NONPARAMETRIC AND GAUSSIAN BIVARIATE TRANSVARIATION THEORY: ITS APPLICATION TO ECONOMICS

Camilo Dagum

Econometric Research Program
Research Memorandum No. 99
June 1968

The research described in this paper was supported in part by the National Science Foundation NSF GS 1840 and in part by Office of Naval Research N00014-67 A-0151-0007 Task No. 047-086.

Princeton University
Econometric Research Program
207 Dickinson Hall
Princeton, New Jersey

NONPARAMETRIC AND GAUSSIAN BIVARIATE TRANSVARIATION THEORY: ITS APPLICATION TO ECONOMICS*

Camilo Dagum
Econometric Research Program
Princeton University

1. Introduction

The theory of transvariation, formulated by Corrado Gini, was first introduced into mathematical statistics in Gini's "Il Concetto di Transvariazione e le sue Prime Applicazioni", published in 1916.

Gini's formulation of this theory grew out of his attempt to develop a proper statistical method for the solution of the following problem: given the sign of the difference between the means or medians of two populations, find a probabilistic statement regarding the sign or the intensity of the difference (the common area, etc.) between two random observations corresponding to each of the populations. The probabilistic statement of each characteristic defines a parameter of transvariation, i.e., the probability of transvariation, the r^{th} intensity of transvariation ($r = 1,2,\ldots$), the space of transvariation and the discriminative value.

For a given sample, the estimation of the parameters of transvariation introduces new random variables. These new variables then require of us an estimation of their corresponding variances and an analysis of their statistical properties.

Under the hypothesis of dependent and independent random variables we can, in general, estimate parameters of transvariation for $i(i=1,\ldots,s)$ distributions

*The subject of this paper was discussed in the Econometric Research Program Seminar under the direction of Professor Oskar Morgenstern and in the Department of Statistics Seminar under the direction of Professor John W. Tukey. The author is very much indebted to them for their interest in the subject and stimulating comments. The author expresses his appreciation to Mr. Peter Kaminsky, a Princeton University student, for his diligence and collaboration in preparing the final English version.

of k (k=1,...,n) dimensions. Working with marginal distributions, we are able, in this manner, to estimate a complete set of parameters for s distribution of n dimensions. For example, when i = 2, we can estimate the k^{th} marginal parameter of transvariation (k=1,2,...,n-1) and the total parameters of transvariation (of the order n). In sum, then, for each parameter of transvariation between two distributions, the number of marginal and total estimations is

$$\sum_{k=1}^{n} \binom{n}{k} = 2^n - 1 \quad .$$

Having considered all of the 2^n-1 combinations between two distributions of n dimensions, we are afforded a broad understanding of the phenomena under analysis.

Parameters of transvariation are a function of the means, variances, and probability distribution function of the populations under analysis.

2. Theoretical and Applied Contributions

2.1 Historical references

Probabilistic statements of the probability and intensity of transvariation were first introduced by Gini [Ref. 21] in 1916. But the idea of transvariation, without a quantitative statement and formal development, had been advanced by earlier scientists. The first known statement on the subject was included in Jacob Bernoulli's classical contribution to the theory of probability. [Refs. 1 and 29]. Bernoulli considered the probability that an elder brother died after his younger one. Further, Quetelet [Ref. 39], Lexis [Ref. 33] and Johansen [Refs. 32 and 41)] pointed out the interest of analyzing the probability that two observations have a difference of sign opposite to the means of the populations under observation industrial sector was, for the purposes of this study, disaggregated into 83 branches. As a consequence, for each year of the study, we were confronted with a four-dimensional variate that allowed us to apply the transvariation theory between two or more distributions and between two or more variates.

3. Notations

We will use the following notation and breviations.

The triple $(\Omega, \mathcal{F}, P)$ is called a probability space. Ω denotes a sample space, wherein a sample point (elementary event) $\omega \in \Omega$, and a set E of sample points is a subset in the sample space, such that $E \subset \Omega$. $\mathcal{F}$ is a σ-field of sets in Ω and P is a probability measure for sets in a σ-field $\mathcal{F}$.

By Λ, we denote a random experiment or random observation (r.e.). In each realization of an r.e. we are interested in the probabilities that certain events will occur. Thus we are interested in the measurement or observation of numerical quantities associated with each r.e. For example, one may be interested in the prices of several commodities in a given market for several periods of time, or the quantities produced for each period in a given economic unit or space. Such numerical quantities are sample realizations of a random variable (r.v.) denoted by a small greek letter $\xi, \eta, \ldots$. An r.v. $\xi = \xi(\omega)$ is a real-valued function defined for all sample points ω of a basic probability space $(\Omega, \mathcal{F}, P)$. The set of values that $\xi(\omega)$ can take for all $\omega \in \Omega$, defines Ω', that is the sample space (range space) of $\xi(\omega)$. The inverse function ξ^{-1} takes every interval $\xi \leq x$ into a measureable ω set. Therefore, if $\mathcal{F}'$ is a σ-field of sets in Ω', the r.v. $\xi(\omega)$ maps the sample points ω in Ω into sample points ξ in Ω' such that, for every Borel set in $E' \in \mathcal{F}'$, there is an event $E \in \mathcal{F}$, $E \subset \Omega$, for which $\xi(\omega) \in E'$. Hence, the inverse image of the set E' is

$$E = \xi^{-1}(E')$$

and

$$P'(E') = P(\xi^{-1}(E')) .$$

Therefore the triple $(\Omega', \mathcal{F}', P')$ defines a new probability space induced from a basic probability space $(\Omega, \mathcal{F}, P)$ by an r.v. $\xi(\omega)$.

The concluding lower case latin letters $t, u, v, \ldots$ are used to symbolize sample observations of random variables corresponding to a realization of an r.e. The cummulative distribution function (c.d.f.) $F(x)$, where

$$F(x) = P(\xi \leq x)$$

specifies the probabilistic behavior of an r.v. ξ. Its corresponding probability density function (p.d.f.) is $f(x)$ for the continuous case, and $p_i = P(\xi = x_i)$ for the discrete case.

In general, ξ is a k-dimensional vector random variable

$$\xi = (\xi_1, \ldots, \xi_k), \quad k=1,2,\ldots,n .$$

Therefore

$$F(x) = F(x_1, \ldots, x_k) = P(\xi_1 \leq x_1, \ldots, \xi_k \leq x_k)$$

and for the sample space Ω' of ξ

$$\Omega' \subset R_k$$

where R_k is a k-dimensional Euclidean space. If nothing is specified, our sample space will be

$$\Omega' = R_k, \quad k=1,2,\ldots,n .$$

4. Bivariate Transvariation Theory

For a given r.e. Λ we observe a bivariate r.v.

(4.1) $$\xi = (\xi_1, \xi_2)$$

where

(4.2) $$x = (x_1, x_2)$$

are sample observations of ξ and

(4.3) $$F(x) = P(\xi \leq x \, , \; \xi_2 \leq x_2)$$

is its c.d.f.

We introduce an arbitrary real parameter in R_2

(4.4) $$\lambda_0 = (\lambda_1, \lambda_2) \in R_2$$

and we assume, without loss of generality,

(4.5) $$\lambda_1 > \lambda_2 \quad \text{and} \quad \lambda = \lambda_2 - \lambda_1 < 0.$$

We now introduce the following linear transformation of the r.v. (4.1)

(4.6) $$\tau = \xi_2 - \xi_1$$

which takes the ample values

(4.7) $$t = x_2 - x_1 \; .$$

Definition of transvariation: We have, by definition, a transvariation, in a given r.e. Λ , between the r.v. ξ_1 and ξ_2 and the arbitrary parameters λ_1 and λ_2, when the difference $x_2 - x_1 = t$ is of opposite sign to the difference $\lambda_2 - \lambda_1 = \lambda$.

Measure of a transvariation: The absolute value of the power

(4.8) $$t^r = (x_1 - x_2)^r$$

defines the r^{th} intensity of a sample transvariation. When $r = 1$ we have an intensity of the order 1 and a measure of this transvariation.

Convention 1: Given a sample of independent replications of an r.e.Λ we then compute as transvariations one half of the number of null differences, i.e., one half of the times we observe $x_1 - x_2 = 0$.

The definition of transvariation is illustrated in Fig. 1. The r.v. $\xi = (\xi_1, \xi_2)$ takes values $x = (x_1, x_2) \in R_2$, where R_2 is the Euclidean plane. Any sample realization of the random experiment Λ such that we have

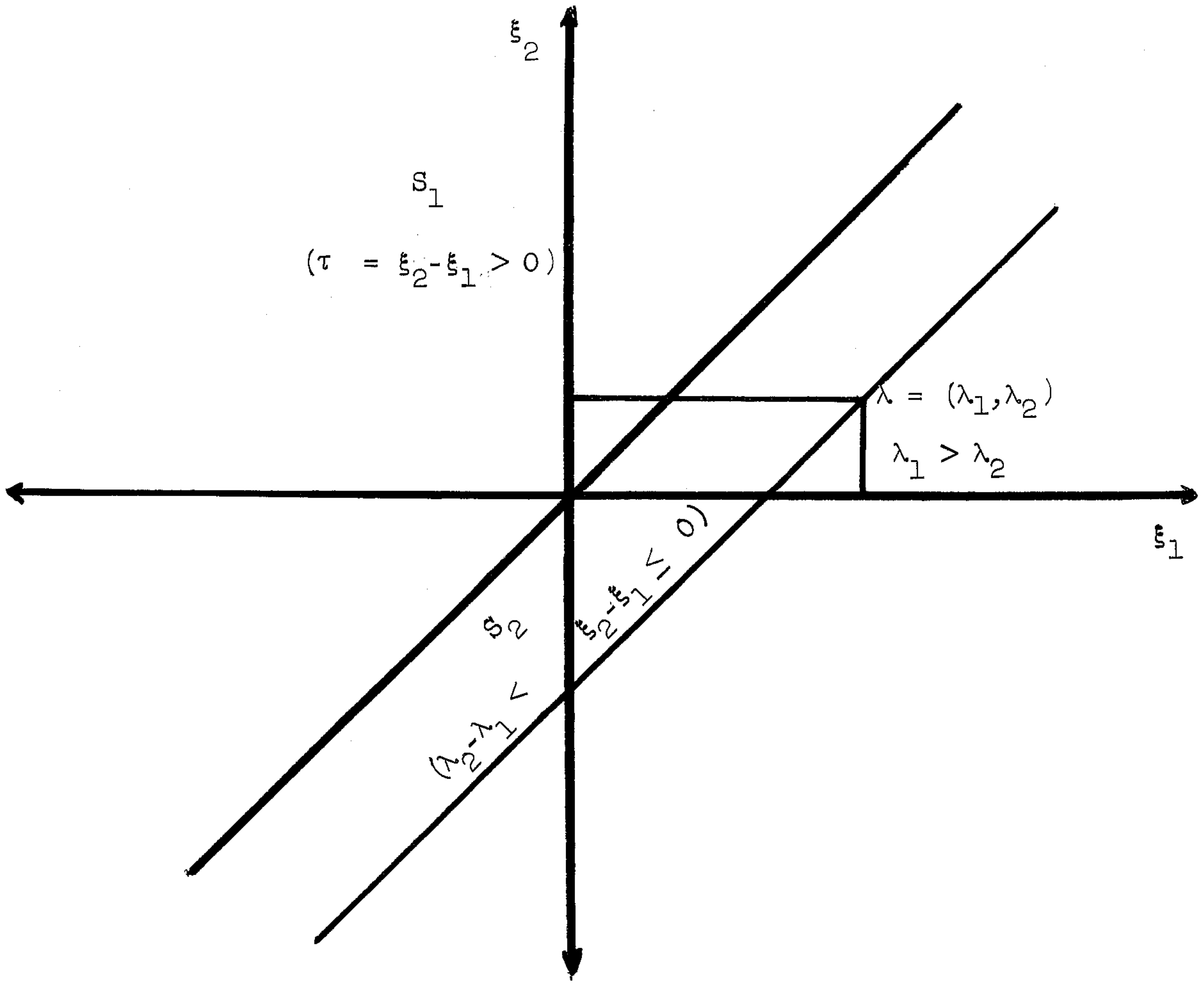

Figure 1

an observation

$$x = (x_1, x_2) \epsilon\, S_1 \tag{4.9}$$

defines a transvariation, given that $\lambda_1 > \lambda_2$. I.e., any elementary event $\omega \in \Omega$ such that $\xi(\omega) \in S_1$ defines a transvariation.

Definition of transvariability: The transvariability between an r.v. $\xi = (\xi_1, \xi_2)$, or its linear transformation $\tau = \xi_2 - \xi_1$, and an arbitrary parameter $\lambda_0 = (\lambda_1, \lambda_2)$, where $\lambda = \lambda_2 - \lambda_1$, is the probability that the r.v. $\tau = \xi_2 - \xi_1$ takes a value $t = x_2 - x_1$ of sign opposite to the parameter $\lambda = \lambda_2 - \lambda_1$.

In figure 1, the probability of an event that fulfills the requirement (4.9) is, by definition, the transvariability between the r.v. (4.1), or it linear transformation (4.6), and the arbitrary parameters (4.5). In symbolic form

$$\begin{aligned} p &= P(\xi_2 > \xi_1 \mid \lambda_2 < \lambda_1) = P(\tau = \xi_2 - \xi_1 > 0 \mid \lambda = \lambda_2 - \lambda_1 < 0) \\ &= \int_{S_1} d\,P = \int_0^\infty d\,G(t) = 1 - G(0) \end{aligned} \tag{4.10}$$

where $G(t)$ is the c.d.f. of the r.v. τ.

LEMMA 1: If the identity function $x_1 = x_2$ is a median straight line, then, by the definition of a median straight line, it is

$$p = \int_{S_1} d\,P = \frac{1}{2} \tag{4.11}$$

i.e., the unit mass is evenly distributed on each halfplane separated by the identity function.

We will find out if there is a maximum of transvariability. Introducing the following real variable vector of the order two

(4.12) $$h_0 = (h_1, h_2) \quad .$$

From (4.5), we then suppose

(4.13) $$h = h_2 - h_1 \leq 0$$

and we may now express the transvariability as a function of h

(4.14) $$p(h) = P(\xi_2 - \xi_1 > h_2 - h_1 | \lambda_2 - \lambda_1 < 0)$$
$$= P(\tau > h | \lambda < 0) = \int_h^\infty d\, G(t) = 1 - G(h) \, .$$

The probability function (4.14) defines a transvariability for all h such that $h > \lambda$. This probability function corresponds, in Fig. 1, to any event to the left of the straight line

(4.15) $$x_1 - h_1 = x_2 - h_2$$

provided that the point λ_0 is to the right. Such is the case because of (4.5), (4.13) and the specification $h > \lambda$. Further, the linear function (4.15) is parallel to the identity function

(4.16) $$x_1 = x_2 \quad .$$

Since (4.14) is a non-decreasing and non-negative set function with h approaching λ, then we have, for the maximum p_M of transvariability

(4.17) $$p_M = \lim_{h \to \lambda} p(h) = P(\tau > \lambda | \lambda < 0) \, .$$

Using set-builder notation, we define (for its corresponding graphical representation, see Fig. 1)

(4.18) $$S_1 = \{(\xi_1, \xi_2) \, | \xi_2 - \xi_1 > 0, \, \lambda_2 - \lambda_1 < 0\} = \{\tau | \, \tau > 0, \, \lambda < 0\}$$

(4.19) $$S_2 = \{(\xi_1, \xi_2) \, | \lambda_2 - \lambda_1 < \xi_2 - \xi_1 \leq 0\} = \{\tau | \, \lambda < \tau \leq 0\}$$

then

(4.20) $$S_1 \cup S_2 = \{\tau | \tau > \lambda, \quad \lambda < 0\} \quad .$$

Therefore, the maximum of transvariability (4.17) is

(4.21) $$p_M = \int_{\lambda}^{\infty} d\, G(t) = \int_{S_1 \cup S_2} d\, P = 1 - G(\lambda)$$

which is a function of parameter λ.

If we, as usual, denote by Ω a sample space (sure event) with elementary events $\omega \in \Omega$, then, the random vector (4.1) is a function $\xi = \xi(\omega)$ from the sample space Ω to the two-dimensional Euclidean space R_2. Therefore

(4.22) $$S_1 \cup S_2 \subset R_2$$

and

(4.23) $$p \leq p_M \leq 1 \quad .$$

$p_M = 1$ if[1] $S_1 \cup S_2 = R_2$, and $p_M = p$ iff (if and only if) one of the following two conditions is fulfilled: i) S_2 is an empty set or ii) S_2 is a set of P-measure zero. A particular case of empty S_2 occurs when $\lambda_1 = \lambda_2$, i.e., $\lambda = 0$.

For the practical utilization of the theory we need a probabilistic concept that will assume values in a fixed interval for any value of the maximum of transvariability. For our purposes, this interval is 0 to 1.

Definition of probability of transvariation: The probability of transvariation between r.v. ξ, or its linear transformation τ, and an arbitrary parameter $\lambda_0 = (\lambda_1, \lambda_2)$, associated with a given r.e. Λ, is, by definition, the ratio between the transvariability and its maximum. I.e.

(4.24) $$P = \frac{p}{p_M} \quad .$$

[1]If R_2 contains no subsets of P-measure zero, then the condition is if and only if.

Some relevant characteristics of the probability of transvariation are the following:

1) Its range is the closed unit interval; i.e.,

$$0 \leq P \leq 1$$

2) It is dimensionless (a measure of dimension zero); therefore, it is independent of the measurement units of the observations. It allows a comparative analysis of different sets of observations;

3) It is a never-increasing set function of λ and

$$\lim_{\lambda \to -\infty} P(\lambda) \doteq p$$

4) If the linear function (Lemma 1)

$$x_1 - \lambda_1 = x_2 - \lambda_2 \tag{4.25}$$

is a median straight line, we have $p_M = \frac{1}{2}$ and $P = 2p$.

<u>Convention 2:</u> If the maximum of transvariability is null and, <u>a fortiori</u>, the transvariability will also be null; we thus assign the value zero to the probability of transvariation.

<u>Definition of moment of transvariation:</u> The expectation of the intensity of the order r defines the r^{th} moment of transvariation. Hence

$$m_2 = E[(\xi_2 - \xi_1)^r \mid \xi_2 > \xi_1, \lambda_2 < \lambda_1] = E(\tau^r \mid \tau > 0, \lambda < 0) \quad . \tag{4.26}$$

The moment of transvariation is then a conditional mathematical expectation of the r^{th} power of the r.v. (4.6), relative to the hypothesis $\xi_2 > \xi_1$, when $\lambda_2 < \lambda_1$.

Following the approach used for maximum of transvariability we have:

Definition of the maximum of the r^{th} moment of transvariation:

(4.27) $$m_{r,M} = \lim_{h \to \lambda} m_r(h) = E[(\tau-\lambda)^r | \tau > \lambda, \quad \lambda < 0] .$$

LEMMA 2: The zero-order moment of transvariation and its maximum is nothing else than the transvariability and its maximum respectively.

The moment of transvariation and its maximum satisfy the following inequalities

(4.28) $$m_r \leq m_{r,m} \leq \beta_{r,\lambda}$$

where

(4.29) $$\beta_{r,\lambda} = E(|\tau - \lambda|^r)$$

i.e., the r^{th} absolute moment of the deviations from λ of the r.v.τ.

LEMMA 3: i) If λ_1 and λ_2 are, respectively, the expectation of the r.v.ξ_1 and ξ_2, then

(4.30) $$\beta_{1\lambda} = 2m_{1,M}; \quad \lambda = E(\xi_2) - E(\xi_1) .$$

ii) If λ_1 and λ_2 are the same as in i) and the bivariate probability density function is symmetric with respect to the point λ in R_2, then

(4.31) $$\beta_{r,\lambda} = 2m_{rM} , \qquad r=0,1,2,\ldots\ldots .$$

The proof of i) follows from the wellknown property of the deviation from the mean of all r.v.'s. The proof of ii) follows from the assumption of symmetry and also, this same assumption of symmetry, allows the extension of (4.31) to r=0, i.e. to the maximum of transvariability, because of the equality between the mean and the median. In that particular case, $m_{0,M} = p_M = \frac{1}{2}$.

Definition of the r^{th} intensity of transvariation: The r^{th} root of the ratio between the r^{th} moment of transvariation (4.26) and its maximum (4.27) defines the r^{th} intensity of transvariation. Then

$$(4.32) \qquad I_r = \left(\frac{m_r}{m_{r,M}} \right)^{\frac{1}{r}} \quad ; \qquad r = 1, 2, \ldots\ .$$

If Lemma 3, ii) is satisfied, it follows from (4.31)

$$(4.33) \qquad I_r = \left(\frac{2\, m_r}{\beta_{r,\lambda}} \right)^{\frac{1}{r}} \quad ; \qquad r = 1, 2, \ldots\ .$$

For $r=1$ it is not necessary that the distribution be symmetric because we may apply Lemma 3, i).

The relevant characteristics of the intensity of transvariation are similar to those of the probability of transvariation. In particular it is a dimensionless measurement and

$$0 \leq I_r \leq 1 \ .$$

Convention 3: If, as in Convention 2, $P(\tau > 0) = P(\tau > \lambda) = 0$, then we assign the value zero to the intensity of transvariation.

5. Transvariation Between Two Independent Random Variables

5.1 The hypothesis of independence allows the introduction of two more parameters of transvariation, namely: i) the area (space, in higher dimensions) and ii) the discriminative value. This hypothesis also allows a simpler mathematical development for r.v.'s of more than two dimensions, as will be seen in a forthcoming paper.

Having dealt, in the preceding section, with the linear transformation (4.6), the hypothesis of independence will not be as relevant as it will be in transvariation in higher dimensions.

If the coordinates ξ_1 and ξ_2 of the r.v. ξ in (4.1) are independent, its c.d.f. in (4.3) becomes

(5.1) $$F(x_1,x_2) = F_1(x_1)\, F_2(x_2)$$

and the c.d.f. of the r.v. τ in (4.6)

(5.2) $$\begin{aligned} G(t) &= E_{x_1}[F_2(x_1+t)] = \int_{-\infty}^{\infty} F_2(x_1+t)\, dF_1(x_1) \\ &= 1 - E_{x_2}[F_1(x_2-t)] = 1 - \int_{-\infty}^{\infty} F_1(x_2-t)\, d\, F_2(x_2) \quad . \end{aligned}$$

From the definitions given in section 4 and from (5.2), the mathematical forms for the parameter of transvariations under the hypothesis of independence follow directly. For the transvariability and its maximum we have, from (4.10), (4.21) and (5.2)

(5.3) $$\begin{aligned} p = 1 - G(0) &= 1 - \int_{-\infty}^{\infty} F_2(x_1)\, d\, F_1(x_1) \\ &= \int_{-\infty}^{\infty} F_1(x_2)\, d\, F_2(x_2) \end{aligned}$$

(5.4) $$\begin{aligned} p_M = 1 - G(\lambda) &= 1 - \int_{-\infty}^{\infty} F_2(x_1+\lambda)\, d\, F_1(x_1) \\ &= \int_{-\infty}^{\infty} F_1(x_2-\lambda)\, d\, F_2(x_2) \quad . \end{aligned}$$

For the r^{th} moment of transvariation and its maximum, from (4.26), (4.27) and (5.2), we obtain

(5.5) $$m_r = \int_0^{\infty} t^r\, d\, G(t) = \int_0^{\infty}\int_{-\infty}^{\infty} t^r\, f_2(x_1+t)\, d\, F_1(x_1)\, dt$$

(5.6) $$m_{r,M} = \int_{\lambda}^{\infty} (t-\lambda)^r\, dG(t) = \int_{\lambda}^{\infty}\int_{\infty}^{\infty} (t-\lambda)^r\, f_2(x_1+t)\, d\, F_1(x_1)\, dt \quad .$$

The probability and intensity of transvariation follow at once, from (4.24), (4.32) and from the preceding results.

5.2 Area of transvariation. Let ξ_1 and ξ_2 be two independent r.v.'s in R_1 with p.d.f. $f_1(x)$ and $f_2(x)$. Let N_1 and N_2 represent the size of the populations corresponding to each of the sample spaces of ξ_1 and ξ_2.

Let η be a new r.v. such that, its p.d.f. is

(5.7) $$f(x) = \frac{N_1 f_1(x) + N_2 f_2(x)}{N_1 + N_2}$$

and the size of the populations from which η is observed or measured following an r.e. is $N = N_1 + N_2$.

We introduce now the function $g(x)$ of the r.v. η defined as follows:

(5.8) $$g(x) = N_1 f_1(x) \quad \text{if} \quad N_1 f_1(x) \leq N_2 f_2(x)$$

(5.9) $$g(x) = N_2 f_2(x) \quad \text{if} \quad N_1 f_1(x) > N_2 f_2(x) .$$

The p.d.f. corresponding to the definitions (5.8) and (5.9) is

(5.10) $$\frac{g(x)}{N} = \frac{g(x)}{N_1 + N_2} .$$

The common frequencies (common area) of the r.v. ξ_1 and ξ_2 is, by definition, the area of transvariation, i.e.

(5.11) $$C = \frac{1}{N} \int_{-\infty}^{\infty} d\, G(x) .$$

Supposing continuity, in Fig. 2 we illustrate this concept for the p.d.f. $f_1(x)$ and $f_2(x)$.

Let S be the common area and, because there is a single point of intersection, z , (5.11) takes the following mathematical expression,

(5.12) $$C = \frac{S}{N} = \frac{N_2}{N} \int_{-\infty}^{z} f_2(x)\, dx + \frac{N_1}{N} \int_{z}^{\infty} f_1(x)\, dx .$$

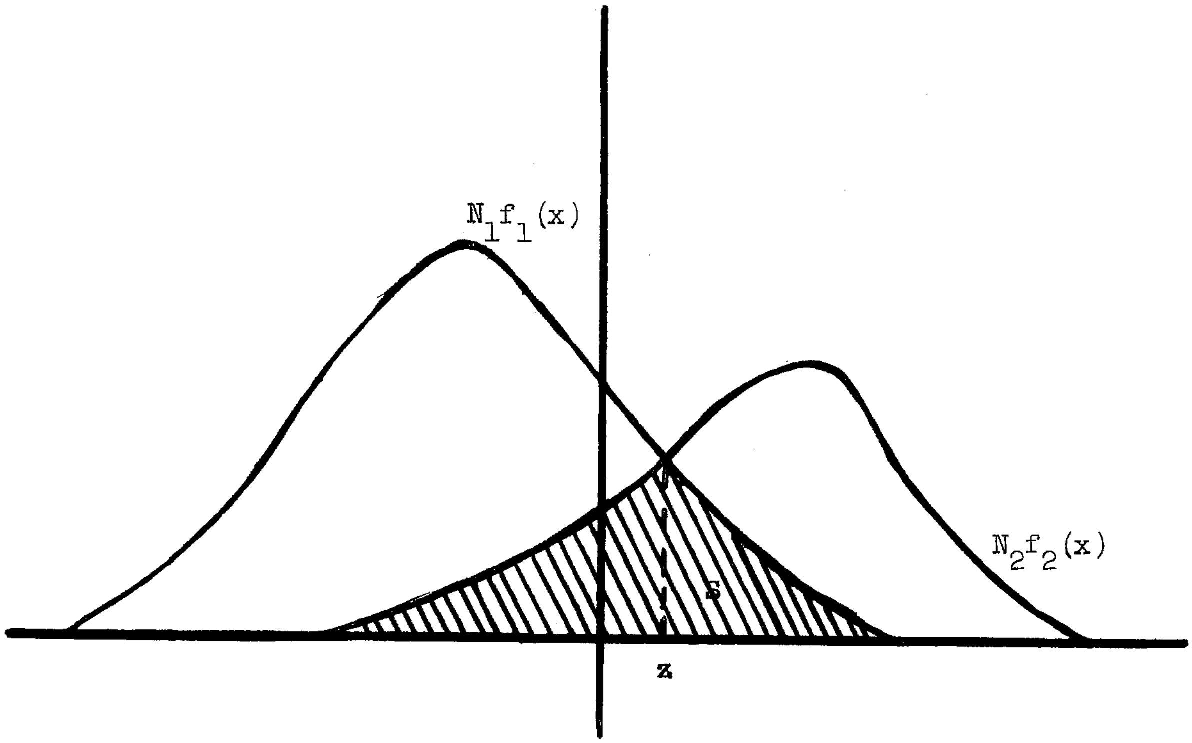

Figure 2

In Table 1 we have an example of a discrete case and, of course, it will be easy to extend the analysis to mixed (discrete-continuous) cases.

TABLE 1

x	$N_1f_1(x)$	$N_2f_2(x)$	g(x)	G(x)
0	3	2	2	2
1	7	7	7	9
2	0	12	0	9
3	4	6	4	13
4	6	0	0	13
5	4	9	4	17
6	1	9	1	18
	25	45	18	

In Figure 3 we have the cummulative frequency distribution of $g(x)$. Applying (5.11), we have $C = \frac{1}{60} G(6) = \frac{18}{60} = 0.30$.

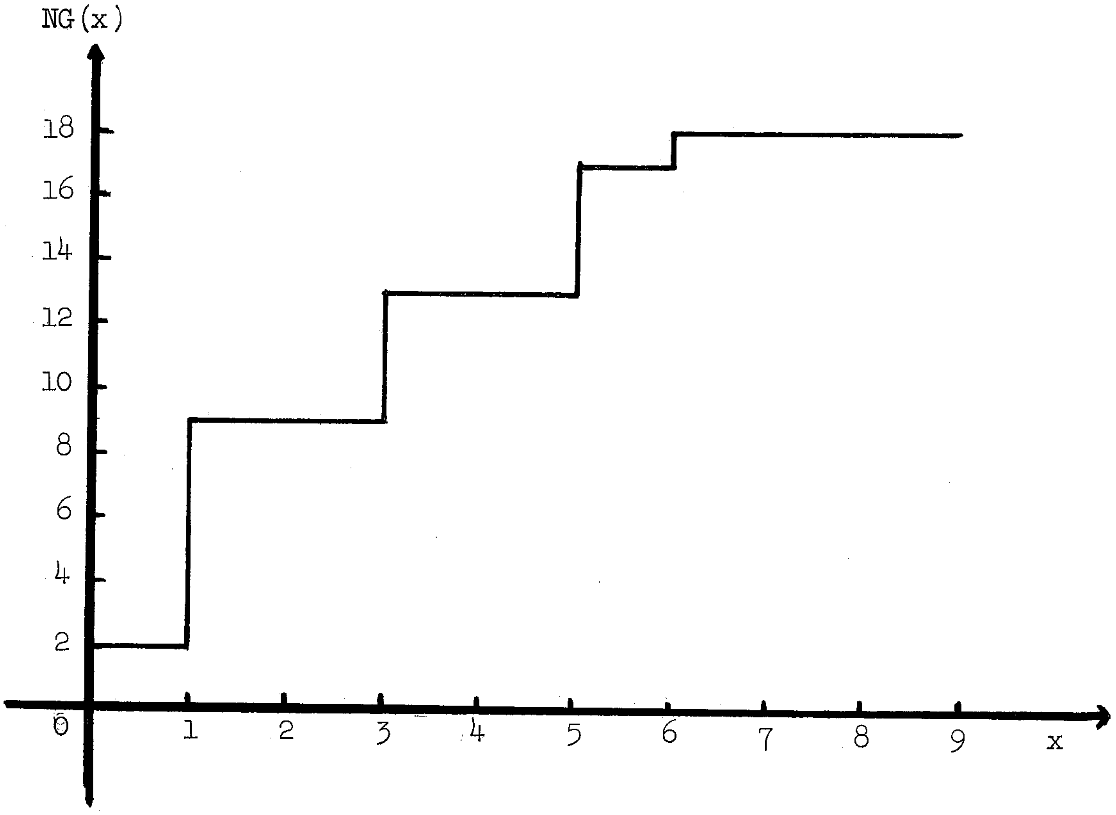

Figure 3

LEMMA 4: The value of C in (5.11) belongs to the interval $0 \leq C \leq \frac{1}{2}$. It takes the maximum value of $\frac{1}{2}$ when $N_1 = N_2$ and ξ_1 and ξ_2 are equivalent r.v.'s.

The proof of this assertion follows from (5.8), (5.9) and (5.10).

Definition of the ratio of the transvariation area: The ratio between the area of transvariation and its maximum, assuming that $N_1 = N_2$ and ξ_1 and ξ_2 are equivalent random variables, defines the ratio of the transvariation area. Therefore

$$H = 2C = \frac{2}{N} \int_{-\infty}^{\infty} d\, G(x) \quad . \tag{5.13}$$

The probability and the area of transvariation are different concepts. But we can observe that, when the probability of transvariation is null the area of transvariation is also null. The reverse of this statement is, quite clearly, false.

The range of (5.13) is the unit closed interval [0,1].

5.3. Discriminative value. Let ξ_1, ξ_2, $f_1(x)$, $f_2(x)$, N_1 and N_2 stand as in section 5.2. Let

$$M_1 = E(\xi_1) > M_2 = E(\xi_2) \tag{5.14}$$

Definition of discriminative value: The value of z of the r.v. ξ_1 and ξ_2 that minimizes the error $\epsilon(x)$ introduced under the assumptions that: 1) ξ_1 takes only values greater than z ; and further 2) that ξ_2 takes on values less than z , are, by definition, the discriminative value between the r.v. ξ_1 and ξ_2 .

Let $n_1(x)$ be the error resultant from the assumption that ξ_1 is greater than x and $n_2(x)$ resultant from the assumption that ξ_2 is less than x .

Hence

$$n_1(x) = N_1 \int_{-\infty}^{x} d\, F_1(t) \tag{5.15}$$

$$n_2(x) = N_2 \int_{x}^{\infty} d\, F_2(t) \quad . \tag{5.16}$$

Therefore, the error function is

$$\epsilon(x) = \frac{n_1(x) + n_2(x)}{N_1+N_2} \quad . \tag{5.17}$$

If we assume that $\epsilon(x)$ is continuous and differenciable, the discriminative value z is one of the roots of the equation

$$N_1 f_1(x) - N_2 f_2(x) = 0 \tag{5.18}$$

that minimizes its corresponding error function, that is, an abscissa of one of the intersection points between both frequency distributions. If there is more than one root, the discriminative value, z, will be defined by the root lowest in absolute value.

We can dismiss the restriction (5.14) and ascertain the value of z that minimizes (5.17) for each of the following alternative hypotheses.

$$n_1(x) = N_1 \int_{-\infty}^{x} d\,F_1(t) \quad ; \quad n_2(x) = N_2 \int_{x}^{\infty} d\,F_2(t) \tag{5.19a}$$

$$n_1(x) = N_1 \int_{x}^{\infty} d\,F_1(t) \quad ; \quad n_2(x) = N_2 \int_{-\infty}^{x} d\,F_2(t) \quad . \tag{5.19b}$$

The value of z that corresponds to the minimum value of $\epsilon(x)$, between hypothesis (5.19a) and (5.19b), will define the discriminative value. Following this criterion and assuming the continuity and differenciability of (5.17), the error of the discriminative value when there is only one intersection point, is then equal to the area of transvariation.

The concepts of area of transvariation, discriminative value and error function, carry on a strong flavor of the Neyman-Pearson theory of testing simple hypotheses.

6. Case of the Bivariate Normal Distribution

6.1. We will suppose now that the r.v. (4.1) has a bivariate normal distribution

(6.1) $$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-p^2}} \exp\left[-\frac{1}{2}Q(x_1,x_2)\right]$$

where (x_1,x_2) is any point in the Euclidean plane R_2, and

(6.2) $$Q(x_1,x_2) = -\frac{1}{(1-p^2)}\left[\frac{(x_1-\alpha_1)^2}{\sigma_1^2} - \frac{2p(x_1-\alpha_1)(x_2-\alpha_2)}{\sigma_1\sigma_2} + \frac{(x_2-\alpha_2)^2}{\sigma_2^2}\right]$$

is a positive definite quadratic form.

Let the parameter (4.4) be

(6.3) $$(\lambda_1,\lambda_2) = (\alpha_1,\alpha_2)$$

and

$$\lambda = \alpha_2 - \alpha_1 < 0 .$$

α_1 and α_2 are the means and medians of ξ_1 and ξ_2 respectively. The variances and covariance are:

(6.5) $$\mathrm{var}(\xi_i) = \sigma_i^2 = \sigma_{ii}, \qquad i=1,2.$$

(6.6) $$\mathrm{cov}(\xi_1,\xi_2) = \rho\,\sigma_1\sigma_2 = \sigma_{12} .$$

Applying the definitions of transvariability (4.10) and moment of transvariation (4.26), we have

(6.7) $$p = P(\xi_2>\xi_1|\alpha_2<\alpha_1) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-p^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{x_2}\exp\left[-\frac{1}{2}Q(x_1,x_2)\right]dx_1dx_2$$

(6.8) $$m_r = E[(\xi_2-\xi_1)^r|\xi_2>\xi_1,\ \alpha_2<\alpha_1]$$
$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-p^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{x_2}(x_2-x_1)^r\exp\left[-\frac{1}{2}Q(x_1,x_2)\right]dx_1dx_2 .$$

After lengthy transformations of (6.7) and (6.8) [Ref. 9] we arrive at their final expressions. We can arrive at the same result working with the linear transformation (4.6) [Ref. 9]. In fact for the r.v. (4.6) we have

(6.9) $$E(\tau) = E(\xi_2 - \xi_1) = \alpha_2 - \alpha_1 = \lambda < 0$$

(6.10) $$\text{var}\ \tau = \sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 .$$

Its p.d.f. is

(6.11) $$g(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(t-\lambda)^2}{2\sigma^2}\right] .$$

By direct application of (4.10) and (4.21), and after the substitution $u = \frac{t-\lambda}{\sigma}$, we have:

(6.12) $$p = P(\tau > 0 \mid \lambda < 0) = 1 - \Phi\left(-\frac{\lambda}{\sigma}\right) = \Phi\left(\frac{\lambda}{\sigma}\right)$$

(6.13) $$p_M = P(\tau > \lambda \mid \lambda < 0) = 1 - \Phi(0) = \frac{1}{2}$$

where

(6.14) $$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{t^2}{2}}\, dt$$

is the c.d.f. of the standardized form of the normal distribution.

By application of (4.24) the probability of transvariation becomes:

(6.15) $$P = 2\Phi\left(\frac{\lambda}{\sigma}\right) .$$

For the moment of transvariation (4.26) and its maximum (4.27) we have (Dagum, Ref. 9)

(6.16) $$m_r = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\lambda}{\sigma}}^{\infty} (\lambda + \sigma u)^r e^{-\frac{1}{2}u^2}\, du =$$

$$= \lambda^r \Phi\left(\frac{\lambda}{\sigma}\right) + \frac{1}{\sqrt{2\pi}} \sum_{s=1}^{r} \binom{r}{s} \lambda^{r-s} \sigma^s G\left(s; -\frac{\lambda}{\sigma}\right) ,$$

(6.17)
$$m_{r,M} = \frac{\sigma^r}{\sqrt{2\pi}} \int_0^\infty u^r e^{-\frac{1}{2}u^2} du = \frac{1}{2} \beta_{r,\lambda} .$$

Where, for $s = 2k$

(6.18)
$$G(2k; -b) = \frac{(2k)!}{2^k\ k!} \sqrt{2\pi}\ \Phi(b) - e^{-\frac{b^2}{2}} \sum_{i=0}^{k-1} \frac{(2i)!\ \binom{2k}{2i}}{2^i\ i!\ \binom{k}{i}} b^{2k-2i-1}$$

and, for $s=2k+1$

(6.19)
$$G(2k+1; -b) = e^{-\frac{b^2}{2}} \sum_{i=0}^{k} 2^i\ i! \binom{k}{i} b^{2k-2i} .$$

In (6.17), $\beta_{r,\lambda}$ is given by

(6.20)
$$\beta_{r,\lambda} = E(|\tau - \lambda|^r) = \frac{2^{\frac{r}{2}} \sigma^r}{\sqrt{\pi}} \Gamma\left(\frac{r+1}{2}\right)$$

which can be written, for r even and odd, respectively

(6.21)
$$\beta_{2k,\lambda} = \mu_{2k} = \frac{(2k)!}{2^k\ k!} \sigma^{2k}$$

(6.22)
$$\beta_{2k+1,\lambda} = \frac{2^{\frac{2k+1}{2}}\ k!}{\sqrt{\pi}} \sigma^{2k+1}$$

where μ_{2k} is the 2k-th central moment.

The r^{th} intensity of transvariation follows at once by the substitution of (6.16) and (6.17) in (4.32).

For $r=1$ we have

(6.23)
$$I_1 = \frac{m_1}{m_{1,M}} = \sqrt{2\pi}\ [\Phi' (\tfrac{\lambda}{\sigma}) + \tfrac{\lambda}{\sigma} p]$$
$$= \sqrt{2\pi}\ [\Phi' (\tfrac{\lambda}{\sigma}) + \tfrac{\lambda}{\sigma} \Phi (\tfrac{\lambda}{\sigma})]$$

where $\Phi'(\cdot)$ is the p.d.f. of the standardized normal distribution.

Since we are dealing with the transvariation of the bivariate normal distribution with respect to the means of each marginal distribution, as was set down in (6.9), and because of the symmetry of the distribution with respect to its center of gravity - the point of coordinates (α_1,α_2) - , Lemma 1 and (4.25) are relevant and applicable; hence

(6.24) $$x_1 - \alpha_1 = x_2 - \alpha_2$$

is a median straight line. Therefore, $p_M = \frac{1}{2}$ is in agreement with the result deduced in (6.13). Likewise, by applying Lemma 3, the maximum of the moment of transvariation for (6.17) is the same as that of (4.31). The definition of $\beta_{r,\lambda}$ was given in (4.29) and its value for the r^{th} absolute central moment of the bivariate normal distribution in (6.20), with their corresponding mathematical result, for r even and odd, in (6.21) and (6.22), respectively.

6.2. Hypothesis of independence. The mathematical result deduced for the probability and intensity of transvariation for the bivariate normal distribution is invariant with respect to the hypothesis of independence. Under this hypothesis, the covariance in (6.6) is null, because $\rho = 0$ and the variance in (6.10) becomes

(6.25) $$\mathrm{var}(\tau) = \sigma^2 = \sigma_1^2 + \sigma_2^2 \quad .$$

Under the hypothesis of independence we can work out two other parameters of transvariation already introduced in section 5, namely, the area of transvariation and the discriminative value. The frequency functions corresponding to each marginal distribution is

(6.26) $$N_1 f_1(x_1) = \frac{N_1}{\sigma_1\sqrt{2\pi}} \exp\{-\frac{(x_1-\alpha_1)^2}{2\sigma_1^2}\} \Longrightarrow F_1(x) = \int_{\infty}^{x} f_1(t)\, dt$$

(6.27) $$N_2 f_2(x_2) = \frac{N_2}{\sigma_2\sqrt{2\pi}} \exp\{-\frac{(x_2-\alpha_2)^2}{2\sigma_2^2}\} \Rightarrow F_2(x) = \int_{-\infty}^{x} f_2(t)\, dt$$

For the area of transvariation as well as for the discriminative value, we are interested in the roots x of the equation

(6.28) $$\frac{N_1}{\sigma_1\sqrt{2\pi}} \exp\{-\frac{(x-\alpha_1)^2}{2\sigma_1^2}\} = \frac{N_2}{\sigma_2\sqrt{2\pi}} \exp\{-\frac{(x-\alpha_2)^2}{2\sigma_2^2}\} .$$

Solving (6.28) we have

(6.29) $$x = \frac{(\alpha_1\sigma_2^2 - \alpha_2\sigma_1^2) \pm \sigma_1\sigma_2\sqrt{\Delta}}{\sigma_2^2 - \sigma_1^2}$$

where

$$\Delta = (\alpha_1 - \alpha_2)^2 - 2(\sigma_2^2 - \sigma_1^2) \log \frac{\sigma_1 N_2}{\sigma_2 N_1} .$$

6.2.1. Area of transvariation. We may analyze the following particular cases (for a more detailed analysis, see Dagum [Ref. 9]).

1) $\sigma_1 \neq \sigma_2$, $\Delta > 0$, Fig. 4:

(6.31) $$C = \frac{1}{N_1+N_2} [(N_2(1 + F_2(x_1) - F_2(x_2)) + N_1(F_1(x_2) - F_1(x_1))]$$

2) $\sigma_1 \neq \sigma_2$, $\Delta < 0$, Fig. 5:

(6.32) $$C = \frac{N_1}{N_1+N_2}$$

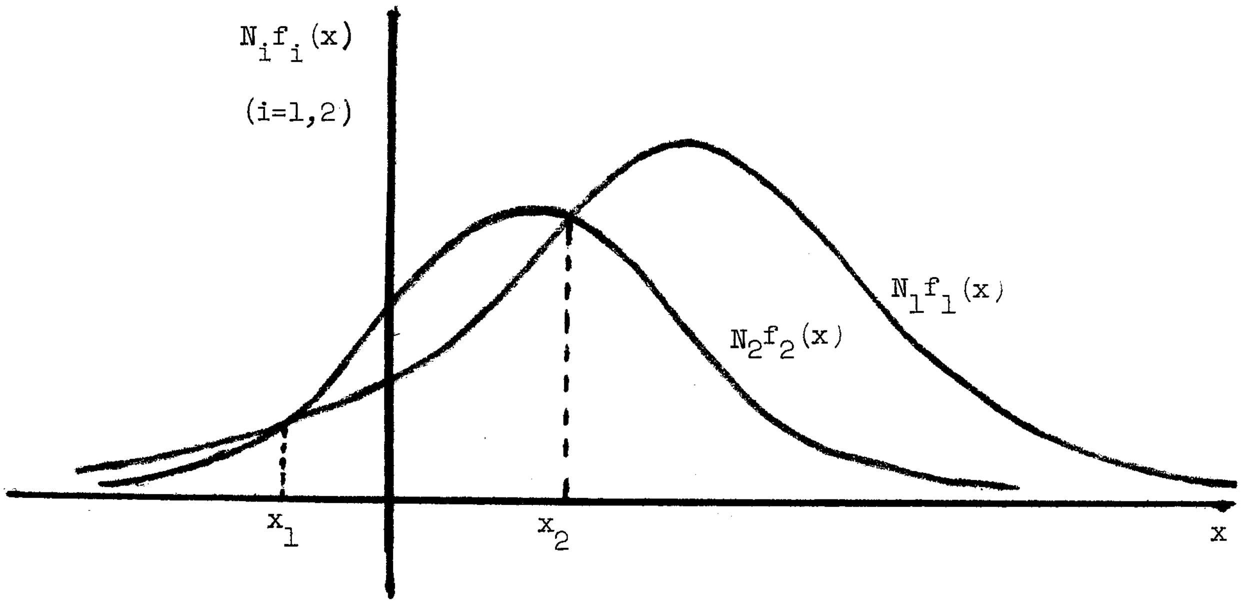

Fig. 4: $\sigma_1 \neq \sigma_2$; $\Delta > 0$

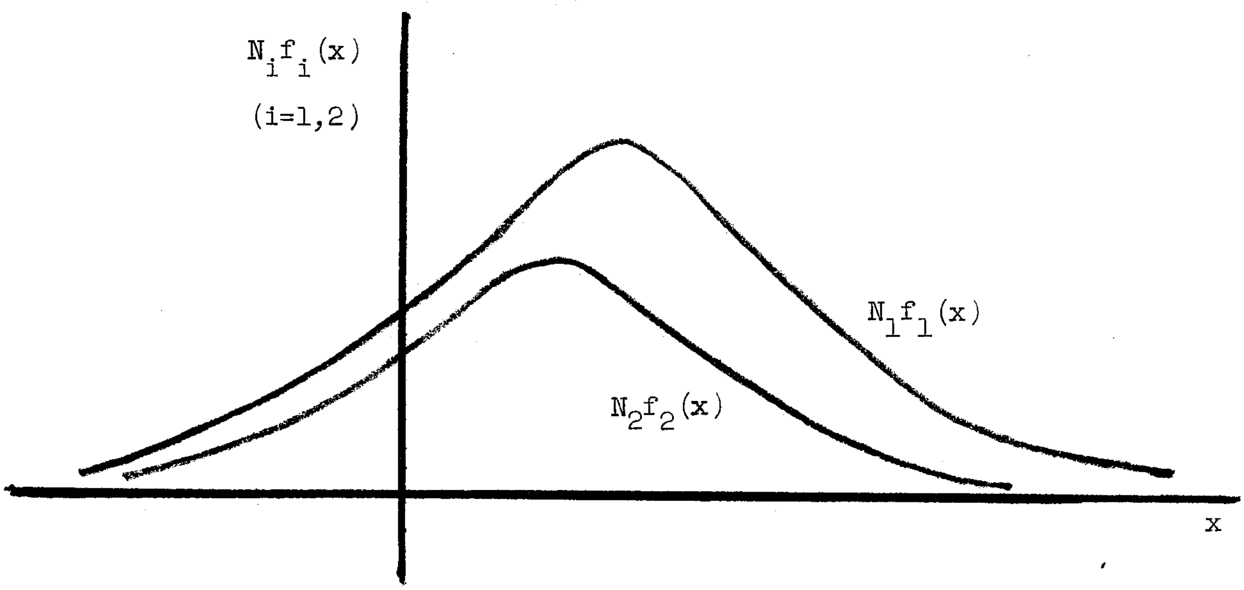

Fig. 5: $\sigma_1 \neq \sigma_2$; $\Delta < 0$

3) $\sigma_1 = \sigma_2$, $\alpha_1 \neq \alpha_2$. Fig. 6:

$$C = \frac{1}{N_1+N_2} \left[N_2 - N_2 F_2(x_1) + N_1 F_1(x_1)\right] \tag{6.33}$$

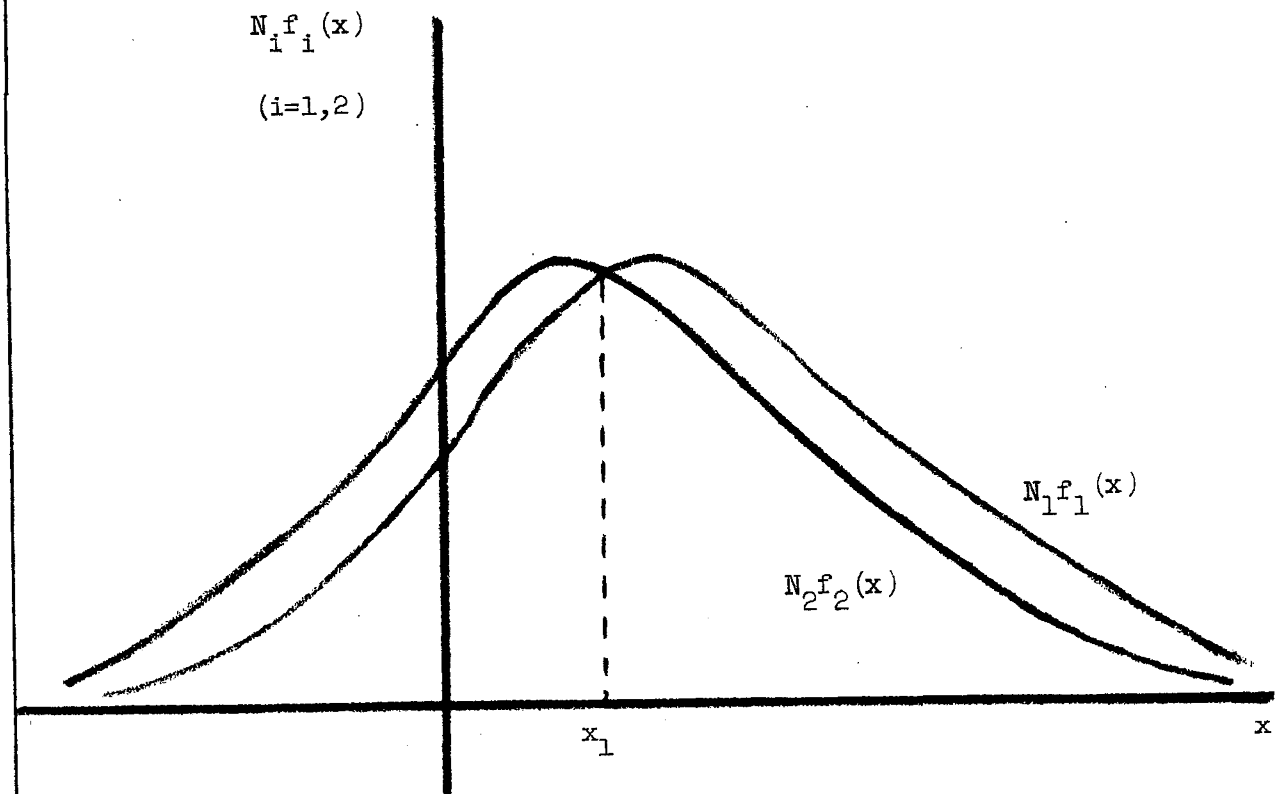

Fig. 6: $\sigma_1 = \sigma_2$; $\alpha_1 \neq \alpha_2$

4) $\sigma_1 = \sigma_2$, $\alpha_1 = \alpha_2$, $N_1 < N_2$:

$$C = \frac{N_1}{N_1+N_2} \tag{6.34}$$

5) $\sigma_1 = \sigma_2$, $\alpha_1 = \alpha_2$, $N_1 = N_2$ is a trivial case of equivalent random variables with populations of the same size. The area of transvariation is, of course, equal to $\frac{1}{2}$.

6.2.2. Discriminative value. In the analysis of the discriminative value we have the following possible cases for the distributions (6.26) and (6.27):

1) They do not intersect.

2) They intersect in only one point.

3) They intersect in two points.

Case 1 gives us a singlular point for the discriminative value.

Case 2 implies $\sigma_1 = \sigma_2$ and $\alpha_1 \neq \alpha_2$ (Fig. 6). Therefore, the discriminative error is:

(6.35)
$$x_1 = \frac{1}{2}(\alpha_1 + \alpha_2) + \frac{\sigma_1^2}{\alpha_1 - \alpha_2} \log \frac{N_2}{N_1} .$$

The error value is then

(6.36)
$$\epsilon(x_1) = \frac{N_1 F_1(x_1) + N_2[1 - F_2(x_1)]}{N_1 + N_2}$$

For the particular and interesting case where $N_1 = N_2$, we have

(6.37)
$$x_1 = \frac{1}{2}(\alpha_1 + \alpha_2)$$

and

$$\epsilon(x_1) = \Phi_1\left(\frac{\lambda}{\sigma\sqrt{2}}\right)$$

where λ is given by (6.4) and σ by (6.25).

Case 3, in conjunction with the alternative hypotheses (5.19a) and (5.19b), give four possible discriminative error values for the two roots. The minimum one defines the discriminative error value and its corresponding root is the discriminative value.

7. Means, Variances and Covariances of the Sample Transvariation Parameters

7.1. Estimators in the discrete case. In applications of transvariation theory we deal with sample observations, i.e., a finite realization of an r.e.Λ. Suppose we have n realizations of an r.e.Λ under constant conditions. In each realization we observe the r.v. (4.1) or its linear transformation (4.6). Hence, a sample observation would be

$$x_h = (x_{1i}, x_{2j}); \quad i=1,\ldots,k_1 \; ; \quad j=1,\ldots,k_2 \; . \tag{7.1}$$

After convenient transformations we have all the ordinates with common subintervals and within the same interval, i.e.

$$x_{1,i+1} - x_{1,i} = x_{2,j+1} - x_{2j} = h \; ; \quad i,j=1,\ldots,k, \tag{7.2}$$

and

$$p_{ij} = p(x_{1i}, x_{2j}) \; . \tag{7.3}$$

By application of the definitions of transvariability, moment of transvariation and their corresponding maximums, we have, for the discrete case

$$p = P(\xi_2 > \xi_1 \mid \lambda_2 < \lambda_1) + \frac{1}{2} P(\xi_1 = \xi_2) \tag{7.4}$$

$$= \sum_{j=\ell}^{k} \sum_{i=1}^{\ell-1} p_{ij} + \frac{1}{2} \sum_{i=1}^{k} p_{ii}$$

$$p_M = P(\xi_2 > \xi_1 - hc \mid \lambda_2 < \lambda_1) + \frac{1}{2} P(\xi_2 = \xi_1 - hc) \tag{7.5}$$

$$= \sum_{j=\ell}^{k} \sum_{i=1}^{\ell-1} p_{i,j-c} + \frac{1}{2} \sum_{i=1}^{k} p_{i,i-c} \; .$$

$$m_r = h^r \sum_{i=\ell}^{k} \sum_{i=1}^{\ell-1} \left(\frac{x_{2j} - x_{1i}}{h} \right)^r p_{ij} \tag{7.6}$$

(7.43) $$\mathrm{var}(p_\Lambda) = \sigma_p^2 = \frac{n}{n-1} E(p_\Lambda - p)^2 = \frac{1}{n-1} \left(pq - \frac{1}{4} p_0\right) .$$

The variance of maximum of transvariability

(7.44) $$\mathrm{var}(p_{M;\Lambda}) = \sigma_{p_M}^2 = \frac{n}{n-1} E(p_{M;\Lambda} - p_M)^2 = \frac{1}{n-1} \left(p_M q_M - \frac{1}{4} p_{-c}\right)$$

The covariance between the transvariability and its maximum

(7.45) $$\mathrm{cov}(p_\Lambda, p_{M,\Lambda}) = \frac{n}{n-1} E(p_\Lambda - p)(p_{M;\Lambda} - p_M) = \frac{1}{n-1}\; p\, q_M .$$

Likewise, from (7.39), (7.40) and (7.41) we find the unbiased estimators for the variances of the moment of transvariation and its maximum, and the covariance between them.

The variance of the r^{th} moment of transvariation is

(7.46) $$\mathrm{var}(m_{r;\Lambda}) = \sigma_{m_r}^2 = \frac{n}{n-1} E(m_{r,\Lambda} - m_r)^2 = \frac{m_{2r} - m_r^2}{n-1}$$

The variance of its maximum is

(7.47) $$\mathrm{var}(m_{r,M;\Lambda}) = \sigma_{m_{r,M}}^2 = \frac{m_{2r,M} - m_{r,M}^2}{n-1} .$$

The covariance between the r^{th} moment of transvariation and its maximum is

(7.48) $$\mathrm{cov}(m_{r;\Lambda}, m_{r,M;\Lambda}) = \frac{m_{\overline{r+r},M} - m_r m_{r,M}}{n-1}$$

($m_{\overline{r+r},M}$ was defined in 7.42).

An unbiased estimator of the variance of probability of transvariation (Dagum, Ref. 9) is

(7.49)
$$\mathrm{var}(P_\Lambda) = \sigma_P^2 = \frac{n}{n-1} E(P_\Lambda - P)^2$$
$$= P^2 \left[\frac{\sigma_p^2}{p^2} + \frac{\sigma_{p_M}^2}{p_M^2} - \frac{2\ \mathrm{cov}(p_\Lambda, p_{M;\Lambda})}{p\ p_M} \right] .$$

The variances and the covariance of the right hand member of (7.49) were obtained in (7.43) to (7.45).

For the particular case where the parameters (λ_1, λ_2) are the medians of their corresponding random variables, we have

(7.50)
$$p_M = q_M = \frac{1}{2}$$

therefore

(7.51)
$$\mathrm{var}(P_\Lambda) = \frac{4}{n-1} \left(p\ q - p^2 - p^2\ p_{-c} - \frac{1}{4} p_0 \right) .$$

Likewise, an unbiased estimator for the variance of the intensity of transvariation is

(7.52)
$$\mathrm{var}(I_{r;\Lambda}) = \sigma_{I_r}^2 = \frac{n}{n-1} E(I_{r;\Lambda} - I_r)^2$$
$$= \frac{I_r^2}{r^2} \left[\frac{\sigma_{m_r}^2}{m_r^2} + \frac{\sigma_{m_{r,M}}^2}{m_{r,M}^2} - \frac{2\ \mathrm{cov}(m_{r;\Lambda}, m_{r,M;\Lambda})}{m_r\ m_{r,M}} \right] .$$

The variances and the covariance of the right hand members of (7.52) were obtained in (7.46) to 7.48).

<u>7.6. Variances and covariances on the hypothesis of normal distributions.</u> The variances and covariances deduced in section 7.5 are of general application, unrestricted with respect to the particular distribution model of its corresponding random variables. In section 7.5, working on the hypothesis of normal distribution,

we substitute the transvariation parameter estimators for the corresponding mathematical results for the bivariate normal distribution obtained in section 6. Therefore, for the particular case of $(\lambda_1,\lambda_2) = (\alpha_1,\alpha_2)$, i.e. transvariation with respect to the means, we have the following.

The variance of transvariability (7.43) becomes, from (6.12)

(7.53) $$\operatorname{var}(p_r) = \frac{1}{n-1}\, pq = \frac{1}{n-1}\, \Phi(\frac{\lambda}{\sigma})\ \Phi(-\frac{\lambda}{\sigma}) .$$

The variance of its maximum (7.44) is, from (6.13) or (7.50)

(7.54) $$\operatorname{var}(p_{M;\Lambda}) = \frac{1}{n-1}\, p_M\, q_M = \frac{1}{4(n-1)} .$$

The covariance between the transvariability and its maximum (7.45) is, from (6.12) and (6.13)

(7.55) $$\operatorname{cov}(p_\Lambda,\ p_{M,\Lambda}) = \frac{1}{n-1}\, p\, q_M = \frac{1}{2(n-1)}\, \Phi(\frac{\lambda}{\sigma}) .$$

Hence, the variance of the probability of transvariation (7.51) is

(7.56) $$\operatorname{var}(P_\Lambda) = \frac{4}{n-1}\, p(q-p) = \frac{4}{n-1}\, \Phi(\frac{\lambda}{\sigma})\, [\Phi(-\frac{\lambda}{\sigma}) - \Phi(\frac{\lambda}{\sigma})] .$$

For the variance of the r^{th} moment of transvariation (7.46) we replace there the corresponding results deduced in (6.16), in association with (6.18) and (6.19).

Likewise, for its maximum (7.47), we use (6.17) in association with its corresponding result in (6.20) to (6.22).

The same results are used for the variance of the r^{th} intensity of transvariation in (7.52).

For purposes of application, it is important that we derive the variance of the first order intensity of transvariation on the hypothesis of normality. From (7.52), in conjunction with (7.46) to (7.48), we obtain

$$(7.57) \qquad \operatorname{var}(I_{1,\Lambda}) = \frac{I_1^2}{n-1}\left(\frac{m_2}{m_1^2} + \frac{m_{2,M}}{m_{1,M}^2} - \frac{2m_{1+1,M}}{m_1 \cdot m_{1,M}}\right)$$

whereas, from (6.16) to (6.23), (6.9), (7.8) and (7.42), we have

$$(7.58) \qquad m_1 = \lambda\, \Phi(\tfrac{\lambda}{\sigma}) + \sigma \Phi'(\tfrac{\lambda}{\sigma})$$

$$(7.59) \qquad m_2 = (\lambda^2 + \sigma^2)\, \Phi(\tfrac{\lambda}{\sigma}) + \lambda\, \sigma\, \Phi'(\tfrac{\lambda}{\sigma})$$

$$(7.60) \qquad m_{1,M} = \frac{1}{2}\beta_{1,\lambda} = \frac{\sigma}{\sqrt{2\pi}}$$

$$(7.61) \qquad m_{2,M} = \frac{1}{2}\beta_{2,\lambda} = \frac{1}{2}\sigma^2$$

$$(7.62) \qquad m_{1+1,M} = |\lambda| m_1 + m_2$$

$$(7.63) \qquad I_1 = \sqrt{2\pi}\,[\Phi'(\tfrac{\lambda}{\sigma}) + \tfrac{\lambda}{\sigma}\, \Phi(\tfrac{\lambda}{\sigma})]$$

which are direct computations derived from a table of the standardized normal distribution.

8. Applications

8.1 Statistical Interpretation. According to the experience of this author, a criterion for deciding whether or not there is a statistically significant difference between two distribution functions would be a significance level of 60 percent for probability and intensity of transvariation. Thus, for a sample estimate of greater than or equal to 60 percent for both parameters, we may conclude that there is no statistically significant difference between the distribution of the two random variables. The converse of this is true for a level of less than 60 percent.

If one estimate is less than 60 percent and the other is greater than or equal to, we need additional statistical evidence to determine the difference, in which case we may use the following supplementary criterion:

If the average of the sample estimates of both probability and intensity of transvariation is less than 60 percent, we may conclude that there is a statistically significant difference between the two corresponding distributions.

These criteria are in agreement with the purpose of the parameters of transvariation, namely, the simultaneous measurement of the effects of the means, variances and probability distributions of the two random variables that generated the sample realizations under analysis.

The applications that follow are intended first of all, to underline the power of transvariation theory as a quantitative method to deal with comparative statics in economic analysis.

8.2. Nonparametric Estimations. The following 11 statistical variables represent applications of transvariation theory to economics[1]:

1) $y_{1(60)i}$ = average weekly hours worked in 1960, for 49 States (including District of Columbia and excluding Alaska and Hawaii) of the United States;

2) $y_{1(66)j}$ = id. id. in 1966;

3) $y_{2(60)i}$ = average hourly earnings in 1960, for 49 States (including District of Columbia and excluding Alaska and Hawaii) of the United States;

4) $y_{2(66)j}$ = id. id. in 1966;

5) u_{1i} = wage indexes of skilled workers in Buenos Aires (Argentina) for twelve occupations in the industrial sector in December 1967 (Base index 1960 = 100);

6) u_{2j} = id. id. for unskilled workers;

7) x_{1i} = price indexes of 33 items of meat and fish included in family basket of the cost of living index in Buenos Aires (Argentina) for 1967 (Base index 1960 = 100);

8) x_{2i} = id. id. to x_{1i} for 31 items of clothing (15 mens' and 16 womens');

9) x_{3i} = id. id. for 19 items of bread and cereals;

10) x_{4i} = id. id. for 24 items of education and culture;

11) x_{5i} = id. id. for 39 items of fruits and vegetables.

---

[1]Sources: United States, Bureau of the Census, Statistical Abstract of the United States, 1967. Argentina, Dirección Nacional de Estadística y Censos, Costo de Vida, E.S. 60, Enero 1968.

Table 8.1 gives the sample mean and size for each of the eleven statistical series.

Table 8.1

Variables	Mean λ_i	Sample size n_i
$y_{1(60)i}$	40.04	49
$y_{1(66)j}$	41.41	49
$y_{2(60)i}$	2.19	49
$y_{2(66)j}$	2.40	49
u_{1i}	631.6	12
u_{2j}	633.9	12
x_{1i}	412.9	33
x_{2i}	510.3	31
x_{3i}	533.8	19
x_{4i}	610.8	24
x_{5i}	617.8	39

Nonparametric estimation were performed for thirteen relevant cases out of the eleven variables of Table 8.1. The thirteen pairs of variables considered together with the means of their differences and the combined number of observations are given in Table 8.2.

Table 8.2

Variables	Mean $\lambda = \lambda_i - \lambda_j$	Combined sample size $n = n_i \times n_j$
$y_{1(60)i} - y_{1(66)j}$	- 1.37	2401
$y_{2(60)i} - y_{2(66)j}$	- 0.21	2401
$u_{1i} - u_{2j}$	- 2.3	144
$x_{1i} - x_{2j}$	- 97.4	1023
$x_{1i} - x_{3j}$	-140.9	627
$x_{1i} - x_{4j}$	-197.9	792
$x_{1i} - x_{5j}$	-204.9	1287
$x_{2i} - x_{3j}$	- 43.5	589
$x_{2i} - x_{4j}$	-100.5	744
$x_{2i} - x_{5j}$	-107.5	1209
$x_{3i} - x_{4j}$	- 57.0	456
$x_{3i} - x_{5j}$	- 64.0	741
$x_{4i} - x_{5j}$	- 7.0	936

The nonparametric sample estimators of transvariability (formula 7.4), maximum of transvariability (formula 7.5), r^{th} moment of transvariation, for $r = 1$ and 2 (formula 7.6), maximum of the r_{th} moment of transvariation, for $r = 1$ and 2 (formula 7.7) and the moment $m_{\overline{1+1},M}$ (formula 7.42, for $r = 1$) were applied to the thirteen pairs of random variables of Table 8.2. Their corresponding estimates are given in Table 8.3.

Table 8.3

Variables	p	p_M	m_1	m_{1M}	m_2	m_{2M}	$m_{1+1,M}$
$y_{1(60)i}-y_{1(66)j}$	0.149	0.514	0.163	0.601	0.401	1.348	0.624
$y_{2(60)i}-y_{2(66)j}$	0.327	0.501	0.104	0.192	0.051	0.114	0.074
$u_{1i}-u_{2j}$	0.549	0.549	41.806	43.068	5450.360	5645.575	5546.514
$x_{1i}-x_{2j}$	0.188	0.528	10.933	44.794	1123.177	5984.101	2188.007
$x_{1i}-x_{3j}$	0.155	0.521	9.991	54.081	921.299	8772.699	2329.064
$x_{1i}-x_{4j}$	0.184	0.646	15.174	96.311	1863.888	20860.150	4866.894
$x_{1i}-x_{5j}$	0.213	0.590	13.581	101.479	1365.905	22261.170	4168.644
$x_{2i}-x_{3j}$	0.370	0.472	37.367	55.578	5665.927	9676.330	7291.372
$x_{2i}-x_{4j}$	0.398	0.782	44.818	128.354	7928.055	29852.770	12432.251
$x_{2i}-x_{5j}$	0.424	0.592	46.127	100.915	7482.945	22961.840	12441.592
$x_{3i}-x_{4j}$	0.487	0.613	70.787	101.571	15162.260	24916.560	19197.110
$x_{3i}-x_{5j}$	0.468	0.575	70.534	103.998	14890.550	25982.760	19404.700
$x_{4i}-x_{5j}$	0.505	0.512	125.957	129.512	64051.720	65839.910	64933.422

By application of (7.43), (7.44), (7.46), (7.47) and (7.48) we may estimate the variances and covariances corresponding to the estimates given in the first four columns of Table 8.3. Their results are given in Table 8.4.

Table 8.4

Variables	$var(p_{\Lambda})$	$var(p_{M,\Lambda})$	cov $(p_{\Lambda}, p_{M,\Lambda})$	$var(m_{1,\Lambda})$	$var(m_{1M,\Lambda})$	cov $(m_{1r}, m_{1M,\Lambda})$
$y_{1(60)i}-y_{1(66)j}$	0.000051	0.000101	0.000030	0.000156	0.000411	0.000219
$y_{2(60)i}-y_{2(66)j}$	0.000092	0.000104	0.000068	0.000017	0.000032	0.000022
$u_{1i} - u_{2j}$	0.001732	0.001732	0.001732	25.892290	26.508518	26.195810
$x_{1i} - x_{2j}$	0.000149	0.000244	0.000087	0.982051	3.891983	1.661737
$x_{1i} - x_{3j}$	0.000209	0.000399	0.000118	1.312260	9.341842	2.857400
$x_{1i} - x_{4j}$	0.000190	0.000289	0.000083	2.065268	14.645161	4.305227
$x_{1i} - x_{5j}$	0.000130	0.000188	0.000068	0.918711	9.302630	2.169877
$x_{2i} - x_{3j}$	0.000396	0.000424	0.000332	7.261340	11.203013	8.868370
$x_{2i} - x_{4j}$	0.000322	0.000230	0.000117	7.966912	18.005515	8.990190
$x_{2i} - x_{5j}$	0.000202	0.000200	0.000143	4.433154	10.577835	6.445929
$x_{3i} - x_{4j}$	0.000549	0.000521	0.000414	22.310949	32.087853	26.389556
$x_{3i} - x_{5j}$	0.000336	0.000328	0.000269	13.399408	20.496268	16.309958
$x_{4i} - x_{5j}$	0.000267	0.000267	0.000264	51.536287	52.477457	52.000376

By application of (4.24), (4.32), (7.49) and (7.52) we estimate the probability and first order intensity of transvariation, their corresponding variances and mean square errors. The results are given in Table 8.5.

Table 8.5

Variables	$P = \frac{p}{p_M}$	$I_1 = \frac{m_1}{m_{1M}}$	$var(P_\Lambda)$	$var(I_{1,\Lambda})$	σ_P	σ_{I_1}
$y_{1(60)i}-y_{1(66)j}$	0.290	0.271	0.000160	0.000186	0.013	0.014
$y_{2(60)i}-y_{2(66)j}$	0.652	0.544	0.000189	0.000053	0.014	0.007
$u_{1i} - u_{2j}$	1.0	0.971	0.0	0.000007	0.0	0.003
$x_{1i} - x_{2j}$	0.356	0.244	0.000425	0.000201	0.021	0.014
$x_{1i} - x_{3j}$	0.297	0.185	0.000641	0.000197	0.025	0.014
$x_{1i} - x_{4j}$	0.285	0.158	0.000398	0.000116	0.020	0.011
$x_{1i} - x_{5j}$	0.361	0.134	0.000322	0.000049	0.018	0.007
$x_{2i} - x_{3j}$	0.784	0.672	0.000610	0.000130	0.025	0.011
$x_{2i} - x_{4j}$	0.509	0.349	0.000430	0.000236	0.021	0.015
$x_{2i} - x_{5j}$	0.715	0.457	0.000285	0.000074	0.017	0.009
$x_{3i} - x_{4j}$	0.794	0.697	0.000586	0.000108	0.024	0.010
$x_{3i} - x_{5j}$	0.813	0.678	0.000352	0.000065	0.019	0.008
$x_{4i} - x_{5j}$	0.987	0.973	0.000026	0.000002	0.005	0.001

The results given in Table 8.5, for the thirteen pairs of random variables studied, in conjunction with the criteria for statistical interpretation given in 8.1, allow the following conclusions:

1) The spatial (by States) distribution of the average weekly hours worked in the United States in 1960 is significantly different from the average observed in 1966. In fact, the probability of

transvariation estimated was equal to 29 percent, with a mean square error equal to 1.3 percent, and the intensity of transvariation of the first order was equal to 27.1 percent with a mean square error equal to 1.4 percent.

2.) The spatial (by States) distribution of the average hourly earnings in the United States in 1960, by application of the supplementary criterion, is significantly different from the average observed in 1966. There is an important group of States with average hourly earnings in 1960 that overlap the average hourly earnings in 1966 of another important group of States.

3) The comparative differences of wage indexes between skilled and unskilled workers in Buenos Aires, for twelve categories of industrial employment in December 1967, with base index 1960 = 100, are not statistically significant (with almost certainty). Therefore, the wage structure by degrees of training was not changed in December 1967 compare with 1960.

4) The remaining ten applications correspond to the five group of items selected from the family basket used in the computations of the cost of living index in Buenos Aires. The first four transvariations, which correspond to the transvariations of the price of mean and fish in comparison with prices in the remaining four categories, demonstrate a significant difference. Hence, the relative prices of meat and fish in 1967 are significantly less than the prices of clothes, bread and cereals, education and culture and fruit and vegetables when compared with a common base of 100 for the year 1960.

The relative prices of clothes were not statistically different when they were compared with the prices of bread and cereals. Compared with fruit and vegetables and with education and culture, they were statistically different (by application of the supplementary criterion of 8.1).

The remaining three cases are not statistically significant.

Therefore, the impact on the realtive prices of the highly inflationary process in Argentina, from 1960 to 1967 (the price indexes grew from 100 in 1960 to approximately 500 in 1967), for the five categories of goods studied, were as follows:

i) unfavorable prices of meat and fish relative to the other four groups;

ii) unfavorable prices of clothes relative to fruit and vegetables and to educational and cultural prices and no significant differences with the other (bread and cereals);

iii) no significant differences of relative prices between bread and cereals, education and culture and fruit and vegetables.

8.3. Estimates under the assumption of Gaussian distributions. Under the assumption of independent Gaussian distributions we can estimate the following parameters:

1) Probability of transvariation;

2) First order intensity of transvariation;

3) Area of transvariation;

4) Discriminative value.

8.3.1. Probability and intensity of transvariation. The parametric estimations of the probability and intensity of transvariation, under the assumption of the Gaussian distributions, were performed for two out of the thirteen non-parametric cases analyzed in 8.2. They are:

1) $y_{1(60)i} - y_{1(66)j}$ $(i,j = 1,\ldots,49)$ = differences of the average weekly hours worked in 1960 compared with 1966 for 49 States (including District of Columbia and excluding Alaska and Hawaii) of the United States;

2) $y_{2(60)i} - y_{2(66)j}$ $(i,j = 1,\ldots,49)$ = differences of the average hourly earnings in 1960 compared with 1966 for 49 States (including District of Columbia and excluding Alaska and Hawaii) of the United States.

We assumed, in both cases, that the variables analyzed were statistically independent. For the variables of average weekly hours this assumption is acceptable. For the variables of average hourly earnings, however, the assumption can be regarded as too strong, because it is expected that both variables will move in the same direction as economic growth, which is a causal variable. Hence, the degree of correlation will be a function of the income distribution structure by States. If it does not change in the 49 States then the correlation coefficient will approach one. If the income distribution structure by States change with different intensity, then the degree of correlation will decrease.

Applying (6.9) and (6.25) to the sample observations, we obtained the information we needed to compute the sample estimates of the probability and intensity of transvariation and their corresponding variances. In Table 8.6

we found the sample means and variances of the random variables considered and, in Table 8.7, the sample estimates of the transvariation parameters and their corresponding variances.

Table 8.6

Variables	Mean	Variance	Mean Square Error
$y_{1(60)i}$	40.04	1.2604	1.12
$y_{1(66)j}$	41.41	1.3237	1.15
$y_{2(60)i}$	2.19	0.1110	0.33
$y_{2(66)j}$	2.40	0.1223	0.35
$y_{1(60)i} - y_{1(66)j}$	-1.37	2.5841	1.61
$y_{2(60)i} - y_{2(66)j}$	-0.21	0.2333	0.48

The probability and intensity of transvariation estimated for the weekly hourly earnings in both cases, nonparametric and Gaussian, are respectively equal. For the average weekly hours worked, this equality holds only for the intensity of transvariation estimation.

8.3.2.. Area of Transvariation. Using the information of Table 8.6 and applying (6.29) and (6.30) in conjunction with (6.4) we can calculate the roots of equation (6.28) and their corresponding standardized deviations that are summed up in Table 8.8.

Table 8.7

	Estimators	$y_{1(60)i}-y_{1(66)j}$	$y_{2(60)i}-y_{2(66)j}$
1.	$\frac{\lambda}{\sigma}$	-0.851	-0.445
2.	$p = \Phi(\frac{\lambda}{\sigma})$	0.198	0.328
3.	σ_p	0.008	0.010
4.	p_M	0.500	0.500
5.	σ_{p_M}	0.010	0.010
6.	$P = 2p$	0.395	0.656
7.	σ_P	0.014	0.014
8.	$m_1 = \lambda \Phi(\frac{\lambda}{\sigma}) + \sigma \Phi'(\frac{\lambda}{\sigma})$	0.177	0.104
9.	$m_2 = (\lambda^2 + \sigma^2) \Phi(\frac{\lambda}{\sigma}) + \lambda \sigma \Phi'(\frac{\lambda}{\sigma})$	0.269	0.054
10.	σ_{m_1}	0.010	0.004
11.	$m_{1,M} = \frac{1}{\sqrt{2\pi}} \sigma$	0.641	0.193
12.	$m_{2,M} = \frac{1}{2} \sigma^2$	1.292	0.117
13.	$\sigma_{m_{1M}}$	0.019	0.006
14.	$m_{\overline{1+1},M} = \lvert\lambda\rvert m_1 + m_2$	0.510	0.077
15.	$I_1 = \sqrt{2\pi} [\Phi'(\frac{\lambda}{\sigma}) + \frac{\lambda}{\sigma} \Phi(\frac{\lambda}{\sigma})]$	0.275	0.540
16.	σ_{I_1}	0.009	0.008

Table 8.8

Estimators	$y_{1(60)i}$ with $y_{1(66)j}$	$y_{2(60)i}$ with $y_{2(66)j}$
$\triangle = (\alpha_1-\alpha_2)^2 - 2(\sigma_2^2-\sigma_1^2)\log\frac{\sigma_1 N_2}{\sigma_2 N_1}$	1.873	0.047
$X_1 = \alpha_1\sigma_2^2 - \alpha_2\sigma_1^2 + \sigma_1\sigma_2\sqrt{\triangle}$	0.955	0.024
$X_2 = \alpha_1\sigma_2^2 - \alpha_2\sigma_1^2 - \sigma_1\sigma_2\sqrt{\triangle}$	-2.580	-0.026
$x_1 = \frac{X_1}{\sigma_2^2-\sigma_1^2}$	-15.077	-2.163
$x_2 = \frac{X_2}{\sigma_2^2-\sigma_1^2}$	40.743	2.318
$t_{11} = \frac{x_1-\alpha_1}{\sigma_1}$	-49.097	-13.054
$t_{12} = \frac{x_2-\alpha_1}{\sigma_1}$	-0.581	-0.240
$t_{21} = \frac{x_1-\alpha_2}{\sigma_2}$	-49.098	-13.057
$t_{22} = \frac{x_2-\alpha_2}{\sigma_2}$	0.622	0.393
$F_1(x_1) = \Phi(t_{11})$	0.0	0.0
$F_1(x_2) = \Phi(t_{12})$	0.280	0.405
$F_2(x_1) = \Phi(t_{21})$	0.0	0.0
$F_2(x_2) = \Phi(t_{22})$	0.733	0.653
C	0.274	0.376
H	0.548	0.752

From Tables 8.6 and 8.8 we have $\sigma_1 \neq \sigma_2$ and $\Delta > 0$; hence, we apply (6.31) in order to estimate the area of transvariation. From Table 8.8 and formulas (6.31) and (5.13) we have, for the variables of the average weekly hours worked

$$C = \frac{1}{2}[1 - F_2(x_2) + F_2(x_1) + F_1(x_2) - F_1(x_1)] = 0.274$$

$$H = 2\,C = 0.548$$

and for the variables of average hourly earnings:

$$C = 0.376$$

$$H = 2\,C = 0.752.$$

8.3.3. Discriminative value. Under the assumption of Gaussian distributions, both cases analyzed intersect in two points. Their corresponding roots x_1 and x_2 , for each case were previously obtained in Table 8.8, when we solved (6.28) for the area of transvariation. Working with the alternative hypotheses (5.19.a) and (5.19b) and the two roots x_1 and x_2 we obtain four discriminative errors. Among them, the minimum one corresponds to hypothesis (5.19.a) where, in both cases analyzed, $x = x_2$. Therefore, for the variable of average weekly hours worked, the discriminative value is

$$x_2 = 40.743 \Rightarrow t_{12} = -0.581 \quad \text{and} \quad t_{22} = 0.622$$

and the discriminative error is

$$\epsilon(x_2) = \frac{F_1(x_2) + 1 - F_2(x_2)}{2} = \frac{\Phi(t_{12}) + 1 - \Phi(t_{22})}{2} = 0.27 .$$

For the average hourly earnings variables, the discriminative value is

$$\epsilon\, x_2 = 2.318 \Rightarrow t_{12} = -0.240 \quad \text{and} \quad t_{22} = 0.393$$

and the discriminative error is

$$\epsilon(x_2) = \frac{\Phi(t_{12}) + 1 - \Phi(t_{22})}{2} = 0.38 .$$

ABSTRACT

This paper is concerned with the bivariate transvariation theory. It presents an historical account of the subject and a development of the theory. It deals with the probability of transvariation and its related concepts, namely, transvariability and its maximum; the r^{th} moment of transvariation, its maximum and the r^{th} intensity of transvariation; area of transvariation and discriminative value. These concepts are developed without parametric constraint and under the assumption of Gaussian distribution. The sample variances and covariances of the transvariation parameters estimators were herein deduced.

The applications were performed on economic variables of the United States, namely, 1) average weekly hours worked and 2) the average hourly earnings in the years 1960 and 1966, distributed by States. It was also applied to economic variables of Argentina, namely, wage indexes for skilled and unskilled industrial workers in Buenos Aires, for December 1967, (base equal 100 in 1960) and price indexes for five groups of commodities in 1967 (base equal 100 in year 1960). These applications allowed very important quantitative conclusions regarding: a) the spatial distribution of average weekly hours worked and hourly earnings in the United States in 1960 compared with 1966; b) the relative wage distribution between skilled and unskilled industrial workers in Buenos Aires (Argentina) and c) the relative price changes among five groups of commodities included in the computation of the cost of living index in Buenos Aires.

The applications underline the fruitfulness of transvariation theory as a quantitative method to deal with comparative statics analysis.

REFERENCES

1. Bernoulli, Jacob (1713): Ars Conjectandi. Opus posthumum. Basileae.

2. Boldrini, Marcello (1919): Su alcune differenze sessuali secondarie nelle dimensioni del corpo umano alla nascita e alle età superiori. Archivio per l'Antropologia e l'Etnologia, vol. XLIX.

3. ________________ (1920): I cadavari degli sconosciuti. Ricerche demografiche e antropologiche sul materiale della Morgue di Roma. Scuola Positiva.

4. ________________ (1920): Differenze sessuali nei pesi del corpo e degli organi umani. Rendiconti della R. Accademia dei Lincei, vol. XXIX, s. V.; fasc. 1-4.

5. ________________ (1921): Gli studi statistici sul sesso: la proporzione dei sessi nelle nascite e i carateri sessuali secondari. Rassegna di studi sessuali.

6. ________________ (1924): Misure interne ed esterne di alcune ossa lunghe nel l'uomo e nella donna. Rendiconti della R. Accademia dei Lincei, vol. XXXIII, s.V.; fasc. 7-10.

7. ________________ (1927): Biometria. Problemi della vita della specie e degli individui. Ed. Cedam, Padova.

8. Castellano, Vittorio (1934): Sullo scarto quedratico medio della probabilità di transvariazione. METRON, vol. XI, n° 4. (Also in Ref. 24).

9. Dagum, Camilo (1960): Teoría de la transvariación, sus aplicaciones a la economía. METRON, Istituto di Statistica, Università di Roma, vol. XX, n° 1-4.

10. ____________ (1960): Transvariación entre mas de dos distribuciones. Studi in Onore di Corrado Gini, vol. I, Università di Roma.

11. ____________ (1960): Transvariazione fra più di due distribuzioni. (Paper in Ref. 24).

12. ____________ (1961): Transvariación en la hipótesis de variables aleatorias normales multidimensionales. Bulletin of the International Statistical Institute, vol. XXXVIII, part. IV, Tokyo.

13. ____________ (1965): Probabilité et intensité de la transvariation dans l'space a N dimensions. Economie Appliquée, No. 4.

14. Dagum, Camilo (1966): Wahrscheinlichkeit und ausmaβ der transvariation im n-dimensionalen raum. Statistische Hefte, 7 Jahrgang, Heft 1|2.

15. Dalhberg, G. (1951): Eine methode zur messung von rassenunterschieden, METRON, Vol. XVI, No. 1-2.

16. De Lucia, Luigi (1952): Transvariazione tra caratteri connessi. Atti della XII riunione della Società Italiana di Statistica. (Also in Ref. 24).

17. De Novellis, Mirella (1960): Ulteriori contributi alla teoria della transvariazione. Gli indici di transvariazione nel caso di distribuzioni che diano luogo a curve iperboliche. (Paper in Ref. 24).

18. ______________ (1960): Applicazioni della teoria della transvariazione allo studio di alcuni problemi di antropometria e biologia. (Paper in Ref. 24).

19. Emanuelli, Filippo (1960): Sulla determinazione della probabilità e dell' intensità di transvariazione tra due variabili casuali nel caso che queste seguano una legge di distribuzione dovuta all' Edgeworth. (Paper in Ref. 24).

20. Filipponi, A. (1953): Sul grado di stabilità nei caratteri Protomagalhaensia Marottai (Sporozoa, Gregarinidae). Rivista di Parassitologia, vol. XIV, No. 3.

21. Gini, Corrado (1916): Il concetto di transvariazione e le sue prime applicazioni. Giornale degli Economisti e Rivista di Statistica. (Also in Ref. 24).

22. ______________ (1942): Per la determinazione delle probabilità di transvariozione tra più gruppi. Atti della V riunione della Società Italiana di Statistica. (Also in Ref. 24).

23. ______________ (1951): Della misura sintetica della transvariazione rispetto ad n caratteri (transvariazione n-dimensionale). Atti della XI riunione della Società Italiana di Statistica. (Also in Ref. 24).

24. ______________ (ed.) (1960): Transvariazione. Memorie di metodologia statistica, Vol. II, a cura di Giuseppe Ottaviani. Roma: Libreria Goliardica.

25. ______________ (1960): Della selezione indiretta. (Paper in Ref. 24).

26. ______________ (1965): Di un' applicazione della transvariazione alla tassonomia (La regola del 75 o/o). METRON, Vol. XXIV, No. 1-4.

27. Gini, C. e G. Livada (1943): Transvariazione a più dimensioni. Atti della VI riunione della Società Italiana di Statistica. (Also in Ref. 24).

28. ______________ (1943): Nuovi comtributti alla teoria della transvariazione. Atti della VII riunione della Società Italiana di Statistica. (Also in Ref. 24).

29. Gini, C. e G. Sonnino (1951): Contributo alla teoria della transvariazione fra seriazioni correlate. Atti della XI riunione della Società Italiana di Statistica. (Also in Ref. 24).

30. Gini, C., C. Viterbo, C. Benedetti e A. Herzel (1959): Problemi di transvariazione inversa. METRON, Vol. XIX, n° 1-2. (Also in Ref. 24).

31. Herzel, Amato (1959): Influenza del raggruppamento in classi sulla probabilità e sull' intensità di transvariazione. METRON, Vol. XIX, n° 1-2. (Also in Ref. 24).

32. Johannsen, W. (1913): Elemente der exakten erblichkeitslehre. Jena.

33. Lexis, W. (1913): Abhandlungen zur theorie der bevölkerungs und moralstatistik. Jena.

34. Livada, Gregorio (1943): Procedimento per el calcolo dell'intensità di transvariazione. Atti della VI riunione della Società Italiana di Statistica.

35. Miani-Calabrese, Donato (1943): La transvariazione rispetto al sesso dei caratteri fisici dell'infanzia. Statistica, octobre-decembre. (Also in Ref. 24).

36. Ottaviani, Guiseppe (1939): Sulla probabilità che une prova su due variabili casuali X e Y varifichi la disuguaglianza X < Y e sul corrispondente scarto cuadratico medio. Giornale dell' Istituto Italiano degli Attuari, anno X, 3-4. (Also in Ref. 24).

37. ________________ (1960): Sulla transvariabilità. (Paper in Ref. 24).

38. ________________ (1960): Alcune considerazioni sul concetto di transvariazione. (Paper in Ref. 24).

39. Quetelet, A. (1846): Lettres sur la théorie des probabilités appliquée aux sciences morales et politiques. Bruxelles.

40. Salvemini, Tommaso (1960): Transvariazione e analisi discriminatoria. (Paper in Ref. 24).

41. Steriotis, Pietro (1960): Contributo alla transvariazione delle distribuzioni gaussiane. (Paper in Ref. 24).

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) PRINCETON UNIVERSITY		2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP
3. REPORT TITLE NONPARAMETRIC AND GAUSSIAN BIVARIATE TRANSVARIATION THEORY: ITS APPLICATION TO ECONOMICS		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Research Memorandum No. 99		
5. AUTHOR(S) (Last name, first name, initial) Dagum, Camilo		
6. REPORT DATE June 1968	7a. TOTAL NO. OF PAGES 57	7b. NO. OF REFS 41
8a. CONTRACT OR GRANT NO. ONR Contract N00014-67 A-0151-0007 b. PROJECT NO. Task No. 047-086 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Research Memorandum No. 99 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
10. AVAILABILITY/LIMITATION NOTICES Distribution of this document is unlimited.		
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Logistics and Mathematical Branch Office of Naval Research Washington, D.C. 20360	
13. ABSTRACT (See Page 53 of text)		

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Transvariation theory						
Mathematical statistics						
Non-parametric methods						
Gaussian Transvariations						
Econometrics methods						
Comparative statics						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.