SOME RESULTS ENVOLVING THE CONCEPTS OF MOMENT GENERATING FUNCTION AND AFFINITY BETWEEN DISTRIBUTION FUNCTIONS. EXTENSION FOR $k-$DIMENSIONAL NORMAL DISTRIBUTION FUNCTIONS

A. DORIVAL CAMPOS
Universidade de São Paulo

We present a function $\rho(F_1, F_2, t)$ which contains Matusita’s affinity and express the «affinity» between moment generating functions. An interesting result is expressed through decomposition of this «affinity» $\rho(F_1, F_2, t)$ when the functions considered are $k-$dimensional normal distributions. The same decomposition remains true for others families of distribution functions. Generalizations of these results are also presented.

Keywords: Affinity, moment generating functions, distance, inner product, multivariate normal distributions, probability density functions, absolutely convergent.

1. INTRODUCTION AND PRELIMINARIES

Let $F_1$ and $F_2$ be two distribution functions defined on $\mathbb{R}$ and let us denote by $f_i(x)$ the probability density function of $F_i$ with respect to a measure $m$ in $\mathbb{R}$, for $i = 1, 2$.

We find in the literature several forms of defining distance between distributions of the same class. Matusita (1955) by making use of the distance function denoted by $d(F_1, F_2)$ and expressed by

$$d(F_1, F_2) = \left\{ (f_1^{1/2}(x) - f_2^{1/2}(x), f_1^{1/2}(x) - f_2^{1/2}(x)) \right\}^{1/2},$$

introduced in the statistical literature the concept of affinity between the distributions $F_1$ and $F_2$ denoted by $\rho_2(F_1, F_2)$ and defined by

$$\rho_2(F_1, F_2) = \left( f_1^{1/2}(x), f_2^{1/2}(x) \right)$$

which is related to $d(F_1, F_2)$ through the expression

$$d^2(F_1, F_2) = 2 \left\{ 1 - \rho_2(F_1, F_2) \right\}$$

where $(f, g)$ denotes the inner product of $f(x)$ and $g(x)$ defined by:

$$(f, g) = \int_{\mathbb{R}} f(x) g(x) \, dm$$


Concrete expressions for the affinity between two multivariate normal distributions were established by Matusita (1966). As a following step, Matusita (1967) extended the notion of affinity to cover the case where there are $r$ distributions involved and established concrete expressions when the $r$ distributions are $k-$dimensional normal.

Our work is characterized by the introduction of the concept of a function, denoted by $P(t)$, functionally expressed through the moment generating functions relative to the distributions considered and another expression denoted by $\rho(F_1, F_2, t)$ that contains as a particular case the affinity between the distribution functions $F_1$ and $F_2$, or in other words, express the «affinity» between the moment generating functions relative to $F_1$ and $F_2$. We also present a result that express the decomposition of $\rho(F_1, F_2, t)$ in a product of two factors identified as the affinity and the moment generating function when $F_1$ and $F_2$ are $k-$dimensional normal distributions. This result is extended to cover the case where there are $r$ $k-$dimensional normal distributions.
In this same way, the concept of a more general function \( D_r(s_1, \ldots, s_r; t) \) is introduced, and the results obtained through it contains those developed in this paper as those ones established by Matusita (1966, 1967a) and Campos (1978).

2. RESULTS

Definition 1. Let \( F_1 \) and \( F_2 \) be two distribution functions belonging to the same class and let \( f_1(x) \) and \( f_2(x) \) their respective probability density functions with respect to a measure \( m \) defined on \( \mathbb{R} \). Let us suppose that there is a scalar \( t, -h \leq t \leq h \) \((h > 0)\) such that the integral below, defined through the inner product, is absolutely convergent.

Now, we define:

\[
P(t) = \exp \{tx/2\} \left\{ f_1^{1/2}(x) - f_2^{1/2}(x) \right\}, \exp \{tx/2\} \left\{ f_1^{1/2}(x) - f_2^{1/2}(x) \right\}
\]

where \((f, g)\) denotes the inner product of \( f(x) \) and \( g(x) \), defined by:

\[
(f, g) = \int_{\mathbb{R}} f(x) g(x) \, dm
\]

From (2.1), we obtain:

\[
P(t) = M_1(t) + M_2(t) - 2\rho(F_1, F_2, t)
\]

where:

\( M_i(t) \) represent the moment generating function for the distribution \( F_i \) whose probability density function is \( f_i(x) \), \( i = 1, 2 \);

and

\[
(2.2) \quad \rho(F_1, F_2, t) = \left( \exp \{tx/2\} f_1^{1/2}(x), \exp \{tx/2\} f_2^{1/2}(x) \right)
\]

From (2.1) and (2.2) we verify that:

i) \( P(0) = d^2(F_1, F_2) \),

ii) \( F_1 = F_2 \) implies \( P(t) = 0 \) for all \( -h \leq t \leq h \)

iii) \( \rho(F_1, F_2, 0) = \rho_2(F_1, F_2) \)

iv) \( F_1 = F_2 = F \) implies \( \rho(F, t) = M(t) \)

where:

\[
d^2(F_1, F_2) = \left( f_1^{1/2}(x) - f_2^{1/2}(x), f_1^{1/2}(x) - f_2^{1/2}(x) \right)
\]

227
and $\rho_2(F_1, F_2)$ is the affinity between the distributions $F_1$ and $F_2$ as defined by Matu-

ta (1966).

**Theorema 1.** Let $F_1$ and $F_2$ be $k$–dimensional nonsingular normal distributions, whose

probability density functions are given by:

$$(2\pi)^{-k/2} |A|^{1/2} \exp \left\{ -1/2(A^{-1}(x - a), x - a) \right\}$$

and

$$(2\pi)^{-k/2} |B|^{-1/2} \exp \left\{ -1/2(B^{-1}(x - b), x - b) \right\},$$

respectively, where:

$x$ is a $k$–dimensional (column vector);

$A$ and $B$ are covariance matrices of degree $K$ and

$a, b$ are $k$–dimensional mean (column) vectors.

In these conditions, we have:

$$\rho(F_1, F_2, t) = \rho_2(F_1, F_2) \cdot M_C(t)$$

where:

$t$ is $k$–dimensional (column) vector and

$M_C(t)$ is the moment generating function of a $k$–dimensional normal distribution with

mean vector $C^{-1}(A^{-1}a + B^{-1}b)$ and covariance matrix $2C^{-1}$, given by

$$M_C(t) = \exp \left\{ (C^{-1}(A^{-1}a + B^{-1}b), t) + (C^{-1}t, t) \right\}$$

with $C = A^{-1} + B^{-1}$.

**Proof:** From (2.2), we have:

$$\rho(F_1, F_2, t) = (2\pi)^{-k/2} |AB|^{-1/4} \int_{\mathbb{R}^k} \exp \left\{ -1/4 Q \right\} dx_1, \ldots, dx_k$$

where:

$$Q = (A^{-1}(x - a), x - a) + (B^{-1}(x - b), x - b) - 4(x, t)$$

228
By working with this algebraic sum of inner products, we obtain:

\[(2.5) \quad Q = ((A^{-1} + B^{-1}) \bar{x}, \bar{x}) - 2(A^{-1}a + B^{-1}b + 2t, \bar{x}) + (A^{-1}a, a) + (B^{-1}b, b)\]

If we define the transformation:

\[y = C^{1/2} \bar{x}\]

with \(C = A^{-1} + B^{-1}\) and Jacobian equal to mod \(|C|^{-1/2}\), (2.5) may be written as follows:

\[Q = (y, y) - 2(C^{-1/2}(A^{-1}a + B^{-1}b + 2t), y) + (A^{-1}a, a) + (B^{-1}b, b)\]

That is,

\[(2.6) \quad Q = Q_1 - (C^{-1}(A^{-1}a + B^{-1}b + 2t), A^{-1}a + B^{-1}b + 2t) + (A^{-1}a, a) + (B^{-1}b, b)\]

where:

\[Q_1 = (y - C^{-1/2}(A^{-1}a + B^{-1}b + 2t), y - C^{-1/2}(A^{-1}a + B^{-1}b + 2t))\]

We have also:

\[(A^{-1}a, a) = (A^{-1}a, C^{-1}A^{-1}a) + (A^{-1}a, C^{-1}B^{-1}a)\]

and

\[(B^{-1}b, b) = (B^{-1}b, C^{-1}A^{-1}b) + (B^{-1}b, C^{-1}B^{-1}b)\]

By using these results in (2.6), we obtain, after same algebraic manipulations:

\[(2.7) \quad Q = Q_1 - (C^{-1}B^{-1}b, A^{-1}a) - (C^{-1}A^{-1}a, B^{-1}b) + (A^{-1}a, C^{-1}B^{-1}a) + (B^{-1}b, C^{-1}A^{-1}b) + (C^{-1}(A^{-1}a + B^{-1}b + 2t), 2t) - (2C^{-1}t, A^{-1}a + B^{-1}b)\]

From (2.7), it follows that:

\[Q = Q_1 + ((B C A)^{-1}b, b - a) - ((A C B)^{-1}a, b - a) - 4(C^{-1}(A^{-1}a + B^{-1}b), t) - 4(C^{-1}t, t)\]
Since $C = A^{-1} + B^{-1}$, we have:

$$A C B = B C A = A + B$$

Or,

(2.8) \[ Q = Q_1 + ((A + B)^{-1}(b-a), b-a) - 4(C^{-1}(A^{-1}a + B^{-1}b), t) - 4(C^{-1}t, t) \]

Using (2.8) in (2.4), we obtain:

(2.9) \[
\rho(F_1, F_2) = (2\pi)^{-k/2} |AB|^{-1/2} \exp \left\{-\frac{1}{4}((A + B)^{-1}(b - a), b - a)\right\} \cdot \int_{\mathbb{R}^k} \exp \left\{ -\frac{1}{4} Q_1 \right\} \cdot \int_{\mathbb{R}^k} \exp \left\{ -\frac{1}{4} Q_1 \right\} d y_1, \ldots, d y_k
\]

We easily verify that:

$$\int_{\mathbb{R}^k} \exp \left\{ -\frac{1}{4} Q_1 \right\} d y_1, \ldots, d y_k = 2^{k/2} (2\pi)^{k/2}$$

Or,

(2.10) \[
\rho(F_1, F_2) = |AB|^{-1/2} \left| \frac{1}{2}(A + B) \right|^{-1/2} \exp \left\{-\frac{1}{4}((A + B)^{-1}(b - a), b - a)\right\} \cdot \exp \left\{ (C^{-1}(A^{-1}a + B^{-1}b), t) + (C^{-1}t, t) \right\}
\]

since

$$|C|^{-1/2} = |A|^{1/2} |A + B|^{-1/2} |B|^{1/2}$$

It follows from theorem demonstrated by Matusita and (2.3) that (2.10) may be written as:

$$\rho(F_1, F_2, t) = \rho_2(F_1, F_2) M_0(t)$$

**Corollary 1.** When $A = B$ it follows that

$$\rho(F_1, F_2, t) = \exp \left\{-1/8(A^{-1}(b - a), b - a)\right\} M_0(t)$$

230
where:

$$M_\alpha(t) = \exp \left\{ \frac{1}{2}(a + b), t \right\} + \frac{1}{2}(A t, t) \right\}$$

**Corollary 2.** If $a = b$ it follows that:

$$\rho(F_1, F_2, t) = \left| A B \right|^{1/4} \left| \frac{1}{2} \left( A + B \right) \right|^{-1/2} \exp \left\{ (a, t) + \left( C^{-1}, t \right) \right\}$$

**Corollary 3.** For $F_1 = F_2 = F$, we have:

$$\rho(F, t) = \exp \left\{ (a, t) + \frac{1}{2}(A t, t) \right\} = M_\alpha(t)$$

where $M_\alpha(t)$ is the moment generating function of $F$.

The conclusion (result) of theorem 1 is naturally generalized for $r$ $k$-dimensional normal distributions. To accomplish this objective, we first generalize the concept of $\rho(F_1, F_2, t)$, by considering $r$ distributions $F_1, \ldots, F_r$, defined over the same space $\mathbb{R}$, with probability density functions $f_1(x), \ldots, f_r(x)$ with respect to a measure on $\mathbb{R}$, and let us suppose that the integral below be absolutely convergent. Then we define:

**Definition 2**

$$\rho(F_1, \ldots, F_r, t) = \int_\mathbb{R} \left\{ \prod_{j=1}^r \exp(tx) f_j(x) \right\}^{1/r} \ d m$$

If $F_1, \ldots, F_r$ denote $r$ $k$-dimensional non singular normal distributions whose probability density functions are given for $j = 1, \ldots, r$ by

$$(2\pi)^{-k/2} |A_j|^{-1/2} \exp \left\{ -1/2(A_j^{-1}(x - a_j), x - a_j) \right\}$$

where $A_j$ is the covariance matrix and $a_j$ the mean vector of $F_j$, respectively, we have the following result whose proof we omit:

**Theorem 2**

$$\rho(F_1, \ldots, F_r, t) = \rho_r(F_1, \ldots, F_r) M_\alpha(t)$$
where

$$\rho_r(F_1, \ldots, F_r) = \left\{ \frac{1}{r} \sum_{j=1}^{r} |A_j^{-1}| \right\}^{1/2} \exp \left\{ -1/2 r \sum_{j=2}^{r} ((A_j D A_j)^{-1} a_j - a_i - \ldots \right\} - \sum_{i=1}^{r-1} ((A_i D A_i)^{-1} a_i - a_i - \ldots \right\}$$

is the affinity between $r$ $k$-dimensional normal (non-singular) distributions obtained by Matusita (1967a) and expressed in another form by Campos (1978); and

$$M_{\alpha}(t) = \exp \left\{ \left( D^{-1} \left( \sum_{j=1}^{r} A_j^{-1} a_j \right), t \right) + \frac{1}{2} \left( r D^{-1} t, t \right) \right\}$$

is the moment generating function of a $k$-dimensional normal distribution with mean $D^{-1} \left( \sum_{j=1}^{r} A_j^{-1} a_j \right)$ and covariance matrix $r D^{-1}$ with $D = \sum_{j=1}^{r} A_j^{-1}$ and $t$ a $k$-dimensional (column) vector.

With the objective of generalizing these results, as those obtained by Matusita (1966, 1967a) or Campos (1978) we introduce the following definition.

**Definition 3.** Let $F_1, \ldots, F_r$ be multivariate distributions defined on the same space $\mathbb{R}$ and let $f_1(x), \ldots, f_r(x)$ be their respective probability density functions. Let us suppose that there are $r$ scalars $s_1, \ldots, s_r$ such that:

$$\sum_{j=1}^{r} s_j = 1 \quad \text{and} \quad 0 \leq s_j \leq 1 \quad \text{for} \quad j = 1, \ldots, r.$$

In these conditions, and if the integral below is absolutely convergent, we define:

$$D_r(s_1, \ldots, s_r, t) = \int_{\mathbb{R}^k} \prod_{j=1}^{r} \exp \left\{ s_j(x, t) \right\} f_j^{s_j}(x) \, dx_1 \ldots dx_k$$
where \( \tilde{t}_j \) is a \( k \)-dimensional (column) vector with components \( t_{j1}, \ldots, t_{jk}, j = 1, \ldots, r \).

If \( F_1, \ldots, F_r \) denote \( k \)-dimensional normal (non-singular) distributions defined as (2.11) we establish the following result:

**Theorem 3**

\[
D_r(s_1, \ldots, s_r; \tilde{t}) = D_r(s_1, \ldots, s_r) M_G \left( \sum_{j=1}^{r} s_j \tilde{t}_j \right)
\]

where:

\[
D_r(s_1, \ldots, s_r) = \left\{ \prod_{j=1}^{r} |A_j|^{-s_j/2} \right\} \left\{ \sum_{j=1}^{r} s_j A_j^{-1} \right\}^{-1/2}
\]

\[
\cdot \exp \left\{ -1/2 \left\{ \sum_{\substack{j=2 \atop 1 \leq i < j}}^{r} (s_i s_j (A_i C A_j)^{-1} a_j - a_i) - \sum_{\substack{i=1 \atop 1 < j \leq r}}^{r} (s_i s_j (A_i C A_j)^{-1} a_j - a_i) \right\} \right\}
\]

and \( M_G \left( \sum_{j=1}^{r} s_j \tilde{t}_j \right) \) is the moment generating function for a \( k \)-dimensional normal distribution with mean vector \( C^{-1} \left( \sum_{j=1}^{r} s_j A_j^{-1} a_j \right) \) and covariance matrix \( C^{-1} \), expressed by:

\[
M_G \left( \sum_{j=1}^{r} s_j \tilde{t}_j \right) = \exp \left\{ \left( C^{-1} \left( \sum_{j=1}^{r} s_j A_j^{-1} a_j \right), \sum_{j=1}^{r} s_j \tilde{t}_j \right) + \right. \]

\[
1/2 C^{-1} \left( \sum_{j=1}^{r} s_j \tilde{t}_j, \sum_{j=1}^{r} s_j \tilde{t}_j \right) \right\}
\]

(2.14)
with $C = \sum_{j=1}^{r} s_j A_j^{-1}$.

**Proof:** By the definition 3, we have:

\begin{equation}
D_r(s_1, \ldots, s_r; t) = (2\pi)^{-k/2} \prod_{j=1}^{r} |A_j|^{-s_j/2} \int_{\mathbb{R}^k} \exp \left\{ -\frac{1}{2} Q \right\} \, dx_1 \ldots dx_k
\end{equation}

where

$$Q = \sum_{j=1}^{r} s_j \left( A_j^{-1} (x - a_j), x - a_j \right) - 2 \sum_{j=1}^{r} s_j (x, t_j)$$

That is,

\begin{equation}
Q = (C \bar{x}, \bar{x}) - 2(b, \bar{x}) - 2(t, \bar{x}) + \sum_{j=1}^{r} \left( s_j A_j^{-1} a_j, a_j \right)
\end{equation}

with

$$b = \sum_{j=1}^{r} s_j A_j^{-1} a_j$$

and

$$t = \sum_{j=1}^{r} s_j t_j$$

After the transformation

$$y = C^{1/2} x$$

whose Jacobian is $|C|^{-1/2}$ and same intermediate steps, (2.16) may be written

\begin{equation}
Q = \left( y - C^{-1/2}(b + t), y - C^{-1/2}(b + t) \right) - \left( C^{-1} b, t \right) -
\end{equation}

$$\left( C^{-1} t, b \right) - \left( C^{-1} t, t \right) + \sum_{j=1}^{r} \left( s_j A_j^{-1} a_j, a_j \right) - \left( C^{-1} b, b \right)$$

234
One may also prove that:

\[
\sum_{j=1}^{r} \left( s_j A_j^{-1} a_j, a_j \right) - \left( C^{-1} b, b \right) = \sum_{i \leq i < j}^{r} \left( s_i s_j (A_i C A_j^{-1} a_i, a_j - a_i) - \right. \\
- \left. \sum_{i = 1}^{r-1} \left( s_i s_j (A_i C A_j^{-1} a_i, a_j - a_i) \right) \right)
\]

On applying this result, (2.17) is expressed as:

(2.18) \[ Q = Q_3 + Q_1 - Q_2 - 2 \left( C^{-1} b, t \right) - \left( C^{-1} t, t \right) \]

where:

\[
Q_3 = \left( y - C^{-1/2}(b + t), y - C^{-1/2}(b + t) \right)
\]

\[
Q_1 = \sum_{i \leq i < j}^{r} \left( s_i s_j (A_i C A_j^{-1} a_i, a_j - a_i) \right) \quad \text{and}
\]

\[
Q_2 = \sum_{i = 1}^{r-1} \left( s_i s_j (A_i C A_j^{-1} a_i, a_j - a_i) \right)
\]

By substitution of (2.18) in (2.15), we obtain:

(2.19) \[ D_r(s_1, \ldots, s_r; t) = (2\pi)^{-k/2} \prod_{j=1}^{r} |A_j|^{-s_j/2}, \]

\[
\cdot \exp\left\{ -\frac{1}{2} (Q_1 - Q_2) \right\} \exp\left\{ (C^{-1} b, t) + \left( \frac{1}{2} C^{-1} t, t \right) \right\}.
\]

\[
\cdot \int_{\mathbb{R}^k} \exp\left\{ -\frac{1}{2} Q_3 \right\} |C|^{-1/2} dy_1 \ldots dy_k
\]

The integral of (2.19) after the transformation

\[ z = y - C^{-1/2}(b + t) \]

becomes equal

\[ |C|^{-1/2}(2\pi)^{k/2} \]

235
By using this result in (2.19), we obtain, in accord with (2.13) and (2.14) the result that we have established through theorem 3, that is,

$$D_r(s_1, \ldots, s_r; t) = D_r(s_1, \ldots, s_r) \cdot M_G \left( \sum_{j=1}^{r} s_j t_j \right)$$

This result is a generalization in the sense of summarize the results established through theorems 1 and 2 as those obtained by Matusita (1966, 1967a) or Campos (1978).

REFERENCES


