( \phi \). A normal prior is \( S-\text{MRR}_2 \) if

\[
V^{-1} = D - \alpha' \alpha
\]

where \( D \) is a positive definite diagonal matrix, say \( D = \text{diag}(d_1, \ldots, d_k) \) with \( d_i > 0 \), and \( \alpha = (\alpha_1, \ldots, \alpha_k) \) with \( \alpha_i \geq 0 \) and \( \sum_{i=1}^{k} \alpha_i^2 d_i^{-1} < 1 \) (Karlin and Rinott [1980b]). In particular, if the correlation matrix for the normal prior distribution is

\[
\begin{pmatrix}
1 & \rho & \cdots & \rho \\
\vdots & \ddots & \vdots & \vdots \\
\rho & \cdots & 1
\end{pmatrix},
\]

where \( \rho \leq 0 \), then the prior p.d.f. is \( S-\text{MRR}_2 \) (Block, Savits, and
Example 2 - Multivariate Bernoulli Trials

Let $y_1, y_2, \ldots$ be a sequence of i.i.d. $k$-dimensional vectors with common multinomial distribution with parameters $n = 1$ and $p = (p_1, \ldots, p_k)$, where $\frac{y_i}{n} = 1$. That is, $y_1, y_2, \ldots$ are ind. and $y_i \sim M(1, p)$ $\forall i = 1, 2, \ldots$. Note that for any finite sequence $y_1, \ldots, y_m$, the vector $X = \sum_{j=1}^{m} y_j = (x_1, \ldots, x_k)$ is a sufficient statistic since the probability mass function of $y_1, \ldots, y_m$ is

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

where $I(p)$ is the indicator function of $\frac{y_i}{n} = 1$. Clearly, $f(x|p) = h(x)L$.

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 $i = 1, \ldots, k$, $E\phi(p_i|x)$ is increasing in $x_i$ for every increasing real function $\phi$.

Suppose that a Dirichlet distribution with parameters each no smaller than unity is chosen to represent our prior opinion about $p$. With this choice, the prior p.d.f. for $p$ is S-MRR$_2$ (see Karlin and Rinott [1980b] or Block, Savits, and Shaked [1980]). Thus, since $p_i x_i$ is PF$_2$ in $p_i$ ($i = 1, \ldots, k$), Theorem 5 applies and by Remark 4, for $i = 1, \ldots, k$, $E\phi(p_i|x)$ is decreasing in $x_j$ whenever $j \neq i$ and $\phi$ increasing. $\square$
Remark

5 - Note that Example 2 includes both the multinomial and the negative multinomial models.

Example 3 - Multivariate Hypergeometric Distribution

The probability mass function in this case may be expressed as

\[ f(x|\theta) = h(x) \prod_{i=1}^{k} \binom{\theta_i}{x_i} I(\theta), \]

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

Suppose that a Dirichlet-Multinomial distribution [denoted by \( DM(n;\alpha) \)] with shape parameters, \( \alpha_i \)'s, each no smaller than unity is chosen to represent our opinion about \( \theta \). In addition, note that \( \binom{\theta_i}{x_i} \) is PF_2 in \( \theta_i \) for every fixed \( x_i \). Thus, Theorem 5 and Remark 4 apply. To conclude this example, recall that the posterior distribution of \( \theta - x \) (the unsampled population) is \( DM(N; \alpha_1 + x_1, \ldots, \alpha_k + x_k) \) whose component means are given by \( \frac{\binom{\alpha_i + x_i}{x_i}}{\sum_{j=1}^{k} \binom{\alpha_j + x_j}{x_j}} \) for \( i = 1, \ldots, k \). Thus, for every \( i = 1, \ldots, k \), \( \text{E}\{\theta_i - x_i|x\} \) 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 $(\theta_1, \theta_2)$. 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 $(r_1, r_2, \alpha)$ 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 $E(\phi(\theta)|x_1, x_2)$ increases in $x_1$ and decreases in $x_2$ for every increasing function $\phi$. The analogous result for $\theta_2$ is obvious. □

Example 5 - The Exponential Family

This application is of a general nature. Consider that the distribution of $x$ belongs to the exponential family. With a proper
reparametrization we may consider $\Theta$ such that

$$f(x|\Theta) = h(x) \exp(\Theta^T x) c(\Theta).$$

Since $e^{\Theta^T x}$ is TP$_2$, Theorem 4 applies and for any increasing $\phi$, $E(\phi(\Theta^T x)|x)$ 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.

References


**Security Classification of This Page**

<table>
<thead>
<tr>
<th>1. Report Number</th>
<th>2. Govt Accession No.</th>
<th>3. Recipient's Catalog Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSU M-545R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFOSR 78-107R</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Title (and subtitle)</th>
<th>5. Type of Report &amp; Period Covered</th>
<th>6. Performing Org. Report Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Influence of the Sample on the Posterior Distribution</td>
<td>technical November, 1981</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. Author(s)</th>
<th>8. Contract or Grant Number(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salwa Fahmy, Carlos A. de B. Pereira, Frank Proschan and Moshe Shaked</td>
<td>F 49620-82-K-0007</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9. Performing Organization Name and Address</th>
<th>10. Program Element, Project, Task Area &amp; Work Unit Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department of Statistics</td>
<td></td>
</tr>
<tr>
<td>Florida State University</td>
<td></td>
</tr>
<tr>
<td>Tallahassee, FL. 32306</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. Controlling Office Name and Address</th>
<th>12. Report Date</th>
<th>13. Number of Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boiling Air Force Base</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington, D.C. 20332</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>14. Monitoring Agency Name &amp; Address (if different from Controlling Office)</th>
<th>15. Security Class. (of this report)</th>
<th>15a. Declassification/Downgrading Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>16. Distribution Statement (of this report)</th>
<th>17. Distribution Statement (of the abstract entered in Block 20, if different from report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution Unlimited</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>18. Supplementary Notes</th>
<th>19. Key Words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monotone Likelihood Ratio Property, Multivariate Totally Positive Functions, Multivariate Reverse Rule Functions, Negative Dependence, Bayesian Operation (Prior to Posterior).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>20. Abstract (Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>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).</td>
</tr>
</tbody>
</table>