CHAPTER 5:

LOGICAL COMPARISON OF NONMETRIC TEMPORAL PATH ANALYSIS AND THE LINEAR MODELS APPROACH

Introduction

In Chapter 4 nonmetric temporal path analysis (NTPA) and the linear models approach (LMA) were compared empirically. The very same data were analyzed with NTPA, correlational analysis and analysis of variance (ANOVA), the latter two being fundamental parametric statistical procedures in the LMA. Other statistical procedures such as multiple correlational analysis, canonical analysis, discriminate analysis, multivariate analysis of variance, factor analysis, etc., were not compared to NTPA since these other procedures are extensions of correlational analysis and ANOVA. Indeed, it has been shown that correlational analysis and ANOVA are essentially the same, in that both are linear, additive models, as discussed in Chapter 1. The empirical comparison of NTPA and the LMA demonstrated that NTPA provided results not obtainable in the LMA, and that NTPA-aggregated data were amenable to the LMA, provided that one is willing to make the additional assumptions required for drawing statistical inferences in the LMA (i.e. primarily the assumption of determinism and the resultant modeling of a relation by a mathematical function).

Both NTPA and the LMA are methodological theories of measurement and verification of theoretical relations. They are theories about theory testing, or more precisely, theories of inductive inference, in contradistinction to deductive and retroductive inference (c.f., Steiner, 1978). The form of inductive inference is given:

1. A is true of b_1, b_2, . . ., b_n.
2. b_1, b_2, . . ., b_n are members of some class B.
3. Hence, A is true of all members of class B. (Steiner, 1978, p. 9)

The form of inductive inference is statistical. Both NTPA and the LMA aim for statistical generalization--i.e., inferring population parameters from sample data, although each can be used descriptively as well. These two approaches differ in the conception of A, above. In NTPA A is conceived as a nonmetric temporal path, whereas in the LMA A is conceived as a mathematical function. Both approaches assume that the b_i's are independent and representative of B. These assumptions are normally fulfilled through random selection of b's from population B. Thus, the essential difference between NTPA and LMA is the conception of A and the manner in which A is empirically verified.

The major question being addressed here is: Is NTPA more adequate than the LMA in the verification of stochastic educational relations? To answer this question, criteria are needed for comparison. Steiner set forth criteria for comparing two theories. One theory (T_1) or meta-theory is more adequate than another (T_2) iff:

1. T_1 is more complete than T_2. This is true iff: 1.1. T_2 is derivable from T_1, and 1.2. T_1 describes relations which are not described by T_2.

2. T_1 is inclusive of T_2. T_1 is reducible to T_2 iff: 2.1. The conjunction of T_1 and R includes T_2, where R is a set of translation rules matching expressions in T_1 and T_2.

3. T_1 is a strong alternative to T_2. This is true iff: 3.1. T_1 includes all those data, phenomena, events or event relations that T_2 includes; 3.2. T_1 and T_2 are empirically inconsistent; and 3.3. T_1 has higher empirical content than T_2. (Adapted from Steiner, cited in Maccia, 1974)

1. NTPA Is More Complete than the LMA

1.1. The LMA is derivable from NTPA. Since a deterministic process is a special case of a stochastic process, this claim should be self-evident. For example, suppose that two classifications are of interest. Let C1 describe academic instruction available to a student and C2 describe student orientation to instruction, where:

Classification	Category
C1. Academic instruction available	$c1_1$. Direct (DI) $c1_2$. Non-direct (NDI)
C2. Student orientation	$c2_1$. Engaged (EN) $c2_2$. Non-engaged (NE)

The Cartesian product of C1 and C2 results in the set of ordered pairs, $\{(c1_1,c2_1), (c1_1,c2_2), (c1_2,c2_1), (c1_2,c2_2)\}$. The probabilities of these ordered pairs may be conveniently represented by a table which illustrates the joint relation between C1 and C2:

	EN	NE	
DI	P(DI & EN)	P(DI & NE)	P(DI)
NDI	P(NDI & EN)	P(NDI & NE)	P(NDI)
	P(EN)	P(NE)	

If in any row of the table the $P(c1_i \& c2_j)$ is equal to one and all other cells in that row have zero probabilities, this would exemplify a deterministic process which can be modeled by the LMA. In other words, the cell probabilities can be deduced from the marginal probabilities even though the cell probabilities may not have been directly measured. In this case, both NTPA and the LMA will yield equivalent results. However, when cell values are not one, zero, or equal, they cannot be correctly deduced from the marginal values, as proved mathematically in Chapter 4. On the other hand, if the cell values are known, then the marginal values can always be determined. Since the LMA deals only with marginal values and NTPA deals with both cell and marginal values, it is patent that the LMA is derivable from NTPA.

1.2. NTPA describes relations which are not described by the LMA. The above example demonstrates that NTPA can describe relationships (i.e., joint or sequential) which cannot be described by the LMA. In particular, if relations are stochastic but nonetheless differentiatable, the LMA cannot verify them, whereas NTPA can.

In summary, NTPA is more complete than the LMA since NTPA can describe both stochastic and deterministic relations, whereas the LMA can only describe the latter.

2. NTPA is Inclusive of the LMA

2.1. The conjunction of NTPA and R includes the LMA, when R is a set of translation rules matching expressions in NTPA and the LMA. NTPA can be reduced to the LMA by applying the formulae of classical probability theory. See Chapters 2 and 4. However, the LMA cannot be reduced to NTPA, except in the special case where joint and/or sequential probabilities are zero, one, or equal. This was proved mathematically in Chapter 4. Put simply, marginal probabilities can be determined by adding corresponding cell probabilities. It is always possible to analyze NTPA aggregrated data with the LMA, providing the LMA assumptions are warranted and fulfilled. The converse does not obtain, except when relations are deterministic or equally indeterminate.

Therefore, it is patent that NTPA is inclusive of the LMA.

3. NTPA is a Strong Alternative to the LMA

3.1. NTPA includes all those data, phenomena, events or event relations that the LMA includes. In Chapter 4 it was shown that both NTPA and the LMA could be used to investigate the relationship between direct instruction and student engagement. Specifically, both appoaches provided estimates of the P(DI), P(NDI), P(EN) and P(NE).

3.2. NTPA and the LMA are empirically inconsistent. In Chapter 4 it was concluded from the LMA that the relationship between direct instruction and student engagement was significantly positive ($\underline{r}$ = .57), but that the proportion of direct instruction accounted for approximately one third of the variance in the proportion of student engagement. From NTPA it was concluded that a very strong positive relation existed between the occurrence of direct instruction and student engagement: $P(EN|DI) = .97$ on the average (S.E. = .003).

3.3. NTPA has higher empirical content than the LMA. This claim was also clearly supported by the results presented in Chapter 4. NTPA permitted analysis of relations that could not be performed in the LMA--e.g., $P(EN|DI)$, $P(EN|NDI)$, $P(NE|DI)$ and $P(NE|NDI)$.

Therefore, NTPA is a strong alternative to the LMA.

Conclusion

Since NTPA is more complete than the LMA, NTPA is inclusive of the LMA and NTPA is a strong alternative to the LMA, it is concluded that NTPA is more adequate than the LMA as a methodological procedure for verification of stochastic relations.

Summary

Extant statistical models commonly used in the social sciences are primarily linear and therefore deterministic. It is common in these models for factors to be independently measured and then related by some mathematical function which provides a good fit to the data--i.e., minimizes errors of prediction. Analysis of variance (ANOVA) and multiple correlational analysis (e.g., regression) are predicated upon an equation for a line surface of the general form:

$$Y = A + BX + E \quad [1]$$

Y is the measure of the variable to be predicted, X is the measure of the predictor, A is a constant term (Y intercept), B is the slope of the line, and E is error of measurement or residual. Empirical relationships are presumed to be linear (or curvilinear by extending the basic form of equation [1]). This is a linear models approach (LMA) to the measurement and verification of relations, since it fundamentally relies on the modeling of a relation by a line surface in n-dimensional space.

In ANOVA, eta-squared or omega-squared estimates the proportional reduction of error (PRE) in predicting the dependent variable(s) given knowledge of the independent variable(s) (Hays, 1973; Kirk, 1968). In correlational analysis, the square of the product moment coefficient is likewise an indication of the PRE. In multiple correlational analysis (regression, path, canonical, discriminate and factor analysis), PRE is similarly estimated (Kerlinger & Pedhazur, 1973).

An alternative approach to the estimation of a relation is to not constrain it by linear and metric assumptions. A relation need not be viewed as functional in the mathematical, set-theoretic sense--i.e., modeled by an equation for a line surface. In set theory, a relation is taken as the Cartesian product of two or more sets of elements. An ordered pair, or more generally n-tuple, symbolizes the specific joining of elements. In information theory, categories in a classification are analogous to elements in a set with the added condition that categories in a classification are mutually exclusive and exhaustive. Likewise, relations among categories can be modeled by the cross product of two or more classifications. By taking information as a characterization of occurrences, observed events or states of affairs can be mapped into categories in classifications (Maccia & Maccia, 1966). When time of occurrence is considered as well, such a mapping constitutes a nonmetric temporal path for each of the _n_-tuples of categories. The path is nonmetric since the mapping is not measurement in the usual sense of assigning a numerical value to a state of affairs--i.e., that the classification into which events are mapped is the set of integer or rational numbers. The path is temporal since simultaneity and sequence in time are also mapped.

When the occurrences of nonmetric temporal paths are observed and enumerated, their relative frequency and duration can be estimated. Since probability theory can be defined in terms of set theory, and information theory can be defined in terms of both probability and set theory, the uncertainty of occurrences of nonmetric temporal paths can be expressed in terms of probabilities. This procedure has been termed, 'nonmetric temporal path analysis' (NTPA). This NTPA methodology was explicated in detail in Chapter 2.

In summary, in the linear models approach (LMA) the strength of a relation is estimated by the linear association between two or more independently measured variables. In NTPA the strength of a relation is estimated by the probability of occurrence of a nonmetric temporal path which specifies the relation. The two procedures differ at the measurement level and in the manner in which the strength of the relation is empirically estimated.

The LMA appears to have evolved from Newtonian mechanics and the physical sciences in which deterministic assumptions and linear modeling have proved to be generally fruitful. Verificational procedures in the social sciences, including educational research, appear to have been adopted from the physical science paradigm. However, educational research involves the study of human beings in a teaching-studenting process. Educational systems are organismic, not mechanistic (Steiner, 1978). Organismic systems are stochastic, not deterministic--i.e., there is uncertainty of final states with respect to any given initial state. If relations among educational system components are indeed stochastic, then the LMA cannot adequately verify them, since it is predicated upon deterministic assumptions. NTPA can be used to verify stochastic relations. Thus, NTPA would appear to be more adequate than the LMA in the verification of stochastic relations.

Formal Definition of NTPA

Classification, C, is defined as a set of categories, c_1 through c_n, which are mutually exclusive and exhaustive:

$$C = \{c_1, c_2, \ldots , c_n\}$$

Categories c_i and c_j are exclusive iff:

$$c_i \& c_j = \{\Phi\} \quad \text{(for all i,j)}$$

Categories c_1, c_2, ... c_n are exhaustive iff:

$$c_i \cup c_j \cup \ldots c_n = \{C\}$$

The probability (or propensity) of c_i is defined:

$$P(c_i) = \frac{m(c_i)}{m(C)} \qquad [2]$$

The measure function, $m(c_i)$, is defined as the frequency of observed occurrences of events coded by category c_i in classification C. Note that:

$$m(C) = m(c_1) + m(c_2) + \ldots + m(c_n)$$

Also,

$$P(\sim c_i) = 1 - P(c_i)$$

$$P(c_i \ \& \ c_j) = 0 \quad \text{(by definition)}$$

$$P(c_i \cup c_j \cup \ldots c_n) = 1 = P(C)$$

$$P(c_i \cup c_j) = P(c_i) + P(c_j)$$

The probability of c_i <u>then</u> c_j is defined:

$$P(c_i, c_j) = \frac{m(c_i, c_j)}{m(c_i)} \qquad [3]$$

The measure function, $m(c_i, c_j)$ is defined as the frequency of observed occurrences of event patterns of the form 'IF c_i, THEN c_j', where c_i occurs at time t_1 and c_j occurs at time t_2 and $t_1 < t_2$, and no other c_n in classification C occurs during the interval t_1 through t_2.

In general:

$$P(c_i, c_j, \ldots, c_m, c_n) = \frac{m(c_i, c_j, \ldots, c_m, c_n)}{m(c_i, c_j, \ldots, c_m)} \qquad [4]$$

Assume classifications C1, C2, ... Cn, where $c1_i$ is a member of C1, $c2_i$ is a member of C2, and cn_i is a member of Cn. The probability of $c1_i$ and $c2_j$ is defined:

$$P(c1_i \ \& \ c2_j) = \frac{m(c1_i \ \& \ c2_j)}{m(C1_i \ \& \ C2_j)} \qquad [5]$$

The measure function, $m(c1_i \ \& \ c2_j)$, for joint classification is defined as the frequency of the observed joint occurrences of events coded as $c1_i$ and $c2_j$, where $c1_i$ begins at time t and $c2_j$ is also occurring at time t. Note also that:

$$m(C1_i \ \& \ C2_j) = \Sigma_i \Sigma_j \ m(c1_i \ \& \ c2_j).$$

The $P(c1_i \ \& \ c2_j \ \& \ \ldots \ cn_m)$ can be likewise defined by extending the above definition for two classifications to n classifications.

The time measure function, $tm(c_i)$, is defined as the total duration of observed events coded as c_i in classification C. The duration of c_i, where c_i begins at t_1 and ends at t_2 is defined:

$$d(c_i) = t_2 - t_1.$$

Thus, the total duration of c_i over a period of observation T is defined:

$$tm(c_i) = \Sigma_t \; d(c_i)_t.$$

Note that $tm(C) = T$ = total duration of the observation = $tm(c_1) + tm(c_2) + \ldots + tm(c_n)$. It should be also noted that the above definitions of probabilities of nonmetric temporal paths, based on frequency measure functions, can be likewise used for estimating probabilities based on time measure functions, simply by substituting the time measure functions (tm) for the frequency measure functions (m) in equations [2] through [5]. Thus, probabilities may be estimated in NTPA by either relative frequency, relative duration, or both, depending on the nature of the inquiry.

An event, $E(S_t, C)$, is defined as a change of state of system S relevant to classification C at time t. A joint event, $E(S_t$, C1 & C2 & ... Cn) is defined as a simultaneous change of one or more states of system S relevant to classifications C1, C2, ... Cn at time t.

Systematic observation of S is the mapping of system events relevant to n classifications into their respective categories as they are observed to occur in time. Hypotheses take the form of queries, specified as nonmetric temporal paths, to be verified by analysis of observational data. The results of queries are in the form of estimated probabilities (propensities, likelihoods) of a system's processes. Thus, hypotheses are assumed to be probabilistic and reflect the likelihood of the occurrence of an individual system's processes and/or system-environment transactions, estimated by measures of relative frequency, relative duration, or both. Generalizations across systems of a given type can be made if appropriate sampling strategies are employed. If the systems are independent, then the probability measures derived from observations of each system can be averaged; a population mean and confidence interval can be estimated for the specified nonmetric temporal path by using extant procedures of statistical inference. The unit of analysis on which this mean is based is the measure of the probability of the specified process in each individual system sampled.

General Remarks about NTPA

The measurement theory implicit in NTPA is standard, with the exception that event relations are enumerated. The path (process, pattern) which indicates the relation among factors of interest is itself nonmetric and temporal--hence, the name, 'nonmetric temporal path analysis'. However, the resulting frequency or duration measure derived from observations of a nonmetric temporal path is metric. The measure derived from observations of the relation is numerical, but the relation itself is not. NTPA differs from the LMA in measurement of relations, where the parts of the relation are measured separately in the LMA and the strength of the relationship is estimated by a statistical measure of association. In short, the LMA relates the measures by a linear or curvilinear function, whereas NTPA measures the relation in terms of the uncertainty of its occurrence, expressed as a probability.

The unit of measure in NTPA is based on the occurrence of events or patterns of events characterized by categories in classifications during systematic observation. To obtain a frequency measure, each event or pattern is assigned a weight of one. Thus, a frequency measure is simply an enumeration of occurrences of events or event patterns. This is no different than the measurement of length, for example, where the units of measurement (e.g., inches) are counted when making an observation of the length of some object. The resulting measurement is the frequency of the units of measurement observed. Duration is measured in NTPA using conventional units of time (seconds) elapsed from the beginning of an event or event pattern to its end.

It should also be noted that the formula for conditional probability does not in general hold for NTPA relative frequency estimates, since there is not necessarily a one-to-one correspondence between event occurrences in different classifications. In other words, the $P(c1_i, c2_j)$ is not equivalent to the $P(c2_j | c1_i)$. However, conditional probabilities may legitmately be constructed from NTPA probabilities based on time measure functions. For example, the $P(c2_j | c1_i)$ can be determined by dividing the $P(c2_j \; \& \; c1_i)$ by the $P(c1_i)$, but only when the latter two probabilities are based on the NTPA time measure (tm) function. Finally, conditional probability should not be conflated with the probability of a sequential occurrence. Conditional probability depends on joint occurrence—it does not matter whether $c1_i$ or $c2_j$ begins first, insofar as both can co-occur. However, the order of occurrence is patently relevant for estimating the probability of a temporal sequence (e.g., $P(c1_i, c2_j)$), regardless of whether a frequency or time measure function is used as a basis of estimation.

An Empirical Comparison of NTPA and the LMA

To illustrate the difference between NTPA and the LMA in measurement and verification of relations, an empirical example was provided in Chapter 4 utilizing data from an observational study of the academic learning time of mildly handicapped students. Twenty-five classrooms were observed for a total of eight to ten hours each over a period of about six months by highly trained observers using the Academic Learning Time Observation System (ALTOS: Frick & Rieth, 1981). For purposes of illustration, only two classifications were discussed, each consisting of two categories.

Classification	Categories
C1. Academic instruction available	$c1_1$. Direct (DI) $c1_2$. Non-direct (NDI)
C2. Student orientation to task	$c2_1$. Engaged (EN) $c2_2$. Non-engaged (NE)

Direct instruction (DI) refers to academic transaction with the target student or group of students of which the target student is a member during an educational activity. From the point of view of a given target student, the source of direct instruction could be the teacher, another person in the classroom such as a peer or an aide, or some other source capable of sending information to <u>and</u> receiving information from that target student (e.g., computer-assisted instruction). If there was no academic transaction with the target student or group containing that student during an educational activity, then the type of instruction was considered to be non-direct (NDI). A target student was considered to be engaged (EN), if she or he appeared to be attending to the academic substance of an educational activity. If not, then she or he was considered to be non-engaged (NE). It should be noted that each of the twenty-five target students was taken as an independent system whose environment was his or her classroom. Thus, the observer's task was to characterize, using categories from the ALTOS, the simultaneous behavior of the system and the system's environment. Both NTPA and the LMA were used to analyze the relationships between system and system environment, for purposes of comparison. Proportion of time was used to estimate the probability of occurrence of categories and selected conditional relations among categories.

LMA results. In the linear models approach (LMA), DI and EN are each considered as random variables. Note that NDI and NE provide no additional information, since NDI and DI, and NE and EN, respectively, are ipsative. The mean P(EN) was .741 (S.D. = .101), and the mean P(DI) was .432 (S.D. = .144). The relation between DI and EN, estimated by the Pearson product moment coefficient, was .57 and statistically significant ($p < .05$). Thus, there appeared to be a positive linear relationship between DI and EN, significantly greater than zero, and represented by the linear regression equation, P(EN)' = .57 + .40P(DI). The standard error of beta was .12. Knowledge of the proportion of DI reduces the uncertainty of the prediction of the proportion of EN by about 32 percent ($r^2 = .32$). It would appear that, although positive, the relationship between DI and EN is not very strong in terms of percent of variance accounted for (i.e., proportional reduction of error). About twice as much variance is unpredictable as is predictable. The relation is symmetrical, since $r_{EN,DI} = r_{DI,EN}$. Thus, the percent of variance of DI accounted for by EN is also 32 percent. In addition, since the sample size is relatively small, the beta weight for the linear regression equation has a relatively large standard error. One would have little confidence in estimating the population beta with such a small sample.

NTPA results. The very same observational data for the 25 systems were also analyzed by the NTPA approach. Joint and conditional distributions were formed for DI & EN, DI & NE, NDI & EN, NDI & NE, EN|DI and EN|NDI. The <u>relation</u>, EN|DI, was considered to be a single random variable. For each system a measure of EN|DI was constructed from the observational data, estimating the probability of student engagement, given that direct instruction was also occurring. The proportion of time that this nonmetric temporal path was observed for each system estimates the probability of the relation for that system. When averaged across independent systems, the mean P(EN|DI) was .967 (S.D. = .029). In other words, given that DI is occurring, students are very likely to be engaged in the educational activity. The odds that students are on-task compared to being off-task during direct instruction are about 19 to 1. On the other hand, in the absence of DI, students are engaged about 57 percent of the time [P(EN|NDI) = .573, S.D. = .142]. A further interpretation is that students are about 13 times more likely to be off-task (NE) when no DI is provided, compared to being off-task when DI is provided.

The relation between DI and EN is asymmetrical, since the P(EN|DI) is not equal to the P(DI|EN): .967 $\neq$ .561 . Furthermore, the standard error of the mean P(EN|DI) is relatively small (.003). The 95 percent confidence interval for the mean P(EN|DI) is .955 to .979. Since the sampling distribution is negatively skewed for a mean proportion which is this close to 1.0, a different method of determining the confidence interval is probably warranted, such as Tchebycheff's inequality, or use of the arcsine transformation, prior to interval estimation.

Discussion

One may argue, such as has Cronbach (personal communication), that NTPA represents a special case of analysis of variance (ANOVA), where the design is sliced in an unconventional manner (i.e., data are aggregated according to NTPA specifications). Hence, the data may be treated descriptively, not experimentally, as a repeated measures ANOVA. While it is true that NTPA more closely resembles ANOVA than it does correlational analysis, the latter two are constrained by assumptions not made in NTPA. Both ANOVA and correlational analysis are predicated upon linear equations (e.g., [1]), hence requiring deterministic assumptions. Such assumptions are not made in NTPA. Since a deterministic relation is a special case of a stochastic relation, NTPA aggregated data can always be subjected to the LMA, providing that deterministic assumptions are warranted and associated statistical assumptions are met. However, the converse does not obtain. Data collected from a linear models perspective cannot ordinarily be subjected to NTPA. This was proved mathematically in Chapter 4. Joint probability estimates are not derivable from singular (marginal) probability estimates, except where the relation is deterministic or equally indeterminate. Sequential probability estimates are likewise not derivable from marginal probability estimates.

NTPA is more adequate than the LMA, since NTPA is more complete than the LMA, NTPA is inclusive of the LMA, and NTPA is a strong alternative to the LMA. NTPA is more complete than the LMA, since the LMA is derivable from NTPA and NTPA describes relations not describable by the LMA. NTPA is inclusive of the LMA because NTPA results can also be treated in the LMA, but not the converse. NTPA is a strong alternative to the LMA.

Both may be used to investigate the same phenomena, but the results of NTPA and the LMA are empirically inconsistent, and NTPA has higher empirical content than the LMA. The latter was exemplified with data from an observational study. In the linear models approach, the strength of the relationship between direct instruction and student engagement was estimated to be .32, using the square of the product moment coefficient, and .79, using eta squared in ANOVA. In NTPA, the probability of student engagement during direct instruction, $P(EN|DI)$, was estimated to be .97, while the probability of student engagement during no direct instruction, $P(EN|NDI)$, was estimated to be .57.

Finally, if a general systems view of education is taken, then the fruitfulness of the NTPA approach to educational inquiry is patent. If the linear models approach has indeed obfuscated the development of knowledge about education through inquiry, then a more adequate research methodology is likely to result in an increase in the rate of development of knowledge of educational systems processes and system-environment transactions. It has been demonstrated that nonmetric temporal path analysis is more adequate than the linear models approach, the latter being the predominate methodology utilized in extant educational research.