CHAPTER 49

Principle of Statistical Analysis for Clinical Research: A Primer

Uday C Ghoshal, Shweta Tripathi, Dipti Chourasia
Department of Gastroenterology
Sanjay Gandhi Postgraduate Institute of Medical Sciences
Lucknow

Introduction

Most clinician would think that why a clinician needs to know about statistical tests? Statistical tests are the pillars on which data analysis for all types of clinical, basic and epidemiological research stand. Those who are not into research also need some understanding of statistical analysis to properly understand strength and weakness of published data to take decision on whether to apply it while clinical decision-making? With the availability of several user-friendly softwares, performance of statistical tests has become a reality even for non-statistician provided he is computer-friendly and understands basic principle of statistical analysis (Fig. 49.1).

In this chapter, we will review some basic concepts related to data and its analysis. The broad headings include, (1) types of data, (2) central tendency and dispersion of data: Mean, median, mode, range, SD, (3) distribution of data: normal (parametric) and not normal (non-parametric), (4) concept of hypothesis test, (5) concept of p value, (6) concept of one tailed or two tailed tests, (7) concept of type I and type II error, (8) concept of power and sample size, (9) paired and unpaired tests, (10) choosing a statistical test for a database, (11) concept of univariate and multivariate analysis, (12) concept of association and causation, (13) correlation and regression, (14) sensitivity, specificity, positive, negative predictive values and diagnostic accuracy, (15) concept of ROC curve, (16) software for statistical analysis, (17) statistics resources on the web.

Types of data

Once we complete a study, we get some data. This is a collection of information or observations. However, in order to infer the result of the study we need to analyze the data using appropriate statistical tests. Before applying any statistical test it is very important to identify the type of data as choice of statistical tests depends on the type of data (Fig. 49. 2).

1. Categorical or nominal or discrete data: Data that can be divided into categories or groups such as male female and can take only discrete values and not decimal are called so. Some examples are given in Fig. 2.

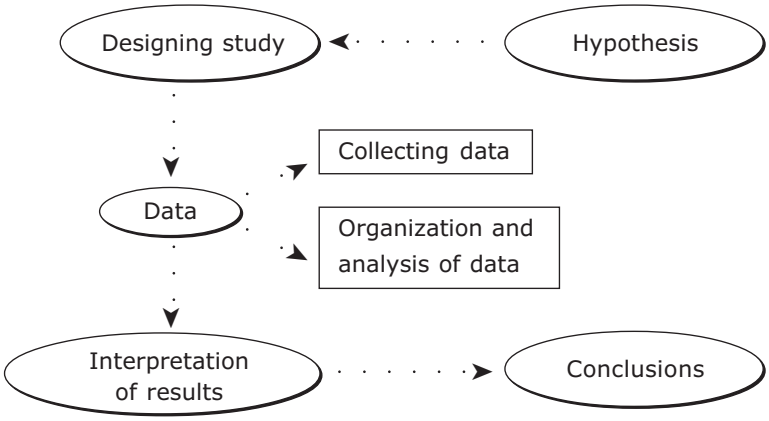


Fig. 49.1 Principle and methods of statistical analysis

Table 49.1 Types of data

Categorical or Nominal or Discrete	Continuous or numerical	
	Scalar	**Ordinal**
■ Gender: Male, female	■ Blood pressure	■ Pain intensity
■ Marital status: Married, single	■ Height	■ Health rating
■ Blood groups: A, B, AB, O	■ Body temperature	■ Disease stage: Mild, moderate, severe

2. Continuous or numerical: Data that can take any value including decimal are called continuous data. A continuous data can be either ordinal or scalar (Fig. 48.2). Data are considered to be ordinal if the values are arbitrarily ranked (put in order) but truly these are not scalar, such as mild, moderate and severe disease. Fig. 48.2 gives some more examples of ordinal data. For an ordinal data the distance between two adjacent ranks may not be same. For example, rank of students in a class is an ordinal data as difference between the marks of students with 1st, 2nd and 3rd rank may not be same. Data that can be measured on a scale is said to be scalar.

3. Parametric or non-parametric: A normally distributed (Gaussian distribution) data is said to be parametric and a data set, which does not have a normal distribution is called as non-parametric. On the basis of frequency curve of data its pattern of distribution can be determined (Fig. 49.3). In general, a data of large sample size (n >50) is normally distributed (parametric) and that of low sample size

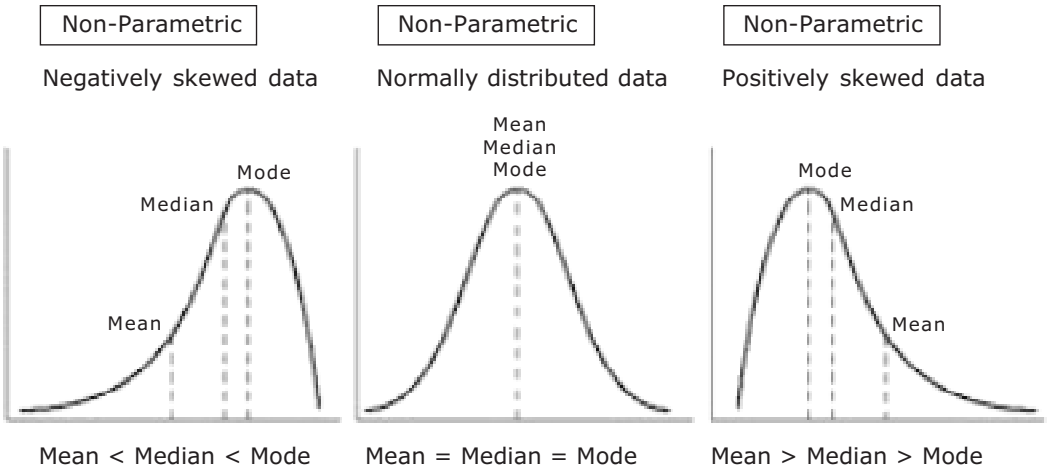


Fig. 49.3 Principle and methods of statistical analysis

(n <30) tends to be non-parametric (skewed). If the values of mean and median are wide apart the data is unlikely to be normally distributed. Though there are several statistical tests to check normalcy of distribution of the data, this is outside the scope of this chapter.

Properties of data

- Central tendency
- Dispersion

Central Tendency and its measures

Central Tendency is defined as "middle value" or typical value of data, which represents the whole data set. It gives us an idea of the "value" around which all the observations in a data set appear to concentrate. The most important and frequently used measures of central tendency are mean, median and mode (Fig. 48.4).

1. **Mean**: It is arithmetic mean or average of a data set. It is equal to sum of all the values divided by total number of values in the data set.

Merits

- It is easy to understand and compute.
- It is based on all the observations in the data set.

Fig. 49.4 Mean, median and mode and effect of outliers

Data Set 1:
3 3 3 5 6 7 8 9 11 13 98 (Outliner 98)
Data Set 2:
3 3 3 5 6 7 8 9 11 13 (No outliner)

Data Set 1	Data Set 2
Mean = 166/11 = 15.09	Mean = 68/10 = 6.8
Median = 7	Median = 6+7/2 = 6.5
Mode = 3	Mode = 3

Demerits

- It cannot be used for qualitative data.
- It is greatly affected by outliers or extreme values in a data set and therefore, may lead to fallacious conclusions (Fig. 49.4).

2. **Median**: It is the middle value of a data set above and below which lie an equal number of data values (if total number of values in the data set is odd), when the values are arranged in ascending or descending order. If total number of values in a data set is even then the median is equal to the arithmetic mean of two middle values.

   Merits
   - It is not affected by extreme values
   - It can be located by inspection

   Demerits
   - It does not depend on all the data.

3. **Mode**: It is the most frequently occurring value of a data set. Its merits and demerits are similar to that of median. In addition, it can be used for qualitative data.

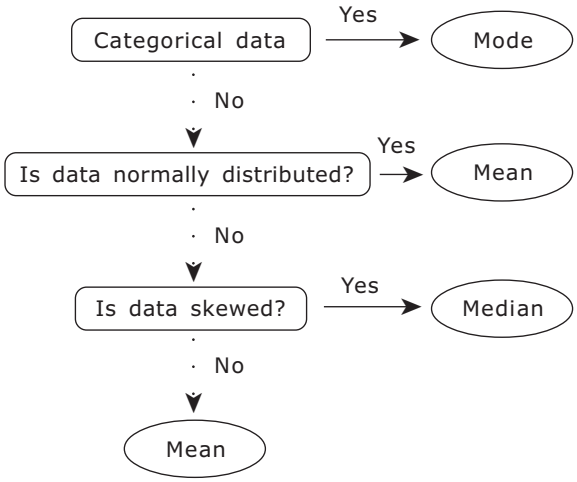


Fig. 49.5 Data and choice of central tendency

The mean, median and mode are affected by distribution or skewness of data (Fig. 48.3). The choice of central tendency therefore depends upon the type and distribution of data (Fig. 49.5).

In general, as the skewness increases the mean and median move away from mode. If mean is less than median, the data is skewed to left and if it is greater than median, the data is skewed to right.

Dispersion (variability) of data

Measures of dispersion or variability of a data give an idea of the extent to which the values are clustered or spread out. In other words, it gives an idea of homogeneity and heterogeneity of data. Two sets of data can have similar measures of central tendency but different measures of dispersion (Fig. 49.6). Therefore, measures of central tendency should be reported along with measures of dispersion.

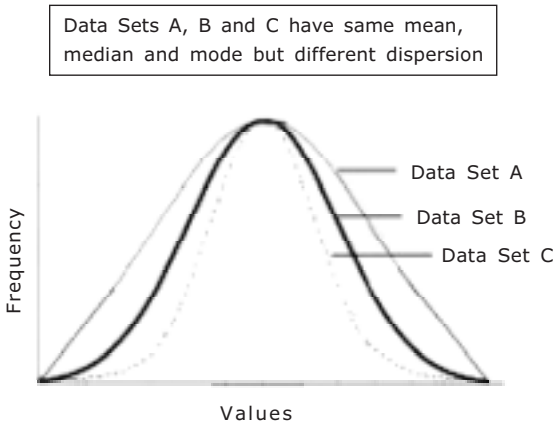


Fig. 49.6 Comparision of central tendency and dispersion of data

Measures of dispersion include:

1. Range: It is the simplest measure of dispersion. It can be represented as the difference between maximum and minimum value or simply as maximum and minimum value. Range is given with median.
2. Mean deviation: It is the average of deviation from arithmetic mean

3. Standard deviation: Standard deviation is always given with mean. It denotes (approximately) the extent of variation of values from the mean. (if the standard deviation is 10, then the values tend to be about 10 units above and below the mean). Mathematically, it is square root of variance. All the values in a data set have a particular deviation score, which is calculated by subtracting mean of the data from the value. Variance is the mean of squares of all the deviation scores of a data. Higher values of variance and standard deviation represent higher variability in the data and vice versa. Zero represents no variability.

Concept of hypothesis test

Hypothesis testing (or significance testing) is to quantify our belief against a particular hypothesis. For example, in a clinical trial for testing a new drug against the current drug,

1. Null hypothesis: (H0) assumes no effect of the given drug (i.e. the new drug is no better, on average, than the current drug; i.e. H0: there is no difference between the two drugs).
2. Alternative hypothesis: (H1) holds that the null hypothesis is not true. The alternative hypothesis relates more directly to the theory we wish to investigate (i.e. the new drug has a significantly different effect, on average, compared to that of the current drug).

A p value of less than 0.05 means that probability of null hypothesis (H0) being correct is less than 5% (less than 5 out of 100 means less than 0.05 out of 1). Let us elaborate on the concept of p value further.

The Concept of probability (P value)

In statistics, P value represents the probability of an outcome. Its value ranges from zero to one. A small P value shows that the difference between two samples is unlikely to be by chance. As per rule $P<0.05$ is considered to be statistically significant difference. However, it is better to report exact P values instead of reporting reference values as "P<0.05" or "P>0.05". Interpretation of significance on the basis of P value is summarized in figure 48.7. A P value of 0.02 means null hypothesis is rejected with more than 98% probability.

P value	Significance	Inerpretation: Evidence against null hypothesis
≥0.1	Not significant	No
≥0.05 to < 0.1	Trend	Suggestive
≤0.01 to <0.05	Significant	Moderate
<0.01	Highly significant	Very strong

Fig. 49.7 Interpretation of P value

One-tailed and two tailed P values

When two groups are compared the P value obtained can be one or two-tailed. Two-tailed p-value is given for two-tailed hypothesis while one-tailed p-value is given for one-tailed hypothesis. A two-sided hypothesis states that there is a difference between the experimental and the control group and the change is possible in either direction (either group can have larger mean). One-sided hypothesis states a specific direction for the change (specified group will have larger mean). For example: If a drug's effect on blood pressure is being investigated then it can lead either to increase or decrease in blood pressure

according to two tailed analysis; in contrast, one tailed analysis assumes that the investigational drug can only reduce blood pressure and hence, if there was increase in blood pressure after administration of the drug, this is ignored. If we expect that the drug can either increase or decrease blood pressure, which is more rational, then two-tailed p-value should be used. Two-tailed P values are usually more appropriate as it it assumes that the investigational drug can either increase or decrease blood pressure and thus eliminates bias. One should choose a one-tail P value only when very convincing evidence is available for assuming that a particular group will always have larger mean or that the change can be only in one direction (E.g. mean desert temperature vs. mean mountain temperature), even before collecting any data. Since such conditions are rarely met in biological situation and therefore, two-tail P values should be preferred over one-tailed P values.

Concept of type I and type II error

There are four ways in which conclusion of the test might relate to the reality (Fig. 48.8):

- True positive and true negative.
- False positive and false negative:

Type I error: To reject the null hypothesis when it is true or false positive error or a (alpha) error. In the above example, type I error would mean that the effects of two drugs were found to be different by statistical analysis, when in fact there was no difference between them.

Type II error: To accept the null hypothesis when it is false or false negative error or ß (beta) error. In the above example, type II error would mean that the effects of two drugs were not found different by statistical analysis, when in fact there was difference.

Concept of power and sample size

The power of a statistical hypothesis test measures the test's ability to reject the null hypothesis when it is actually false - that is, to make a correct decision.

Statistical power

= rejection of H0 when H1 is true
= making a correct decision
= 1 - ß

The maximum power a test can have is 1, the minimum is 0. Ideally we want a test to have high power, close to 1. Increasing the sample size is the best way to increase the power of a statistical test.

Determining sample size is a very important issue because samples that are too large may waste time, resources and money, while samples that are too small may lead to inaccurate results. In order to calculate the sample size required, we have to have some idea of the results we expect in the study. We need to decide the followings before calculating the sample size:

- Power of the study that we want
- Level of significance that we want (p value)
- Variability of observation. e.g.: the standard deviation, if we have a numerical variable
- Smallest effect of interest-the magnitude of the effect that is clinically important that we do not want to overlook. This is often a difference (e.g. difference in means or proportions). Sometimes it is expressed as a multiple of the standard deviation of the observations (the standardized difference).

Thus, the standardized difference =

$$\frac{\text{Target difference}}{\text{Standard deviation}}$$

For example, in a group of septic shock patients treated with a new drug and traditional drug, the mean arterial pressure was 95 and 81 mmHg, respectively, corresponding to a difference of 14 mmHg. And the standard deviation for the mean arterial pressure was 18 mmHg. Thus the standardized difference will be 14/18 = 0.78. However, in practice the standard deviation is unlikely to be known in advance, but it may be possible to estimate it from other similar studies in comparable populations, or perhaps from a pilot study.

Ways to calculate sample size

Quick formulae: These exist for particular power values and significance levels for some hypothesis tests (e.g. Lehr's formulae).

For the unpaired t-test and Chi-squared test, Lehr's formula for calculating sample size of a power of 80% and a two-sided significance level of 0.05, the required sample size in each group is:

$$\frac{16}{(\text{Standardized difference})^2}$$

Note that a numerator of 21 (instead of 16) relates to a power of 90%.

Calculating sample size for clinical trials:
Based on the two types of results, clinical trials may give either qualitative or quantitative data or both. While calculating the sample size, we consider percentage of the valid outcomes (example: survivors and deaths) in case of qualitative data and mean of the valid outcomes (example: level of a biological parameter in on two different treatments) in case of quantitative data.

Paired and unpaired tests

Two measurements or values of a variable in the same subject before and after an intervention are called paired data (Fig. 48.8). Measurements or values of a variable in subjects who are not related to each other (e.g. in cases vs. controls) are called unpaired data (Fig. 49.8).

Paired data	Unpaired data
■ Blood pressure after and before exercise	■ Blood pressure in drug injected and control mice
■ Bed occupancy in a hospital in winter vs. the summer	■ Gene frequency in cancer patients vs controls
■ Blood sugar before and after lactose ingestion	■ Blood sugar in diabetic and non-diabetic patients
■ H. pylori positivity before and after PPI treatment	■ H. pylori positivity in gatric cancer and duodenal ulcer patients

Fig. 49.8 Paired and unpaired data

Choosing a statistical test for a database

One a study has been completed, data have been tabulated into the spreadsheet, we need to decide what statistical test needs to be performed for different variables tabulated into the spreadsheet. Choice of different tests depends on the following factors: (1) whether the variable is categorical or continuous? (2) Whether the data in that variable is normally distributed (parametric) or not normally distributed (non-parametric)? (3) whether the data is paired or unpaired? Fig. 48.9 and Fig. 48.10 show the choice of various statistical tests based on the above parameters.

Fig. 49.9 How to choose statistical tests for your data

Parametric test	Equivalent non-parametric test	Reason for the test	Example
One sample (paired t test)	Wilcoxon sign rank test	Compares two sets of observation on a single sample	To compare blood sugar levels before and after meals
2 samples (Unpaired t-test)	Mann-Whitney U test	Compares two independent samples drawn from a same group of individuals	To compare heights of boys and girls of the same class.
One way analysis of variance (F test) using total sum of squares	Kruskall-Wallis analysis of variance by ranks	For comparison of continuous data with three or more observations	To compare blood sugar levels 1-h, 2-h and 3-h after meals
Two way analysis of variance	Two way analysis of variance by ranks	As given above, but it also tests the effect (interaction) of two different covariates	To determine if, blood sugar levels 1-h, 2-h & 3-h after meals differs in male and female
No direct equivalent	Chi squared test	Tests the null hypothesis that the proportion of discontinuous variables estimated from two or more independent samples are the same	To assess whether students are more likely to pass if the earnings in their family is more
McNemar's test	No direct equivalent	Tests the null hypothesis that the proportions estimated from a paired sample are the same	To compare sensitivity & specificity of two different diagnostic tests in same sample
Pearsons's correlation coefficient (product moment)	Spearman's rank correlation coefficient	To assess strength of straight line association between two continuous variables	To assess whether serum calcium level is related to the serum PTH level in osteoporosis patients
Regression by least squares method	No direct equivalent	Describes numerical relation between two quantitative variables, allowing one value to be predicted from the other	To evaluate, how serum glucose levels varies with age
Multiple regression by least squares method	No direct equivalent	Describes numerical relation between a dependent variable & several predictor variables (covariates)	To observe whether and to what degree an individual's age, body weight and cholesterol determines their blood pressure

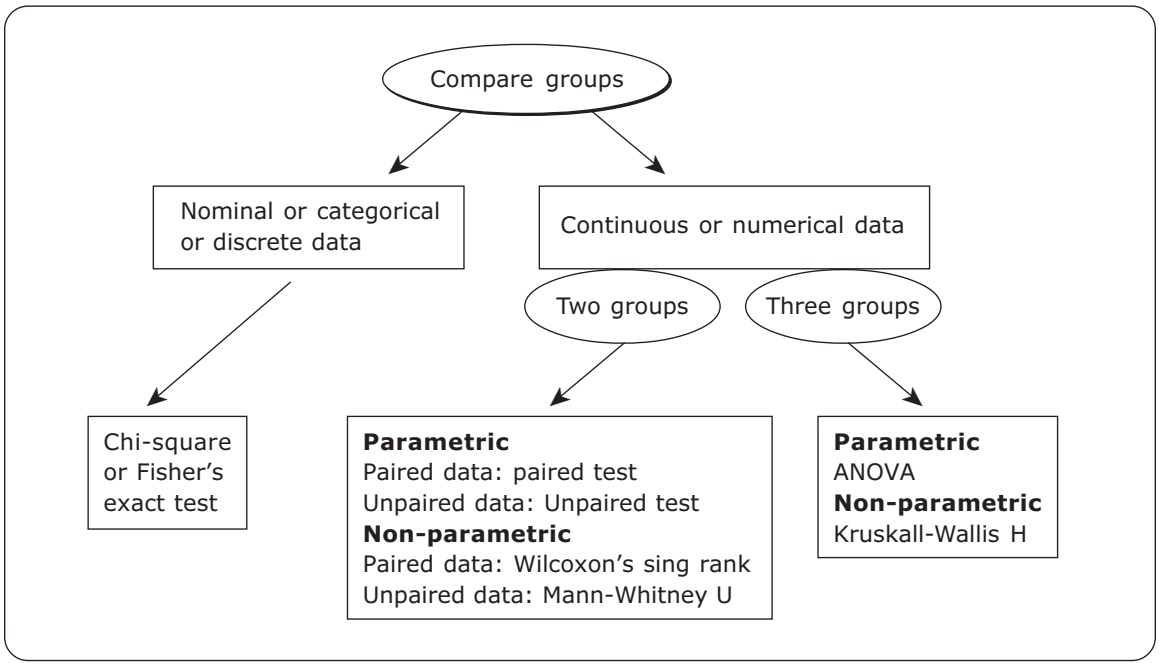


Fig. 49.10 How to choose statistical tests for your data

Concept of univariate and multivariate analysis

Till now, we have discussed about univariate analysis only, in which one analyzes whether one variable is associated with another or not? Such association does not necessarily mean causation. For example, in Fig. 49.9 we have given an example in which the question is whether student is more likely to pass if their family earning is more? If this is so, we can't necessarily conclude that higher earning causes higher chance of passing examination. It may be related to better facilities for studies, better environment for studies, better schooling, better libraries, better teachers, better health, and better awareness. Multivariate analysis simultaneously tests effect of all these factors on performance in the examination and helps in inferring which are the independent factors that result in better performance in the examination. There are several methods for multivariate analysis, which are outside the scope of this chapter, which is merely a primer for basic statistical analysis.

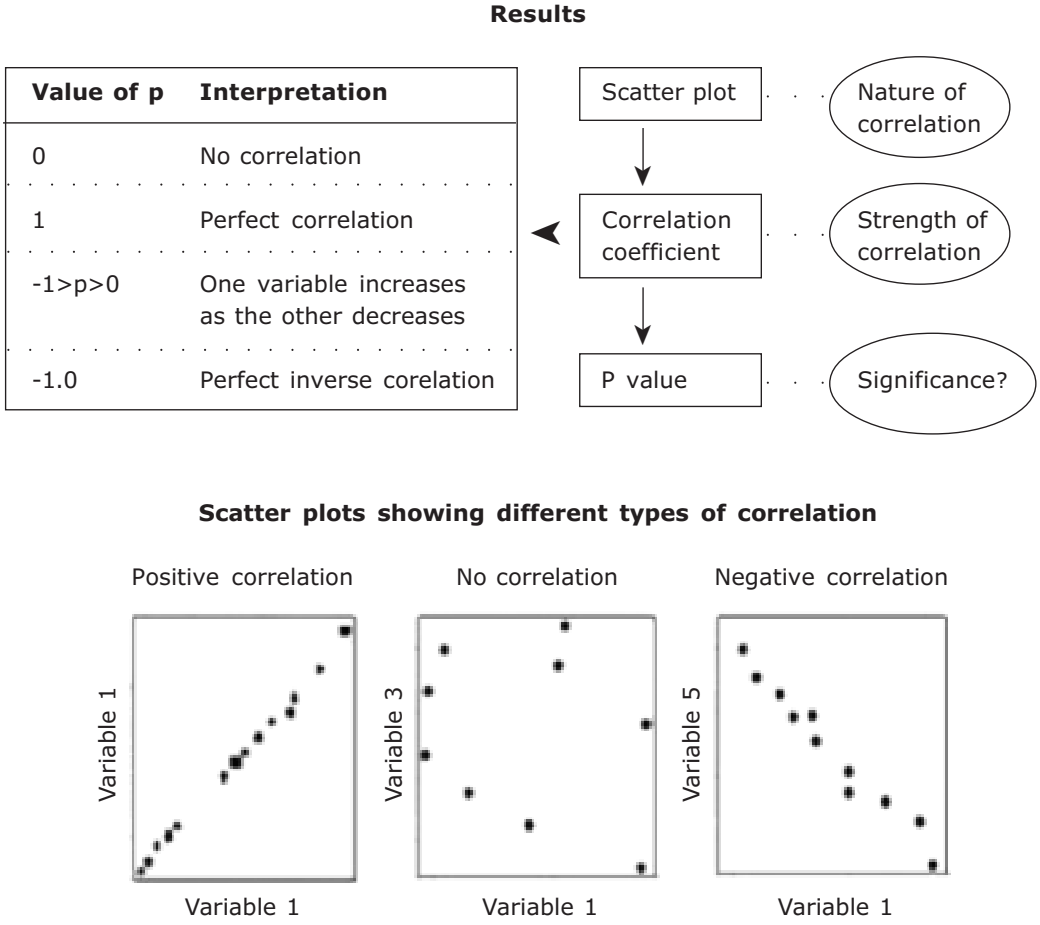
Concept of association and causation

Significant p values by the above analyses suggest association between the two or more than two variables. However, however strong may be the relationship, it does not prove causation. For example, presence of an association between X and Y tells us nothing about either the presence or the direction of causality. To demonstrate that X has caused Y, we need to have more that an association. Sir Austin Bradford Hill developed some criteria that should be met before causation is assumed.

Correlation and Regression

Correlation is relationship between the two sets of continuous data; for example relationship between fasting blood sugar and body weight or relationship between height and body weight. Correlation statistics is used to determine the extent to which two independent variables are related and yields a number called coefficient of correlation. Correlation coefficient may be positive or negative and may vary from -1 to +1 (Fig. 49.11). Positive correlation means that values of two different variables increase and decrease together (direct relationship). For example, speed of running and pulse rate correlates positively. Negative correlation means that if value of one variable decreases then value of the other variable increases (inverse relationship). For example, age and number of scalp hair may correlates negatively. The strength of a correlation is determined by absolute value of correlation coefficient; closer is the value to 1, stronger is the correlation. For example, a correlation of -0.9 indicates an inverse relationship between two

Fig. 49.11 Interpretation of correlation coefficient (?) and correlation statistics results



variables and shows a stronger relationship than that associated with a correlation of +0.2 or -0.5. Correlation between two variables is shown by scatter plot (Fig. 15). P value in a correlation statistics indicates whether the correlation (or no correlation) observed is real or by chance. Interpretation of results of correlational statistics is summarized in Fig. 49.11.

Correlation analysis is important because it can be used to predict values of one variable on the basis of value of other variable. A correlation does not mean causation but it also does not mean absence of causation, that is, if two variables exhibit strong correlation then one of the variables may cause the other. Correlation data is therefore not sufficient evidence for causation. Pearson correlation is applied for parametric data while Spearman correlation is applied for non-parametric data. Combined effect of a group of variable upon a variable not included in the group is called as multiple correlation.

Regression analysis is used to predict the values of a quantitative dependent variable based on the values of one or more independent variables. In simple regression analysis, there is one quantitative dependent variable and one independent variable. In multiple regression analysis, there is one quantitative dependent variable and two or more independent variables. For example, one may derive a formula to predict liver span (dependent variable) from the height of a person (independent variable).

Linear regression statistics finds the best-fit line (line of regression) that predicts dependent variable from independent variable (Fig. 49.12). Linear regression statistics is applied to data where independent variable is continuous measured variable such as concentration. If the independent variable is categorical (e.g. present vs. absent) then logistic regression is used.

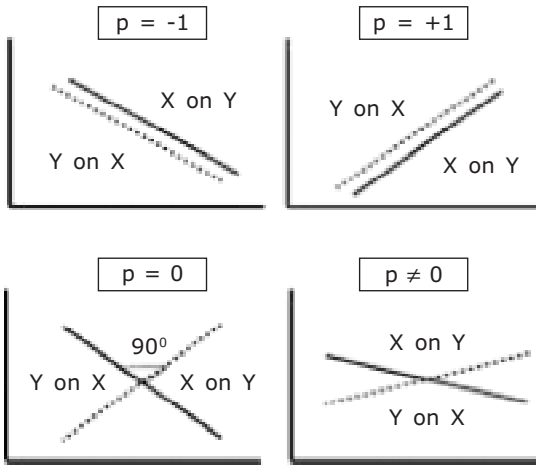


Fig. 49.12 Interpretation of lines of regression

Sensitivity and specificity

An ideal test is one that would clearly distinguish between presence and absence of disease in question. But, in actual practice, false positives, false negatives and errors do occur. A test is valid if it detects most people with the target disorder, excludes most without the disorder, and if a positive test usually indicates that the disorder is present. Fig. 48.13 shows a 2x2 table summarizing relationship between a diagnostic test and actual presence of disease. (Fig. 49.13)

Of the n number of individuals studied, A+C have the disease.

Therefore, **prevalence** (P) of the disease = (A+C) / n

Of the A+C individuals who have the disease, A have positive test results (true positive) and C have negative test results (false negative). Of

	Disease		
Test result	**Present**	**Absent**	**Total**
Positive	A	B	A + B
Negative	C	D	C + D
Total	A + C	B + D	n = A+B+C+D

$$+PV = \frac{A}{A+B} \qquad -PV = \frac{D}{C+D}$$

$$Se = \frac{A}{A+C} \quad Sp = \frac{D}{B+D} \quad P = \frac{A+C}{n} \quad DA = \frac{A+D}{n}$$

Fig. 49.13 Relationship between diagnostic test and actual presence of disease

the B + D individuals who do not have the disease, D have negative test results (true negative) and B have positive test results (false positive).

Sensitivity (Se) = Proportion of the individuals with the disease who are correctly identified by the test. A sensitive test will hardly miss individuals with the disease.
Sensitivity = A / A+C

Specificity (Sp) = Proportion of the individuals without the disease who are correctly identified by the test. A specific test will rarely misclassify individuals without the disease as diseased.

Specificity = D / B+D

These are usually expressed as percentage.
For conditions that are easily treatable, we prefer a high sensitivity; for those that are serious and untreatable, we prefer a high specificity in order to avoid making a false positive diagnosis.

Positive, negative predictive values and diagnostic accuracy

Positive predictive value (+PV) = Proportion of individuals with a positive test result who have the disease.

Positive predictive value = A / A+B

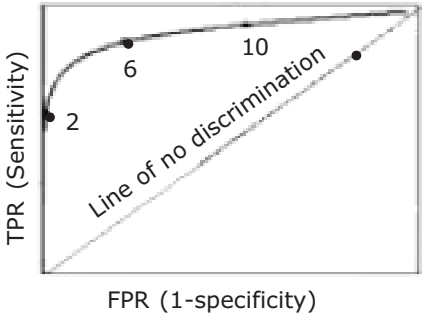
Negative predictive value (-PV) = Proportion of individuals with a negative test who do not have the disease.
Negative predictive value = D / C+D

Diagnostic accuracy (DA): Accuracy is the proportion of all test results (positive and negative) that are correct.

Diagnostic accuracy = A+D / A+B+C+D

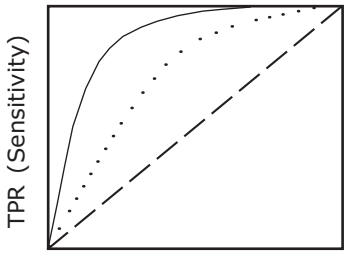
Predictive values are dependent on the prevalence of the disease in the population being studied. In populations where the disease is common, the positive predictive value will be much

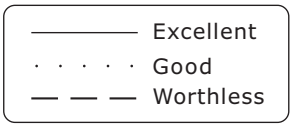
Cut off	Sensitivity or TRP	Specificity	1-specificity or FPR
2	0.56	0.99	0.01
6	0.78	0.81	0.19
10.	0.91	0.42	0.58



Points above the line of discrimination indicate good results, while points below the line indicate wrong results

Interpretation of ROC curves





Nature of cure: *Qualitative* test accuracy
- Closer the curve follows the left-hand and the top border more accurate the test
- Closer the curve comes to line of discrimination, less accurate the test

Area under the curve: *Measure of* test accuracy (grading)
- 0.9-1 = excellent (A)
- 0.8-0.9 = good (B)
- 0.7-0.8 = fair (C)
- 0.6-0.7 = poor (D)
- 0.5-0.6 = fail (F)

Fig. 49.14 ROC curve

Fig. 49.16 Softwares for statistical analysis

Software	Main functions
SPSS	Specializing in data mining, and data analysis
Excel	Basic statistics
Mstat	Design of experiments (RCB, CRD, Lattice), print labels, and collect, organize, and analyze the data
Prism 4.0	Basic biostatistics, curve fitting and scientific graphing, easily organize, analyze, and graph repeated experiments; pick appropriate statistical tests and interpret the results
NCSS	Statistical analysis and graphics system
SAS	Business Intelligence and Analytics
ARC (1.06)	Statistical Analysis Tool for Regression Problems
ASSISTAT (7.3)	Regression and Variance Analysis, Statistical Tests
EPI INFO (3.3.2)	Epidemiological Statistics
ESTA+ (3.1.4) SSP (2.75) WINSTATS (2007)	Descriptive Statistics
EULER (2.4)	Matlab Clone with basic Statistical functions
G7 (7.32)	Regression Analysis
IRRISTAT (4.3)	Basic Statistical Analysis of Experimental Data aimed primarily at the Analysis of Data from Agricultural Field Trials
OPEN EPI (2005)	Epidemiologic Statistics for Public Health
SALSTAT (BETA)	Application Designed for Scientific Statistical Analysis
STATCALC (2.0)	Probability Calculator
STATEASY (0.4)	Multivariate Statistics
MOREPOWER (4.0) PS (2.1)	Power and Sample Size Calculator
BIPLOT (1.1)	Excel Add-In for Multivariate Analysis
XLSTATISTICS (5.76)	Set of Microsoft Excel Workbooks for Statistical Analysis of Data

higher than in populations where the disease is rare. The converse is true for negative predictive values.

Receiver Operating Characteristic curve

Receiver operating characteristic (ROC) is a plot of sensitivity vs. 1 - Specificity or plot of true positive rate (TPR) against the false positive rate (FPR), for the different possible cut offs (threshold values) of a diagnostic test (Fig. 18). It shows the relationship between sensitivity and specificity (any increase in sensitivity is accompanied by a decrease in specificity). ROC curve gives an idea of accuracy of a test (efficiency of the test to discriminate between true positive and true negative). The area under the curve gives the measure of text accuracy. Area of ROC curves is calculated by complex mathematical models but can be obtained easily by various computer programs. The interpretation of ROC curves is given in Fig. 49.14.

Statistics resources on the web

There are several resources in the web to learn biostatistics. Some of these URL are given below (Fig. 49.7)

Table 49.7 Statistics resources on the web

- http://www.anu.edu.au/nceph/surfstat/surfstat-home/surfstat.html
- http://www.stat.ufl.edu/vlib/statistics.html
- http://faculty.vassar.edu/lowry/VassarStats.html
- http://graphpad.com/quickcalcs/index.cfm
- http://www.bmj.com/statsbk/
- http://www.ats.ucla.edu/STAT/
- http://archives.math.utk.edu/topics/statistics.html
- http://www.ifigure.com/math/stat/testing.htm
- http://en.wikipedia.org/wiki/Statistics
- http://www.statsoft.com/textbook/stathome.html

Further reading

1. Trisha Greenhalgh. How to read a paper: The basics of evidence based medicine. BMJ publishing group, BMA House, Tavistock Square, London, UK; 1998. (First published 1997, Third impression 1998, Reprinted in India 1999)
2. Anthony N Glaser. High-yield Biostatistics. Lippincott Williams and Wilkins, 227 East Washington Square, Philadelphia, USA; 1995. (Copyright 1995, Reprinted in India 2000)
3. Robert H Fletcher, Suzanne W. Fletcher and Edward H. Wagner. Clinical epidemiology: The essentials. 3rd ed. Lippincott Williams and Wilkins, 351 West Canadian Street Baltimore, Maryland, USA; 1996.
4. Carolyn M Hicks. Research Methods for Clinical Therapists. 3rd Edition. Churchill Livingstone, Robert Stevenson House, 1-3 Baxter's Place, Leith Walk, Edinburgh EH1 #AF, UK. 1999.