
UNIT 19 COMPUTER DATA ANALYSIS

Structure

- 19.1 Introduction
 - 19.2 Objectives
 - 19.3 What is SPSS?
 - 19.4 Get Yourself Acquainted with SPSS
 - 19.5 Menu Commands and Sub-commands
 - 19.6 Basic Steps in Data Analysis
 - 19.7 Defining, Editing and Entering Data
 - 19.8 Data File Management Functions
 - 19.8.1 Merging Data Files
 - 19.8.2 Aggregate Data
 - 19.8.3 Split File
 - 19.8.4 Select Cases
 - 19.9 Running a Preliminary Analysis
 - 19.9.1 Six Characteristics of a Dataset
 - 19.9.2 Data Transformation
 - 19.9.3 Exploring Data
 - 19.9.4 Graphical Presentation of Data
 - 19.9.5 Scatterplots and Histograms
 - 19.10 Understanding Relationship Between Variables
 - 19.10.1 The Mean Procedure
 - 19.10.2 Linear Regression
 - 19.10.3 Curve Estimation
 - 19.11 Non-parametric Tests
 - 19.12 SPSS Production Facility
 - 19.13 Statistical Analysis System (SAS)
 - 19.14 NUDIST
 - 19.15 Let Us Sum Up
 - 19.16 Unit-end Activities
 - 19.17 Suggested Readings
 - 19.18 Answers to Check Your Progress
- Appendix

19.1 INTRODUCTION

In earlier units, we provided you with a detailed understanding of how quantitative and qualitative data are analysed manually. Although some of us still carry out analysis of data manually, the advent of sophisticated computer software has made data analysis more convenient and easier. Earlier, the software which could only be run on large mainframe computers can now be run with considerable ease on the PCs. SPSS is one such software which is used in educational research. You can analyze large and computer data files with thousands of variables on your PC without compromising the quality and the precision of analysis.

In this unit, we will introduce you to the software for quantitative and qualitative data analysis. We will provide in this details of SPSS package which is comparatively more popular among research students for quantitative data analysis. We will also introduce in this unit the Statistical Analysis System (SAS), another software for quantitative data analysis. We will introduce a software called NUDIST for qualitative data analysis.

19.2 OBJECTIVES

After going through this Unit, you should be able to:

- explain describe the main features of the SPSS;
- write about as well as use the data management operations and techniques of analysis using SPSS;
- acquire skills in the use of SPSS for basic statistical analysis with a special focus on the measures of central tendency, dispersion, correlation and regression; and
- present the data and the SPSS results graphically.

19.3 WHAT IS SPSS ?

SPSS* is one of the leading desktop statistical packages. It is an ideal companion to the database and spreadsheet, combining many of their features as well as adding its own specialized functions. SPSS for windows is available, as a base module and a number of optional add-on enhancements are also available. Some versions present SPSS as an integrated package including the base and some important add-on-modules.

SPSS Professional Statistics provides techniques to examine similarities and dissimilarities in data and to classify data, identify underlying dimensions in a data set. It includes procedures for cluster, k-cluster, discriminating factor, multi-dimensional scaling, and proximity and reliability analysis.

SPSS Advanced Statistics includes procedures for logistic regression, log-linear analysis, multivariate analysis and analysis of variance. This module also includes procedures for constrained non-linear regression, probit, Cox and actuarial survival analysis.

SPSS Tables creates a high quality presentation—quality tabular reports including stub and banner tables and display of multiple response data sets. The new features include pivot tables, a valuable tool for presentation of selected analytical output tables.

SPSS Trends performs comprehensive forecasting and time series analysis with multiple curve fitting models, smoothing models and methods for estimation of autoregressive functions.

SPSS categories performs conjoint analysis and optional scaling procedures, including correspondence analysis.

SPSS Chaid provides simplified tabular analysis of categories data, develops predictive models, screens out extraneous predictor variables, and produces easy to

* SPSS is registered trademark of the SPSS Corporation, USA.

read tree diagrams that segment a population into sub-groups that share similar characteristics.

Recently, the SPSS Corporation announced the release of SPSS version 8.0. Many new add-on products have also been launched in the recent months. You can consult the SPSS World Wide Web site for the latest developments and additions to the computing power SPSS. Technical support is also available to the registered user at the SPSS site. The SPSS Web site is <http://www.spss.com>. Select white papers on SPSS applications in major disciplines are also available on this site.

SPSS version 7.5 for Windows is now available with most users across the globe. The present unit discusses some of the commonly used data management techniques and statistical procedures using SPSS 7.5. Since new features are added almost daily, you are advised to check for these details on the currently installed version of SPSS on your computer and also consult the user manuals before undertaking complex type of data analysis. The on-line help is also available. There may be some procedures and syntax related changes from one version to another. We will attempt to provide you with procedures that are most commonly used with SPSS Release 7.5. In case these are not available on your version of SPSS, please consult the relevant SPSS authorized representative or the WWW site of the SPSS corporation.

19.4 GET YOURSELF ACQUAINTED WITH SPSS

The SPSS for Windows can be run from Windows 98 through Windows XP operating systems. Unix, Mac and mainframe versions of the SPSS software are also available. The illustrations in this Unit are based on SPSS version for Window 95/98/NT operating systems.

Starting SPSS

The SPSS for Windows uses graphical environment, descriptive menus and simple dialog boxes to do most of the work. It produces three type of files, namely data files, chart files and text files.

To start SPSS, click the start button on your Computer². On the start menu that appears, click *Program*. Another menu appears on the right of the *start* menu. If there is an entry marked SPSS, that's the one you want to click. If there isn't, click the program group where SPSS was installed and an entry marked SPSS will appear. Click the SPSS 7.5 entry. You will know when the SPSS has started and SPSS Data Editor window appears. To begin with, the SPSS data editor window will be empty and a number of menus will appear on the top of the window. You will start the operations by loading a data set or by creating a new file for which data is to be entered from the data editor window. The data can also be imported from other programs like Dbase, ASCII, Excel and Lotus.

Existing SPSS

Make sure that all SPSS and other files are saved before quitting the program. You should exit the software by shutting off the program by selecting Exit SPSS command from the file menu of the SPSS Data Editor window. In case of unsaved files, the SPSS will prompt you to save or discard the changes in the file.

² It is assumed that a proper licensed and valid version of SPSS is already installed on the computer you are working with.

Saving data and other files

Many types of file can be saved using 'save' or 'save as' command. Various types of files used in SPSS are: Data, Syntax, Chart or Output. Files from spreadsheets or other databases can also be imported by following the appropriate procedure. Similarly, an SPSS file can be saved as a spreadsheet or in dbase format. Select the appropriate save type command and save the file. The SPSS data files are saved with .sav as the secondary name. Though SPSS files could be given any name, the use of reserved words and symbols is to be avoided in all types of file names.

Printing of data and output files

The contents of SPSS data files, Output Navigator files and Syntax Files can be printed using the standard 'Print' command. The SPSS uses the default printer for printing. In the case of network printers, an appropriate printer should be selected for printing the output. It is suggested that Ink jet or Laser jet printers should be used for printing graphs and charts. Tabular data can be easily printed using a Dot matrix Printer.

Operating Windows in SPSS

There are seven types of Windows in SPSS which are frequently referred to during the data management and analysis stages. These are:

Data Editor

As mentioned earlier, the data editor window opens automatically as soon the SPSS gets loaded. To begin with, the data editor does not contain any data. The file containing the data for analysis has to be loaded with the help of the 'file' menu sub-commands by using various options available for this purpose. The contents of the active data file are displayed in the data editor window. Only one data editor window will be active at a time. No statistical operations can be performed until some data is loaded into data editor.

Output Navigator

All SPSS messages, statistical results, tables and charts are displayed in the output navigator. The output in the navigator Window can be edited and saved for future reference. The Output Navigator opens automatically, the first time some output is generated. The user can customize the presentation of reports and tables displayed in the Output Navigator. The output can be directly imported into reports prepared under word processing packages, and the output files are saved with an extension xxxx.spo.

Pivot Tables

The output shown in the Output Navigator can be modified in many ways using the Edit and Pivot Table Option, which can be used to edit text, swap rows and column, add colour, prepare custom made reports/output, create and display selectively multi-dimensional tables. The results can be selectively hidden and shown using features available in Pivot Tables.

Graphics

The Chart Editor helps in switching between various types of charts, swapping of X-Y axis, changing colour and providing facilities for presenting data and results through various type of graphical presentations.

It is useful for customizing the charts to highlight specific features of the charts and map.

Text Editor

The text output not displayed in the Pivot Tables can be modified with the help of Text Editor. It works like an ordinary Text Editor. The output can be saved for future reference or sharing purposes

Syntax Editor

The Syntax Editor can be opened and closed like any other file using the File Open/ New command. The use of Syntax file is recommended when the same type of analysis is to be performed at frequent intervals of time or on a large number of data files. Using Syntax File for such purposes automates complex analysis and also avoids errors due to frequent typing of the same command. The commands can be pasted on the Syntax files using a particular command and pastes buttons from the menu. Experienced user can directly type the commands in the Syntax window. To run the syntax, select the commands to be executed and click on the run button at the top of the syntax window. All or some selected commands from the Syntax File will be executed. The Syntax File is saved as xx.sps.

Script Editor

This facility is normally used by the advanced users. It offers fully featured programming environment that uses the Sax BASIC language and includes a Script Editor, Object Browser, Debugging features and context sensitive help. Scripting allows you to automate tasks in SPSS including:

- Automatically customizing output
- Open and save data files
- Display and manipulate SPSS dialog boxes
- Run data transformation and statistical procedures using SPSS command Syntax.
- Export charts as graphics files in a variety of formats.

The present module will not go into the details of the advanced features of SPSS including scripting.

Check Your Progress

Notes : a) Space is given below for your answer.

b) Compare your answer with the one given at the end of this unit.

1. Write any three features of SPSS package.

.....

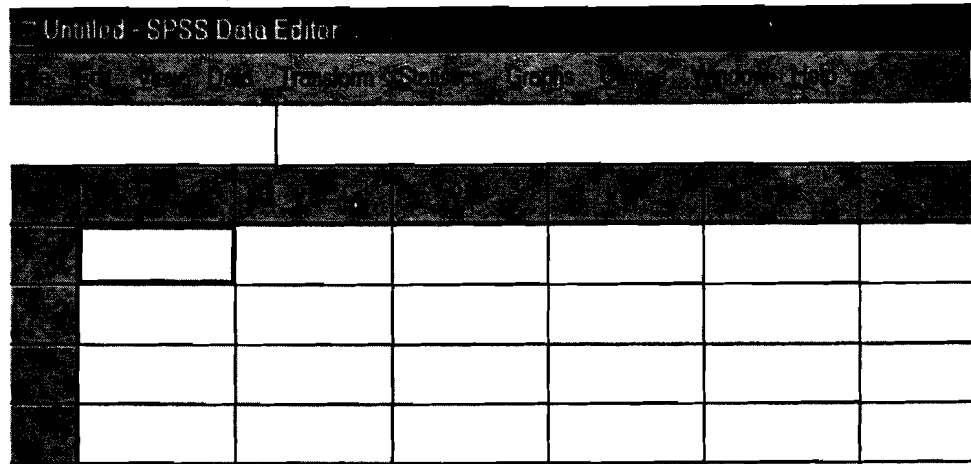
2. How many types of windows are there in SPSS and what are those windows?

.....

19.5 MENU COMMANDS AND SUB-COMMANDS

Most of the commands can be executed by making appropriate selections from the menu bar. Some additional commands and procedures are available only through the Syntax Windows. The SPSS user manuals provide a comprehensive list of commands, which are not available through menu driven options. If you want a comprehensive overview of the basics of SPSS, there is an on-line tutorial, as extensive help of SPSS is available by using the 'Help' menu command. The CD version of the software contains an additional demo module.

Since SPSS is menu driven, each Window has its own menu bar. While some of the menu bars are common, the others are specific to a particular type of Window. We will present below the menu and sub-menus of the Data Editor window. You may consult the SPSS manuals for other types of menu and sub-menu commands.



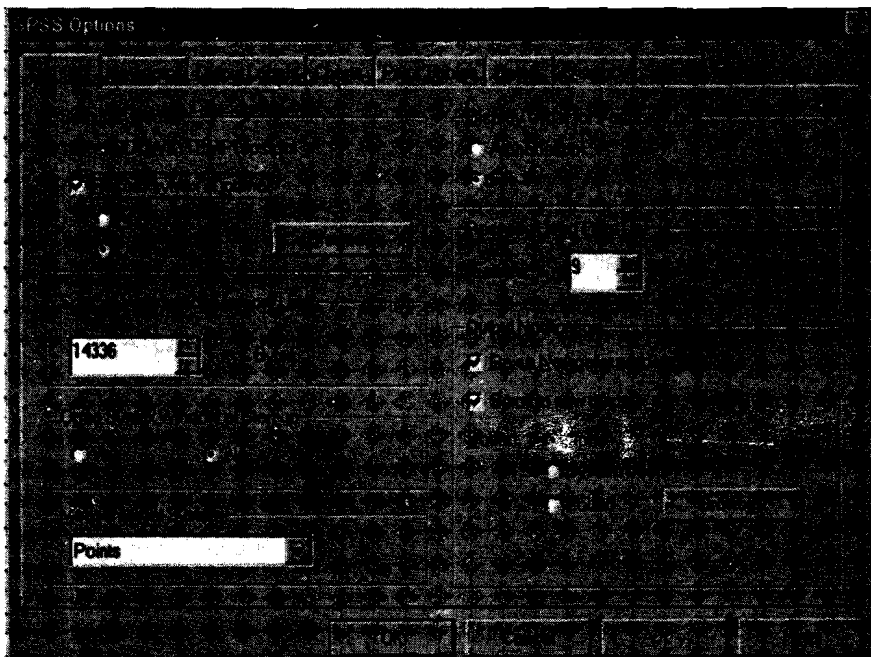
The following table shows the Data Editor Menus. Each commands in the main menu has a number of sub-command.

| Menu | Function/sub-commands |
|-------------|--|
| File | : Open and save data file, to import data created in other formats like Lotus, Excel, Dbf etc. Print control functions like page setup, printer setup and associated functions. ASCII data can also be read into SPSS. Data capture command is used to import data from RDMS structures. |
| Edit | : These functions are similar to those available in general packages. These include undo, redo cut, copy, paste, paste special, find, find and replace. Option setting for the SPSS are controlled through Edit menu. |
| View | : Customize tool bars, Fonts grid and display of data, displays option for showing value labels. |
| Data | : This is very important menu as far as management of data is concerned. Variable definition, inserting new variables, transposing templates, aggregating and merging of data files, splitting data files for specific analysis are some important commands in Data Menu. |

- Transform** : Compute new variables, recode, random number generation, ranking, time series data transformation, count and missing value analysis are undertaken using Transform Command.
- Statistics** : As the name implies, Statistics Menu incorporates statistical procedures. Frequency distribution, cross-tabulations, comparison of means, correlation, simple and multiple regression, ANOVA, Log linear regression, discriminate analysis, canonical analysis, factor analysis, non-parametric tests and time series analysis are undertaken using Statistics menu.
- Graphs** : Includes options for generating various type of custom made graphics like bar, pie, area, X-Y and high-low charts, pareto, control charts, box-plots, histograms, P-P and Q-Q charts and time series representation of data.
- Utilities** : Information about variables, information on working a data file, run scripts and define sets are some of the important functions carried out through Utilities command.
- Window** : Windows menu are used to switch between SPSS windows.
- Help** : Context specific help through dialog boxes, demo of the software, and information about the software are some of the important options under Help command. It provides a connection for the SPSS home page. The statistical coach included in the help module is very useful in understanding various stages of executing a procedure.

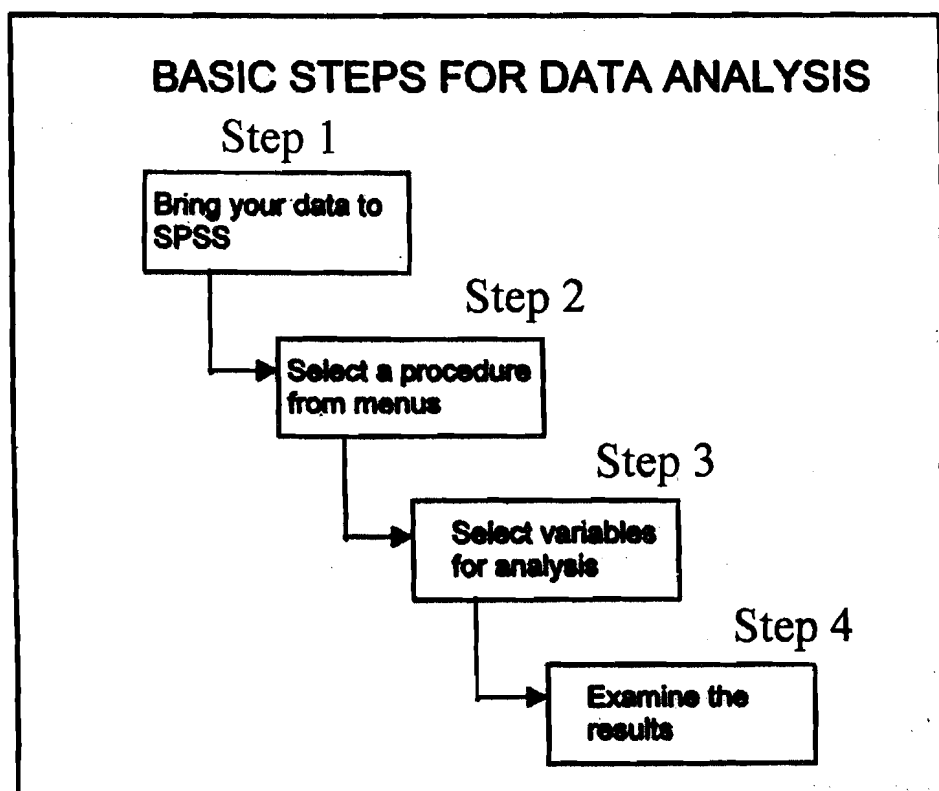
Setting the Options

The SPSS provides a facility for setting up of the user-defined options. Use the Edit menu and then select Options. The following types of optional setting are allowed in SPSS. Make the appropriate changes to set the options according to your choice.



19.6 BASIC STEPS IN DATA ANALYSIS

There are four basic steps involved in data analysis using SPSS. These are shown in the following figure.



Bring your data into SPSS: You can bring your data into SPSS in the following ways:

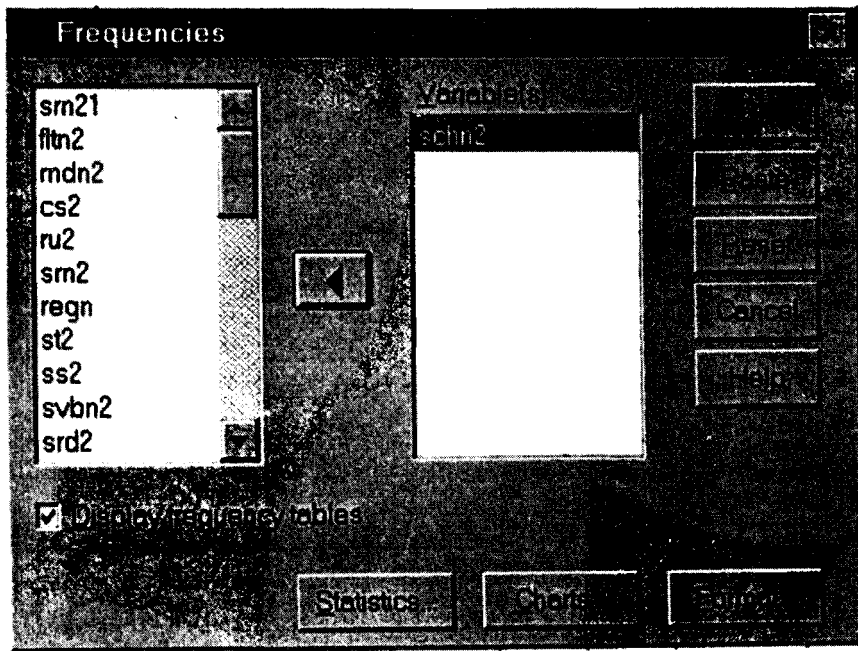
- Enter data directly into SPSS Data Editor.
- Open previously saved SPSS data file
- Read a spreadsheet data into SPSS data editor.
- Import data from DBF files
- Import data from RDBMS packages like Access, Oracle, Power, Builder, etc.

Select a Procedure from Menu: Before embarking on a statistical analysis, it is advised that you are clear as to what analysis is to be performed. Select the corresponding procedure to work on the data or create charts or tables using the selected procedure.

The command could either be directly executed or pasted on a Syntax Window. As mentioned earlier, pasting the command on the Syntax Window will be useful for undertaking batch processing or for subsequent use, especially where the same type of repetitive analysis required. Pasting the command will not lead to its execution. The command has to be selected and executed using the run command.

Select the variables: All the variables in the active file are listed each time a dialog box is opened. Select the appropriate variables for the selected procedure. Selection of at least one variable is necessary to run a statistical procedure. The

variables may be numeric, string, date or logical. You should be aware that string variables cannot be manipulated to the same extent as the numeric variables.



Run the Procedure and Examine the Output: After completing the selection process for the procedure and the variables, execute the SPSS command. Most of the commands are executed by clicking OK on the dialog box. The processor will execute the procedures and produce a report in the Output Navigator.

19.7 DEFINING, EDITING AND ENTERING DATA

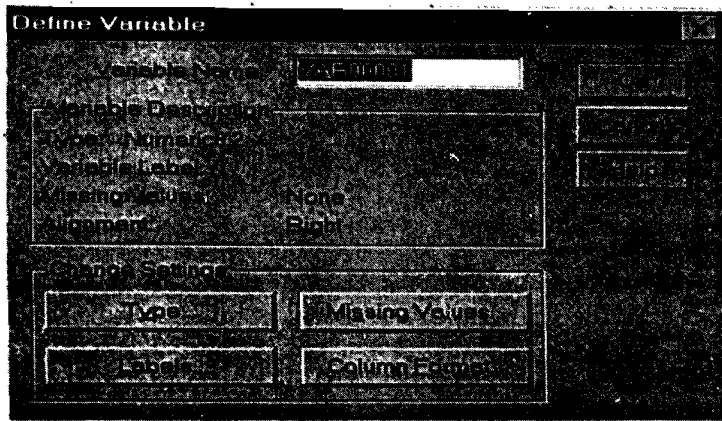
As mention earlier, there are many options for creating SPSS data files. The data can either be directly entered through Data Editor or imported from spreadsheets, ASCII file and other RDBMS packages like Oracle and Access. The data is arranged in the form of rows and column in the Editor Window. The rows refer to the observations or cases and the columns to the variables. Each cell is defined as the intersection of a row and a column and refers to the value of a particular variable for a specific case/observation. While defining data, it is important to identify a primary key which is unique for each observation/case.

Variable Definition

Before entering the data into SPSS, it is advised that you define your variables. Such a definition will be very helpful for data entry and analysis stages. The following information about each variable is provided to define it:

- A name for the variable (upto 8 characters only)
- A description (label)
- A series of labels which explain the value entered (value labels)
- A declaration as to which values are non-valid and should be excluded from the statistical analysis and other operations (missing values). The information is important to understand the response pattern and also to specify the observations which should be excluded from the analysis.

Define variable (provide relevant information asked in the dialog box).



Data can be entered directly using SPSS Data Editor window. However, if the data is large, you are advised to use a data entry package. The data can also be edited/changed in the data editor window. To change the value in any cell, bring the cursor to the particular cell, enter the new value and press enter. New variables can also be added and the existing variables can be deleted in the Data Editor Window.

19.8 DATA FILE MANAGEMENT FUNCTIONS

SPSS is very flexible as far as management of data files is concerned. While only one file can be opened for analysis at a time, the SPSS provides flexibility in merging multiple data files with the same structure into one single data file, merging files to add new variables, partially select the cases for analysis, make group of data based on certain characteristics and use different weights for different variables. Some of these functions are discussed below. Groups of data can also be defined to facilitate the analysis of the most commonly referred variables (see utilities and data commands).

19.8.1 Merging Data Files

Researchers are often faced with a situation where data from different files are to be merged or a limited number of variables from large complex data files are required. The following types of facility are available for merging files using SPSS.

Adding variables: Adding variables is useful when two data files contain the information about the same case but on different variables. For example, the teachers database may contain two files, one having the educational qualifications and the other having the names of the courses taught. Both the files could be combined to analyze the variables available in them. The data on a key and unique variable from both the files can be combined easily. The key variables must have the same name in both the data files. Both the data files should be sorted on the common key variable.

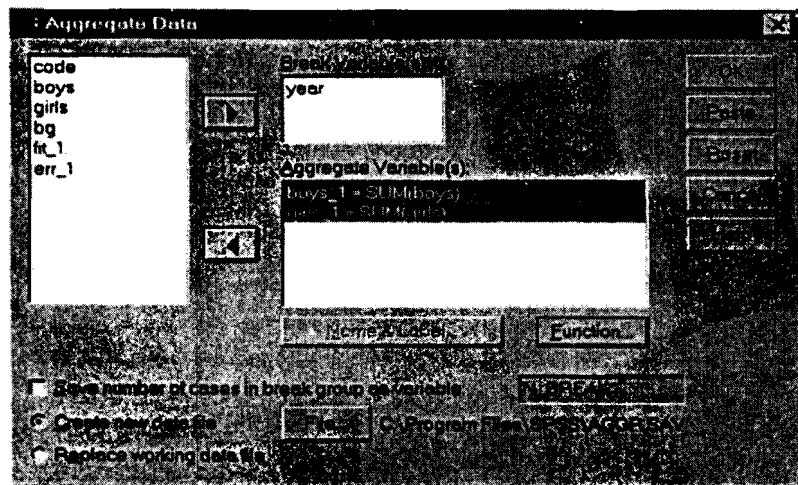
Adding cases: This option is used when the data from two files having the same variables are to be combined. For example, you may record the same information for students in different study centers in India abroad. The data can be merged to create a centralized database by using Add cases command.

| Break Variable | Mean | Count |
|-------------------|------|-------|
| code | | |
| boys | 6.80 | 7 |
| girls | 6.98 | 7 |
| bg | 6.99 | 7 |
| fit_1 | 6.99 | 7 |
| err_1 | 6.99 | 7 |
| Aggregate | | |
| Orthogonal Design | 6.56 | 6 |
| Split File | 6.42 | 6 |
| Select Cases | 6.31 | 6 |
| Weight Cases | | |

19.8.2 Aggregate Data

Aggregate Data command combines groups of cases into a single summary case and creates a new aggregate data file. Cases are aggregated, based on the value of one or more grouping variables. The new (aggregated) file contains one record for each group. The aggregate file could be saved with a specific name to be provided by you. Otherwise, the default name is aggregate.sav. For example, the data on learners, achievement could be aggregated by sex, state and region.

A number of aggregate functions are available in the SPSS. These include sum, mean, number of cases, maximum value, minimum value, standard deviation and first and the last value. Other summary functions include percentage and fractions below and above a particular cut-off user-defined value.



19.8.3 Split File

The researcher is often interested in the comparison of a summary and other statistics based on certain group behaviour. For example, in a study of learning achievement, the researcher may be interested in comparing the mean scores for students belonging to different sex groups. The sex is taken as a grouping variable. Multiple grouping variables can also be selected. A maximum of eight grouping variables can be defined. Cases need to be sorted out by grouping variables. Two options are available for comparative analysis. These are: compare groups and organize output by groups. The split file is available under Data menu for making such comparisons.

19.8.4 Select Cases

Select case command can be used for selecting a random sub-sample or sub-group of cases based on a specified criteria that includes variables and complex expression. The following criteria are used for Select Case command.

Select if (condition is satisfied)

Variable values and their range

Date and time range

Arithmetic expressions

Logical expressions

Functions

Row numbers

Following the Select Case command, the unselected cases can either be deleted or temporarily filtered. Deleted cases are removed from the active file and cannot be recovered. You should be careful while selecting Delete option. Filtered option will be deleted temporarily. When the Select Case option is on, it is indicated in the Data Editor window.

Check Your Progress

Notes : a) Space is given below for your answer.

b) Compare your answer with the one given at the end of this unit.

3. What are the basic steps in data analysis ?

.....

.....

.....

.....

.....

.....

4. What is a split file ?

.....

.....

.....

.....

19.9 RUNNING A PRELIMINARY ANALYSIS

Before running advanced statistical analysis it is important that you understand the salient features of your data. Use of statistical applications on a data set, the behaviour of which is not known, can give misleading conclusions. The following section explains the six characteristics which must be examined for a given data set before attempting an advanced analysis.

19.9.1 Six Characteristics of a Dataset

One strong argument for using computers and graphical presentation of the data is the advantage of viewing the data in a variety of ways. Preliminary exploration of data and its graphical presentation helps attain these objectives. The following characteristics will help you in deciding on the best plan for data management, analysis and presentation. SPSS includes commands for analyzing of data along the following lines.

Shape: The shape of the data will be the main factor in determining what set of summary statistics best explains the data. Shape is commonly categorised as symmetric, left-skewed or right-skewed, and an uni-modal, bi-modal or multi-modal. Frequency distribution, plots and graphical presentation of data histogram, P-P, Q-Q, scatter, Box-Plot are illustrative of the techniques that can be used for determining the shape of a data set. It is important that the user should have enough knowledge of the properties of various statistical distributions, their graphical presentations, characteristics and limitations.

Location: Location is simpler and more descriptive than measures of central tendency. Common measures of location are the mean and the median. Measures of central tendency also can be calculated for various sub-groups of a data set.

Spread: This measure describes the amount of variation in the data. Again approximate value is sufficient initially, with the measure of spread being informed by the shape of the data, and its intended use. Common measures of spread are variance, standard deviation and inner-quartile range. Percentile range is another measure which is used for measurement of dispersion.

Outliers: Outliers are data values that lie away from the general cluster of values. Each outlier needs to be examined to determine if it represents a possible value from the population being studied, in which case it should be retained, or if it is non representative (or an error) in which case it should be excluded. You should properly weigh and carefully examine the behaviour of outliers before accepting or rejecting of an observation/case. The best choice to display when looking for outlier is Box-plot. Range, i.e., maximum and minimum values can also be used to examine the behaviour of outliers.

Clustering: Clustering implies that data tend to bunch around certain values. Clustering shows most clearly on a dot-plot. Histogram, stem and leaf analysis are also important procedures to examine the clustering pattern of a data set.

Association and relationship: Researchers often look for associative characteristics or similarities and dissimilarities in the behaviour of some variables. For example, achievement scores and hours of study may be positively correlated whereas the teacher motivation and drop-out rate may be negatively associated with each other. Correlation coefficient is the most commonly used measure for understanding the nature and magnitude of association between two variables.

You should be clear that association does not imply relationship. A relationship is defined by the cause and effect type of link. Normally, there is one dependent variable and one or more than one independent variable in the cause and effect relationship. Cause and effect relationship is captured through regression analysis.

The analysis of data along the above lines provides considerable insight into the nature of data and also helps researchers in understanding key relationships between variables. It is assumed that the relationships are of linear type. Non-linear relationships can also be examined using non-linear techniques of analysis and also using data transformation techniques.

19.9.2 Data Transformation

Data transformation is a very useful aspect of SPSS. Using data transformation, you can collapse categories, recode the data and create new variables based on complex equations and conditional statements. Some of the functions are detailed below:

Compute variable

- Compute value for numeric and string variables
- Create new variables or replace the value of existing variables. For the new variable, you can specify the variable type and label.
- Compute value selectively for sub-sets of data based on logical conditions.
- Use built-in functions, statistical functions, distribution functions and string functions.

Recode variables

Recoding of variables is an important characteristic of data management using SPSS. Many continuous and discrete variables need to be recoded for meaningful analysis. Recoding can be done either within the same variable or a new variable can be generated. Recoding in the same variable will replace and original values for this purpose. Recoding in a new variable will replace the old values with new values. The following example illustrates the need and use of recoding variables.

A survey of the primary school was conducted in Delhi. Along with other variables, information on the type of management was also collected. The management code was designed as follows:

1. Government
2. Local bodies
3. Private aided
4. Private unaided
5. Others

Lets us assume that a comparative analysis of the government and the private management schools is to be undertaken. This will be done by combining categories 1 and 2 and also 3 and 4. This can be achieved by recoding the management code as 1 (for 1 and 2 categories) and 2 for 3 and 4 categories into a new variable.

Assuming that a database on primary schools in Delhi is available, the enrolment analysis could be attempted by making suitable categories, i.e. schools with less than 50 students, 51-150,15-250 and more than 250 students. This could be achieved by recoding the enrolment variable into a new variable 'category'. If at a later stage in the analysis, it is found that a new category is to be introduced, it can again be achieved by recoding the enrolment data.

Count

Count is an important command available in SPSS and is used for counting occurrences of the same value(s) in a list of variables within the same case. For example, a survey might contain a list of books purchased (yes/no) by the students. You could count the number of 'yes' response, or a new variable can be generated which gives the value of count indicating the number of books bought.

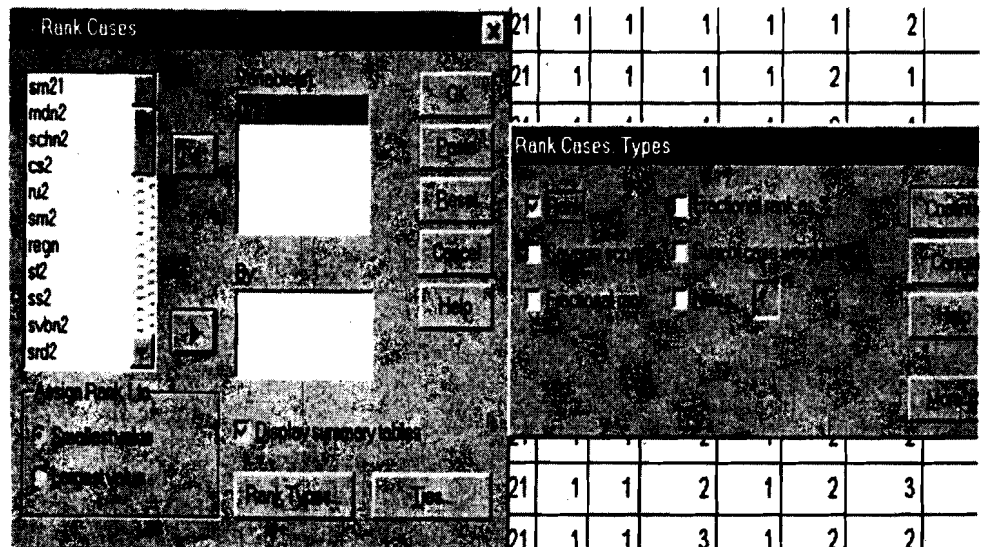
Procedure to run count command

- Choose Transform from the main menu
- Chose count
- Enter the name of a target variable (variable where the count value will be stored).
- Select two or more variables of the same type (numeric or string)
- Click define variable and specify which value(s) to be counted.
- Click OK after the selection has been made.

In survey on learners’ achievement, the answer code to each question in language and mathematics could be recorded for each student. The codes could be ‘1’ for the correct answer ‘2’ for the wrong answer and ‘3’ for no reply. Count command can then be used to count the number of correct answers.

Rank Cases

Rank Cases command can be used to rank observation in an ascending or a descending order. Other options available for ranking cases are shown in the right hand panel of the following figure.



19.9.3 Exploring Data

The Frequencies procedure provides statistics and graphic displays that are useful for describing many type of variables. Frequency counts, simple and cumulative percentages, mean, median and mode, sum, standard deviation, range minimum and maximum values, standard error of the mean, skewness and kurtosis, bar charts and pie charts, and histograms are some of the methods used to explore the data before a sophisticated and advanced analysis is undertaken.

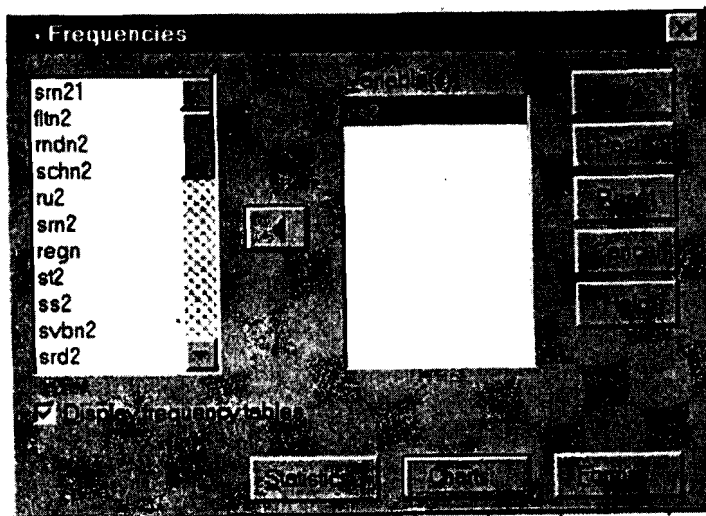
If you want to compare summary statistics for each of the several groups of cases, use split file on the ‘Data’ menu. Use of Explore, Summarize or means procedure is recommended for initial exploration of data. Use the following commands to obtain frequencies:

From the menu choose:

Statistics

Summarize

Frequencies



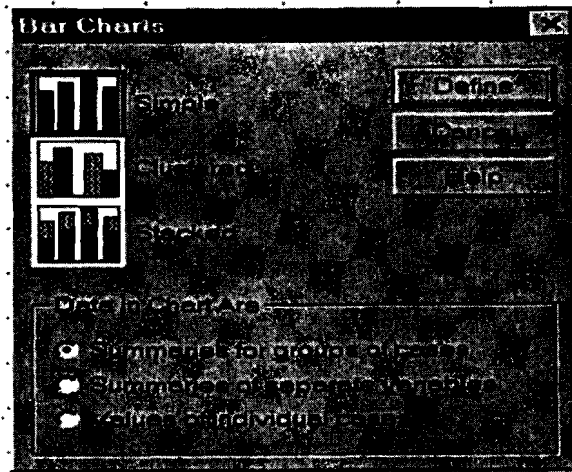
Use the Statistics and Charts sub-commands (as shown in the above figure) to select the desired features. More than one variable could be selected for frequency distribution. You must remember, that before attempting frequency distribution, recoding of continuous type of variables will be necessary.

19.9.4 Graphical Presentation of Data

SPSS offers extensive facilities for viewing the data and its key features in high-resolution charts and plots. From the main menu, select Graphs and the following screen appears. Various types of Graph that can be drawn using SPSS are indicated in the sub-commands.

| d150u2 - SPSS Data Editor | | | | | | | |
|---------------------------|------|-------|------|----|--|-------|----|
| 1:sm21 | | | | | | | |
| | sm21 | fln2 | mdn2 | rc | | | |
| 1 | 9999 | 82001 | 42 | | | 71112 | 21 |
| 2 | 0 | 82001 | 42 | | | 71112 | 21 |
| 3 | 0 | 82001 | 42 | | | 71112 | 21 |
| 4 | 0 | 82001 | 42 | | | 71112 | 21 |
| 5 | 9999 | 82001 | 42 | | | 71112 | 21 |
| 6 | 1002 | 82001 | 42 | | | 71117 | 21 |
| 7 | 1000 | 82001 | 42 | | | 71117 | 21 |
| 8 | 1001 | 82001 | 42 | | | 71117 | 21 |
| 9 | 1007 | 82001 | 42 | | | 71117 | 21 |

Select a chart type from the Graph menu. This opens a chart dialog box as shown below:

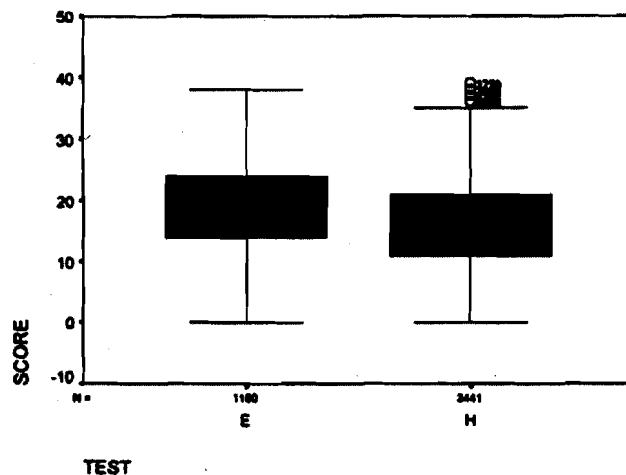


After the appropriate selections have been made, the output is displayed in Output Navigator window. The chart can be modified by a double click on any part of the chart. Some typical modifications include the following:

- Edit axis titles and labels and footnotes
- Change Scale (X-Y)
- Edit the legend
- Add or modify a title
- Add annotation
- Add an outer frame

Another important category of charts is High-Low which are often used to represent variables like maximum and minimum temperature in a day, stock market behaviour or other similar variables.

Box-plot and Error Bar charts helps you to visualize distribution and dispersion. Box-plot displays the median and quartiles and special symbols are used to identify outliers, if any. Error Bar charts displays the mean and confidence intervals or standard errors. To obtain a box-plot, choose Box-plot from the Graphs menu. The simple box-plot for mean scores obtained in English and Hindi is shown in the following diagram:



The above figure shows that there were a large number of outliers in the case of Hindi scores as compared to English. The outliers were along the higher side. This shows that many students were scoring very high marks. The size (numbers) of cases are shown along the X-axis. The boxes show the median and the quartile values for both the tests.

19.9.5 Scatterplots and Histograms

Scatterplots highlight the relationship between two quantitative variables by plotting the actual values along X-Y axis. The scatterplots are useful to examine the actual nature of relationship between these variables. This could be either linear or non-linear in form. To help visualize the relationship, you can add a simple linear or a quadratic regression line. A 3-D scatter plot adds a third variable in the relationship. You can rotate the two dimensional projection of the three dimensions to delineate the underlying patterns. In order to obtain a scatter plot, select Scatter from the Graphs option.

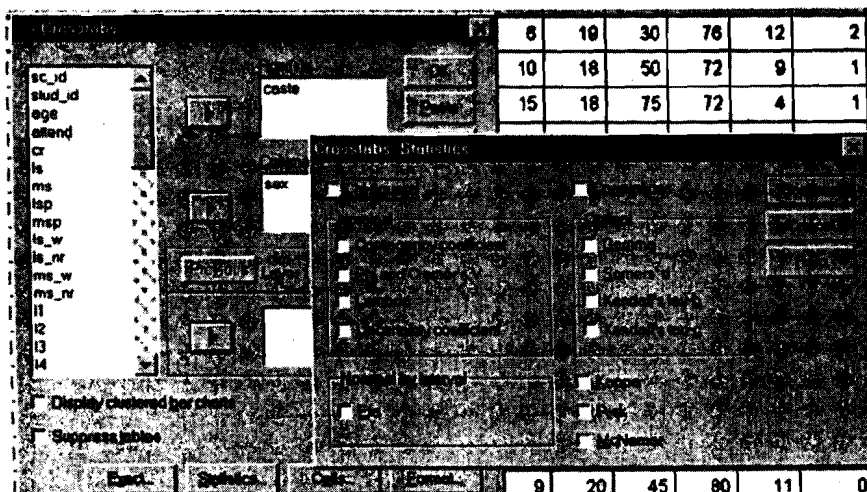
A histogram will be obtained by selecting histogram option from the Graphs menu. The variable for which a histogram is to be obtained should be selected from the dialog box. The normal curve can also be displayed along with the histogram to visually see the extent of similarity between the actual distribution of values and the normal curve.

Pareto and control charts are used to analyze and improve the quality of an ongoing process. You may refer to the SPSS manuals for use of these techniques.

19.10 UNDERSTANDING RELATIONSHIP BETWEEN VARIABLES

The foregoing details focused on the techniques of analysis describing the behaviour of individual variables. However, most of the research studies require relationship between two or more variables to be examined. For example, one may be interested in questions like, "do the achievement scores of boys and girls in the same class differ?"

Cross-tabulation is the simplest procedure to describe a relationship between two or more categories of variables. Cross-tabulation is useful for any type of categorical variable, especially, when the categories are small and mutually exclusive. Some variables could be aggregated into convenient categories by using the Recode command.



The cells in the standard two-way frequency table display the counts or the number of cases falling into the categories distinguished by the row and column variables. There is no category showing the missing data in a two-way classification.

The SPSS also provides for a number of options while displaying the results of cross-tabulations. These relate to percentage distribution of frequencies/cases in terms of row total, column total and grand total. Any or all of these options can be selected. Each of the options can be selected depending upon the objective of analysis. The following table shows the distribution of students by their sex in a simple study.

The simple output of the cross-tabulation procedure is shown below.

Cast * Sex Cross Tabulation

| Count | | Sex | | Total |
|-------|---------|------|--------|-------|
| | | Male | Female | |
| Caste | SC/ST | 204 | 322 | 528 |
| | OBC | 86 | 136 | 222 |
| | General | 1256 | 1274 | 2530 |
| Total | | 1546 | 1732 | 3278 |

19.10.1 The Mean Procedure

The Mean procedure is very useful tool when the average value of a variable is to be computed based on sub-division of the data into groups based on the value of some other variable. For example, you may be required to compute average achievement score of children based on their age. Or you may be required to compute average monthly income of the respondents by their occupations and experience. While the mean procedure has immense use in understanding the sub-group behaviour, it also suffers from certain limitations. It cannot be used in the case of categorical variables. Moreover, the sub-groups should have a reasonably large number of value for the mean value to be representative. The specifications for a subgroup average are:

Place the continuous variables in the 'Dependent list'

Place the categorical variables in the 'Independent list'.

The mean scores in Mathematics by caste groups as obtained using the Mean procedure are given below:

Report

| | | |
|----------------|---------|-------|
| Score Math | | |
| Mean | SC/ST | 17.76 |
| | OBC | 17.51 |
| | General | 20.09 |
| | Total | 19.54 |
| N | SC/ST | 526 |
| | OBC | 222 |
| | General | 2530 |
| | Total | 3278 |
| Std. Deviation | SC/ST | 6.40 |
| | OBC | 6.20 |
| | General | 5.11 |
| | Total | 5.50 |

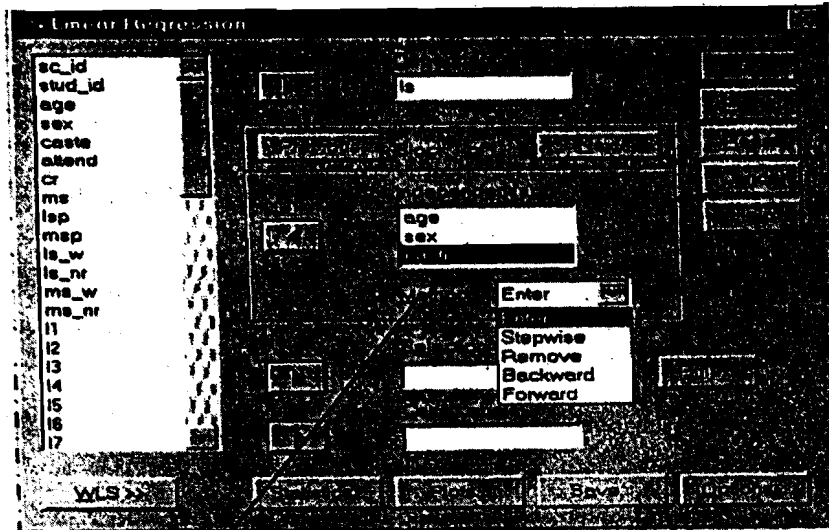
How do you predict the sales of ice cream in the coming summer season? What are the important determinants of achievement in government schools? Is there any relationship between educational attainment and per capita income of a household? These are the types of question which are often asked by development planners and policy analysts. Regression analysis is a technique to address some of these questions.

Linear regression is the most commonly used procedure for the analysis of a cause and effect relationship between one dependent variable and a number of independent variables. The dependent and independent variables should be quantitative. Categorical variables like sex and religion should be recorded to dummy (binary) variables and other types of contrast variables. An important assumption of the regression analysis is that the distribution of the dependent variable is normal. Moreover, the relationship between the dependent and all the independent variables should be linear and all observations should be independent of each other.

SPSS provides extensive scope for regression analysis using various types of selection processes.

The method of selecting of independent variables for linear regression analysis is an important choice which the researcher should consider before running the analysis. You can construct a variety of regression models from the same set of variables by using different methods.

You can enter all the variables in a single step or enter the independent variables selectively.



Variable selection method:

It allows you to specify how independent variables are entered into the regression analysis. The following options are available:

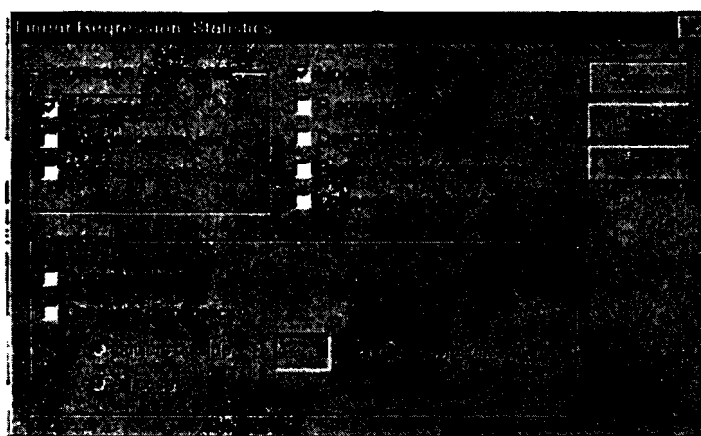
- Enter: To enter all the variables in a single step, select Enter option.
- Remove: To remove the variables in a block in a single step.
- Forward: It enters one variable at a time based on the selected criterion.

- **Backward:** All variables are entered in the first instance and then one variable is removed at a time on the selected criterion.
- **Stepwise:** Stepwise variable entry and removal examines the variables in the block at each step for entry and removal. This is a forward step procedure.

All the variables must pass the tolerance criterion to be entered in the equation, regardless of the entry method specified. The default tolerance limit is 0.0001. A new variable will not be entered if it causes the tolerance of another variable already entered to be dropped below the tolerance limit.

Linear Regression Statistics

The following statistics are available on linear regression models. Estimates and Model Fit are the two options which are selected by default.



Regression coefficient: The Estimates option displays regression coefficient, B , standard error, standard coefficient beta, t -value, and two tailed significance level of t . Covariance matrix displays a variance-covariance matrix of regression coefficients with covariance off the diagonal and variance off the diagonal. A correlation matrix will also be displayed.

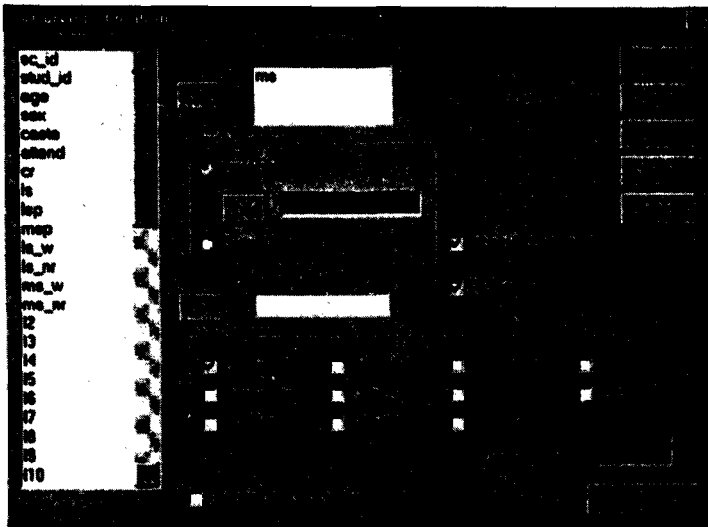
Model fit: The variables entered and removed from the model are displayed. Goodness of fit statistics, R-square, multiple R, and adjusted R square, standard error of the estimate and an analysis of variance table is displayed.

If other options are ticked, the statistics corresponding to each of the options are also displayed in the output Navigator.

If the data does not show linear relationship and the transformation procedure does not help, try using Curve Estimation procedure

19.10.3 Curve Estimation

There are many situations when the researcher is not sure about the nature of the curve that fits in a given data set. In such cases, Curve Estimation command is used to fit various types of curve on a given data. After examining the results can decide on the best fit equation. The SPSS includes 11 curve estimation regression models. A separate model is produced for each dependent variable. It is recommended that before running the curve estimation procedure, you should examine the graphical output to ascertain how the independent and dependent variables appear to be related to each other. The linear relationship assumes that the dependent variable will be normally distributed.



A scatter plot of learning achievement may reveal that the relationship between the mean score and the time spent on a task is linearly related. You might like to fit a linear model to the data and check the validity of the assumption of the model. It is quite possible that a non-linear model may give the best fit.

19.11 NON-PARAMETERIC TESTS

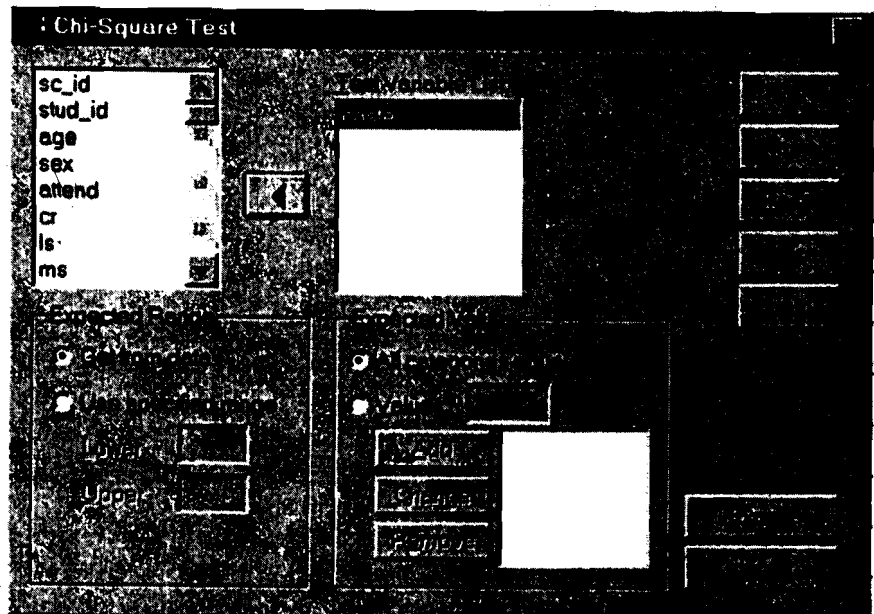
The non-parametric test procedure provides several tests that do not require assumptions about the shape of the underlying distribution. These include the following most commonly used test:

- Chi-square test
- Binomial test
- Run test
- One sample kolmogorov Seminov test
- Two independent Sample tests
- Tests for several independent samples
- Two related sample tests
- Tests for several related samples.

Here, we shall discuss the procedure for Chi-square test only. You are advised to consult the SPSS users' manual and other statistical books for detailed discussion on the other tests.

Chi-Square

Chi-square test is the most commonly used test in educational research. This test compares the observed and the expected frequencies in each cell/category to test either that all categories contain the same proportion of values or that each category contains a user specified proportion of values.



Consider that a bag contains red, white and yellow balls. You want to test the hypothesis that the bag contains all types of balls in equal proportion. To obtain Chi-square test, choose Chi-square from Non-parametric tests in the Statistics command. Select one or more variables. Each variable produces a separate output.

By default, all categories have equal expected values as shown in the above figure. Categories can have user specified proportion also. In order to provide user specific expected values, select the values option and add the user expected values. The sequence in which the values are entered is very important in this case. It corresponds to the ascending order of the category values of the test variable.

19.12 SPSS PRODUCTION FACILITY

The SPSS Production facility provides the ability to run SPSS in an automated mode. SPSS runs unattended and uninterrupted and terminates after executing the last command. Production mode is useful if you run the same set of time-consuming analysis periodically.

The SPSS Production facility uses command syntax file to tell SPSS about the commands to be executed. We have already discussed the important features of the command syntax. The command syntax file can be edited in a standard text editor.

To run the SPSS Production facility, quit the SPSS if it is already running. SPSS Production facility cannot be run when SPSS is running. Start SPSS Production program from the start window of window 2000/XP. Specify the syntax file that you want to use in the production job. Click Browse to select the Syntax File. Save the production file job. Run the production file job at any time.

19.13 STATISTICAL ANALYSIS SYSTEM (SAS)

Like the SPSS, the Statistical Analysis System (SAS) package calculate descriptive statistic of your choice e.g., Mean, Standard Deviation etc. SAS is available for both mainframe and personal computers. It is strong in its treatment of data, in clarity of its graphics and in certain business applications. The various statistical procedures carried out by SAS are always preceded by the word PROC which

stand for procedure. The most commonly used SAS statistical procedures are as follows: (Sprinthall et.al, 1991).

- PROC MEANS: Descriptive statistics (mean, standard deviation, maximum and minimum values and so on).
- PROC CORR: Pearson correlation between two or more variables.
- PROC t-TEST: t-test for significant difference between the means of two groups.
- PROC ANOVA: Analysis of variance for all types of designs (one way, two-way and other).
- PROC FREQ: Frequency distribution for one or more variables.

As pointed out by Klieger (1984) SAS package is comparatively more difficult to use due to its procedural complexities. For greater details on SAS package you are advised to consult the books by Klieger and Sprinthall.

19.14 NUDIST

Computer programs help in the analysis of qualitative data, especially in understanding a large (say 500 or more pages) text database, Studies using large databases such as ethnographies with extensive interviews, computer programs provide an invaluable aid in research.

NUDIST (Non-numerical unstructured data indexing, searching and theorizing) program was developed in Australia in 1991. This package is used for qualitative analysis of data. Here we present briefly the main features of this package. This software requires, 4 megabytes of RAM and atleast 2 megabytes space for data files in your PC or MAC. In your PC it operates under windows (Creswell 1998).

As a researcher this software will help you to provide the following.

1. Storing and organizing files: First establish document files and store information with the NUDIST programme. Document files consist of transcript from an interview, notes of observation or any article scanned from a newspaper.
2. Searching for themes: Tag segments of text from all the documents that relate to a single idea or theme. For example, distance learners, in a study on effectiveness of distance education talk about the role of academic counsellors. The researcher can create a node in NUDIST as 'Role of Academic Counsellors'. Researcher will select text in the transcript where learners have talked about this role and merge it into role of Academic Counsellors. Information can be retained in this node and researcher can take print in different ways in which learners talk about the role of Academic Counsellors.
3. Crossing themes: Taking the same example of roll of counsellors, the researcher can relate this node to other nodes. Suppose the other node is qualifications of counsellor. There are two categories like Graduate and Post Graduate. The researcher will ask NUDIST to cross the two categories, role of counsellors and qualification of counsellors to see for example whether there is any relation between graduate counsellor and their roll than the post graduate counsellor and their roll. NUDIST software generates information for a matrix with information in the cells reflecting different perspectives.
4. Diagramming: In this package, once the information is categorized, categories are identified. These categories are developed into nine visual picture of the

categories that display their inter connectedness. This is called a tree diagram in NUDIST software. Tree diagram is a hierarchical tree of categories where root node is the top and parents and siblings in the tree. This tree diagram is a useful device for discussing the data analysis of quantitative research in conference.

5. Creating a template: In a qualitative research, at the beginning of data analysis, the researcher will create a template which is apriori code book for organizing information.

For further details on NUDIST software you may like to consult the following.

Kelle, E.(ed.), *Computer-aided Qualitative Data Analysis*, Thousand Oaks, CA: Sage, 1995.

Tesch, R., *Qualitative Research: Analysis Types and Software Tools*, Bristol, PA: Falmer, 1990.

Check Your Progress

Notes : a) Space is given below for your answer.

b) Compare your answer with the one given at the end of this unit.

5. Name the six characteristics of data sheet.

.....

.....

.....

.....

.....

6. What is NUDIST?

.....

.....

.....

.....

19.15 LET US SUM UP

The foregoing details examined the various types of statistical application of the SPSS in data management, presentation and analysis. The discussion was based on the assumption that you have a basic understanding of the statistical methods. It was highlighted that the researchers must try to explore the data using various simple but powerful statistical techniques. In this connection, six characteristics of the data were examined for exploring it fully. The procedures involved in the use of various statistics were also discussed in detail. Procedures for running regression analysis to understand the relationship between variables were also discussed. Those of you who are comfortable with the basic statistical procedures in SPSS can explore the advanced features, including those aimed at automating the statistical analysis using the SPSS Production facility and also the use of scripting in data analyzing.

19.16 UNIT-END ACTIVITIES

1. Obtain quantitative data for any kinds of research work with the help of research tools and analyse the data using SPSS package.
2. Analyse the data collected for your disseration work using SPSS package.

19.17 SUGGESTED READINGS

SPSS Base 7.5 for Windows, User's Guide, SPSS Inc.

SPSS Base 7.5 Application Guide, SPSS Inc.

SPSS Advanced Statistics 7.5, SPSS Inc.

A number of white papers dealing with various applications are available on the SPSS website: www.spss.com. This site is updated regularly with new materials. Advanced user may like to obtain/download the relevant materials for their use.

Creswell, John, W. (1998): *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Sage Publications, Inc., International Educational and Professional Publisher, USA.

Klieger, D.M.(1984): *Computer Usage for Social Scientists*. Newton, Mass: Allyn and Bacon.

Sprinthall, Richard C., (et.al.) (1991): *Understanding Educational Research*. New Jersey: Prentice Hall.

19.18 ANSWERS TO CHECK YOUR PROGRESS

1.
 - a) SPSS table creates a high quality presentation in tabular form.
 - b) SPSS Trends performs comprehensive forecasting and time series analysis.
 - c) SPSS categories performs conjoint analysis and optional scaling procedures.
2. There are seven types of windows in SPSS. These are data editor, output navigator, pivot tables, graphics, text editor, syntax editor and script editor.
3.
 - i) Bring your data into SPSS
 - ii) Select a procedure from Menu
 - iii) Select the variables
 - iv) Run the procedure and examine the output
4. Split file is meant for comparing a summary and other statistics based on certain group behaviour.
5.
 - i) shape, ii) Location, iii) Spread, iv) Outliers, v) Clustering vi) Association and relationship.
6. NUDIST (Non-numerical Unstructured Data Indexing, Searching and Theorizing) is a software package for qualitative data analysis.

APPENDIX

Table A : Fractional parts of the total area (taken as 10,000) under the normal probability curve, corresponding to distances on the base line between the mean and successive points laid off from the mean in units of standard deviation.

| x | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|-----|--------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0 | 0000 | 0040 | 0080 | 0120 | 0160 | 0199 | 0239 | 0279 | 0319 | 0359 |
| 0.1 | 0398 | 0438 | 0478 | 0517 | 0557 | 0596 | 0636 | 0675 | 0714 | 0753 |
| 0.2 | 0793 | 0832 | 0871 | 0910 | 0948 | 0987 | 1026 | 1064 | 1103 | 1141 |
| 0.3 | 1179 | 1217 | 1255 | 1293 | 1331 | 1368 | 1406 | 1443 | 1480 | 1517 |
| 0.4 | 1554 | 1591 | 1628 | 1664 | 1700 | 1736 | 1772 | 1808 | 1844 | 1879 |
| 0.5 | 1915 | 1950 | 1985 | 2019 | 2054 | 2088 | 2123 | 2157 | 2190 | 2224 |
| 0.6 | 2257 | 2291 | 2324 | 2357 | 2389 | 2422 | 2454 | 2486 | 2517 | 2549 |
| 0.7 | 2580 | 2611 | 2642 | 2673 | 2704 | 2734 | 2764 | 2794 | 2823 | 2852 |
| 0.8 | 2881 | 2910 | 2939 | 2967 | 2995 | 3023 | 3051 | 3078 | 3106 | 3133 |
| 0.9 | 3159 | 3186 | 3212 | 3238 | 3264 | 3290 | 3315 | 3340 | 3365 | 3389 |
| 1.0 | 3413 | 3438 | 3461 | 3485 | 3508 | 3531 | 3554 | 3577 | 3599 | 3621 |
| 1.1 | 3643 | 3665 | 3686 | 3708 | 3729 | 3749 | 3770 | 3790 | 3810 | 3830 |
| 1.2 | 3849 | 3869 | 3888 | 3907 | 3925 | 3944 | 3962 | 3980 | 3997 | 4015 |
| 1.3 | 4032 | 4049 | 4066 | 4082 | 4099 | 4115 | 4131 | 4147 | 4162 | 4177 |
| 1.4 | 4192 | 4207 | 4222 | 4236 | 4251 | 4265 | 4279 | 4292 | 4306 | 4319 |
| 1.5 | 4332 | 4345 | 4357 | 4370 | 4383 | 4394 | 4406 | 4418 | 4429 | 4441 |
| 1.6 | 4452 | 4463 | 4474 | 4484 | 4495 | 4505 | 4515 | 4525 | 4535 | 4545 |
| 1.7 | 4554 | 4564 | 4573 | 4582 | 4591 | 4599 | 4608 | 4616 | 4625 | 4633 |
| 1.8 | 4641 | 4649 | 4656 | 4664 | 4671 | 4678 | 4686 | 4693 | 4699 | 4706 |
| 1.9 | 4713 | 4719 | 4726 | 4732 | 4738 | 4744 | 4750 | 4756 | 4761 | 4767 |
| 2.0 | 4772 | 4778 | 4783 | 4788 | 4793 | 4798 | 4803 | 4808 | 4812 | 4817 |
| 2.1 | 4821 | 4826 | 4830 | 4834 | 4838 | 4842 | 4846 | 4850 | 4854 | 4857 |
| 2.2 | 4861 | 4864 | 4868 | 4871 | 4875 | 4878 | 4881 | 4884 | 4887 | 4890 |
| 2.3 | 4893 | 4896 | 4898 | 4901 | 4904 | 4906 | 4909 | 4911 | 4913 | 4916 |
| 2.4 | 4918 | 4920 | 4922 | 4925 | 4927 | 4929 | 4931 | 4932 | 4934 | 4936 |
| 2.5 | 4938 | 4940 | 4941 | 4943 | 4945 | 4946 | 4948 | 4949 | 4951 | 4952 |
| 2.6 | 4953 | 4955 | 4956 | 4957 | 4959 | 4960 | 4961 | 4962 | 4963 | 4964 |
| 2.7 | 4965 | 4966 | 4967 | 4968 | 4969 | 4970 | 4971 | 4972 | 4973 | 4974 |
| 2.8 | 4974 | 4975 | 4976 | 4977 | 4977 | 4978 | 4979 | 4979 | 4980 | 4981 |
| 2.9 | 4981 | 4982 | 4982 | 4983 | 4984 | 4984 | 4985 | 4985 | 4986 | 4986 |
| 3.0 | 4986.5 | 4986.9 | 4987.4 | 4987.8 | 4988.2 | 4988.6 | 4988.9 | 4989.3 | 4989.7 | 4990.0 |
| 3.1 | 4990.3 | 4990.6 | 4991.0 | 4991.3 | 4991.6 | 4991.8 | 4992.1 | 4992.4 | 4992.6 | 4992.9 |
| 3.2 | 4993.129 | | | | | | | | | |
| 3.3 | 4995.166 | | | | | | | | | |
| 3.4 | 4996.631 | | | | | | | | | |
| 3.5 | 4997.674 | | | | | | | | | |
| 3.6 | 4998.409 | | | | | | | | | |
| 3.7 | 4998.922 | | | | | | | | | |
| 3.8 | 4999.277 | | | | | | | | | |
| 3.9 | 4999.519 | | | | | | | | | |
| 4.0 | 4999.683 | | | | | | | | | |
| 4.5 | 4999.966 | | | | | | | | | |
| 5.0 | 4999.9971333 | | | | | | | | | |

Table B : Ordinates of the normal probability curve expressed as fractional parts of the mean ordinate, Y.

| <i>x</i> | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.0 | 00000 | 99995 | 99980 | 99955 | 99920 | 99875 | 99820 | 99755 | 99685 | 99596 |
| 0.1 | 99501 | 99396 | 99283 | 99158 | 99025 | 98881 | 98728 | 98565 | 98393 | 98211 |
| 0.2 | 98020 | 97819 | 97609 | 97390 | 97161 | 96923 | 96676 | 96420 | 96156 | 95882 |
| 0.3 | 95600 | 95309 | 95010 | 94702 | 94387 | 94055 | 93723 | 93382 | 93024 | 92677 |
| 0.4 | 92312 | 91399 | 91558 | 91169 | 90774 | 90371 | 89961 | 89543 | 89119 | 88688 |
| 0.5 | 88250 | 87805 | 87353 | 86896 | 86432 | 85962 | 85488 | 85006 | 84519 | 84060 |
| 0.6 | 83527 | 83023 | 82514 | 82010 | 81481 | 80957 | 80429 | 79896 | 79359 | 78817 |
| 0.7 | 78270 | 77721 | 77167 | 76610 | 76048 | 75484 | 74916 | 74342 | 73769 | 73193 |
| 0.8 | 72615 | 72033 | 71448 | 70861 | 70272 | 69681 | 69087 | 68493 | 67896 | 67298 |
| 0.9 | 66689 | 66097 | 65494 | 64891 | 64287 | 63683 | 63077 | 62472 | 61865 | 61259 |
| 1.0 | 60653 | 60047 | 59440 | 58834 | 58228 | 57623 | 57017 | 56414 | 55810 | 55209 |
| 1.1 | 54607 | 54007 | 53409 | 52812 | 52214 | 51629 | 51027 | 50437 | 49848 | 49260 |
| 1.2 | 48675 | 48092 | 47511 | 46933 | 46357 | 45783 | 45212 | 44644 | 44078 | 43516 |
| 1.3 | 42956 | 42399 | 41845 | 41294 | 40747 | 40202 | 39661 | 39123 | 38569 | 38058 |
| 1.4 | 37531 | 37007 | 36487 | 35971 | 35459 | 34950 | 34445 | 33944 | 33447 | 32954 |
| 1.5 | 32465 | 31980 | 31500 | 31023 | 30550 | 30082 | 29618 | 29158 | 28702 | 28251 |
| 1.6 | 27804 | 27361 | 26923 | 26489 | 26059 | 25634 | 25213 | 24797 | 24385 | 23978 |
| 1.7 | 23575 | 23176 | 22782 | 22392 | 22008 | 21627 | 21251 | 20879 | 20511 | 20148 |
| 1.8 | 19790 | 19436 | 19086 | 18741 | 18400 | 18064 | 17732 | 17404 | 17081 | 16762 |
| 1.9 | 16448 | 16137 | 15831 | 15530 | 15232 | 14939 | 14650 | 14364 | 14083 | 13806 |
| 2.0 | 13534 | 13265 | 13000 | 12749 | 12483 | 12230 | 11981 | 11737 | 11496 | 11259 |
| 2.1 | 11025 | 10795 | 10570 | 10347 | 10129 | 09914 | 09702 | 09495 | 09290 | 09090 |
| 2.2 | 08892 | 08698 | 08507 | 08320 | 08136 | 07956 | 07778 | 07604 | 07433 | 07265 |
| 2.3 | 07199 | 06939 | 06780 | 06624 | 06471 | 06321 | 06174 | 06029 | 05888 | 05750 |
| 2.4 | 05614 | 05481 | 05350 | 05222 | 05096 | 04973 | 04852 | 04734 | 04618 | 04505 |
| 2.5 | 04394 | 04285 | 04179 | 04074 | 03972 | 03873 | 03775 | 03680 | 03586 | 03494 |
| 2.6 | 03405 | 03317 | 03232 | 03148 | 03066 | 02986 | 02908 | 02831 | 02757 | 02684 |
| 2.7 | 02612 | 02542 | 02474 | 02408 | 02343 | 02280 | 02218 | 02157 | 02098 | 02040 |
| 2.8 | 01984 | 01929 | 01876 | 01823 | 01772 | 01723 | 01674 | 01627 | 01581 | 01536 |
| 2.9 | 01492 | 01449 | 01408 | 01367 | 01328 | 01288 | 01252 | 01215 | 01179 | 01145 |
| 3.0 | 01111 | 00819 | 00598 | 00432 | 00309 | 00219 | 00153 | 00106 | 00073 | 00050 |
| 4.0 | 00034 | 00022 | 00015 | 00010 | 00006 | 00004 | 00003 | 00002 | 00001 | 00001 |
| 5.0 | 00000 | | | | | | | | | |

Table C : Table of Critical Values of t.

| | <i>Degrees of Freedom</i> | | <i>Level of Significance</i> | | |
|------|---------------------------|-------|------------------------------|-------|--|
| | | | | | |
| | 0.10 | 0.05 | 0.02 | 0.01 | |
| 1 | 6.34 | 12.71 | 31.82 | 63.66 | |
| 2 | 2.92 | 4.30 | 6.96 | 9.62 | |
| 3 | 2.35 | 3.18 | 4.54 | 5.84 | |
| 4 | 2.13 | 2.78 | 3.75 | 4.60 | |
| 5 | 2.02 | 2.57 | 3.36 | 4.03 | |
| 6 | 1.94 | 2.45 | 3.14 | 3.71 | |
| 7 | 1.90 | 2.36 | 3.00 | 3.50 | |
| 8 | 1.86 | 2.31 | 2.90 | 3.36 | |
| 9 | 1.83 | 2.26 | 2.82 | 3.25 | |
| 10 | 1.81 | 2.23 | 2.76 | 3.17 | |
| 11 | 1.80 | 2.20 | 2.72 | 3.11 | |
| 12 | 1.78 | 2.18 | 2.68 | 3.06 | |
| 13 | 1.77 | 2.16 | 2.65 | 3.01 | |
| 14 | 1.76 | 2.14 | 2.62 | 2.98 | |
| 15 | 1.75 | 2.13 | 2.60 | 2.95 | |
| 16 | 1.75 | 2.12 | 2.58 | 2.92 | |
| 17 | 1.74 | 2.11 | 2.57 | 2.90 | |
| 18 | 1.73 | 2.10 | 2.55 | 2.88 | |
| 19 | 1.73 | 2.09 | 2.54 | 2.86 | |
| 20 | 1.72 | 2.09 | 2.53 | 2.84 | |
| 21 | 1.72 | 2.08 | 2.52 | 2.83 | |
| 22 | 1.72 | 2.07 | 2.51 | 2.82 | |
| 23 | 1.71 | 2.07 | 2.50 | 2.81 | |
| 24 | 1.71 | 2.06 | 2.49 | 2.80 | |
| 25 | 1.71 | 2.06 | 2.48 | 2.79 | |
| 26 | 1.71 | 2.06 | 2.48 | 2.78 | |
| 27 | 1.70 | 2.05 | 2.47 | 2.77 | |
| 28 | 1.70 | 2.05 | 2.47 | 2.76 | |
| 29 | 1.70 | 2.04 | 2.46 | 2.76 | |
| 30 | 1.70 | 2.04 | 2.46 | 2.75 | |
| 35 | 1.69 | 2.03 | 2.44 | 2.72 | |
| 40 | 1.68 | 2.02 | 2.42 | 2.71 | |
| 45 | 1.68 | 2.02 | 2.41 | 2.69 | |
| 50 | 1.68 | 2.01 | 2.40 | 2.68 | |
| 60 | 1.67 | 2.00 | 2.39 | 2.66 | |
| 70 | 1.67 | 2.00 | 2.38 | 2.65 | |
| 80 | 1.66 | 1.99 | 2.38 | 2.64 | |
| 90 | 1.66 | 1.99 | 2.37 | 2.63 | |
| 100 | 1.66 | 1.98 | 2.36 | 2.63 | |
| 125 | 1.66 | 1.98 | 2.36 | 2.62 | |
| 150 | 1.66 | 1.98 | 2.35 | 2.61 | |
| 200 | 1.65 | 1.97 | 2.35 | 2.60 | |
| 300 | 1.65 | 1.97 | 2.34 | 2.59 | |
| 400 | 1.65 | 1.97 | 2.34 | 2.59 | |
| 500 | 1.65 | 1.96 | 2.33 | 2.59 | |
| 1000 | 1.65 | 1.96 | 2.33 | 2.58 | |
| | 1.65 | 1.96 | 2.33 | 2.58 | |

| | | | | | | | | | | |
|----|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| 12 | 4.75 9.33 | 3.88 6.93 | 3.49 5.95 | 3.26 5.41 | 3.11 5.06 | 3.00 4.82 | 2.85 4.50 | 2.69 4.16 | 2.50 3.78 | 2.30 3.36 |
| 13 | 4.67 9.07 | 3.80 6.70 | 3.41 5.74 | 3.18 5.20 | 3.02 4.86 | 2.92 4.62 | 2.77 4.30 | 2.60 3.96 | 2.42 3.59 | 2.21 3.16 |
| 14 | 4.60 8.86 | 3.74 6.51 | 3.34 5.56 | 3.11 5.03 | 2.96 4.69 | 2.85 4.46 | 2.70 4.14 | 2.53 3.80 | 2.35 3.43 | 2.13 3.00 |
| 15 | 4.54 8.68 | 3.68 6.36 | 3.29 5.42 | 3.06 4.89 | 2.90 4.56 | 2.79 4.32 | 2.64 4.00 | 2.48 3.67 | 2.29 3.29 | 2.07 2.87 |
| 16 | 4.49 8.53 | 3.63 6.23 | 3.24 5.29 | 3.01 4.77 | 2.85 4.44 | 2.74 4.20 | 2.59 3.89 | 2.42 3.55 | 2.24 3.18 | 2.01 2.75 |
| 17 | 4.45 8.40 | 3.59 6.11 | 3.20 5.18 | 2.96 4.67 | 2.81 4.34 | 2.70 4.10 | 2.55 3.79 | 2.38 3.45 | 2.19 3.08 | 1.96 2.65 |
| 18 | 4.41 8.28 | 3.55 6.01 | 3.16 5.09 | 2.93 4.58 | 2.77 4.25 | 2.66 4.01 | 2.51 3.71 | 2.34 3.37 | 2.15 3.01 | 1.92 2.57 |
| 19 | 4.38 8.18 | 3.52 5.93 | 3.13 5.01 | 2.90 4.50 | 2.74 4.17 | 2.63 3.94 | 2.48 3.63 | 2.31 3.30 | 2.11 2.92 | 1.88 2.49 |
| 20 | 4.35 8.10 | 3.49 5.85 | 3.10 4.94 | 2.87 4.43 | 2.71 4.10 | 2.60 3.87 | 2.45 3.56 | 2.28 3.23 | 2.08 2.86 | 1.84 2.42 |
| 21 | 4.32 8.02 | 3.47 5.78 | 3.07 4.87 | 2.84 4.37 | 2.68 4.04 | 2.57 3.81 | 2.42 3.51 | 2.25 3.17 | 2.05 2.80 | 1.81 2.36 |
| 22 | 4.30 7.94 | 3.44 5.72 | 3.05 4.82 | 2.82 4.31 | 2.66 3.99 | 2.55 3.75 | 2.40 3.45 | 2.23 3.12 | 2.03 2.75 | 1.78 2.30 |
| 23 | 4.28 7.88 | 3.42 5.66 | 3.03 4.76 | 2.80 4.46 | 2.64 3.94 | 2.53 3.71 | 2.38 3.41 | 2.20 3.07 | 2.00 2.70 | 1.76 2.26 |

| | | | | | | | | | | |
|----|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| 24 | 4.26 7.82 | 3.40 5.61 | 3.01 4.72 | 2.78 4.22 | 2.62 3.90 | 2.51 3.67 | 2.36 3.36 | 2.18 3.93 | 1.98 2.66 | 1.73 4.21 |
| 25 | 4.24 7.77 | 3.38 5.57 | 2.99 4.68 | 2.76 4.18 | 2.60 3.86 | 2.49 3.63 | 2.34 3.32 | 2.16 2.99 | 1.96 2.62 | 1.71 2.17 |
| 26 | 4.22 7.72 | 3.37 5.53 | 2.98 4.64 | 2.74 4.14 | 2.59 3.82 | 2.47 3.59 | 2.32 3.29 | 2.15 2.96 | 1.95 2.58 | 1.69 2.13 |
| 27 | 4.21 7.68 | 3.35 5.49 | 2.96 4.60 | 2.73 4.11 | 2.57 3.78 | 2.46 3.56 | 2.30 3.26 | 2.13 2.93 | 1.93 2.55 | 1.67 2.10 |
| 28 | 4.20 7.64 | 3.34 5.45 | 2.95 4.57 | 2.71 4.07 | 2.56 3.75 | 2.44 3.53 | 2.29 3.23 | 2.12 2.90 | 1.91 2.52 | 1.65 2.06 |
| 29 | 4.18 7.60 | 3.33 5.42 | 2.93 4.54 | 2.70 4.04 | 2.54 3.73 | 2.43 3.50 | 2.28 3.20 | 2.10 2.87 | 1.90 2.49 | 1.64 2.08 |
| 30 | 4.17 7.56 | 3.32 5.39 | 2.92 4.51 | 2.69 4.02 | 2.53 3.70 | 2.42 3.47 | 2.27 3.17 | 2.09 2.84 | 1.89 2.47 | 1.62 2.01 |
| 35 | 4.12 7.42 | 3.26 5.27 | 2.87 4.40 | 2.64 3.91 | 2.48 3.59 | 2.37 3.37 | 2.22 3.07 | 2.04 2.74 | 1.83 2.1 | 1.57 1.90 |
| 40 | 4.08 7.31 | 3.23 5.18 | 2.84 4.31 | 2.61 3.83 | 2.45 3.51 | 2.34 3.29 | 2.18 2.99 | 2.00 2.66 | 1.79 2.29 | 1.52 1.82 |
| 45 | 4.06 7.23 | 3.21 5.11 | 2.81 4.25 | 2.58 3.77 | 2.42 3.45 | 2.31 3.23 | 2.15 2.94 | 1.97 2.61 | 1.76 2.23 | 1.48 1.75 |
| 50 | 4.03 7.17 | 3.18 5.06 | 2.79 4.20 | 2.56 3.72 | 2.40 3.41 | 2.29 3.19 | 2.13 2.89 | 1.95 2.56 | 1.74 2.18 | 1.44 1.68 |
| 60 | 4.00 7.08 | 3.15 4.98 | 2.76 4.13 | 2.52 3.65 | 2.37 3.34 | 2.25 3.12 | 2.10 2.82 | 1.92 2.50 | 1.70 2.12 | 1.39 1.60 |
| 70 | 3.98 7.01 | 3.13 4.92 | 2.74 4.07 | 2.50 3.60 | 2.35 3.29 | 2.23 3.07 | 2.07 2.78 | 1.89 2.45 | 1.67 2.07 | 1.35 1.53 |

| | | | | | | | | | | |
|------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| 80 | 3.96 6.96 | 3.11 4.88 | 2.72 4.04 | 2.49 3.56 | 2.33 3.26 | 2.21 3.04 | 2.06 2.74 | 1.88 2.42 | 1.65 2.03 | 1.31 1.47 |
| 90 | 3.95 6.92 | 3.10 4.85 | 2.71 4.01 | 2.47 3.53 | 2.32 3.23 | 2.20 3.01 | 2.04 2.72 | 1.86 2.39 | 1.64 2.00 | 1.28 1.43 |
| 100 | 3.94 6.90 | 3.09 4.82 | 2.70 3.98 | 2.46 3.51 | 2.30 3.21 | 2.19 2.99 | 2.03 2.69 | 1.85 2.37 | 1.63 1.98 | 1.26 1.39 |
| 125 | 3.92 6.84 | 3.07 4.78 | 2.68 3.94 | 2.44 3.47 | 2.29 3.17 | 2.17 2.95 | 2.01 2.66 | 1.83 2.33 | 1.60 1.94 | 1.21 1.32 |
| 150 | 3.90 6.81 | 3.06 4.75 | 2.66 3.91 | 2.43 3.45 | 2.27 3.14 | 2.16 2.92 | 2.00 2.63 | 1.82 2.31 | 1.59 1.92 | 1.18 1.27 |
| 200 | 3.89 6.76 | 3.04 4.71 | 2.65 3.88 | 2.42 3.41 | 2.26 3.11 | 2.14 2.89 | 1.98 2.60 | 1.80 2.28 | 1.57 1.88 | 1.14 1.21 |
| 300 | 3.87 6.72 | 3.03 4.68 | 2.64 3.85 | 2.41 3.38 | 2.25 3.08 | 2.13 2.86 | 1.97 2.57 | 1.79 2.24 | 1.55 1.85 | 1.10 1.14 |
| 400 | 3.86 6.70 | 3.02 4.66 | 2.63 3.83 | 2.40 3.37 | 2.24 3.06 | 2.12 2.85 | 1.96 2.56 | 1.78 2.23 | 1.54 1.84 | 1.07 1.11 |
| 500 | 3.86 6.69 | 3.01 4.65 | 2.62 3.82 | 2.39 3.36 | 2.23 3.05 | 2.11 2.84 | 1.96 2.55 | 1.77 2.22 | 1.54 1.83 | 1.06 1.08 |
| 1000 | 3.85 6.66 | 3.00 4.63 | 2.61 3.80 | 2.38 3.34 | 2.22 3.04 | 2.10 2.82 | 1.95 2.53 | 2.76 2.20 | 1.53 1.81 | 1.03 1.04 |
| | 3.84 6.64 | 2.99 4.60 | 2.60 3.78 | 2.37 3.32 | 2.21 3.02 | 2.09 2.80 | 1.94 2.51 | 1.75 2.18 | 1.52 1.79 | |

Degree of freedom for smaller mean square

Table E: X² Table. P gives the probability of exceeding the tabulated value of x² for the specified number of degrees of freedom (df). The values of x² are printed in the body of the table.

| <i>df.</i> | <i>0.95</i> | <i>0.90</i> | <i>0.80</i> | <i>0.70</i> | <i>0.50</i> | <i>0.30</i> | <i>0.20</i> | <i>0.10</i> | <i>0.05</i> | <i>0.02</i> | <i>0.01</i> |
|------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 1 | 0.00393 | 0.0158 | 0.0642 | 0.148 | 0.455 | 1.074 | 1.642 | 2.706 | 3.841 | 5.412 | 6.635 |
| 2 | 0.103 | 0.211 | 0.446 | 0.713 | 1.386 | 2.408 | 3.219 | 4.605 | 5.991 | 7.824 | 9.210 |
| 3 | 0.352 | 0.584 | 1.005 | 1.424 | 2.366 | 3.665 | 4.642 | 6.251 | 7.815 | 9.837 | 11.345 |
| 4 | 0.711 | 1.064 | 1.649 | 2.195 | 3.357 | 4.878 | 5.989 | 7.779 | 9.488 | 11.668 | 13.277 |
| 5 | 1.145 | 1.610 | 2.343 | 3.000 | 4.351 | 6.064 | 7.289 | 9.236 | 11.070 | 13.388 | 15.086 |
| 6 | 1.635 | 2.204 | 3.070 | 3.828 | 5.348 | 7.231 | 8.558 | 10.645 | 12.592 | 15.083 | 16.812 |
| 7 | 2.167 | 2.833 | 3.822 | 4.671 | 6.346 | 8.383 | 9.803 | 12.017 | 14.067 | 16.622 | 18.475 |
| 8 | 2.733 | 3.490 | 4.594 | 5.527 | 7.344 | 9.524 | 11.030 | 13.362 | 15.507 | 18.168 | 20.090 |
| 9 | 3.325 | 4.168 | 5.380 | 6.393 | 8.343 | 10.656 | 12.242 | 14.684 | 16.919 | 19.679 | 21.666 |
| 10 | 3.940 | 4.865 | 6.179 | 7.267 | 9.342 | 11.781 | 13.442 | 15.987 | 18.307 | 21.161 | 23.209 |
| 11 | 4.575 | 5.578 | 6.989 | 8.148 | 10.341 | 12.899 | 14.631 | 17.275 | 19.675 | 22.618 | 24.725 |
| 12 | 5.226 | 6.304 | 7.807 | 9.034 | 11.340 | 14.011 | 15.812 | 18.549 | 21.026 | 24.054 | 26.217 |
| 13 | 5.892 | 7.042 | 8.634 | 9.926 | 12.340 | 15.119 | 16.985 | 19.812 | 22.362 | 25.472 | 27.688 |
| 14 | 6.571 | 7.790 | 9.467 | 10.821 | 13.339 | 16.222 | 18.151 | 21.064 | 23.685 | 26.873 | 29.141 |
| 15 | 7.261 | 8.547 | 10.307 | 11.721 | 14.339 | 17.322 | 19.311 | 22.307 | 24.996 | 28.259 | 30.578 |

| | | | | | | | | | | | |
|----|--------|--------|--------|---------|--------|--------|--------|--------|--------|--------|--------|
| 16 | 7.962 | 9.312 | 11.152 | 12.624 | 15.338 | 18.418 | 20.465 | 23.542 | 26.296 | 26.633 | 32.000 |
| 17 | 8.672 | 10.085 | 12.002 | 13.531 | 16.338 | 19.511 | 21.615 | 24.760 | 27.587 | 30.995 | 33.409 |
| 18 | 9.390 | 10.865 | 12.857 | 14.440 | 17.338 | 20.601 | 22.760 | 25.869 | 28.869 | 32.346 | 34.805 |
| 19 | 10.117 | 11.651 | 13.716 | 15.352 | 18.338 | 21.689 | 23.900 | 27.204 | 30.144 | 33.687 | 36.191 |
| 20 | 10.851 | 12.443 | 14.578 | 16.266 | 19.337 | 22.775 | 25.038 | 28.412 | 31.410 | 35.020 | 37.566 |
| 21 | 11.591 | 13.240 | 15.445 | 17.182 | 20.337 | 23.858 | 26.171 | 29.615 | 32.671 | 36.343 | 38.932 |
| 22 | 12.338 | 14.041 | 16.314 | 18.101 | 21.337 | 24.939 | 27.301 | 30.813 | 33.924 | 37.659 | 40.289 |
| 23 | 13.091 | 14.848 | 17.187 | σ19.021 | 22.337 | 26.018 | 28.429 | 32.007 | 35.172 | 38.968 | 41.638 |
| 24 | 13.848 | 15.659 | 18.062 | 19.943 | 23.337 | 27.096 | 29.553 | 33.196 | 36.415 | 40.270 | 42.980 |
| 25 | 14.611 | 16.473 | 18.940 | 20.867 | 24.337 | 28.172 | 30.675 | 34.382 | 37.652 | 41.566 | 44.314 |
| 26 | 15.379 | 17.292 | 19.820 | 21.792 | 25.336 | 29.246 | 31.795 | 35.563 | 28.885 | 42.856 | 45.642 |
| 27 | 16.151 | 18.114 | 20.703 | 22.719 | 26.336 | 30.319 | 32.912 | 36.741 | 40.113 | 44.140 | 46.963 |
| 28 | 19.928 | 18.939 | 21.588 | 23.647 | 27.366 | 31.391 | 34.027 | 37.916 | 41.337 | 45.9 | 48.278 |
| 29 | 17.708 | 19.768 | 22.475 | 24.577 | 28.336 | 32.461 | 35.139 | 39.087 | 42.557 | 46.69 | 49.588 |
| 30 | 18.493 | 20.599 | 23.364 | 25.508 | 29.386 | 33.530 | 36.250 | 40.256 | 43.773 | 47.962 | 50.892 |

Table F: Conversion of a Pearson r into a corresponding Fisher's z coefficient.*

| r | z | r | z | r | z | r | z | r | z | r | z |
|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|
| .25 | .26 | .40 | .42 | .55 | .62 | .70 | .87 | .85 | 1.26 | .950 | 1.83 |
| .26 | .27 | .41 | .44 | .56 | .63 | .71 | .89 | .86 | 1.29 | .955 | 1.89 |
| .27 | .28 | .42 | .45 | .57 | .65 | .72 | .91 | .87 | 1.33 | .960 | 1.95 |
| .28 | .29 | .43 | .46 | .59 | .68 | .74 | .95 | .89 | 1.42 | .970 | 2.09 |
| .30 | .31 | .45 | .48 | .60 | .69 | .75 | .97 | .90 | 1.47 | .975 | 2.18 |
| .31 | .32 | .46 | .50 | .61 | .71 | .76 | 1.00 | .905 | 1.50 | .980 | 2.30 |
| .32 | .33 | .47 | .51 | .62 | .73 | .77 | 1.02 | .910 | 1.53 | .985 | 2.44 |
| .33 | .34 | .48 | .52 | .63 | .74 | .78 | 1.05 | .915 | 1.56 | .990 | 2.65 |
| .34 | .35 | .49 | .54 | .64 | .76 | .79 | 1.07 | .920 | 1.59 | .995 | 2.99 |
| .35 | .37 | .50 | .55 | .65 | .78 | .80 | 1.10 | .925 | 1.62 | | |
| .36 | .38 | .51 | .56 | .66 | .79 | .81 | 1.13 | .930 | 1.66 | | |
| .37 | .39 | .52 | .58 | .67 | .81 | .82 | 1.16 | .935 | 1.70 | | |
| .38 | .40 | .53 | .59 | .68 | .83 | .83 | 1.19 | .940 | 1.74 | | |
| .39 | .41 | .54 | .60 | .69 | .85 | .84 | 1.22 | .945 | 1.78 | | |

* r 's under 25 may be taken as equivalent to z 's.

Table G: Table of Critical Values of K_p in the Kolmogorov-Smirnov Two-Sample Test. (Small Samples)

| N | <i>One-tailed test</i> | | <i>Two-tailed test</i> | |
|-----|------------------------|----------------|------------------------|----------------|
| | $\alpha = .05$ | $\alpha = .01$ | $\alpha = .05$ | $\alpha = .01$ |
| 3 | 3 | - | - | - |
| 4 | 4 | - | 4 | - |
| 5 | 4 | 5 | 5 | 5 |
| 6 | 5 | 6 | 5 | 6 |
| 7 | 5 | 6 | 6 | 6 |
| 8 | 5 | 6 | 6 | 7 |
| 9 | 6 | 7 | 6 | 7 |
| 10 | 6 | 7 | 7 | 8 |
| 11 | 6 | 8 | 7 | 8 |
| 12 | 6 | 8 | 7 | 8 |
| 13 | 7 | 8 | 7 | 9 |
| 14 | 7 | 8 | 8 | 9 |
| 15 | 7 | 9 | 8 | 9 |
| 16 | 7 | 9 | 8 | 10 |
| 17 | 8 | 9 | 8 | 10 |
| 18 | 8 | 10 | 9 | 10 |

| N | One- tailed test | | Two-tailed test | |
|----|------------------|----------------|-----------------|----------------|
| | $\alpha = .05$ | $\alpha = .01$ | $\alpha = .05$ | $\alpha = .01$ |
| 19 | 8 | 10 | 9 | 10 |
| 20 | 8 | 10 | 9 | 11 |
| 21 | 8 | 10 | 9 | 11 |
| 22 | 9 | 11 | 9 | 11 |
| 23 | 9 | 11 | 10 | 11 |
| 24 | 9 | 11 | 10 | 12 |
| 25 | 9 | 11 | 10 | 12 |
| 26 | 9 | 11 | 10 | 12 |
| 27 | 9 | 12 | 10 | 12 |
| 28 | 10 | 12 | 11 | 13 |
| 29 | 10 | 12 | 11 | 13 |
| 30 | 10 | 12 | 11 | 13 |
| 35 | 10 | 12 | 11 | 13 |
| 40 | 11 | 14 | 13 | |

Table H: Table of Critical Values of D in the Kolmogorov-Smirnov Two-Sample Test.
Large samples: two-tailed test

| Level of significance | Value of D so large as to call for reflection of null hypothesis at the indicated level of significance where D - maximum $S_{n_1}(X) - S_{n_2}(X)$ |
|-----------------------|---|
| 10 | $1.22 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$ |
| .05 | $1.36 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$ |
| .025 | $1.48 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$ |
| .01 | $1.63 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$ |
| .005 | $1.73 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$ |
| .001 | $1.95 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$ |

Table I : Table of Probabilities Associated with Values as Small as Observed Values of x in the Binomial Test.

Given in the body of this table are one-tailed probabilities under null hypothesis for the binomial test when $P = Q = 1/2$.

| x N | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 5 | .031 | .188 | .500 | .812 | .969 | | | | | | | | | | | |
| 6 | .016 | .109 | .344 | .656 | .891 | .984 | † | | | | | | | | | |
| 7 | .008 | .062 | .227 | .500 | .773 | .938 | .992 | † | | | | | | | | |
| 8 | .004 | .035 | .145 | .363 | .637 | .855 | .965 | .996 | † | | | | | | | |
| 9 | .002 | .020 | .090 | .254 | .500 | .746 | .910 | .980 | .998 | † | | | | | | |
| 10 | .001 | .011 | .055 | .172 | .377 | .623 | .828 | .945 | .989 | .999 | † | | | | | |
| 11 | | .006 | .033 | .113 | .274 | .500 | .726 | .887 | .967 | .994 | † | | | | | |
| 12 | | .003 | .019 | .073 | .194 | .387 | .613 | .806 | .927 | .981 | .997 | † | | | | |
| 13 | | .002 | .011 | .046 | .133 | .291 | .500 | .709 | .867 | .954 | .989 | .998 | † | | | |
| 14 | | .001 | .006 | .029 | .090 | .212 | .395 | .605 | .788 | .910 | .971 | .994 | .999 | † | | |
| 15 | | | .004 | .018 | .059 | .151 | .304 | .500 | .696 | .849 | .941 | .982 | .996 | † | | |
| 16 | | | .002 | .011 | .038 | .105 | .227 | .402 | .598 | .773 | .895 | .962 | .989 | .998 | † | |
| 17 | | | .001 | .006 | .025 | .072 | .166 | .315 | .500 | .685 | .834 | .928 | .975 | .994 | .999 | † |
| 18 | | | .001 | .004 | .015 | .048 | .119 | .240 | .407 | .593 | .760 | .881 | .952 | .985 | .996 | .999 |
| 19 | | | | .002 | .010 | .032 | .084 | .180 | .324 | .500 | .676 | .820 | .916 | .968 | .990 | .988 |
| 20 | | | | .001 | .006 | .021 | .058 | .132 | .252 | .412 | .588 | .748 | .868 | .942 | .979 | .994 |
| 21 | | | | .001 | .004 | .013 | .039 | .095 | .192 | .332 | .500 | .668 | .808 | .905 | .961 | .987 |
| 22 | | | | | .002 | .008 | .026 | .067 | .143 | .262 | .416 | .584 | .738 | .857 | .933 | .974 |
| 23 | | | | | .001 | .005 | .017 | .047 | .105 | .202 | .339 | .500 | .661 | .798 | .895 | .953 |
| 24 | | | | | .001 | .003 | .011 | .032 | .076 | .154 | .271 | .419 | .581 | .729 | .846 | .924 |
| 25 | | | | | | .002 | .007 | .022 | .054 | .115 | .212 | .345 | .500 | .655 | .788 | .855 |

† 1.0 or approximately 1.0.

Table J: Table of Critical Values of T in the Wilcoxon Matched-Pairs Signed-Ranks Test.

| N | Level of Significance for one-tailed test | | |
|----|---|-----|------|
| | .025 | .01 | .005 |
| | Level of Significance for two-tailed test | | |
| | .05 | .02 | .01 |
| 6 | 0 | — | — |
| 7 | 2 | 0 | — |
| 8 | 4 | 2 | 0 |
| 9 | 6 | 3 | 2 |
| 10 | 8 | 5 | 3 |
| 11 | 11 | 7 | 5 |
| 12 | 14 | 10 | 7 |
| 13 | 17 | 13 | 10 |
| 14 | 21 | 16 | 13 |
| 15 | 25 | 20 | 16 |
| 16 | 30 | 24 | 20 |
| 17 | 35 | 28 | 23 |
| 18 | 40 | 33 | 28 |
| 19 | 46 | 38 | 32 |
| 20 | 52 | 43 | 38 |
| 21 | 59 | 49 | 43 |
| 22 | 66 | 56 | 49 |
| 23 | 73 | 62 | 55 |
| 24 | 81 | 69 | 61 |
| 25 | 89 | 77 | 68 |

Table K: Table of Critical Values of Pearson Product Correlation at the .05 and .01 Levels of Significance.

| Degrees of freedom (N-2) | .05 | .01 | Degrees of Freedom (N-2) | .05 | .01 |
|--------------------------|------|-------|--------------------------|------|------|
| 1 | .997 | 1.000 | 24 | .388 | .496 |
| 2 | .950 | .990 | 25 | .381 | .487 |
| 3 | .878 | .959 | 26 | .374 | .478 |
| 4 | .811 | .917 | 27 | .367 | .470 |
| 5 | .754 | .874 | 28 | .361 | .463 |
| 6 | .707 | .834 | 29 | .355 | .456 |
| 7 | .666 | .798 | 30 | .349 | .449 |
| 8 | .632 | .765 | 35 | .325 | .418 |
| 9 | .602 | .735 | 40 | .304 | .393 |
| 10 | .576 | .708 | 45 | .288 | .372 |
| 11 | .553 | .684 | 50 | .273 | .354 |
| 12 | .532 | .661 | 60 | .250 | .325 |
| 13 | .514 | .641 | 70 | .233 | .302 |
| 14 | .497 | .623 | 80 | .217 | .283 |
| 15 | .482 | .606 | 90 | .205 | .267 |
| 16 | .468 | .590 | 100 | .195 | .254 |
| 17 | .456 | .575 | 125 | .174 | .228 |
| 18 | .444 | .561 | 150 | .159 | .208 |
| 19 | .433 | .549 | 200 | .138 | .181 |
| 20 | .423 | .537 | 300 | .113 | .148 |
| 21 | .413 | .526 | 400 | .098 | .128 |
| 22 | .404 | .515 | 500 | .088 | .115 |
| 23 | .396 | .505 | 1000 | .062 | .081 |

Table L: Table of Critical Values of Spearman Rank Order Correlation at .05 and .01 Levels of Significance (One-tailed test).

| <i>N</i> | .05 | .01 | <i>N</i> | .05 | .01 |
|----------|------|-------|----------|------|------|
| 5 | .900 | 1.000 | 16 | .425 | .601 |
| 6 | .829 | .943 | 18 | .399 | .564 |
| 7 | .714 | .893 | 20 | .377 | .534 |
| 8 | .643 | .833 | 22 | .359 | .508 |
| 9 | .600 | .783 | 24 | .343 | .485 |
| 10 | .564 | .764 | 26 | .329 | .465 |
| 12 | .506 | .712 | 28 | .317 | .448 |
| 14 | .456 | .645 | 30 | .306 | .432 |

Table M: Values of r_s taken as the Cosine of an Angle.

| <i>Angle</i> | <i>Cosine</i> | <i>Angle</i> | <i>Cosine</i> | <i>Angle</i> | <i>Cosine</i> |
|--------------|---------------|--------------|---------------|--------------|---------------|
| 0° | 1.000 | 41° | .755 | 73° | .292 |
| | | 42 | .743 | 74 | .276 |
| 5 | .996 | 43 | .731 | 75 | .259 |
| | | 44 | .719 | 76 | .242 |
| 10 | .985 | 45 | .707 | 77 | .225 |
| | | 46 | .695 | 78 | .208 |
| | | 47 | .682 | 79 | .191 |
| 15 | .966 | 48 | .669 | 80 | .174 |
| 16 | .961 | 49 | .656 | | |
| 17 | .956 | 50 | .643 | 81 | .156 |
| 18 | .951 | | | 82 | .139 |
| 19 | .946 | 51 | .629 | 83 | .122 |
| 20 | .940 | 52 | .616 | 84 | .105 |
| | | 53 | .602 | 85 | .087 |
| 21 | .934 | 54 | .588 | | |
| 22 | .927 | 55 | .574 | | |
| 23 | .921 | 56 | .559 | 90 | .000 |
| 24 | .914 | 57 | .545 | | |
| 56 | .906 | 58 | .530 | | |
| 26 | .899 | 59 | .515 | | |
| 27 | .891 | 60 | .500 | | |
| 28 | .883 | | | | |
| 29 | .875 | 61 | .485 | | |
| 30 | .866 | 62 | .469 | | |
| | | 63 | .454 | | |
| 31 | .857 | 64 | .438 | | |
| 32 | .848 | 65 | .423 | | |
| 33 | .839 | 66 | .407 | | |
| 34 | .829 | 67 | .391 | | |
| 35 | .819 | 68 | .375 | | |
| 36 | .809 | 69 | .358 | | |
| 37 | .799 | 70 | .342 | | |
| 38 | .788 | | | | |
| 39 | .777 | 71 | .326 | | |
| 40 | .766 | 72 | .309 | | |

Table N: A few .05 Level of Significance Values from the F_{max} Distribution Table

| Degrees of Freedom for Largest Group (N-1) | | 2 | 3 | 4 |
|--|----|------|------|------|
| | 9 | 4.03 | 5.34 | 6.31 |
| | 10 | 3.72 | 4.85 | 5.67 |
| | 12 | 3.28 | 4.16 | 4.79 |
| | 15 | 2.86 | 3.54 | 4.01 |