Australian Journal of Basic and Applied Sciences, 5(5): 1094-1107, 2011
ISSN 1991-8178

Sequential Statistical Analysis Approach (SSAA) Towards Contingency Framework Purification in Behavioral Research and Practice

Iraj Malakmohammadi

Department of Agricultural Extension Education, University of Tehran.

Abstract: Theoretical frameworks are common expected outcomes of the qualitative part of mixed method studies on the basis of which conceptual frameworks and structural equation models are being designed quantitatively. Usually, an amalgamation of vague theoretical models, a considerable number of variables (observed and latent), intervening variables, and inappropriate statistical procedures and tests bring behavioral research to nowhere. Researchers must be able to use statistics effectively to organize, evaluate, analyze data, and apply the appropriate statistical tests. This article introduces a "50-step strategic road map" consisting of three sequential compartments: initial, intermediate (inferential), and advanced (modeling) along with their components. This sequential statistical approach assists young researchers and graduate students by meeting their basic need for appropriate statistical tests and by generating much more realistic outcomes.

Key words: theoretical framework, statistical test, structural equation modeling, sequential statistical analysis

INTRODUCTION

Since the time Fincher (1991) stressed that 'research on the substantive issues and concerns of higher education is handicapped by higher education's lack of status and recognition as an academic discipline and/or a professional specialty', the issue of understanding research as a disciplined inquiry is still a stealth, especially in higher education.

Research is a disciplined inquiry that can take numerous forms. Statistical analysis is relevant only for those researches where the information collected is represented by numbers. Ramanathan (2002) and Whiney (2005) stress that numerical information is called data and the sole purpose of statistics is to manipulate and analyze the data. Statistics, then, is a set of mathematical techniques used by social scientists to organize and manipulate data for the purpose of answering questions and testing theories. However, there should be some systematic way of organizing and analyzing data and, at this point, statistics becomes very helpful. But, as indicated by Windish and Diener-West (2006) and Govindrajulue (2004), there are a few, if any, references to the use of sequential statistics in the literature. Although choosing the right statistical test for a particular set of data appears to be an overwhelming task, to Wheater and Cook (2000), particularly if such decisions are rendered after the data are collected, what is overwhelming really is the sequences and placements of statistical tests to understand their role and mission in the first place. Wheater and Cook (2000) believe that the investigator is definitely responsible for the choice of statistical methods used. Therefore, the researcher must be able to use statistics effectively to organize, evaluate, and analyze the data (Whitney 2005) and to apply the proper statistical tests.

To ease the dilemma, it is helpful to identify the statistical test, as stated by Hoffman (2004), which is a procedure for deciding whether an assertion (e.g., a hypothesis) about a quantitative feature of a population is true or false. There are a few cautionary steps to follow in selecting a statistical test in educational research first, because of the high number of variables involved and second, because of the involvement of a considerable number of latent (unobserved) (Vermunt and Magidson 2003), hidden (nominal variable in a regression is the nominal variable that groups together two or more observations) (Moyulsky 1995), and discrete (variable that cannot take on all values within the limits of the variable) variables. There is a high risk for errors. To avoid type I and type II errors, a statistical test can be applied when it is robust (i.e., strong), eligible (i.e., the right test in the right place), appropriate (i.e., at the right position regarding the research purpose), and suitable to the research design. Also, the statistical test should be powerful (i.e., the probability that a test will produce a significant difference at a given significance level). The power of the test is equal

Corresponding Author: Iraj Malakmohammadi, Department of Agricultural Extension Education, University of Tehran.
E-mail: amalek@ut.ac.ir

to the probability of rejecting the null hypothesis when it is untrue i.e., making the correct decision. It is 1 minus the probability of a type II error. The true differences between the populations compared, the sample size, and the significance level chosen affect the power of a statistical test.

According to Dixon (2003), in whatever test you choose, it is important to think about and justify the choice. That justification can be as simple as "I didn't see any complications in the data".

Watt and Berg (2002) stress that the choice of the correct statistical test depends on the definition of the variables, particularly upon their level of measurement. It also depends on the research design used and the nature of the hypotheses: are they comparative or related; is there more than one independent variable?

Considering the number of variables is also critical in the choice of statistical tests. There are a remarkable number of latent variables involved in educational research along with explicit variables; many independent variables are not independent of one another in the real situation; and, more importantly, many researchers do not really know how to select and apply appropriate statistical tests to handle these variables accurately. Therefore, the main purpose of this article is to offer a procedure that will first refine variables and then identify sustained variables reasonably contributing to the research pattern or model as the limitations of each statistical test do not permit researchers to apply many variables they prefer with no statistical logic.

The main problem in conducting MS or PhD and even other onward academic or institutional research is that researchers either do not feel their crucial role in choosing and applying appropriate statistical tests or they are not fully aware of the process of choosing and applying the right test for the right situation (variable combination). Consequently, choosing and applying the statistical tests seems to become a mere copy-and-paste process for which no initial arguments and intellectual reasoning are offered.

Taking off from what Healey (2005) has revealed about the wheel of science being made of four initial parts (theory, hypotheses, observations, and empirical generalization) and using these to respond to an aforementioned problem, the author prepared this analytical article by focusing on the sequential statistical analysis approach (SSAA) to assist researchers in the social, educational, and behavioral fields to, first, purify and, second, promote their research findings. As shown in Fig. 1, while SSAA is central to the model to manage the circulation of the wheel of science, a qualitative or quantitative variable refining process is done at every two initial stages of the science wheel to identify the appropriate and eliminate the inappropriate variables that are not directly measurable (latent variables) in the process of generating knowledge. Another aim is to identify variables that have undesirable impacts or effects on the process and modify these according to certain criteria.

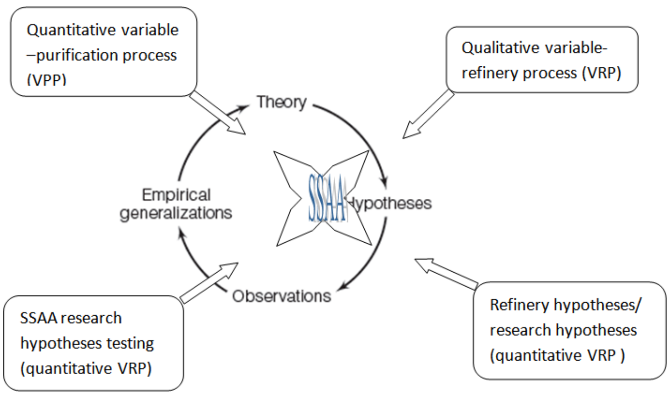

Fig. 1: Wheel of science (adapted from Healey 2005).

Joining the above discussion, Watt and Berg (2002) suggest that answers to the following six questions be given in order to identify the correct statistical test:

1. How many independent variables covary with the dependent variable?
2. At what level of measurement is the independent variable?
3. What is the level of measurement of the dependent variable?
4. Are the observations independent or dependent?
5. Are you comparing populations to populations, a sample to a population, or two or more samples?
6. Is the hypothesis being tested a comparison or a relationship?

It can be assumed from Watt and Burg's inquiry list that they were basically focusing on predictive and comparative research schemes and here their inquiries stop; obviously, this is not enough when what the researcher wants is to investigate the accuracy of the hypothetical multivariable model developed for the research. Complementarily, modeling or constructive research schemes require adequate responses to a few more leading questions as follows:

7. How many variables explicit and latent are actually involved in the research?
8. How can interfering (unwanted) variables be identified and eliminated from the study?
9. Are any premade latent variables already identified to be measured and/or being identified and measured as combined new variable/s?
10. Is there a set of more than one dependent variable being predicted from a set of more than one independent variables?
11. Are respondents or variables being grouped?
12. Is any hypothetical model being tested?

This article introduces a 50-step strategic roadmap to sequential statistical analysis approach (SSAA), which consists of the following three sequential compartments and their components that will enable one to respond to the above basic questions. This will guide researchers in choosing the appropriate statistical test for predictive, comparative, and modeling research and in constructing much more realistic model/s with certain variables and specific arrangement. Note that, in this article, "sequential statistical analysis" indicates the application of appropriate statistical methods in sequence (from one threshold to the next) from the very beginning of the research project to its last stage, which is quite different from "sequential procedures" that Clifton (1967) and Govindarajulu 2004 have clearly identified i.e., " a procedure where the experimenter has the option of looking at the sequence of observations one at a time and decides whether to stop sampling and take a decision or to continue sampling and make a decision later."

The three SSAA compartments and their stages are as follows:

A. Initial phase
1. Variable mining and measurement
2. Variable reduction (refinery)
3. Variables or respondents grouping (exploratory factor analysis [EFA])

B. Intermediate (inferential) phase
1. Variables and groups identification
2. Hypotheses development
3. Hypotheses testing (choosing the appropriate statistical test)

C. Advanced (modeling) phase
1. Regression (multiple and multivariable)
2. Structural equation modeling (SEM), including path analysis and confirmatory factor analysis

Each one of the stages in the above phases is composed of a few steps through which general and specific criteria for selecting and applying statistical tests are being discussed. This paper is intended to be a concise educational tool and guide for choosing the right statistical test/s and generating more realistic models for use in behavioral research generally, and, within that, more specifically, for use in agricultural extension education.

A. Initial Phase:
1. Variable Mining and Measurement:

This entails listing of variables involved in the study and measuring them after a scrutiny of some general research notions as research problem, research question, nature of the data, and research design.

Causes associated with the research problem. Choosing a statistical test depends on the number of variables (dependent/ independent) involved in predictive or comparative research project.

Reality in the public sector is complex. Often, there may be several possible causes associated with a problem; likewise, there may be several factors necessary for a solution. Complex statistical applications are needed that can

- deal with interval and ratio level variables,
- assess causal linkages, and
- forecast future outcomes.

In fact, the answers to questions of comparison and association are provided by quite different statistical tests. Ordinary least square linear regression is the most widely used type of regression for predicting the value of one dependent variable from the value of one independent variable. It is also widely used for predicting the value of one dependent variable from the values of two or more independent variables. When there are two or more independent variables, it is called multiple regression. In addition to this, there are a considerable number of factors involved in selecting appropriate and legitimate statistical tests. Some of the more general and specific ones are identified below.

Kind of research. To justify a test (choosing the right test), a researcher should ask himself/herself two questions: what kind (type) of data has been collected (do I have)? What are my objectives (goals)? (Moyulsky 1995, Wadsworth 2005, Dinove 2008).

Some research is conducted to describe a phenomenon. This sort of research will tend to involve the use of descriptive statistics. Other researches are exploratory. With this sort of research, the researcher is likely to be searching for patterns and relationships in the data. Statistical procedures that reveal patterns or automatically try to fit models of relationships are often useful for this kind of research. Finally, a great deal of research involves testing hypotheses about patterns of causation in the population. For this sort of research, to enable one to infer from his/her population data, procedures that use significance tests must be employed (Bruin 2006). Moreover, in the case of exploring relations between variables, there are two options. When there is no dependent variable and the researcher intends to measure the association between variables, then correlation coefficients can be applied. But, when prediction is intended and at least one dependent variable is involved in the study, then regression analysis would be appropriate.

Research question. As Bruin (2006) has indicated, the analysis must relate to the research question and this may dictate the techniques to be used. To match statistics with research questions, Marion (2004) reveals that the domain of research designs is divided into three categories of research questions: descriptive, differences, and relationship. A descriptive research question seeks to identify and describe some phenomenon. A difference research question asks if there are differences between groups on some phenomenon. A relationship question asks if two or more phenomena are related in some systematic manner.

The aim of the analysis. In social-behavioral research where many variables are involved, it would not be that easy to choose suitable statistical tests to manipulate all the variables. Therefore, by applying EFA, the researcher will have the chance to reduce the number of variables involved to a few defined variable by applying vertical EFA. But, if the researcher is to identify a theoretical variable, then he may utilize a non-vertical EFA.

Kind of variable. The choice of a statistical test depends on the nature of the variable: whether it is continuous/discrete, measurement/categorical (Wheater and Cook 2000).

Variable measurement. Statistical tests are specifically designed to be used on variables measured at a certain level and it is therefore essential that, when choosing a statistical procedure, you are certain at which level the variable you intend to analyze is measured. There are four levels of measurement: nominal, ordinal, interval, or ratio (Kaminsky 2008).

Once we have identified the independent and dependent variables, our next step in choosing a statistical test is to identify the scale of measurement of the variables. The scale of measurement of the independent variable helps us to determine which statistical procedure within the broad category is appropriate (Windish 2006).

2. Variable Reduction (Refinery):

Statistical tests and procedures, specifically, are divided according to the number of variables that they are designed to analyze. Therefore, when choosing a test, it is important to consider how many variables are being analyzed. One set of tests is used on single variables (often referred to as descriptive statistics), a second set is used to analyze the relationship between two variables, and a third set is used to model multivariable relationships (i.e., relationships between three or more variables). To be specific, the number of dependent variables (DV) and independent variables (IV) also matter in choosing appropriate statistical tests very seriously. Before choosing and applying a statistical test, the researcher should be confident about the true number of variables (DVs and IVs) on one hand and the strength of the statistical test on the other hand.

To find out the true number of variables, two strategic key issues must be identified. The first is the number of IVs (causes) initiating the research problem and the second is the level (if there is one DV) and/or number of DVs affected by the IV/s. There are a few variable reduction procedures to be implemented by the researcher so he/she can come up with the optimum IVs in the research, prior to involving all variables in the hypotheses-testing process. Some of these procedures are as follows.

Validity and reliability. Although validity and reliability tests aim to legitimize research instruments (in terms of the most commonly used questionnaires), here in this article, they are being introduced as the initial stage of the variable refinery process. It is commonly known that validity is not measurable; rather it is technically recognized and practically approved through an actual measurement of the variable in a real sense and actually realized by experiencing it in the research measurement and data process mechanism. There are a few ways of measuring the reliability of a variable–e.g.,the Kuder Richardson reliability tests and Cronbach's alpha coefficient. Although Cronbach's alpha provides a diagnostic list of variables to help researchers eliminate certain variables with low coefficients in order to raise the overall correlation coefficient of the items (variables) involved in a measurement, but more specifically to simultaneously assess and control acquiescence and social desirability in the questionnaire items (Ferrando 2009), a more sustained atmosphere is created whenstrategic decisions are made as to whether to retain or eliminate these variables or questions. This can be considered the first stage of the variable refinery process (VRP).

Coefficient of variability (CV). The second stage in VRP is CV application. The CV is usually calculated by dividing standard deviation over the sample mean of the variable for different purposes such as finding diversity, consistency, accountability, priority setting, or even ranking variables. For the first purpose, the lower the CV, the more diverse the variable, and for the second and third, the lower the VC, the more consistent and accountable the variable in the research can be. Strategically, in plant and animal breeding, variables with the highest CV are interesting to the researcher, but, in contrast to behavioral research, variables with higher CVs may be eliminated based on the preference of the researcher and research status quo, while variables with lower CVs will be retained.

Correlation matrix: The third stage in the VRP process involves the correlation matrix. At this stage, those variables with no significant correlation coefficient in the matrix can be left out of the research process. Consequently, variables with statistically significant correlation coefficient may be more legitimate and subject to further investigations in the research process, although there is no guarantee that these variables will have a statistically significant impact or effect on the dependent variable/s in the regression process(see the fifth stage of SRCIST as explained below).

3. Variable or Respondent Grouping (EFA):

Exploratory factor analysis (EFA). Factor analysis, through which some specific variables contribute to create specific factors, is identified as the fourth stage in the VRP. In using this technique, the goal is to represent those things that are related to one another by a more general term such as a factor (Salkind 2008). Each factor represents several different variables and the factors turn out to be more efficient than individual variables at representing outcomes in certain studies. In appreciating factor analysis, the variables that do not significantly contribute to building up factors can be eliminated from the variable list. There are two purposes in the factor analysis system: either to reduce the number of variables, for which vertical factor analysis is being applied, or to make theoretical variables by applying non-vertical factor analysis. Both methods can be used by researchers simultaneously but separately. What is most common in behavioral research is vertical factor analysis, either to validate the research instrument or just to group some variables without any analytical intention (due to the nature of factor analysis that is not applicable in hypotheses testing). However, factor analysis can be applied to explore latent variables and their components or create factors as newly combined variables.

B. Intermediate (inferential) Phase:

1. Variables and Groups:

Independent/dependent or exogenous/ endogenous variables. As Hill and Lewicki (2007) observed, the terms 'dependent' and 'independent' variable apply mostly to experimental research where some variables are manipulated. In this sense, they are "independent" from the initial reaction patterns, features, intentions, etc. of the subjects. Some other variables are expected to be "dependent" on the manipulation of experimental conditions–hat is to say, they depend on "what the subject will do" in response. Somewhat contrary to the nature of this distinction, these terms are also used in studies where we do not literally manipulate independent variables but only assign subjects to "experimental groups" based on some pre-existing properties of the subjects. Also, in terms of data in a research, they are either independent (as required in the analysis of variance) or dependent (as required in multiple regression analysis). Then, the test used should be determined by the data. The choice of test for matched or paired data (Kaminsky 2008) is important in choosing the appropriate statistical test.

To return to the naming of the variables, Streiner (2005) indicates that, in path analysis (and in its more sophisticated counterpart, structural equation modeling), we completely avoid the confusion about IVs and DVs by the simple expedient of not using those labels. Rather, we use the terms exogenous and endogenous variables:

Exogenous variables have straight arrows emerging from them and none pointing to them (except from error terms).

Endogenous variables have at least one straight arrow pointing to them.

The rationale for these terms is that the causes of (or factors that influence) exogenous variables are determined outside the model that we are examining, whereas the factors affecting endogenous variables exist within the model itself.

Number of groups being compared. Usually statistical tests are classified based on the number of groups to be compared. That is, when two groups are being compared, then a group of statistical tests would be helpful in looking at other criteria for choosing the appropriate statistical tests. But, when there are more than two groups to compare, then another group of statistical tests should be applied. Unfortunately, in majority of the cases, graduate students do not consider this role and, unintentionally, apply inappropriate tests.

Kind of groups being compared. What is basically and commonly neglected in graduate research is that all groups involved in the study are being considered as independent groups, while all or some of them may be dependent in their nature. Actually, there are two possibilities in differentiating groups under study–as either independent of or dependent on each other. When comparing groups which are dependent on one another then, a specific series of statistical tests can be recommended. Other than that, when groups are independent of one another, some other statistical tests would be suitable to do the comparison.

2-Hypotheses Development (Null and Research Hypotheses) and Hypothesis Testing:

Main hypothesis. In response to the question of 'what is a hypothesis,' Graveter and Forzano (2008) explain that, in the context of science, a hypothesis is a statement that describes or explains a relationship between or among variables. A specific testable prediction that is driven from the hypothesis is called a research hypothesis. Also, a statistical hypothesis test is defined by Lehmann and Romano (2005) as a method of making statistical decisions using experimental data. In terms of selecting a statistical test, the most important question is "what is the main study hypothesis?" In some cases, there is no hypothesis; the investigator just wants to "see what is there." For example, in a prevalence study, there is no hypothesis to test, and the size of the study is determined by how accurately the investigator wants to determine the prevalence. If there is no hypothesis, then there is no statistical test. It is important to decide a priori which hypotheses are confirmatory (that is, are testing some presupposed relationship), and which are exploratory (are suggested by the data). No single study can support a whole series of hypotheses (Kaminsky 2008).

3. Hypotheses Testing (Choosing the Appropriate Statistical Test):

P value/effect size. One major issue in selecting a statistical test is the p value. If the it is significant, then the researcher usually does not care that some other test would give a smaller (more significant) p value. If the p value is not significant, then the researcher usually considers whether there is a better test (Philip Dixon 2003). This is often true with graduate students whose hypotheses are mostly rejected; they usually try some other statistical tests to possibly change their results. In this case,a type I error is made—obtaining a "significant" result when in fact there is no difference.

Perhaps the most commonly recommended alternative to solely interpreting p values is to determine the magnitude of effect, more commonly known as "effect size." An effect size measure is an indicator of the association that exists between two or more variables. An exception to this is Cohen's *d,* which is a measure of the *distance* between means. These definitions translate into how much variance in one variable is accounted for by knowledge of another variable (Denis 2003, McCloskey 2008, Graveter and Forzano 2008).

Sample size and complexity of data. Sample size, complexity of data, and number of observations are important in choosing the statistical test (NN 2007). Parametric methods and statistics rely on a set of assumptions about the underlying distribution to give valid results. In general, they require the variables to have a normal distribution. When sample size is small (below 30 observations), non-parametric tests (McDonald 2008), instead of parametric tests, are applicable.

When a study reaches a conclusion of "no statistically significant difference," it should not necessarily be concluded that the treatment was ineffective. It is possible that the study missed a real effect because either a small sample was used or the data were quite variable. In this case, a type II error was made—obtaining a "not significant" result when in fact there is a difference.

When interpreting the results of an experiment that found no significant difference, you need to ask yourself how much power the study had to find various hypothetical differences if they existed. The power depends on the sample size and the amount of variation within the groups, where variation is quantified by the standard deviation (SD) (Moyulsky 1995).

As adapted from Anderson-Cook and Dorai-Raj (2003), type I and type II errors are illustrated in Table 1, as well as the formulating power of a test statistic.

Table 1: Sources of type I and type II errors in hypothesis testing (adapted from Anderson-Cook and Dorai-Raj 2003).

State of nature		
Decision	Null is true	Alternative is true
Reject null	Type I error (α)	Correct decision
Do not reject null	Correct decision	Type II error (β)

Type I: rejecting the null hypothesis when the null hypothesis is true.
Type II: failing to reject the null hypothesis when the null hypothesis is false. The probabilities of these errors are defined as follows:

$$\begin{aligned} \alpha &= P(\text{Type I error}) &&= P(\text{rejecting } H_0 | H_0 \text{ is true}) \\ \beta &= P(\text{Type II error}) &&= P(\text{failing to reject } H_0 | H_0 \text{ is false}) \\ Power &= 1-\beta &&= P(\text{rejecting } H_0 | H_0 \text{ is false}) \end{aligned}$$

Power will be the probability of being in the lower right cell of the table above. For a fixed Type I error rate (α) the goal of constructing and testing a hypothesis is to maximize *Power*. The purpose of these applets is to show how α, β and *Power* are related to each other.

Central limit theorem. Based on the central limit theorem, as sample size increases, the shape of the sampling distribution of the test statistic tends to become normal, even if the distribution of the variable that is being tested is not normal.

In practice, this can be applied to test statistics calculated from more than 30 observations. The smaller the sample size, the less you can get out of your data. Standard error is inversely related to sample size, so the larger your sample, the smaller the standard error, and the greater chance you will have of identifying statistically significant results in your analysis (McDonald 2008).

Number of independent hypotheses or multiple comparisons. Interpreting an individual P value is easy. If the null hypothesis is true, the P value is the chance that random selection of subjects would result in a difference (or correlation or association...) as large as (or larger than) that observed in the study. If the null hypothesis is true, there is a 5% chance of randomly selecting subjects such that the trend is statistically significant. However, many scientific studies generate more than one P value. Some studies in fact generate hundreds of P values. Interpreting multiple P values can be difficult. If you test several independent null hypotheses and leave the threshold at 0.05 for each comparison, there is a greater than 5% chance of obtaining at least one "statistically significant" result by chance (Moyulsky 1995, Wadsworth 2005).

Paired or unpaired. When comparing two groups, it should be decided when to use a paired test. When comparing three or more groups, the term 'paired' is not apt and the term 'repeated measures' is used instead. In this case, an unpaired test to compare groups when individual values are not paired or matched with one another is required. Select a paired or repeated-measures test when values represent repeated measurements on one subject (before and after an intervention) or measurements on matched subjects. The paired or repeated-measures test is also appropriate for repeated laboratory experiments run at different times, each with its own control (Moyulsky 1995).

Parametric/nonparametric. Choosing the right test to compare measurements is a bit tricky, as you must choose between two families of tests: parametric and nonparametric. Many statistical tests are based on the assumption that the data are sampled from a Gaussian distribution. These tests are referred to as parametric tests. Tests that do not make assumptions about the population distribution are referred to as nonparametric tests (Motulsky 1995).

Choosing between parametric and nonparametric tests is sometimes easy. You should definitely choose a parametric test if you are sure that your data are sampled from a population that follows a Gaussian distribution (at least approximately) (Motulsky 1995, Dinove 2008, McDonald 2008). Once you assume that a population is distributed in Gaussian (bell shape), statistical tests let you make inferences about the mean (and other properties) of the population. Most commonly used statistical tests assume that the population is Gaussian (bell shape) (Moyulsky 1995).

A nonparametric test should be selected definitely in three situations:

- The outcome is a rank or a score and the population is clearly not Gaussian.

- Some values are "off the scale," that is, too high or too low to measure. Even if the population is Gaussian, it is impossible to analyze such data with a parametric test since you don't know all of the values. Nonparametric tests only know about the relative ranks of the values, it would not matter that you did not know all the values exactly.
- The data are measurements and you are sure that the population is not distributed in a Gaussian manner (Dixon 2003).

Choosing the appropriate statistical test. Table 2 summarizes the process of choosing appropriate statistical tests. This table was originally adapted by the UCLA Academic Technology Services, Statistical Consulting Group from Choosing the Correct Statistic developed by James D. Leeper.

However, in some cases, as having one DV and two or more IVs with independent groups and with ordinal or interval nature of DV, suitable statistical tests are lacking. Also, when there is one interval IV and one DV with interval and nominal nature in one case, and ordinal or interval nature in another case, correlation and nonparametric correlation are recommended, respectively, but this does not sound quite right because of very rare conditions that correlation may imply causation as explained by Huck (2009). Moreover, given the fact that, when there are no IVs in an analysis, obviously, there would not be any DVs too; the case is not considered as causal, canonical correlation and factor analysis when there are just two sets or two and more DVs in the study (asindicated at the end of this table, this does not suit this table and this place).

Table 2: Guidelines to choosing the correct statistical tests through application of SAS, Stata, and SPSS.

Number of dependent variables	Nature of independent variables	Nature of dependent variable(s)	Test(s)	How to SAS	How to Stata	How to SPSS
1	0 IVs (1 population)	Interval & normal	One-sample t-test	SAS	Stata	SPSS
		Ordinal or interval	One-sample median	SAS	Stata	SPSS
		Categorical (2 categories)	Binomial test	SAS	Stata	SPSS
		Categorical	Chi-square goodness-of-fit	SAS	Stata	SPSS
	1 IV with 2 levels (independent groups)	Interval & normal	2 independent sample t-test	SAS	Stata	SPSS
		Ordinal or interval	Wilcoxon Mann-Whitney test	SAS	Stata	SPSS
		Categorical	Chi-square test	SAS	Stata	SPSS
			Fisher's exact test	SAS	Stata	SPSS
	1 IV with 2 or more levels (independent groups)	Interval & normal	One-way ANOVA	SAS	Stata	SPSS
		Ordinal or interval	Kruskal Wallis	SAS	Stata	SPSS
		Categorical	Chi-square test	SAS	Stata	SPSS
	1 IV with 2 levels (dependent/matched groups)	Interval & normal	Paired t-test	SAS	Stata	SPSS
		Ordinal or interval	Wilcoxon signed ranks test	SAS	Stata	SPSS
		Categorical	McNemar	SAS	Stata	SPSS
	1 IV with 2 or more levels (dependent/matched groups)	Interval & normal	One-way repeated measures ANOVA	SAS	Stata	SPSS
		Ordinal or interval	Friedman test	SAS	Stata	SPSS
		Categorical	Repeated measures logistic regression	SAS	Stata	SPSS
	2 or more IVs (independent groups)	Interval & normal	Factorial ANOVA	SAS	Stata	SPSS
		Ordinal or interval	???	???	???	???
		Categorical	Factorial logistic regression	SAS	Stata	SPSS
	1 interval IV	Interval & normal	Correlation	SAS	Stata	SPSS
			Simple linear regression	SAS	Stata	SPSS
		Ordinal or interval	Nonparametric correlation	SAS	Stata	SPSS
		Categorical	Simple logistic regression	SAS	Stata	SPSS
	1 or more interval IVs and/or 1 or more categorical IVs	Interval & normal	Multiple regression	SAS	Stata	SPSS
			Analysis of covariance	SAS	Stata	SPSS
		Categorical	Multiple logistic regression	SAS	Stata	SPSS
			Discriminant analysis	SAS	Stata	SPSS
2 or more	1 IV with 2 or more levels (independent groups)	Interval & normal	One-way MANOVA	SAS	Stata	SPSS
2 or more	2 or more	Interval & normal	Multivariate multiple linear regression	SAS	Stata	SPSS
2 sets of 2 or more	0	Interval & normal	Canonical correlation	SAS	Stata	SPSS
2 or more	0	Interval & normal	Factor analysis	SAS	Stata	SPSS

C. Advance (Modeling) Phase:

According to Bartholomew (1997), a model is

- an abstraction of the real world in which the relevant relations between the real elements are replaced by similar relations between mathematical entities.
- a set of assumptions about the relationship between the parts of the system. Its adequacy is judged by the success with which it can predict the effects of changes in the system.

Any model describing human behavior should be formulated in stochastic terms. When it comes to the solution of the model, it may be described as using a deterministic approximation. The greater simplicity of the deterministic version of a model may also make it easier to grasp the nature of the phenomenon in question.

What is discussed, so far, constitutes the prerequisites for modeling technique, regardless of the model being stochastic or deterministic, in which purified and appropriate variables are being identified and fit into an equation (model).

The simplest model that is capable of prediction is multiple and multivariable regression.

1. Regression:

Calculate linear regressions only if one of the variables (X) is likely to precede or cause the other variable (Y). Definitely choose linear regression if you manipulate the X variable. It makes a big difference which variable is called X and which is called Y, as linear regression calculations are not symmetrical with respect to X and Y (Motulsky 1995).

While regression analysis is the simplest modeling procedure in analytical research, it is recommended here as the fifth stage of VRP due to its capability to identify and measure latent variable/s in the study through a mathematical model. Of course, variables (predictors) involved in predicting latent variables (indicators) can be factors, or original variables that are significantly involved in building up factors through the fourth VRP stage. Either one of these should be specified prior to regression analysis.

The general purpose of multiple regression (the term was first used by Pearson in 1908) is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable**.** Multiple regression is a seductive technique: "plug in" as many predictor variables as you can think of and usually at least a few of them will come out significant (Statistica 2008).

This is practically confirmed by the author when, regardless of the high number of predictors in the equation in many instances, usually, only five to seven variables come out of the calculations to be significant. However, it is possible that the independent variables could obscure each other's effects (Palmer ND). The number of independent variables in the equation should be limited by two factors. First, the independent variables should be included in the equation only if they are based on the researcher's theory about what factors influence the dependent variable. Second, variables that do not contribute very much to explaining the variance in the dependent variable (i.e., to the total R^2), should be eliminated (Palmer ND).

Many difficulties tend to arise when there are more than five independent variables in a multiple regression equation. One of the most frequent is the problem of two or more independent variables being highly correlated to one another. This is called multicollinearity. If a correlation coefficient matrix with all the independent variables indicates correlations of 0.75 or higher, then there may be a problem with multicollinearity.

When two variables are highly correlated, they are basically measuring the same phenomenon. When one enters into the regression equation, it tends to explain most of the variance in the dependent variable that is related to that phenomenon. This leaves little variance to be explained by the second independent variable. Signs of multicollinearity include the following:

1) None of the t ratios of the coefficients are statistically significant, but the F test for the equation as a whole is significant;
2) Adding an additional independent variable to the equation radically changes either the size or the sign (plus or minus) of the coefficients associated with the other independent variables.

If multicollinearity is discovered, the researcher may drop one of the two variables that are highly correlated, or simply leave them in and note that multicollinearity is present (Statistica 2008).

Definitely, therefore, the more careful the refined variable process, the lower the covariance between predictors, and hence, the higher the chance of having powerful variables to predict (explain the variance of the independent variable, otherwise the estimates of the regression line are probably very unstable and unlikely to replicate if one were to do the study over again). However, multiple regression functions as a mediator along the VRP chain of exploring, among many predictors, including considerable latent variables, those variables that can accurately predict variance in the dependent variable and let the accurate variables get into the main gate of SEM (i.e., path analysis).

A common factor for all these methods is that they share three limitations: (a) the population of a sample model structure (at least in the case of regression-based approach); (b) the assumption that all variables can be considered observable; and (c) the conjecture that all variables are measured without error, which may limit their applicability in some research situations.

2. Structural Equation Modeling (SEM):

Raykov and Markoulides (2006), to explain the main reason that SEM is widely used by researchers, indicate that “traditional regression analysis effectively ignores potential measurement error in the explanatory (predictor, independent) variables; as a consequence, regression results can be incorrect and possibly entail misleading substantive conclusions. Garson (2008) believes that “SEM grows out of and serves purposes similar to multiple regression, but in a more powerful way, which takes into account the modeling of interactions, nonlinearities, correlated independents, measurement error, correlated error terms, multiple latent independents each measured by multiple indicators, and one or more latent dependents, also each with multiple indicators. SEM may be used as a more powerful alternative to multiple regression, path analysis, factor analysis, time series analysis, and analysis of covariance. That is, these procedures may be seen as special cases of SEM, or, to put it another way, SEM is an extension of the general linear model (GLM) of which multiple regression is a part."

Haenlein and Kaplan (2004) too, by referring to Gefen, Straub and Boudreau (2000), name regression analysis as a first-generation technique, which analyzes only one layer of relationships among multiple independent and dependent variables. At the same time, they recommend SEM as a second-generation technique that allows simultaneous modeling of relationships among multiple independent and dependent constructs. Moreover, by referring to Diamantopoulos (1994), they stress that one no longer differentiates between dependent and independent variables but distinguishes between exogenous (variables not explained by the postulated model, i.e., independent variables) and endogenous (variables that are explained by relationships contained in the model) latent variables. In terms of the effect of the explanatory variables over the dependent variables, Raykov and Markoulides (2006) observe that SEM enables researchers to readily develop, estimate, and test complex multivariable models as well as to study both direct and indirect effects of variables involved in a given model. The combination of direct and indirect effects makes up the total effect of an explanatory variable on a dependent variable.

In terms of approaches, Haenlein and Kaplan (2004) have indicated two approaches for estimating the parameters of SEM, namely, the covariance-based approach and the variance-based approach, but Garson (2008) views SEM as a confirmatory rather than an exploratory procedure, using one of three approaches: strictly confirmatory approach, alternative model approach, and model development approach. And, in terms of types, Raykov and Markoulides (2006) consider four types of SEM: path analysis model, confirmatory factor analysis model, structural regression model, and latent change model.

The first and second models that are most commonly used by researchers are discussed below as a way of concluding this discussion.

Path analysis (PA). To Salkind (2008), path analysis basically examines the direction relationships through the postulation of some theoretical relationships between variables and then tests to see if the direction of these relationships is substantiated by the data. But, to Streiner (2005), path analysis is an extension of multiple regression. It goes beyond regression in that it allows for the analysis of more complicated models. In particular, it can examine situations in which there are several final dependent variables and those in which there are “chains” of influence, in that variable A influences variable B, which in turn affects variable C. Although, despite its previous name of “causal modeling,” Streiner does not believe in path analysis as to establish causality or even to determine whether a specific model is correct; rather, it can only determine whether the data are consistent with the model. However, it is extremely powerful for examining complex models and for comparing different models to determine which one best fits the data. As with many techniques, path analysis has its own unique nomenclature, assumptions. Regarding this, Salkin (2008) stresses that one of the most interesting uses of path analysis is a technique called SEM, which is used to present results in a graphical presentation of the relationship among all the different factors under consideration.

Potentially, it can be concluded, however, that path analysis, sequentially, infrastructures rather than prerequisite SEM, as does regression (multiple and multi-variable) analysis to PA.

Confirmatory factor analysis. Taking advantage of the model development approach, researchers can test the hypothesis of factors and factor loading through confirmatory factor analysis (CFA) as the threshold stage for SEM. The CFA is usually employed to examine patterns of interrelationships among several latent constructs. According to Raykov and Markoulides (2006), “no specific directional relationships are assumed between constructs, only that they are potentially correlated to one another". The starting point of CFA is a very demanding one, requiring that the complete details of a proposed model be specified before it is fitted to the data." The latter statement by Raykov and Maroulides was more clearly explained by Stapleton (1997) when he described CFA as "a theory-testing model as opposed to a theory-generating method like exploratory factor analysis. In CFA, the researcher begins with a hypothesis prior to the analysis. This model specifies which variables will be correlated with which factors, and which factors are correlated. The process of CFA is described, and it is emphasized that it is important to realize that more than one model may accurately describe the data and that a number of fit indices should be used to determine the fit of the various models. Methods that may increase the fit of the researcher's model to the data are described."

Strategic roadmap to sequential statistical analysis approach (SSAA) The bottom line of a realistic and truly applicable causal research is a justifiable model. The model would not be rationally justified without going through SSAA,that is, firstly, a variable is sustained in the SEM process when it is found to be valid (justified to be measured), consistent (very low coefficient of variability), and accountable (eligible to involve in modeling process) through the VRP (Malekmohammadi 2008).

Complimenting what is already provided in Table 2, theno sequence or hierarchical placement of statistical tests in this tableaims to fill up initial, intermediate, and advance compartments prior to applying SEM as the final step in SSAA processes in Table 3. First, this will diminishmisconceptions among young researchers and graduate students who are after accurate application of statistical methods, and second, this will lower their stat phobia by leading them towards a 50-strategic-sequential-statistic-roadmap for choosing and applying appropriate statistical tests, interpreting their findings, and implementing scientific analysis more realistically in a creative research enterprise.

Table 3: Strategic roadmap to sequential statistical analysis approach (SSAA)

A- INITIAL PHASE	
Variable mining and measurement process	
1-	Identify research problem/s
2-	Specify research question/s
3-	Articulate research objective/s
4-	Review related research and literature (RRRL)
5-	Provide theoretical contingency table (TCT) to show resources, issues, and their frequencies
6-	Select high frequency issues (HFI) in TCT
7-	Design theoretical framework (TF) embodying HFIs and their realized relations
8-	Configure specific research method and materials (RMM) to investigate and test TF
9-	Construct research instrument to collect data Variable reduction (refinery)
10-	Look at the validity of a measure(Turpin 2004).
11-	Test the reliability of RI (pilot research instrument) . Check the reliability of a measure (Turpin 2004), calculate Cronbach's α coefficient
12-	Eliminate items with low reliability coefficient, if applicable, from RI
13-	Define research population (RP) and sampling procedure
14-	Collect the data (Statistica 2008)
15-	Assess each variable separately first (obtain measures of central tendency and dispersion; frequency distributions; graphs); is the variable normally distributed? (Statistic 2008)
16-	Calculate the variables' coefficient of variability (CV)
17-	Eliminate variables with higher CV, if applicable
18-	Identify sustained variables in the study
19-	Develop a correlation matrix
20-	Eliminate variables with no significant correlation coefficient, if applicable
21-	State refinery hypotheses (test variables against sample's characteristics)
22-	Eliminate variables highly affected by sample characteristics, if applicable
Variable or respondent grouping process	
23-	Apply R-type Exploratory Factor Analysis (EFA) (either orthogonal type if factors are not correlated, or oblique type if factors are correlated) to find out hypothetical factors (latent variable/s) and variables capable of building factor/s
24-	Eliminate variables with lower than 1 eigen value
25-	Identify new grouped (factor/s) variables (basically latent)
26-	Compare factor analysis output with theoretical model to identify compatible variables
27-	Design conceptual model by compatible variables
B- INTERMEDIATE (INFERENTIAL) PHASE	
- Hypotheses development	
29-	State the research hypothesis (RH) (Statistica 2008)
30-	State the null hypothesis (NH) (Statistica 2008)
31-	Assess the relationship of each independent variable, one at a time, with the dependent variable (calculate the correlation coefficient; obtain a scatter plot); are the two variables linearly related (Turpin 2004, Statistica 2008)?are responding groups independent?
Variable and group identification	

Table 3: Continue

32- Identify variables' nature (scale) and role (IV/DV) and groups' essence (two/more than two and independent/dependent) Hypotheses testing (choosing appropriate statistical test)
33- Choose specific P value/s to test the null hypotheses by appropriate statistical tests
Choose appropriate statistical test for prediction and/or comparison based on the information in Table 1, corresponding to the null hypotheses in the study
34- Test the null hypotheses
C- ADVANCE (MODELING) PHASE
Regression (multiple and multivariable)
35- Design regression analysis (multiple and/or multiple) to measure indicator/s and latent variable/s
36- Identify independent variable/s and dependent variable in the hypothetical multiple regression equation
37- Calculate F statistic to realize the significance of the equation as a whole
38- Eliminate equation/s with no significant F value (for the whole regression equation)
39- Calculate and examine appropriate measures of association and tests of statistical significance for each coefficient (Statistica 2008)
40- Eliminate predictors with no significant R value (when F value for the equation is significant)
41- Regress each explanatory variable against a constant and the remaining explanatory variables. There should be k–1 values for VIF. If any of them is high, then MC is indicated. It can be concluded that the higher VIF or the lower the tolerance index, the higher the variance of $\hat{\beta}_i$ and the greater the chance of finding βi insignificant, which means that severe MC effects are present (Statistica 2004). A general rule is that the VIF should not exceed 10 (Belsley et al 1980). MC might still be present and hence the next step is to regress each explanatory variable against all the other right hand side variables and compute the tolerance (1–R2) and VIF (Ramanathan ND, Ramathan 2002).
42- Reject or accept the research hypothesis (Turpin 2004)
43- Eliminate variables with insignificant coefficients, but one at a time to find the superior model (Ramanathan ND)
Structural Equation Modeling (SEM)
44- Apply SEM utilizing EQS, LISREL or M*plus* to figure out the final contingency framework (Raykov and Markoulides 2006) Path analysis
45- Explain the practical implications of findings for further investigation through path analysis as threshold for SEM
46- Apply path analysis (PA) and draw the …..
47- Compare the PA outcome with conceptual research framework to argue challenges Confirmatory Factor Analysis (CFA)
48- Identify and explain endogenous and exogenous variables
49- State the SEM hypotheses
50- Apply confirmatory factor analysis (CFA) to test the revised hypothetical model (theoretical model or final path analysis model) and revise the modeling hypothesis based on the CFA outcome (if necessary) for the closest possible arrangements to the real situation and make a final decision about the research contingency model

REFERENCES

Anderson-Cook, M., S. Dorai-Raj, 2003. Making the Concepts of Power and Sample Size Relevant and Accessible to Students in Introductory Statistics Courses using Applets. Journal of Statistics Education [Online], 11(3).

Bartholomew, J., 1998. Social Process. London School of Economics and Political Science.Data and sources of collinearity. New York: John Wiley & Sons.

Belsley, D.A., E. Kuh and R.E. Welsch, 1980. Regression diagnostics: Identifying influential??? 2nd edition

Bruin, J. 2006. Newtest: Command to Compute New Test. UCLA: Academic Technology Services, Statistical Consulting Group. http://www.ats.ucla.edu/stat/stata/ado/analysis/.

Clifton, B., 1969. Sequential Statistics Applied to Radiation Measurement. Health Physics: the Radiation Safety Journal, 17(5). Health Physics Society.

Danoff-Burg, J., 2002. Statistical Tests: Choosing and Understanding the Appropriate Statistical Test for Ecological Studies. From http://www.columbia.edu/itc/cerc/seeu/ atlantic/ statistics. html. Accessed in May 2009.

Denis, D., 2003. Alternatives to Null Hypothesis Significance Testing. Journal of Theory and Science 4(1). The International Consortium for the Advancement of Academic Publication, Athabasca, Canada.

Dinove, I., 2008. Choose the Right Test. From http://www..stat.ucla.edu Accessed in May 2009.

Dixon, P., 2003. Choosing Statistical Methods. Iowa State University, USA. From http://www.bcb.iastate.edu/faculty/Dixon_CV.pdf . Accessed in March 2009.

Ferrando, P., 2009. A General Factor-Analytic Procedure for Assessing Response Bias in Questionnaire Measures. Structural Equation Modeling: A Multidisciplinary Journal, 16(2): 364-381. DOI: 10.1080/10705510902751374.

Fincher, C., 1991. The Possibilities and Actualities of Disciplined Inquiry. Research in Higher Education. 32(6), 625-650. DOI 10.1007, BF00974735.

Garson, D., 2008. 'Structural Equation Modeling' from Statnotes: Topics in Multivariate Analysis. North Carolina State University, Retrieved May 25, 2009.http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm.

Govindrajulue, Z., 2004. Sequential Statistics. World Scientific Publishing Inc. NJ, USA, pp: 1.

Graveter, F. and L. Forzano, 2008. Research Methods for the Behavioral Science. Third ed., pp: 19-22. Thomson/Wadsworth, ISBN0495091456, 9780495091455.

Haenlein, M., A. Kaplan, 2004. A Beginner's Guide to Partial Least Squares Analysis. Lawrence Erlbaum Associates Inc. Understanding Statistics, 3(4): 283-297.

Healey, J., 2005. A Tool for Social Research. 7th ed. Wasdworth, Thomson Learning Academic Resource Center, CA, USA.

Hill, T., P. Lewicki, 2007. Statistics Methods and Applications. StatSoft, Tusla, OK.

Hoffman, H., 2004. Internet Glossary of Statistical Terms in: Statistics Explained. www.animatedsoftware.com/statglos/statglos.htm).

Huch, S., 2009. Statistical Misconceptions. Taylor and Francis Group, NY,USA, pp: 43.

Kaminsky, J., 2008. Qualitative and Quantitative Analysis. Kwantlen Polytechnic University. BNS Nursing Program. Greater Vancouver, British Columbia, Canada.

www.virtualcurriculum.com/N4120/LA9_Figure8.pdf. Accessed on March 2009.

http://www.ats.ucla.edu/stat/sas/notes2/ Accessed in March, 2009.

Lehmann, E., J. Romano, 2005. Testing Statistical Hypotheses (3rd ed.). New York: Springer. ISBN 0387988645.

Malakmohammadi, I., 2008. Variables Refinery Process to Ensure Research Unbiasedness (Validity) and Invariance (Reliability) in Agricultural Extension and Education. American Journal of Agricultural and Biological Sciences, 3(1): 342-347.

Marion, R., 2004. The Whole World of Deduction Research Skills for New Scientists. The University of Texas Medical Branch, Texas, USA.

McDonald, J.H., 2008. Handbook of Biological Statistics, Choosing a Statistical Test. Sparky House Publishing, Baltimore, Maryland, pp: 282-287.

Motulsky, H., 1999. Analyzing Data with GraphPad Prism. GraphPad Software Inc., San Diego, CA.

Moyulsky, H., 1995. Intuitive Biostatics: Choosing a Statistical Test. Oxford University Press, Inc.

McCloskey, D., 2008. The Cult of Statistical Significance. Ann Arbor: University of Michigan Press. ISBN 0472050079.

NN., 2007. Statistical Test: More Complicated Discriminate.

PHY310: Lecture 14. http://nngroup.physics.sunysb.edu/~mcgrew/phy310/lectures/phy310-lecture-14-2007.pdfnngroup.physics.sunysb.edu/~mcgrew/phy310/lectures/phy310-lecture-14-2007.pdf

NN., 2009. Introduction to SAS. UCLA: Academic Technology Services, Statistical Consulting Group. http://www.ats.ucla.edu/stat/sas/notes2/ (accessed July 20, 2009).

Palmer, M. ND. Multiple Regression, 2009. Ordination Method for Ecologists. Oklahoma State University. http://ordination.okstate.edu/MULTIPLE.htm. Accessed.

Ramanathan, R., 2002. Introductory Econometrics with application. Fifth ed. Harcourt College Publishers. SENGAGE Learning. Florence, KY.

Ramanathan, R., 2008. More on Multicollinearity (MC) Variance Inflation Factor (**VIF**) http://www.econ.ucsd.edu/~rramanat/MoreonMC.pdf.

Raykov, T., G. Marcoulides, 2006. A First Course in Structural Equation Modeling. Lawrence Erlbaum Association Inc. Publishers. NJ, USA., pp: 117-118.

Salkind, N., 2008. Statistics for People Who (Think They) Hate Statistics. Excel 2007 Edition. Sage Publication Inc. CA, USA., pp: 326.

Stapleton, C., 1997. Basic Concepts and Procedures of Confirmatory Factor Analysis. Paper presented at the Annual Meeting of the Southwest Educational Research Association. Austin, TX, January 23-25.

Statistica, 2008. Multiple Regression. Electronic Textbook. Statsoft, Inc. http://www.statsoft.com/textbook/stmulreg.html.

Streiner, D., 2005. Research Methods in Psychiatry. Finding Our Way: An Introduction to Path Analysis. The Canadian Journal of Psychiatry, 50: 115-122.

Thompson, B., 2004. Exploratory and Confirmatory Factor Analysis: Understanding concepts and applications. Washington, D.C.: American Psychological Association.

Turpin, G., 2004. Which test? A Clinical Psychologist's Online Guide to Selecting a Statistical Test. Center for Applied Social and Psychological Development. Canterbury Christ Church, UK.

Vermunt, J. and J. Magidson, 2003. Statistical Innovation. Encyclopedia of Social Science Research Methods, Sage Publications.

Wadsworth, 2005. Choosing the Correct Statistical Test. Wadsworth Cengage Learning. Florence, KY. http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/stat_workshp/chose_stat/chose_stat_01.html.

Watt, J. and S. Berg, 2002. Research Methods for Communication Science. Amazon.com

Wheater, P. and P. Cook, 2000. Using Statistics to Understand the Environment. Routledge, NY, USA.

Whitney, 2005. Statistics: A Tool for Social Research. 7th Ed. Thomson Learning, Inc. Thomson Wadsworth. Belmont, CA 94002-3098, USA.

Windish, D., M. Diener-West, 2006. A Clinician-educator's Roadmap to Choosing and Interpreting Statistical Tests. Journal of General Internal Medicine, 21 (9, 8), 1009-1009. DOI 10.1111/j.1525-1497.2006.00390.x.