13. TIME SERIES

13.1 Introduction

A time series is a set of observations collected at specified times, usually at equal intervals. Such series have an important place in the field of Economic and Business Statistics. The series generally relates to prices, consumptions and productions of commodities, money in circulation, bank deposits and bank clearings, sales in a department store etc. From close observation, it is evident that the values of the variable under consideration changes from time to time (days, weeks, months, years etc.) The fluctuations are affected not by a single force but the net effect of multiplicity of forces putting it up and down. If the forces are at equilibrium the series would remain constant. For example, the retail prices of a particular commodity are influenced by a number of factors viz, the crop yield which further depends on weather conditions, irrigation facilities, fertilizers used, transport facilities, consumers demand etc.

The main object of analysis of time series is to isolate these forces which creates the movements and to determine the extent of the forces individually. Analysis of such movements is of great values in many connections one of which is the problem of forecasting further movements. It should thus come as no surprise that many industries and governmental agencies are vitally concerned with this important subject.

13.2 Components of Time Series

The forees at work affecting a time series can be broadly classified into the following categories :

i) Secular Trend or Trend

ii) Seasonal Movements.

iii) Cyclical Movements.

iv) Irregular Movements.

The value of a time series y_t at any time t is required as the result of the combined impact of above four components. We shall try to explain the components one by one.

i) **Trend** : It refers to the general directions of the data indicating the increase or decrease during a long period of time. This is true in most of the data provided in Business and Economic Statistics. For example, the upward tendency of the data is generally seen in human population, agricultural productions, currency in circulations etc. while the downward tendency may be seen in data of births and deaths, epidemics etc. because of advancement of medical science, better medical facilities, literacy, and higher standard of living.

The term long period of time cannot be defined exactly because in some cases long period may be a week, in bactorial population study, we count after every five minutes for a week but in case of agricultural production even 24 months shows no secular change. Therefore, the values should cover a period depending on the practical suitability.

ii) **Seasonal Movements :** Seasonal movements are periodic and regular in a time series with period less than one year. In a time series seasonal movements refer to identical or almost identical patterns of movements during corresponding months of succesive years. The very name suggest that weather plays an important role in such movements and may be attributed for mainly following two cases namely (a) those resulting from natural forces for example the production of certain commodities say paddy, sugar and eggs depend upon season and (b) those resulting from man made conventions for example sales and profit in a department store shows a marked rise during (i) marriage, (ii) annual festival like Eid, Durga Puja, Christmas, Baisakhi Purnima etc. the phenomenon relates to man made conventions.

Here also the time period varies from situation to situation because the hourly number of telephone calls crossing a particular switchboard refers to hourly rhythm whereas the number of accidents in a busy street refers daily rhythm and similarly the issue of books from a library refers to weekly rhythm.

(iii) Cyclical Movements : They generally refer to the long term oscillations about a trend line. The cycles may or maynot be periodic and the period of oscillation is usually more than one year. The cyclical movements are the so called business cycles representing intervals of prosparity, recession, depression and recovery and may last from seven to eleven years.

iv) **Irregular Movements :** Apart from the components mentioned above, the series contain another factor called irregular fluctuations which are purely random, erratic, unforeseen, unpredictable and are due to some irregular circumstances which are beyond the control of human hand but at the same time are a part of our system such as earthquakes, wars, floods, famines, revolutions, epidemics etc. In some cases the effect of irregular movements may not be significant but sometimes may be so intense as to result some other movements.

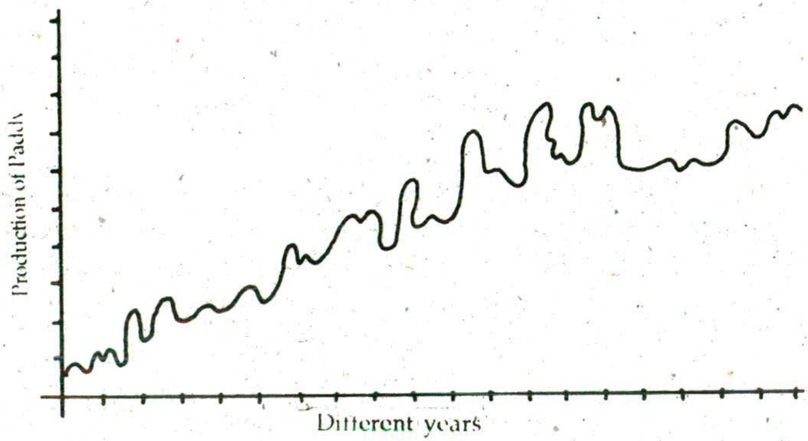


Fig. 13.1 Showing different components.

13.3 Analysis of Time Series

The important problems in time series analysis are :

i) To identify the components whose effect of interaction causes the movement of a time series.

ii) To isolate, study, analyse and measure them independently by holding other components constant. We assume that the time series at a time t is a products of the components $T_{t,}$ $S_{t,}$ C_t and I_t. Symbolically

$$Y_t = T_t \times S_t \times C_t \times I_t \qquad ....(13.1)$$

It may be mentioned that some statisticians prefer to consider Y_t as the sum of four components. But in our study we will consider as given in (13.1)

Measurement of Trend : Trend can be measured by the following methods : i) Graphic Method ii) Method of Semi-Average iii) Method of Curve fitting by the principle of Least squares. iv) Method of Moving Average.

i) Graphic Method : In this method we, at first, plot the points Y_t for various values of t and a free hand smooth curve is drawn enabling us to form an idea about the general trend of the series. The method does not involve any complex mathematical technique and can be used to describe all types of trend, linear or non-linear.

The main drawbacks of the method are : a) Different trend curves can be drawn by different persons from same set of data due to personal biasness. Therefore, skilled and experienced statisticains should be employed to minimise the bias. b) It does not enable us to measure trend.

ii) Method of Semi-Average : In this method the period is divided in to two parts and the average values are calculated for each part. These two values are plotted against their respective time which gives the trend line. The line thus obtained can be extended in both the directions, giving us a trend line. For even number of years the partition can be done easily while for odd number of years the middle most year is usually omitted and the partition into two parts is done.

As compared with graphic method the obvious advantage of this method is its objectivity in the sense that everyone who applies it would get the same results. It is readily comprehensible as compared to the method of moving average and method of least squares (discussed next).

The drawbacks of this method are i) the method assumes linear relationship which may not exist ii) arithmetic mean used for average may also be criticised.

iii) Method of Curve fitting by the Principle of Least Squares : This method is very popular and widely used for fitting mathematical functions to a given set of data. Determining the mathematical functions are usually done by plotting the data graphically or by theoretical understanding of the mechanism of the variable change. The various types of curves that may be used to discribe the given data in practice are as follows :

a) A Straigth line : $Y_t = a + bt$

b) Second degree parabola : $Y_t = a + bt + ct^2$

c) kth degree polynomial : $Y_t = a_o + a_1t + a_2t^2 + ... + a_kt^k$

d) Exponential : $Y_t = ab^t$.

e) Growth curves : $Y_t = a + bc^t$ (Modified exponential curve)

$Y_t = abc^t$ (Gompertz curve)

$Y_t = \frac{k}{1 + e^{a + bt}}$ (Logistic curve)

For simplicity sake, we discuss below the method of fitting a straight line by least squares. For other curves we recommend advanced books on Applied Statistics for the interested readers.

As stated above, the equation for a straight line is $Y_t = a + bt$ ....(13.2)

The sum of squares of the deviations of the observed values of Y_t from the estimated values defined in (13.2) is given by, $S = \sum_t (y_t - a - bt)^2$

Differentiating S with respect to a and b seperately and equating to zero we have two normal equations as follows.

$$\begin{aligned} \Sigma y_t &= na + b\Sigma t \\ \Sigma ty_t &= a\Sigma t + b\Sigma t^2 \end{aligned} \qquad ....(13.3)$$

Solving (13.3) we have the estimate of a and b which gives the desired trend line of (13.2).

The method of curve fitting by the principle of L. S. method is an important and popular method of trend analysis specially when some one is interested in making projections for future times. The form of the mathematical function is generally determined by plotting the time series graphically or by previous experiences and the reliability of the projection mainly depends on it.

Example 13.1 Find the linear trend values of average monthly sales of a certain company during the years 1946-1956

Years	1946	1947	1948	1949	1950	1951
Sales	138.875	174.150	201.483	189.525	218.633	233.367

Year	1952	1953	1954	1955	1956
Sales	254.033	262.225	255.533	286.342	307.250

Solution :

Table-13.1
Calculation of trend values

Year	t	t′	Sales y_t	t'^2	$t'y_t$	Trend values
1946	0	—5	138.875	25	—694.375	154.97
1947	1	—4	174.150	16	—696.600	169.82
1948	2	—3	201.483	9	—604.449	184.67
1949	3	—2	189.525	4	—379.050	199.52
1950	4	—1	218.633	1	—218.633	214.37
1951	5	0	233.367	0	0.000	229.22
1952	6	1	254.033	1	254.033	244.07
1953	7	2	262.225	4	524.450	258.92
1954	8	3	255.533	9	766.599	273.77
1955	9	4	286.342	16	1,145.368	288.62
1956	10	5	307.250	25	1,536.250	303.47
Total	55	0	2,521.416	110	1633.593	2,521.42
Mean	5	0	229.22			229.22

We know $b = \frac{\Sigma t'Y_t}{\Sigma t'^2} = 14.8508$

$a = \overline{Y} - b\,\overline{X} = 154.966$

The linear trend equation is $\hat{Y}_t = 154.966 + 14.8508t.$

iv) Method of Moving Average : It consists in measurement of trend by smoothing out the fluctuation of the series by means of moving average of period m in a series of successive averages of m terms at a time starting from the 1st term. Thus the first average is the mean of first m terms, the second is the mean of m terms from second to (m + 1) th term, the third is the mean of m term from third to (m + 2) th term and so on.

If m is odd = (2k + 1) say, the moving average is placed against the mid value of the time interval it covers i. e. against t = k + 1 and if m is even = 2k say, then the average is placed in between t = k and t = k + 1 and the final centred value is obtained by taking the average of two successive terms. The graphical representation of the average values against time gives us the trend.

For determining the period of moving average which eliminate the oscillatory movement completely, we have to consider the period of oscillation graphically before starting to calculate. If the data depict different cycles which vary in amplitude and period, in such cases the appropriate period of moving average should be equal to or greater than the mean period of the cycles in the data. In such case the moving average does not wipe out the cyclical movement completely and hence a nice picture of general trend connot be given.

The main drawbacks of this methods is that, it does not provide the trend values of all the terms because we cannot have values for the first k and the last k terms of the series and for those values, the forcasting cannot be done.

The method of moving average gives a correct picture of the long terms trend of the series if the trend is linear or approximately linear and the oscillatory movement are regular in period and amplitude. If the trend is not linear moving average introduce bias in the trend values.

Example 13.2 Assume a four-yearly cycle, calculate the trend by the method to moving average from the following data relating to production of tea in a certain tea Eastate.

Year	Production (kg)	Year	Production (kg)
1961	464	1966	540
1962	515	1967	557
1963	518	1968	571
1964	467	1969	586
1965	502	1970	612

Solution : For calculation of trend by the moving average method we prepare Table-13.2.

Table-13.2

Year	Production kg	4-yearly moving total	4-yearly moving average	4-yearly moving average (centred).
1961	464			
1962	515			
		1964	491.0	
1963	518			495.7
		2002	500.50	
1964	467			503.6
		2027	506.75	
1965	502			511.6
		2066	516.50	
1966	540			529.5
		2170	542.50	
1967	557			553.0
		2254	563.50	
1968	571			572.5
		2326	581.50	
1969	586			
1970	612			

4-yearly moving average (centred) values indicate the required trend values for corresponding years.

Measurements of Seasonal Movements : Sesonal variations are usually seen in business and economic data, their measurements are necessary to isolate them to determine the effect of seasons on the size of the variable. The elimination of the seasonal movements are necessary to study as for what would be the value of the series if there were no seasonal swings. The determination of the seasonal variations are necessary for increasing business efficiency and for a smooth production programme.

It is important to note that for studing the seasonal movements, the data must be given for parts of year, such as monthly or quarterly, weekly, daily or hourly. The important method of measuring seasonal movements are :

i) Method of simple average.

ii) Ratio to trend method.

iii) Ratio to moving average method.

iv) Link relative method.

i) **Method of simple average :** It is a very simple method of measuring seasonal movements and the following steps are generally taken :

a) Arrange the data by years and months (or quarters if quarterly data are given).

b) Calculate the average $\overline{x_i}$ (i = 1, 2,........12) for the ith month for all the years [i = 1 represents January, i = 2 represents February etc.]

c) Complete the average $\overline{x}$ of the monthly averages i.e. $\overline{x} = \frac{1}{12} \sum_{i=1} x_i$

d) Seasonal indices for different months are obtained by expressing monthly averages as percentage of $\overline{x}$. Thus

$$\text{Seasonal index for ith month} = \frac{\overline{x_i}}{\overline{\overline{x}}} \times 100$$

It must be remembered that the total of the seasonal indices is 12 x 100 = 1200 for monthly data and 4 x 100 = 400 for quarterly data.

This method is applicable if the data is free from trend and cyclical movements. Since the business data usually have trends. This method, though simple, is not of much practically usefull.

Example 13.3 The data below give the average quarterly prices of a commodity for four years.

Year	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
1967	40.3	44.8	46.0	48.0
1968	50.1	53.1	55.3	59.5
1969	47.2	50.1	52.1	55.2
1970	55.4	59.0	61.6	65.3

Calculate the seasonal movements by the method of simple average.

Solution : Assuming that the trend is absent in the above data, the difference in the averages of various quarters (if there is any) will be due to seasonal changes.

Table-13.3

Computation of seasonal indeces.

Year	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
1967	40.3	44.8	46.0	48.0
1968	50.1	53.1	55.3	59.5
1969	47.2	50.1	52.1	55.2
1970	55.4	59.0	61.6	65.3
Total	193.0	207.0	215.0	228.0
Average	48.25	51.75	53.75	57.00
Seasonal Index	91.57	98.21	102.01	108.18

$$\text{The average of the averages} = \frac{48.25 + 51.75 + 53.75 + 57.00}{4} = 52.69$$

$$\text{Seasonal index for first quarter} = \frac{48.25}{52.69} \times 100 = 91.57$$

$$\text{Seasonal index for Second quarter} = \frac{51.75}{52.69} \times 100 = 98.21$$

$$\text{Seasonal index for third quarter} = \frac{53.75}{52.69} \times 100 = 102.01$$

$$\text{Seasonal index for fourth quarter} = \frac{57}{52.69} \times 100 = 108.18$$

ii) Ratio to trend method : This method is an improvement over simple average method and is based on the assumption that seasonal movement for any given month is constant factor of the trend. The method consists of the following steps :

a) Obtain the trend values by fitting curves to the time series data.

b) Express the original data as percentages of the trend values assuming the multiplicative model given in (13.1). Thus these percentages will contain seasonal, cyclical and irregular movements.

c) The cyclic or irregular movements are then wiped out by taking mean or median of the percentages for different months or quarters.

d) Finally the indices, obtained in step (c) are adjusted to a total 1200 for monthly data or 400 for quarterly data by multiplying them throughout by a constant k given by

$$k = \frac{1200}{\text{total of the indices}} \text{ and } k = \frac{400}{\text{total of the indices}}$$

for months and quarterly data respectivly.

This method is purposeful if the cyclical movements are absent in the series and the obvious advantage of this method is that the measurements are available for each month or quarter and no loss of date is seen as in moving averages.

iii) Ratio to moving average method : Since moving average eliminates the periodic movement for monthly data a 12-month moving average should completely eliminate the seasonal movements if they are of constant pattern and intensity. This method consists of the following steps :

a) Calculate the centred 12-month moving average of the values which gives the estimate of the combined effect of trend and cyclic movements.

b) Express the original data (except for 6 months at the begining and 6 months at the end) as percentages of these using average which will represent the seasonal and irregular movements.

c) The preliminary seasonal indices are now obtained by removing the irregular movements by taking average like arithmetic mean or median of the percentages of different months or quarters.

d) The sum of these indices = S, will not be equal to 1200 for months or 400 for quarters in general so the values are adjusted by a constant factor $\frac{1200}{S}$ for months or $\frac{400}{S}$ for quarters and thus the resultant values will give us the desire indices of the seasonal movements.

This method is the most statisfactory, flexible and widely used but the only disadvantage is that the indices of 1st 6 months and last 6 months connot be obtained in case of 12-month moving average.

Example 13.4 The number of business letters posted in different months during the year 1971 to 1973 are given below. Apply ratio to moving average method to asertain seasonal movements.

Year and Months	No of letters posted	Year and Months	No of letters posted	Year and Months	No of letters Posted
1971		1972		1973	
January	90	January	100	January	110
February	85	February	89	February	93
March	70	March	74	March	78
April	60	April	62	April	66
May	55	May	55	May	58
June	45	June	47	June	40
July	30	July	30	July	35
August	40	August	43	August	45
September	70	September	65	September	72
October	120	October	127	October	130
November	115	November	118	November	118
December	118	December	120	December	124

Solution :

Table-13.4
Calculating of ratio to moving average

(1) Year & Month		(2) No of letters Posted	(3) 12-Point Moving Total	(4) 12-Point M.A. $\frac{Col-3}{12}$	(5) 12-Point M.A. Centred	(6) Ratio to M.A. $\frac{Col-2}{Col-5} \times 100$
1971	January	90				
	February	85				
	March	70				
	April	60				
	May	55				
	June	45				
			898	74.83		
	July	30			75.3	39.8
			908	75.67		
	August	40			75.8	52.8
			912	76.00		
	September	70			76.2	91.9
			916	76.33		
	October	120			76.4	157.1
			918	76.50		
	November	115			76.5	150.3
			918	76.50		
	December	118			76.6	154.0
			920	76.66		
1972	January	100			76.7	130.4
			920	76.66		
	February	89			76.8	115.9
			923	76.91		
	March	74			76.7	96.5
			918	76.50		
	April	62			76.8	80.7
			925	77.16		
	May	55			77.2	71.2
			928	77.33		

(1) Year & Month	(2) No of letter Posted	(3) 12-Point Moving Total	(4) 12-Point M.A. $\frac{\text{Col-3}}{12}$	(5) 12-Point M.A. Centred	(6) Ratio to $\frac{\text{Col-2}}{\text{Col-5}}$ x 100
June	47			77.4	60.7
		930	77.50		
July	30			77.9	38.5
		940	78.33		
August	43			78.5	54.8
		944	78.66		
September	65			78.8	82.5
		948	79.00		
October	127			79.2	160.4
		952	79.33		
November	118			79.5	148.4
		955	79.58		
December	120			79.3	151.3
		948	79.00		
1973 January	110			79.2	138.9
		953	79.41		
February	93			79.5	117.0
		955	79.58		
March	78			79.9	97.6
		962	80.16		
April	66			80.3	82.2
		965	80.41		
May	58			80.4	72.1
		965	80.41		
June	40			80.6	49.6
		969	80.75		
July	35				
August	45				
September	72				
October	130				
November	118				
December	124				

Computation of Adjusted Seasonal indices

Table-13.5

Month	1971	1972	1073	Seasonal Indices = (A. M)	Adj. Seasonal Indices (Seasonal Indices x CF)
January		130.4	138.9	134.7	135.0
February		115.9	117.0	116.5	116.7
March		96.5	97.6	97.1	97.3
April		80.7	82.2	81.5	81.7
May		71.2	72.1	71.7	71.8
June		60.7	49.6	55.2	55.3
July	39.8	38.5		39.2	39.3
August	52.8	54.8		53.8	53.9
September	91.9	82.5		87.1	87.3
October	157.1	160.4		158.8	159.1
November	150.3	148.4		149.4	149.7
December	154.0	151.3		152.7	153
Total				1197.7	1200.1

Here, Correction factor (C. F.) for obtaining adjusted Seasonal indices

$$\frac{1200}{1197.7} = 1.0019$$

iv) Link relative method : This method is based on averaging the link relatives and also known a Pearson's method. Link relative of one time period is the value expressd as a percentage of the preceeding time period. The time period may be month for monthly data and quarter for quarterly data.

$$\text{Link relative (L. R.) for any month} = \frac{\text{Current month's figure}}{\text{Previous month's figure}} \times 100 \quad ....(13.4)$$

This method consists of the following steps :

a) Transform the original data into link relatives by the formula given in (13.4)

b) Take the average by mean or median of these link relatives for each month (quarter) if the data are monthly (quarterly).

c) Convert the average link relatives to chain relatives on the basis of the first month (quarter). chain relative (C. R) for any time period is obtained on multiplying the link relative of that time period by the chain relative of the preceding time period and dividing by 100. Thus for monthly data, the chain relative for the first month i. e. January is taken to be 100.

$$\text{C. R. for February} = \frac{\text{L. R. of February x C. R. of January}}{100}$$

$$= \text{L. R. of February. [Since C.R. of January=100]}$$

$$\text{C. R. for March} = \frac{\text{L. R. or March x C. R. of February}}{100} \text{ and similarly}$$

$$\text{C.R. for December} = \frac{\text{L. R. of December x C. R. of November}}{1000}$$

Now taking the C. R. of December as a base a new C.R. of January is calculated by $\frac{\text{L. R. of January x C. R. of December}}{100}$

Usually this will not be 100 due to trend and so all the C.R. values are to be adjusted by subtracting the correction factor 'd' for each chain relative where $d = \frac{1}{12}$ [New C. R. of January - 100].

Then assuming linear trend the correction factor for February March........December is d. 2d........11d repectively.

e) Finally, adjust the corrected chain relativs to total 1200 by expressing the corrected chain relatives as percentages of their arithmetic mean. The resultant values give the adjusted monthly indices of seasonal movements.

The average of the link relatives contain both trend and cyclic components and is effective if the data show a constant rate of growth. With the help of this method complete data analysis can be done which is not possible in case of moving average method but the only disadvantage of this method may be noted is due to its complexity in calculations.

Example 13.5 The data below give the average quarterly prices of a commodity for five years. Calculate the seasonal indices by the method of link relatives :

	Year				
Quarter	1966	1967	1968	1969	1970
I	30	35	31	31	34
II	26	28	29	31	36
III	22	22	28	25	26
IV	31	36	32	35	33

Solution :

Calculation of seasonal indices by the method of link relatives

Table-13.6

Year	Link Relative			
	First Quarter	Second Quarter	Third Quarter	Fourth Quarter
1966	—	86.7	84.6	140.9
1967	112.9	80.0	78.6	163.6
1968	86.1	93.5	96.6	114.3
1969	96.9	100.0	80.7	140.0
1970	97.1	105.9	72.2	126.9
Arithmetic Mean	98.25	93.22	82.54	137.14
Chain Relatives	100	$\frac{100 \times 93.22}{100}$ $= 93.22$	$\frac{93.22 \times 82.54}{100}$ $= 76.95$	$\frac{76.95 \times 137.14}{100}$ $= 105.4$
Adjusted Chain Relatives	100	92.345	75.20	102.775
Seasonal Indices	100	99.75	81.23	111.00

Measurement of Cyclic Movement : Using the multiplicative model $Y_t = T_t \times S_t \times C_t \times I_t$ stated in (13.1) the residual approach of determining the cyclic movements C_t is essentially consists in eliminating the two known components T_t and S_t obtained earlier and then removing the irregular movements by averaging.

The systematic steps are as folllows :

i) Calculate trend values T_t by moving average method and seasonal movements S_t preferably by ratio to moving average method.

ii) Divide Y_t by $T_t \times S_t$.

iii) The resulting value gives C_t, the cyclic components.

The use of moving average for suitable period averages out the irregular movements automatically.

Measurement of Irregular Movement : Estimatión of irregular movements can be achieved by dividing the original data Y_t by T_t, S_t and C_t given in (13.1). In practical field the values corresponding to irregular movements are seen to be very small in magnitude and often follow the normal distribution i.e. small deviations occur with large frequency and large deviations occur with small frequency.

13.4 Comparability of Data

One of the main objectives of time series analysis is to compare data for different months or different years. In that case one must be very careful when comparing the data and be sure that such comparison is justified. For example, in comparing the data of March with that of February we must realize that March has 31 days whereas February has 28 or 29 days. Similarly comparing the data of February for different years we must remember that the number of days in February in leap years is 29 rather than 28 days. The number of working days in different months of same year or different years are not same due to different causes like holidays, strikes, lay-off etc.

In practice, no difinite rule is followed for making adjustment due to such variations. The need for such adjustment is left to the discretion of the investigator.

13.5 Forecasting

From the above description and mathematical treatments one may have the assistance of the problem of forecasting time series. However, it must be understood that the mathematical treatments of data do not solve all the problems. Coupled with the common sense, experience, ingenuity and good judgment of an investigator, such mathematical analysis can neverthless be of value for long range and short range forecasting.

13.6 Summary of Fundamental Steps in Time Series Analysis

1. The reliability of the time series data analysis must be ensured by making all possible efforts. For example, if one wants to forecast a given time series, it may be helpful to obtain related time series as well as other information. If necessary, adjust the data for valid comparison.

2. Pictorial representation of the time series may be done for understanding the presence of long term trend, seasonal and cyclical movements qualitatively.

3. Long-term trend curve or line may be constructed for obtaining trend values by using the free hand method, semi-average method, least squares method and method of moving average.

4. If seasonal movements are present, seasonal indices are to be determined and deseasonalize the data finally.

5. The deseasonalized data are adjusted for trend and the resulting data contains only cyclical and irregular movements. A moving average of appropriate time period may be calculated to remove irregular movements and thus the data reveal the cyclical movements only.

6. The cyclical movements obtained in steps 5 may be graphically represented for noting any periodicites which may occur.

7. By combining the results of steps 1-6 and using any other available informations make a forecast and if possible discuss sources of error and their magnitude.

14. INTERPOLATION

14.1 Introduction

Interpolation may be defined as the art of determining the intermediate values of a function from a set of given values of a function. In so many occasions, we have to face the situation of dealing with functions whose analytical form may be totally unknown or else is of such a nature that cannot be operated readily. In either case, it is desireable to replace the given function by a simpler form. This operation of replacing a given function by a simple one constitute interpolation in the broad sense. So the general problem of interpolation consisting of representing a function known or unknown in a form in advance with the aid of given values which this function takes for definite values of the independent variable. Thus let $y = f(x)$ be a function given by the values $y_o, y_1, y_2, ...y_n$ which takes for the values $x_o, x_1, x_2, ...x_n$ of the independent variable x, and let $\varphi(x)$ denote and arbitrary simpler function which takes the same values of f(x) for the values of $x_o, x_1, x_2, ...x_n$. Then if f(x) is replaced by $\varphi(x)$ over a given interval, the process constitute interpolation and the function $\varphi(x)$ gives the formula of interpolation. In practical problems we usually choose for $\varphi(x)$, the simplest function, which will represent the given function over the interval in question.

In solving the problem of interpolation we are in a position to find values of the independent variable either at the begining part or at the end or at the middle part of the function. We will try to solve these problems one by one stating different types of difference tables of dependent variable.

14.2 Newton's Formula for Forward Interpolations

In this type of interpolation we require diagonal difference table. A short descriptions of the table is as follows :

Let $y_o, y_1, y_2, ...y_n$ denote a set of values of any function $y = f(x)$ then $y_1 - y_o$, $y_2, -y_1, y_3 - y_2.....y_n - y_{n-1}$, are called the first differences of the function y.

Denoting these differences respectively by $\Delta y_o, \Delta y_1, \Delta y_2, \ldots\ldots \Delta y_{n-1}$, the second differences can be defined as $\Delta^2 y_o = \Delta y_1 - \Delta y_o = y_2 - 2y_1 + y_o$,

$\Delta^2 y_1 = \Delta y_2 - \Delta y_1 = y_3 - 2y_2 + y_1$

etc. and likewise the third differences are

$\Delta^3 y_o = \Delta^2 y_1 -$ $3y_2 + 3y_1 - y_o$, $\Delta^3 y_1 = \Delta^2 y_2 - \Delta^2 y_1 = y_4 - 3y_3 + 3y_2 - y_1$

and so on. The above type of differences can be presented in a tabular form as given in Table - 14.1.

Table - 14.1

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$	$\Delta^6 y$	$\Delta^7 y$
x_o	y_o							
		Δy_o						
x_1	y_1		$\Delta^2 y_o$					
		Δy_1		$\Delta^3 y_o$				
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_o$			
		Δy_2		$\Delta^3 y_1$		$\Delta^5 y_o$		
x_3	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$		$\Delta^6 y_o$	
		Δy_3		$\Delta^3 y_2$		$\Delta^5 y_1$		$\Delta^7 y_o$
x_4	y_4		$\Delta^2 y_3$		$\Delta^4 y_2$		$\Delta^6 y_1$	
		Δy_4		$\Delta^3 y_3$		$\Delta^5 y_2$		
x_5	y_5		$\Delta^2 y_4$		$\Delta^4 y_3$			
		Δy_5		$\Delta^3 y_4$				
x_6	y_6		$\Delta^2 y_5$					
		Δy_6						
x_7	y_7							

This table is called a diagonal difference table.

For getting a formula for forward interpolation we are to find suitable polynomial for replacing any given function over a given interval. Let $y = f(x)$ denote a function which takes the values $y_o, y_1, y_2, \ldots\ldots y_n$ for equidistant values of $x_o, x_1, x_2, \ldots\ldots x_n$ of the independent variable x and let $\varphi(x)$ denote a polynominal of the nth degree. This polynomial may be written in the form ;

$$\varphi(x) = a_o + a_1(x - x_o) + a_2(x - x_o)(x - x_1) + a_3(x - x_o)(x - x_1)(x - x_2) + a_4(x - x_o)(x - x_1)(x - x_2)(x - x_3) + \ldots\ldots$$

$$\ldots\ldots + a_n(x - x_o)(x - x_1)(x - x_2)\ldots\ldots(x - x_{n-1}). \qquad \ldots\ldots(14.1)$$

We shall now determine the co-efficients $a_o, a_1, a_2, \ldots\ldots a_n$ so as to make

$\varphi(x_o) = y_o, \varphi(x_1) = y_1, \varphi(x_2) = y_2 \ldots\ldots \varphi(x_n) = y_n$, substituting in (14.1) the successive values $x_o, x_1, x_2, \ldots\ldots x_n$ for x at the same time putting $\varphi(x_o) = y_o$, $\varphi(x_1) = y_1$ etc. and remembering that $x_1 - x_o = h$, $x_2 - x_o = 2h$ etc.

we have, $y_o = a_o$ or $a_o = y_o$

$$y_1 = a_o + a_1(x_1 - x_o) = y_o + a_1 h$$

$$\therefore a_1 = \frac{y_1 - y_o}{h} = \frac{\Delta y_o}{h}$$

$$y_2 = a_o + a_1(x_2 - x_o) + a_2(x_2 - x_o)(x_2 - x_1) = y_o + \frac{y_1 - y_o}{h} 2h + a_2 (2h)(h).$$

$$\therefore a_2 = \frac{(y_2 - 2y_1 + y_o)}{2!h^2} = \frac{\Delta^2 y_o}{2!h^2}$$

$$y_3 = a_o + a_1(x_3 + x_o) + a_2(x_3 - x_o)(x_3 - x_1) + a_3(x_3 - x_1)(x_3 - x_2)$$

$$= y_o + \frac{y_1 - y_o}{h}(3h) + \frac{(y_2 - 2y_1 + y_o)}{2!h^2}(3h)(2h) + a_3(3h)(2h)(h)$$

$$\therefore a_3 = \frac{(y_3 - 3y_2 + 3y_1 - y_o)}{6h^3} = \frac{\Delta^3 y_o}{3!h^3} \text{ and so on and finally } a_n = \frac{\Delta^n y_o}{n!h^n}$$

Now subsituting these values of $a_o, a_1, a_2, \ldots\ldots a_n$ in (14.1) we get,

$$\varphi(x) = y_o + \frac{\Delta y_o}{h}(x - x_o) + \frac{\Delta^2 y_o}{2h^2}(x - x_o)(x - x_1) + \frac{\Delta^3 y_o}{3!h^3}(x - x_o)(x - x_1)(x - x_2)$$

$$+\dots\dots\dots + \frac{\Delta^n y_o}{n!h^n}(x - x_o)(x - x_2)\dots\dots(x - x_n)$$

$$= y_o + \Delta y_o\frac{(x - x_o)}{h} + \frac{\Delta^2 y_o}{2!}\frac{(x - x_o)}{h} + \frac{\Delta^3 y_o}{3!}\frac{(x - x_o)}{h}\frac{(x - x_1)}{h}\frac{(x - x_2)}{h}$$

$$+\dots\dots\dots+\frac{\Delta^n y_o}{n!}\frac{(x - x_o)}{h}\frac{(x - x_1)}{h}\frac{(x - x_2)}{h}\dots\dots\frac{(x - x_{n-1})}{h} \qquad \dots\dots.(14.2)$$

Now putting $\frac{x - x_o}{h} = u$, $x = x_o + hu$ then since $x_1 = x_o + h$, $x_2 = x_o + 2h$ etc. we have, $\frac{x - x_1}{h} = \frac{x - (x_o - h)}{h} = \frac{x - x_o}{h} - \frac{h}{h} = (u - 1)$,

$\frac{x - x_2}{h} = \frac{x - (x_o + 2h)}{h} = \frac{x - x_o}{h} - \frac{2h}{h} = (u - 2)$ and similarly

$\frac{x - x_{n-1}}{h} = \frac{x - (x_o + (n - 1)h)}{h} = \frac{x - x_o}{h} - \frac{(n - 1)h}{h} = (u - n + 1)$

substituting in (14.2) these values of $\frac{(x - x_o)}{h}, \frac{(x - x_1)}{h}$ etc. we get,

$$\varphi(x) = \varphi(x_o + hu) = g(u) = y_o + u\Delta y_o + \frac{u(u - 1)}{2!}\Delta^2 y_o + \frac{u(u - 1)(u - 2)}{3!}\Delta^3 y_o$$

$$+\dots\dots\dots+\frac{u(u - 1)(u - 2)\dots\dots(u - n + 1)}{n!}\Delta^n y_o. \qquad \dots\dots\dots(14.3)$$

This is the form in which Newton's formula for forward interpolation is usually written. In fact the formula is used mainly for interpolating the values of y near the begining of a set of tabular values and for extrapolating value of y a short distance backward from y_o.

Example 14.1 From the following table, find the number of students who obtained less than 45 marks :

Marks	No. of Students
30—40	31
40—50	42
50—60	51
60—70	35
70—80	31

Solution : We prepare a difference table as given in Table - 14.2.

Table - 14.2

Marks less than	No. of Students (cum. freq.)	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
40	31				
		42			
50	73		9		
		51		—25	
60	124		—16		37
		35		12	
70	159		—4		
		31			
80	190				

We have to estimate the number of students obtaining less than 45 marks.

Here $x = 45, x_o = 40$ and $h = 10 \therefore u = \frac{x - x_o}{h} = \frac{45 - 40}{10} = 0.5$

We know, $\varphi(x) = g(u) = y_o + u\Delta y_o + \frac{u(u-1)}{2!}\Delta^2 y_o + \frac{u(u-1)(u-2)}{3!}\Delta^3 y_o$

$+ \frac{u(u-1)(u-2)(u-3)}{4!}\Delta^4 y_o$

$= 31 + .5 \times 42 + \frac{.5(.5-1)}{2 \times 1} \times 9 + \frac{.5(.5-1)(.5-2)}{3 \times 2 \times 1} \times (-25)$

$+ \frac{.5(5-1)(.5-2)(.5-3)}{4 \times 3 \times 2} \times 37$

$= 31 + 21 - 1.125 - 1.5625 - 1.4452 = 48$ (app.)

Therefore, we concluded that about 48 students obtained less than 45 marks.

14.3 Newton's Formula for Backward Interpolation

In this type of interpolation we required horizantal difference table. A short description of the horizontal difference table is as follows : We consider the function $y = f(x)$ as in diagonal table and written the differences $(y_1 - y_o) = \Delta_1 y_1$ against y_1, $(y_2 - y_1) = \Delta_1 y_2$ against y_2 etc, $\Delta_1 y_2 - \Delta_1 y_1 = \Delta_2 y_2$ against y_2 and so on the difference table can be furnised as given in Table - 14.3.

Table - 14.3

x	y	$\Delta_1 y$	$\Delta_2 y$	$\Delta_3 y$	$\Delta_4 y$	$\Delta_5 y$	$\Delta_6 y$	$\Delta_7 y$
x_o	y_o							
x_1	y_1	$\Delta_1 y_1$						
x_2	y_2	$\Delta_1 y_2$	$\Delta_2 y_2$					
x_3	y_3	$\Delta_1 y_3$	$\Delta_2 y_3$	$\Delta_3 y_3$				
x_4	y_4	$\Delta_1 y_4$	$\Delta_2 y_4$	$\Delta_3 y_4$	$\Delta_4 y_4$			
x_5	y_5	$\Delta_1 y_5$	$\Delta_2 y_5$	$\Delta_3 y_5$	$\Delta_4 y_5$	$\Delta_5 y_5$		
x_6	y_6	$\Delta_1 y_6$	$\Delta_2 y_6$	$\Delta_3 y_6$	$\Delta_4 y_6$	$\Delta_5 y_6$	$\Delta_6 y_6$	
x_7	y_7	$\Delta_1 y_7$	$\Delta_2 y_2$	$\Delta_3 y_7$	$\Delta_4 y_7$	$\Delta_5 y_7$	$\Delta_6 y_7$	$\Delta_7 y_7$

This table is called a horizontal difference table. Inspecting these two table we have seen that $\Delta^3 y_1 = y_4 - 3y_3 + 3y_2 - y_1 = \Delta_3 y_4$ and in general the relation between the Δ's affected exponents and those affected with suscripts is $\Delta^m y_k = \Delta_m y_{m+k}$ (going forward from y_k)

or, $\Delta_m y_n = \Delta^m y_{n-m}$ (going backward from y_n)

where m denotes the order to differences and k and n the number of the tabulated values.

The backward formula is used for interpolating a value of y near the end of polynomial φ(x) in the following form :

$$\varphi(x) = a_o + a_1(x - x_n) + a_2(x - x_n)(x - x_{n-1}) + a_3(x - x_n)(x - x_{n-1})(x - x_{n-2})$$

$$+ a_4(x - x_n)(x - x_{n-1})(x - x_{n-2})(x - x_{n-3}) + \ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots + a_n(x - x_n)(x - x_{n-1})(x - x_{n-2})\ldots\ldots\ldots\ldots(x - x_1) \qquad \ldots\ldots\ldots(14.4)$$

Now we determine the co-efficients a_o, a_1, a_2, so as to make $\varphi(x_n) = y_n$, $\varphi(x_{n-1}) = y_{n-1}$ etc. Now substituting in (14.4) the values of x_n, x_{n-1} etc. for x and at the same time putting $\varphi(x_n) = y_n$, $\varphi(x_{n-1}) = y_{n-1}$ etc. we have, $y_n = a_o$ or, $a_o = y_n$

$$y_{n-1} = a_o + a_1(x_{n-1} - x_n) = y_n + a_1(-h)$$

$$\therefore a_1 = \frac{y_n - y_{n-1}}{h} = \frac{\Delta_1 y_n}{h}$$

$$y_{n-2} = a_o + a_1(x_{n-2} - x_n) + a_2(x_{n-2} - x_n)(x_{n-2} - x_{n-1})$$

$$= y_n + \frac{y_n - y_{n-1}}{h}(-2h) + a_2(-2h)(-h)$$

$$\therefore a_2 = \frac{y_n - 2y_{n-1} + y_{n-2}}{2!h^2} = \frac{\Delta_2 y_n}{2!h^2}$$ Similarly,

$$a_3 = \frac{\Delta_3 y_n}{3!h^3},\ a_4 = \frac{\Delta_4 y_n}{4!h^4} \ldots\ldots\ldots a_n = \frac{\Delta_n y_n}{n!h^n}$$

substituting these values of a_o, a_1, a_2 etc. in (14.4) we have,

$$\varphi(x) = y_n + \frac{\Delta y_n}{h}(x - x_n) + \frac{\Delta_2 y_n}{2!h^2}(x - x_n)(x - x_{n-1}) + \frac{\Delta_3 y_n}{3!h^3}(x - x_n)(x - x_{n-1})(x - x_{n-2})$$

$$+ \frac{\Delta_4 y_n}{4!h^4}(x - x_n)(x - x_{n-1})(x - x_{n-2})(x - x_{n-3}) + \ldots\ldots\ldots$$

$$+ \frac{\Delta_n y_n}{n!h^n}(x - x_n)(x - x_{n-1}) \ldots\ldots\ldots (x - x_1)$$

This formula can be simplied as follows :

$$\varphi(x) = y_n + \Delta_1 y_n \frac{(x - x_n)}{h} + \frac{\Delta_2 y_n}{2!}\frac{(x - x_n)}{h}\frac{(x - x_{n-1})}{h}$$

$$+ \frac{\Delta_3 y_n}{3!}\frac{(x - x_n)}{h}\frac{(x - x_{n-1})}{h}\frac{(x - x_{n-2})}{h}$$

$$+ \ldots\ldots\ldots + \frac{\Delta_n y_n}{n!}\frac{(x - x_n)}{h}\frac{(x - x_{n-1})}{h}\ldots\ldots\ldots\frac{(x - x_1)}{h} \qquad \ldots\ldots\ldots\ldots(14.5)$$

Now we put, $\frac{x - x_n}{h} = u$ or, $x = x_n + hu$

then since $x_{n-1} = x_n - h$, $x_{n-2} = x_n - 2h$ etc. we have,

$$\frac{(x - x_{n-1})}{h} = \frac{x - (x_n - h)}{h} = \frac{x - x_n}{h} + \frac{h}{h} = (u + 1)$$

$$\frac{(x - x_{n-2})}{h} = \frac{x - (x_n - 2h)}{h} = \frac{x - x_n}{h} + \frac{2h}{h} = (u + 2)$$ and similarly

$$\frac{(x - x_1)}{h} = \frac{x - [x_n - (n - {}_1)h]}{h} = \frac{x - x_n}{h} + \frac{(n - {}_1)h}{h} = u + n - 1$$

Substituting in (14.5) the values of $\frac{x - x_n}{h}$, $\frac{x - x_{n-1}}{h}$ etc. We have,

$$\varphi(x) = \varphi(x_n + hu) = \psi(u) = y_n + u\Delta_1 y_n + \frac{u(u+1)}{2!}\Delta_2 y_n$$

$$+ \frac{u(u+1)(u+2)}{3!}\Delta_3 y_n + \frac{u(u+1)(u+2)(u+3)}{4!}\Delta_4 y_n$$

$$\ldots\ldots + \frac{u(u+1)(u+2)\ldots\ldots\ldots\ldots(u+n-1)}{n!}\Delta_n y_n \qquad \ldots\ldots\ldots(14.6)$$

This is the required Newton's formula for backward interpolation. In fact the formula is used mainly for interpolating the values of y near the end of set of tabular values and for extrapolating the values of y at a short distance ahead of y_n.

Example 14.2 Estimate the probable number of persons earning between Tk. 80—90 per day from the following data :

Income in Taka per day	No. of persons
0—20	120
20—40	145
40—60	200
60—80	250
80—100	150

Solution : The difference table is arranged as given in Table - 14.4 with cumulative frequency.

Table - 14.4

Income in Taka not more than x	No. of persons (y) (cum. freq.)	Δy	$\Delta_2 y$	$\Delta_3 y$	$\Delta_4 y$
20	120				
40	265	145			
60	465	200	55		
80	715	250	50	—5	
100	865	150	—100	150	—145

We have to estimate the number of persons getting not more than Tk. 90

Here $x = 90$, $\therefore u = \dfrac{90 - 100}{20} = -0.5$ and $y_n = 865$

We know, $\varphi(x) = G(u) = y_n + u\Delta y_n + \dfrac{u(u+1)}{2!}\Delta_2 y_n + \dfrac{u(u+1)(u+2)}{3!}\Delta_3 y_n$

$$+\frac{u(u+1)(u+2)(u+3)}{4!}\Delta_4 y_n$$

$$= 865 + (-0.5) \times 150 + \frac{(-0.5)(0.5)}{2 \times 1} \times (-100) + \frac{(-0.5)(0.5)(1.5)}{3 \times 2 \times 1} \times (-150)$$

$$+\frac{-0.5 \times 0.5 \times 1.5 \times 2.5}{4 \times 3 \times 2\ 1}(-145) = 865 - 75 + 12.5 + 9.375 + 5.664 = 817.54 \text{ (app)}$$

Therefore, the number of persons earning less than Tk. 90 is about 817.

From the problem, the number of persons earning less than Tk. 80 is 715.

Therefore, the number of persons earning Tk. 80—90 is 817 - 715 = 102.

The process of computing the value of a function outside the range of given set of value is called extrapolation. The above formula (14.3) and (14.6) should be used with caution but if the function is know to run smoothly near the ends of the range of given values and if h, the interval is taken as small as it should be, we are usually safe in extrapolating for a distance h outside the range of the given values.

The formulas (14.3) and (14.6) are called respecrively Newton's forward and backward formula used for interpolating the values near the begining and the end of a given series. For interpolation near the difference table, central difference formulas are generally prefered. These formulas employ differences lying as nearly as possible on a horizontal line through y_o, considered at the centre, in a diagonal difference table.

Two such formulas namely Stirling's and Bessel's are generally considered which will be discussed one after another. To derive the above two formulas, we need Gauss's central difference formulas. Moreover, Gauss's central difference formulas are used when the independent variables or arguments are not equidistant.

14.4 Gauss's Central Difference Formulas

For considering Gauss's formulas we have to consider a new type of difference table called divided difference table which is described as below:

Let $y_o, y_1, y_2 .....y_n$ denote functional values corresponding to any **values** x_o, x_1 $x_2 .....x_n$ of the independent variable or argument. Then the divided differences of y in ascending order are defined as follows :

First order divided differences :

$$\frac{y_1 - y_o}{x_1 - xo} = \delta(x_1, x_o), \frac{y_2 - y_1}{x_2 - x_1} = \delta(x_2, x_1), \frac{y_3 - y_2}{x_3 - x_2} = \delta(x_3, x_2) \text{ etc.}$$

Second order differences

$$\frac{\delta (x_2, x_1) - \delta (x_1, x_o)}{x_2 - x_o} = \delta (x_2, x_1, x_o) \text{ and } \frac{\delta(x_3, x_2) - \delta (x_2, x_1)}{x_3 - x_1} = \delta (x_3 x_2 x_1) \text{ etc.}$$

Third order differences

$$\frac{\delta(x_3, x_2, x_1) - \delta(x_2; x_1, x_o)}{x_3 - x_o} = \delta(x_3, x_2, x_1, x_o) \text{ and}$$

$$\frac{\delta(x_4, x_3, x_2) - \delta(x_3, x_2, x_1)}{x_4 - x_1} = \delta (x_4, x_3, x_2, x_1) \text{ etc.}$$

Similarly fourth order and higher differences can be stated. The differences are generally placed in between, the variables and the table can be furnished as given in Table - 14.5.

Table - 14.5

x	y	δ	δ^2	δ^3	δ^4
x_o	y_o				
		$\delta(x_1, x_o)$			
x_1	y_1		$\delta(x_2, x_1, x_o)$		
		$\delta(x_2, x_1)$		$\delta(x_3, x_2, x_1, x_o)$	
x_2	y_2		$\delta(x_3, x_2, x_1)$		$\delta(x_4, x_3, x_2, x_1, x_o)$
				$\delta(x_4, x_3, x_2, x_1)$	
		$\delta(x_3, x_2)$			$\delta(x_5, x_4, x_3, x_2, x_1)$
x_3	y_3		$\delta(x_4, x_3, x_2)$		
				$\delta(x_5, x_4, x_3, x_2)$	
		$\delta(x_4, x_3)$			
x_4	y_4		$\delta(x_5, x_4, x_3)$		$\delta(x_6, x_5, x_4, x_3, x_2)$
		$\delta(x_5, x_4)$		$\delta(x_6, x_5, x_4, x_3)$	
x_5	y_5		$\delta(x_6, x_5, x_4)$		
		$\delta(x_6, x_5)$			
x_6	y_6				

This type of difference table given in Table - 14.5 is usually used to interpolate values of a function y = f(x) when the values of x are not equidistant. The divided differences are symmetric functions of their arguments and has a definite relation with the simple differences stated earlier.

Here we have, $\delta\,(x_1, x_o) = \dfrac{y_1 - y_o}{x_1 - x_o} = \dfrac{y_o - y_1}{xo - x_1} = \delta\,(x_o, x_1)$

Also $\delta\,(x_1, x_o) = \dfrac{y_1}{x_1 - x_o} - \dfrac{y_o}{x_1 - x_o} = \dfrac{y_1}{x_1 - x_o} + \dfrac{y_o}{x_o - x_1}$

$$\delta(x_2, x_1, x_o) = \frac{\delta(x_2, x_1) - \delta(x_1\, x_o)}{x_2 - x_o}$$

$$= \frac{1}{(x_2 - x_o)}\left[\frac{y_2}{x_2 - x_1} + \frac{y_1}{x_1 - x_2} - \left(\frac{y_1}{x_1 - x_o} + \frac{y_o}{x_o - x_1}\right)\right]$$

$$= \frac{y_2}{(x_2 - x_o)\,(x_2 - x_1)} + \frac{y_1}{(x_1 - x_o)\,(x_1 - x_2)} + \frac{y_o}{(y_o - y_1)\,(x_o - x_2)}$$

Similarly,

$$\delta(x_3, x_2, x_1, x_o) = \frac{y_3}{(x_3 - x_o)\,(x_3 - x_1)\,(x_3 - x_2)} + \frac{y_2}{(x_2 - x_o)\,(x_2 - x_1)\,(x_2 - x_3)}$$

$$+ \frac{y_1}{(x_1 - x_o)\,(x_1 - x_2)\,(x_1 - x_3)} + \frac{y_o}{(x_o - x_1)\,(x_o - x_2)\,(x_o - x_3)}$$

Thus the right hand member of the above equations remain unchanged when any two values of x are interchanged and the corresponding y's are also interchanged. This means that the divided differences remains unchanged regardles of how much its arguments are interchanged.

Thus, δ(1.4, 7) = δ(1, 7, 4) = δ(4, 1, 7) = δ(7, 1, 4) etc.

It can also be proved by mathematical induction that

$$\delta(x, x_o, x_1 \ldots\ldots x_n) = \frac{y}{(x - x_o)\,(x - x_1) \ldots\ldots (x - x_n)} + \frac{y_o}{(x_o - x)\,(x_o - x_1) \ldots\ldots (x_o - x_n)}$$

$$+ \frac{y_1}{(x_1 - x)\,(x_1 - x_o)\,(x_1 - x_2) \ldots\ldots (x_1 - x_n)} + \ldots\ldots\ldots$$

$$\ldots\ldots + \frac{y_n}{(x_n - x_o)\,(x_n - x_1)\,(x_n - x_2) \ldots\ldots (x_n - x_{n-1})} \qquad \ldots\ldots(14.7)$$

14.5 Relation Between Divided Difference and Simple Difference

The relation between divided differences and simple differences can be stated by starting with a set of functional values corresponding to equidistant values of the arguments considering equidistant values of the arguments the divided difference table can be furnished as in Table - 14.6.

Table - 14.6

x	y	δy	$\delta^2 y$	$\delta^3 y$
x_o	y_o			
		$\frac{y_1 - y_o}{h}$		
$x_o + h$	y_1		$\frac{y_2 - 2y_1 + y_o}{2h.h}$	
		$\frac{y_2 - y_1}{h}$		$\frac{y_3 - 3y_2 + 3y_1 - y_o}{3h.2h.h}$
$x_o + 2h$	y_2		$\frac{y_3 - 2y_2 + y_1}{2h.h}$	
		$\frac{y_3 - y_2}{h}$		$\frac{y_4 - 3y_3 + 3y_2 - y_1}{3h.2h.h}$
$x_o + 3h$	y_3		$\frac{y_4 - 2y_3 + y_2}{2h.h}$	
		$\frac{y_4 - y_3}{h}$		$\frac{y_5 - 3y_4 + 3y_3 - y_2}{3h.2h.h}$
$x_o + 4h$	y_4		$\frac{y_5 - 2y_4 + y_3}{2h.h}$	
		$\frac{y_5 - y_4}{h}$		
$x_o - 5h$	y_5			

Comparing the difference of columns of the Table - 14.6 with the diagonal difference table given in Table - 14.1,

$$\text{We have, } \delta y = \frac{y_1 - y_o}{h} = \frac{\Delta y}{h};\ \delta^2 y = \frac{y_2 - 2y_1 + y_o}{2h.h} = \frac{\Delta^2 y}{2!h^2}$$

$$\delta^3 y = \frac{y_3 - 3y_2 + 3y_1 - y_o}{3h.2h.h} = \frac{\Delta^3 y}{3!h^3} \text{ and in general, } \delta^n y_k = \frac{\Delta_n y_k}{n!h^n},\ k = 0, 1, 2 \ldots\ldots$$

That is, it follows that the nth divided difference of a polynomial of the nth degree are constant and hence the (n + 1)th divided differences of a polynomial of the (n + 1) th degree is zero.

a) Gauss's Eorward Formula : The general Newton's interpolation formula for unequal interval can be written starting from first divided difference as

$$\frac{y - y_o}{x - x_o} = \delta(x, x_o)$$

or, $y = y_o + (x - x_o)\,\delta(x, x_o)$ ....(A)

From second divided difference formula, $\dfrac{\delta(x, x_o) - \delta(x_o, x_1)}{x - x_1} = \delta(x, x_o, x_1)$

or, $\delta(x, x_o) = \delta(x_o, x_1) + (x - x_1)\,\delta(x, x_o, x_1)$

Substituting the value of $\delta(x, x_o)$ in (A) are have

$$y = y_o + (x - x_o)\,[\delta(x_o, x_1) + (x - x_1)\,\delta(x, x_o, x_1)]$$

$$= y_o + (x - x_o)\,\delta(x_o, x_1) + (x - x_o)(x - x_1)\,\delta(x, x_o, x_1)$$

Similarly proceeding we have the Newton's general interpolation formula

as $y = y_o + (x - x_o)\,\delta(x_o, x_1) + (x - x_o)(x - x_1)\,\delta(x_o, x_1, x_2)$

$+ (x - x_o)(x - x_1)(x - x_2)\,\delta(x_o, x_1, x_2, x_3) + ...$

$+ (x - x_o)(x - x_1).........(x - x_{n-1})\,\delta(x_o, x_1, x_2, ...x_n)$

$+ (x - x_o)(x - x_1).........(x - x_n)\,\delta(x, x_o, x_1, ...x_n)$ ....(14.8)

Here it seems worth while to state that all terms in the right hand side of (14.8) have the dimensions of y, regardless of the nature of x and of the units in which x is expressed. This formula may be termed as Newton's formula for unequal intervals.

In the general, if we put in (14.8) $x_o = x_o$, $x_1 = x_o + h$, $x_2 = x_o - h$,

$x_3 = x_o + 2h$, $x_4 = x_o - 2h$, $x_5 = x_o + 3h$; $x_6 = x_o - 3h$ we have,

$y = y_o + (x - x_o)\,\delta(x_o, x_o + h) + (x - x_o)(x - x_o - h)\,\delta(x_o, x_o + h, x_o - h)$

$+ (x - x_o)(x - x_o - h)(x - x_o + h)\,\delta(x_o, x_o + h, x_o - h, x_o + 2h)$

$+ (x - x_o)(x - x_o - h)(x - x_o + h)(x - x_o - 2h)\,\delta(x_o, x_o + h, x_o - h, x_o + 2h, x_o - 2h)$

$+$.............

Now we put $u = \dfrac{x - x_o}{h}$ or, $x - x_o = hu$, then

$y = y_o + hu\,(x_o, x_o + h) + hu(hu - h)\,\delta(x_o - h, x_o, x + h)$

$+hu(hu - h)\,(hu + h)\,\delta(x_o - h, x_o, x_o + h, x_o - 2h)$

$+ hu(hu - h)\,(hu + h)\,(hu - 2h)\,\delta(x_o - 2h, x_o - h, x_o, x_o + h, x_o + 2h)$

$+ hu(hu - h)\,(hu + h)\,(hu - 2h)\,(hu + 2h)\,\delta(x_o - 2h, x_o - h, x_o + h, x_o + 2h, x_o + 3h) + \ldots\ldots\ldots$ ....(14.9)

From the relationship between devided differences and simple differences we have,

$$\delta(x_o, x_o + h) = \frac{\Delta y_o}{h},\quad \delta(x_o - h, x_o, x_o + h) = \frac{\Delta^2 y_{-1}}{2!h^2},$$

$$\delta(x_o - h, x_o, x_o + h, x_o + 2h) = \frac{\Delta^3 y_{-1}}{3!h^3},$$

$$\delta(x_o - 2h, x_o - h, x_o, x_o + h, x_o + 2h) = \frac{\Delta^4 y_{-2}}{4!h^4} \text{ and so on.}$$

Substituting these into (14.9) and cancelling the powers of h in each term we have, $y = y_o + u\Delta y_o + u(u - 1)\frac{\Delta^2 y_{-1}}{2!} + u(u^2 - 1)\frac{\Delta^3 y_{-1}}{3!}$

$+ u(u^2 - 1)\,(u - 2)\frac{\Delta^4 y_{-2}}{4!} + u(u^2 - 1)\,(u^2 - 2^2)\frac{\Delta^5 y_{-2}}{5!} + \ldots\ldots\ldots$ ....(14.10)

This is Gauss's forward formula.

b) Gauss's Backward Formula : This formula can be derived easily by substituting $x_o = x_o$, $x_1 = x_o - h$, $x_2 = x_o + h$, $x_3 = x_o - 2h$, $x_4 = x_o + 2h$, $x_5 = x_o - 3h$, $x_6 = x_o + 3h$, $x_7 = x_o - 4h$, $x_8 = x_o + 4h$ etc. in (14.8)

So that, $y = y_o + (x - x_o)\,\delta(x_o, x_o - h) + (x - x_o)\,(x - x_o + h)\,\delta(x_o, x_o - h, x_o + h)$

$+ (x - x_o)\,(x - x_o + h)\,(x - x_o - h)\,\delta(x_o, x_o - h, x_o + h, x_o - 2h)$

$+ (x - x_o)\,(x - x_o + h)\,(x - x_o - h)\,(x - x_o + 2h)\,\delta(x_o, x_o - h, x_o + h, x_o - 2h, x_o + 2h) + \ldots\ldots\ldots$ ....(14.11)

Now we put, $u = \frac{x - x_o}{h}$ or $x = x_o + hu$ then,

$y = y_o + hu\,\delta(x_o - h, x_o) + hu(hu - h)\,\delta(x_o - h, x_o, x_o + h)$

$$+ hu\,(hu + h)\,(hu - h)\,\delta(x_o - h, x_o, x_o + h, x_o - 2h)$$

$$+ hu\,(hu + h)\,(hu - h)\,(hu + 2h)\,\delta(x_o - 2h, x_o - h, x_o, x_o + h, x_o + 2h) + \ldots\ldots (14.12)$$

From the relationship between divided differences and simple differences we have,

$$\delta(x_o - h, x_o) = \frac{\Delta y_{-1}}{h};\ \delta(x_o - h, x_o, x_o + h) = \frac{\Delta^2 y_{-1}}{2!h^2},$$

$$\delta(x_o - 2h, x_o - h, x_o, x_o + h) = \frac{\Delta^3 y_{-2}}{3!h^3},$$

$$\delta(x_o - 2h, x_o - h, x_o, x_o + h_1, x_o + 2h) = \frac{\Delta^4 y_{-2}}{4!h^4} \text{ and so on.}$$

Substituting the above values in (14.12) and cancelling the powers of h in several terms we get,

$$y = y_o + u\Delta y_{-1}\, u(u+1)\frac{\Delta^2 y_{-1}}{2!} + u(u^2 - 1)\frac{\Delta^3 y_{-2}}{3!} +$$

$$u(u^2 - 1)\,(u + 2)\frac{\Delta^4 y - 2}{4!} + u(u^2 - 1)\,(u^2 - 2^2)\frac{\Delta^5 y - 3}{5!} + \ldots\ldots\ldots \quad \ldots\ldots\ldots(14.13)$$

which is known as Gauss's backward formula

c) A Third Gauss's Formula : For the derivation of Bessel's formula we need a third central-difference formula that starts with y_1 and runs parallel to the backward formula (14.13). To derive such a formula we advance the subscripts of x and y in (14.10) by one unit, calling $x_1 - h = x_o$ and change the u's by putting k = 1 in the general formula

$$u - k = \frac{x - x_k}{h} \text{ and thus we get the relation } u - 1 = \frac{x - x_1}{h}$$

or, $x - x_1 = hu - h$.

These changes amount to advancing all subcripts in (14.13) by one unit and replacing u by u - 1 which reduces (14.13) to

$$y = y_1 + (u - 1)\,\Delta y_o + u(u - 1)\,\frac{\Delta^2 y_o}{2} + u(u - 1)\,(u - 2)\frac{\Delta^3 y_{-1}}{3!}$$

$$+ u(u^2 - 1)\,(u - 2)\,\frac{\Delta^4 y_{-1}}{4!} + u(u^2 - 1)\,(u - 2)\,(u - 3)\,\frac{\Delta^5 y_{-2}}{5!} + \ldots\ldots \quad \ldots\ldots\ldots(14.14)$$

which is the required third Gauss's formula.

Stirling's Interpolation Formula : Taking mean of the formulas (14.10) and (14.13) by adding them and dividing the sums throughout by 2, we get,

$$y = y_o + u\frac{(\Delta y - 1 + \Delta y_o)}{2} + \frac{u^2}{2}\Delta^2 y_{-1} + \frac{u(u^2{}_{-1})}{3!}\;\frac{(\Delta^3 y_{-1} + \Delta^3 y_{-2})}{2}$$

$$+\frac{u^2(u^2{}_{-1})}{4!}\Delta^4 y_{-2} + \frac{u(u^2 - 1)(u^2 - 2^2)}{5!}\frac{(\Delta^5 y_{-3} + \Delta^5 y_{-2})}{2} + \ldots\ldots \qquad \ldots\ldots(14.15)$$

which is known as Stirling's formula. It is to be noted that it goes horizontally through y_o.

Bessel's Interpolation Formula : Taking the mean of the formulas (14.10) and (14.14) by adding them and dividing throughout by 2 we have,

$$y = \frac{y_o + y_1}{2} + (u - \frac{1}{2})\,\Delta y_o + \frac{u(u-1)}{2!}\frac{(\Delta^2 y_{-1} + \Delta^2 y_o)}{2} + \frac{u(u - \frac{1}{2})(u - 1)}{3!}\;\Delta^3 y_{-1}$$

$$+\frac{u(u^2 - 1)(u - 2)}{4!}\;\frac{(\Delta^4 y_{-2} + \Delta^4 y_{-1})}{2}$$

$$+\frac{u(u - \frac{1}{2})(u^2 - 1)(u - 2)}{5!}\Delta^5 y_{-2} + \ldots\ldots\ldots\ldots \qquad \ldots\ldots\ldots\ldots(14.16)$$

which is the required Bessel's formula. It follows a horizontal line mid-way between y_o and y_1 in the difference table.

14.6. Interpolation with Unequal Intervals of the Arguments

Newton's forward and backward interpolation formulas are usually used when the values of the arguments are given at equidistant intervals. It is sometimes inconvenient or even impossible to obtain values of a functions at equidistant values of the arguments. Two such formulas are Newton's formula for unequal intervals of the argument and Lagrange's formula. Newton's formula for unequal intervals (14.8) is shown earlier while describing the Gauss's formulas. Here the Lagrange's formula is given below:

Lagrange's Interpolation Formula : Let f(x) denote a polynomial of the nth degree which takes values $y_o, y_1, y_2, \ldots\ldots\ldots y_n$ when x has the values $x_o, x_1, x_2 \ldots\ldots\ldots\ldots x_n$ respectively. Thus it is shown that the (n + 1) th differences of this polynomial is zero. Hence we know that $\delta(x, x_o, x_1, x_2, \ldots\ldots\ldots x_n) = 0$ and thus (14.7) becomes,

$$\frac{y}{(x - x_o)(x - x_1)(x - x_2)........(x - x_n)} +$$

$$\frac{y_o}{(x_o - x)(x_o - x_1)(x_o - x_2)......(x_o - x_n)}$$

$$+ \frac{y_1}{(x_1 - x)(x_1 - x_o)(x_1 - x_2)..........(x_1 - x_n)} + ...$$

$$+ \frac{y_n}{(x_n - x)(x_n - x_o)(x_n - x_1)............(x_n - x_{n-1})} = 0.,$$

Transposing to the right hand side all terms except the first, we have,

$$\frac{y}{(x - x_o)(x - x_1)(x - x_2).......(x - x_n)} =$$

$$\frac{y_o}{(x - x_o)(x_o - x_1)(x_o - x_2).........(x_o - x_n)}$$

$$+ \frac{y_1}{(x - x_1)(x_1 - x_o)(x_1 - x_2).........(x_1 - x_n)} + ...$$

$$+ \frac{y_n}{(x - x_n)(x_n - x_o)(x_n - x_1)(x_n - x_2)......(x_n - x_{n-1})}$$

Solving for y and thus cancelling the common factors $(x - x_o)$, $(x - x_1)$, $(x - x_2)$,......$(x - x_n)$ in the several terms we have,

$$y = \frac{(x - x_1)(x - x_2)........(x - x_n)}{(x_o - x_1)(x_o - x_2).......(x_o - x_n)} y_o +$$

$$\frac{(x - x_o)(x - x_2)........(x - x_n)}{(x_1 - x_o)(x_1 - n_2)..........(x_1 - x_n)} y_1$$

$$+ \frac{(x - x_o)(x - x_1)(x - x_3).........(x - x_n)}{(x_2 - x_o)(x_2 - x_1)(x_2 - x_3)........(x_2 - x_n)} y_2 + ...$$

$$+ \frac{(x - x_o)(x - x_1)(x - x_2).........(x - x_{n-1})}{(x_n - x_o)(x_n - x_1)(x_n - x_2).........(x_n - x_{n-1})} y_n \qquad ;............(14.17)$$

This is Lagrange's formula and is seen to give y_o, y_1, y_2,.......y_n when $x = x_o$, x_1, x_2..........x_n respectively. The values of the independent variable may or may not be equidistant.

Since Lagrange's formula is merely a relation between two variables, either of which may be considered as independent variable. Therefore, it is evident that by considering y as the independent variable we can write a formula giving x as a function of y. Hence by interchanging x and y in (14.17) we have,

$$x = \frac{(y - y_1)(y - y_2)(y - y_3)\ldots\ldots(y - y_n)}{(y_o - y_1)(y_o - y_2)(y_o - y_3)\ldots\ldots(y_o - y_n)} x_o$$

$$+ \frac{(y - y_o)(y - y_2)(y - y_3)\ldots\ldots(y - y_n)}{(y_1 - y_o)(y_1 - y_2)(y_1 - y_3)\ldots\ldots(y_1 - y_n)} x_1$$

$$+ \frac{(y - y_o)(y - y_1)(y - y_3)\ldots\ldots(y - y_n)}{(y_2 - y_o)(y_2 - y_1)(y_2 - y_3)\ldots\ldots(y_2 - y_n)} x_2 + \ldots \quad \ldots$$

$$+ \frac{(y - y_o)(y - y_1)(y - y_2)\ldots\ldots(y - y_{n-1})}{(y_n - y_o)(y_n - y_1)(y_n - y_2)\ldots\ldots(y_n - y_{n-1})} x_n \quad \ldots\ldots\ldots(14.18)$$

From the above discussion, it is clear that the chief uses of Lagrange's formula are two, 1) to find any value of a function when the given values of the independent variable are not equidistant and 2) to find the value of independent variable corresponding to a given value of the function. Therefore, the second problem may be solved by means of (14.18). In this sense Lagrange's formula may be considered as a technique of inverse interpolation.

15. STATISTICAL QUALITY CONTROL

15.1 Introduction

Statistical quality control is a topic on applied business statistics, now widely used in developed countries as well as in developing countries. In short, by it, we usually mean various statistical methods used for maintainence of quality in a certain continuous process of manufactured products. In any manufacturing process it is not possible to produce goods of exactly same quality. We can expect certain variation which is quite natural, some of the variations are due to chance causes which cannot be prevented. This type of variations are allowable, sometimes the variations are due to wrong process which can be detected and prevented, such variations are usually called preventable.

15.2 Purpose of Statistical Quality Control

The main purpose of statistical quality control is to device statistical methods for seperating allowable variations from preventable variations so that necessary steps may be taken as quickly as possible operating in the process. Usually attemps are made to weed out systematic causes of variation as soon as possible so that the actual variation may be supposed to be random causes only.

There are two types of problems usually seen in statistical quality control. In one type of problem, our aim is to control the proportion of defective items in the manufacturing process so that the size is not very large. This is known as process control. In other type of problem, we like to ensure that lots of manufactured goods do not contain very large proportion of defective items. This is known as lot control.

These two problems are distinct in nature because even when the process is in control the proportion of defective products for the entire out-put over a long period will not be large, even though an individual lot of an item may not be of satisfactory quality.

To tackle this problem, process control is achieved mainly through the techniques of control charts whereas the product or lot control is achieved by sampling inspection.

15.3 Types of Quality Measures

In statistical quality control we mainly use two types of quality measures. By quality, we mean any characteristics of the finished product, intermidiate product or of raw material which is of interest.

Many quality characteristics can be measured quantitatively. For example, the diameter of bobbin case,length of screw, tensile strength of a yarn, life of an electrical bulb etc. All these are continuous variable. Sometimes the variable may be of discrete type, for example, the number of defects in a fixed length cloth piece, number of defective pins in a box of pins etc. Sometimes the quality characteristics may be of qualitative nature which can be classified according to mutually exclusive categories only. For example ; a bolt which does not fit the nut is termed as defective, a soap which does not attain the specific size or smell is considered as defective etc.

We usually consider different types of control charts for the above types of quality measures.

15.4 The Concept of Control Chart Technique

The central idea of control chart technique is the division of the observation into some rational sub groups. The variation within the elements of such sub groups is due to chance causes only and the elements are supposed to be identical and belong to a single homogeneous population. The sub-groups usually maintain the order of the production. If the sub-group size is large enough, we sometimes draw sample from it which is usually termed as sub-sample. The use of such sub-groups is to detect some assignable causes of variation that come and go. If the variation is less, the different sub groups are then considered to be identical and the process is assumed to be in control otherwise the process would be considered to be out of control. All these phenomena are detected with the help of control charts.

Control chart technique is a particular diagramatic method of making this comparison and thus deciding whether the process is or is not affected by systematic variation.

First of all, we concentrate our attention on some parameters of the distribution. Let θ be the parameter and t be the estimate or statistic. If the process is in control then θ must be same or almost same from one sub-group to other sub-group and the fluctuations in the values of t from sample to sample should be due to random variation only. Suppose E (t) = μ_t and v(t) = σ_t^2 then for any value of t lying outside the limits $\mu_t - 3\sigma_t$ and $\mu_t + 3\sigma_t$ indicates the presence of systematic variation. From central limit theorem, the argument is that, in case t is normally distributed then

Prob $\{|t - \mu_t| \leq 3\sigma_t\} = 0.9973$ (app).

Evenwhen t is non-normal, we have from Chebyshev's inequality,

Prob $\{|t - \mu_t| \leq 3\sigma_t\} = \frac{8}{9} = 0.8889$ (app.)

Thus, if the observed value of t_i lies between the limits $\mu_t - 3\sigma_t$ and $\mu_t + 3\sigma_t$, it indicates that the ith sample is almost free from assignable causes, otherwise the process from which the ith sample is drawn is suspected to be out of control.

The whole procedure is usually done in a graph by means of horizontal lines where time periods of collection of sub-samples from sub groups are shown in abscissa and the statistic t are plotted as ordinate. $\mu_t - 3\sigma_t$ is given by lower control limit (LCL), $\mu_t + 3\sigma_t$ is given by upper control limit (UCL) and the mean value μ_t gives the central line (CL)

15.5 Shewhart Control Chart

The Shewhart control chart techniques consist of inspecting a fixed or nearly fixed number of articles at regular interval of time during the whole production process, measuring the associate statistics and then plotting them as ordinates on horizontal chart (hypothetical here) like as follows :

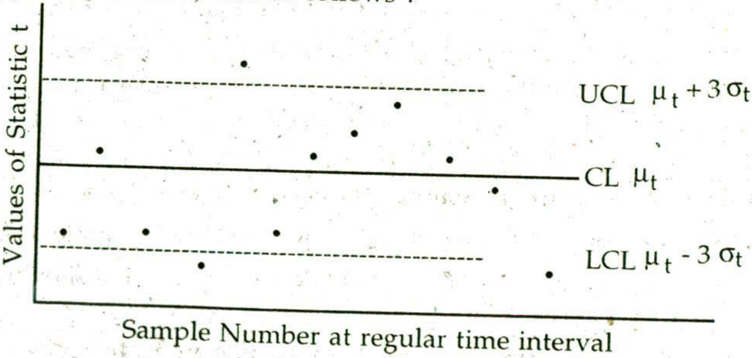


Fig. 15.1 Hypothetical Control Chart

From the above chart it is seen that the process has been out of control in the 4th, 5th and 11th sample points since the 5th sample statistic falls out side the upper control limit and the 4th and the 11th sample statistics are outside the lower control limit. Also by observing closely to the figure it is evident that there is a trend of ups and downs of the values of the statistics from the central line. Along the central line, a series of high values and a series of low values indicate the presence of assignable causes where as alternate high and low values along the central line is the sign of the process that is in control.

The limits on a control chart based on $\mu_t + 3\sigma_t$ and $\mu_t - 3\sigma_t$ are known as 3σ control limits. This achieves control over two types of error usually known as first kind and second kind of error. And in our case we can (1) look for trouble when there is no trouble (2) failing to look for trouble when there are some troubles. 3σ limits i. e. $\mu_t \pm 3\sigma$ are mostly used in United States and $\mu_t \pm 3.09\ \sigma_t$ limits are usually used by the British.

15.6 Types of Control Charts

There are two types of control charts namely (1) Control chart with respect to given standard (2) Control chart with no standard given. Let us describe them one by one.

1) Control Chart with Respect to Given Standard : Here our main purpose is to find out whether the observed values of mean, $\overline{x}$; standard deviation, s ; range R ; fraction defective p and number of defective items c for samples of n items differ from the respective given values of $\overline{x}'$; s´ ; R´; p´ and c´ by amount greater than the values caused by chance. The standard values may be either, provided by the authority as some desired or aimed at values designated by specification or some economic standard of levels provided by experience.

2) Control Chart with no Standard Given : Here we want to find out whether the observed values of $\overline{x}$; s ; R ; p and c for sample of size n vary amongst themselves by amounts greater than what should be attributed to chance.

In both the above cases, the size of the sample from different sub-groups should be small, equal or nearly equal from sub-group to sub-group and should be taken at an equal or small interval of time. For control charts for attributes we should take a sufficiently large sample since the diagnostic power of such charts is much less than that of charts for variables.

15.7 Control Charts for Mean

Case 1 (Standard Given) : Let us suppose that we have drawn a random sample of size n from each sub-group. We know from central limit theorem that E($\bar{x}$) = μ and V($\bar{x}$) = σ^2/n where $\bar{x}$, μ, σ^2 etc. are respectively sample mean, population mean and variance. We also assume that the samples drawn from each sub-group are mutually independent. According to the given condition that the values of μ and σ are specified as μ' and σ' respectively, then the central limits for mean for given standard values will be given by-

$LCL = \mu' - 3\,\sigma'/\sqrt{n} = \mu' - A\sigma'$

$CL = \mu'$ and

$UCL = \mu' + 3\sigma'/\sqrt{n} = \mu' + A\sigma'$

where $A = 3/\sqrt{n}$ which is given in Appendix-1 for different values of n, the sample size.

Now the process is to draw control chart with the help of the above formulas and for any mean value of the sub-sample lying outside the control limits, the process will be called out of control otherwise the process is said to be within control.

Example 15.1 The following data relate to the life of 7 sub samples of 6 electric bulbs each drawn at random of 4 hourly batches from the production process in a day. Draw control chart for mean when the standard value of mean and standard deviation of life length are given as 625 hrs and 65.98 hrs respectively and comment.

Sample No	Life time (hours)					
1	620	687	666	769	839	686
2	501	585	524	585	655	668
3	673	701	686	567	622	660
4	646	626	572	628	632	743
5	495	984	659	643	660	640
6	634	755	625	582	685	555
7	619	710	664	693	773	534

Solution : To draw control chart for mean for given standard we have to calculate the following :

For given mean = 625 and standard deviation = 65.98 the control limits are :

Here $A = 3/\sqrt{n} = 3/\sqrt{6} = 1.22$

Therefore, $LCL = 625 - 1.22 \times 65.98 = 625 - 80.50 = 544.50$

$CL = 625$ and

$UCL = 625 + 80.50 = 705.50$

The total life hours and means are shown below :

Sample No	Total	Mean.
1	4267	711.17
2	3518	586.33
3	3909	651.50
4	3847	641.17
5	4080	680.00
6	3836	639.33
7	3993	665.50

All these mean values are plotted on the chart as follows :

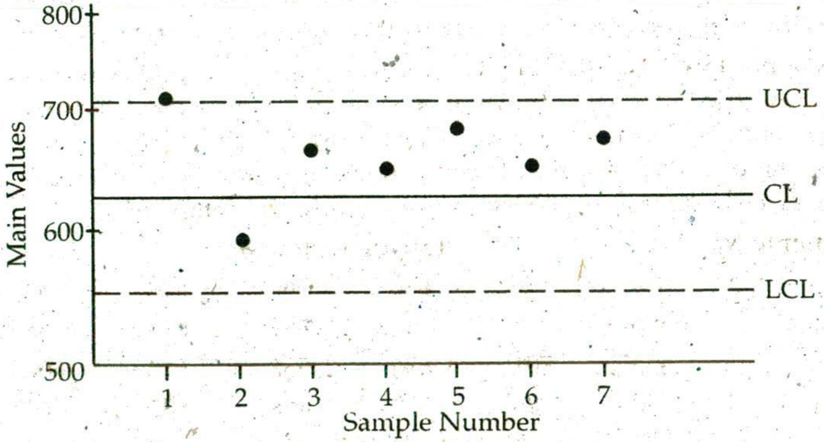


Fig. 15.2 Control chart for mean

Comment : From the control chart, it is seen that a sample point corresponding to the first one is out side the control limits, therefore, the process may be called to be out of control.

Case 2.1 Standard not Given (use the Estimate of Standard Deviation) : Let there be m sub-group from which the successive sub-sample means are $\bar{x}_1 . \bar{x}_2, \bar{x}_3 \ldots\ldots\ldots \bar{x}_m$ and also the successive standard deviations are $s_1, s_2, \ldots\ldots s_m$. Since in this case, the mean μ and standard deviation σ are not specified, they are to be estimated from the sample values and they are as follows :

$$(\bar{\bar{x}}) = \sum_{i=1}^{m} \bar{x}_i/m \quad \text{and} \quad \bar{s} = \sum_{i=1}^{m} s_i/m$$

which are the pooled mean and mean of the sample standard deviations respectively. We know the relations, $E(\bar{\bar{x}}) = \mu$ and

$E(\bar{s}) = C_2\sigma$ (Valid for a normal variable x)

where $C_2 = \dfrac{\Gamma(n/2)}{\Gamma\left(\dfrac{n-1}{2}\right)}\sqrt{\dfrac{2}{n}}$ for n, the sample size, is available in Appendix - 1.

Thus the estimate of μ and σ can be obtained as follows:

$$\hat{\mu} = \bar{\bar{x}} \text{ and } \hat{\sigma} = \frac{\bar{s}}{C_2}.$$

Now using these two estimates, the control limits can be written as

$$LCL = \bar{\bar{x}} - 3\,\bar{s}/C_2\sqrt{n} = \bar{\bar{x}} - A_1\,\bar{s}$$

$$CL = \bar{\bar{x}} \text{ and}$$

$$UCL = \bar{\bar{x}} + 3\,\bar{s}/C_2\sqrt{n} = \bar{\bar{x}} + A_1\,\bar{s}$$

where $A_1 \doteq 3/C_2\sqrt{n}$ which is given in Appendix - 1 for various values of n, the sample size.

Example 15.2 Construct a control chart for mean using the estimate of standard deviation from the data given in Example 15.1 and comment.

Solition : To draw the control chart for mean with the estimated value of standard deviation we calculate the following:

Sample No	Total	Mean	Standard Denation
1	4267	711.17	72.14
2	3518	586.33	61.39
3	3909	651.50	45.07
4	3847	741.17	51.05
5	4080	680.00	147.66
6	3836	639.33	65.98
7	3993	665.50	75.00

Now $\overline{\overline{x}}$ =653.57, $\overline{s}$ =74.04 and A_1 = 1.41

Therefore, LCL = 653.57 – 1.41 x 74.04 = 653.57 – 104.40 = 549.17

CL = 653.57 and

UCL = 653.57 + 104.40 = 757.97

All the mean values are plotted on the chart as follows:

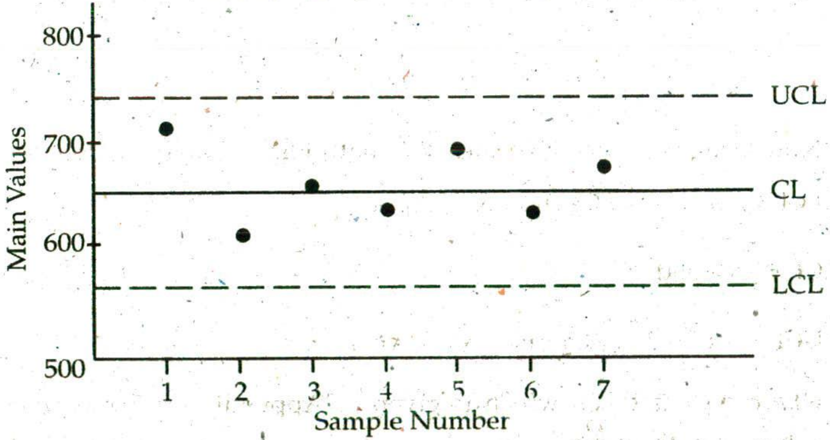


Fig 15.3 Control chart for mean.

Comment : Since all the points are lying within the control limits, the process seems to be within control.

Case 2.2 Standard not Given (Use the estimate of Range): Let there be m sub samples drawn from m sub groups and the sucessive sample means are $\bar{x}_1, \bar{x}_2 \ldots\ldots\ldots\ldots \bar{x}_m$ and also the successive sample ranges are $R_1, R_2 \ldots\ldots\ldots\ldots R_m$. Since in this case, the value of μ and σ are not specified. These are to be estimated from the sample values which are as follows:

$$\text{Let } \bar{\bar{x}} = \sum^{m} \bar{x}_i/m \quad \text{and} \quad \bar{R} = \sum^{m} R_i/m$$

which are the pooled mean and the mean of the sample ranges respectively. The relations $E(\bar{\bar{x}}) = \mu$ and $E(\bar{R}) = d_2\sigma$ (valid for normal variable x) where d_2 is also a function of n, the sample size but not as simple as C_2. The value of d_2 is available in Appendix-1 for different value of n. Now the estimates of μ and σ can be obtained as follows:

$$\hat{\mu} = \bar{\bar{x}} \quad \text{and} \quad \hat{\sigma} = \frac{\bar{R}}{d_2}.$$

Therefore, finally using the range values we have the control limits as follows:

$$LCL = \bar{\bar{x}} - 3\frac{\bar{R}}{d_2\sqrt{n}} = \bar{\bar{x}} - A_2\bar{R}$$

$$CL = \bar{\bar{x}} \quad \text{and}$$

$$UCL = \bar{\bar{x}} + 3\frac{\bar{R}}{d_2\sqrt{n}} = \bar{\bar{x}} + A_2\bar{R}$$

where $A_2 = \frac{3}{d_2\sqrt{n}}$ which is given in the Appendix–1 for various values of n, the sample size.

Example 15.3 Construct a control chart for mean using the estimate of standard deviation from range from the data given in Example 15.1 and comment.

Solution : To draw control chart for mean with the estimated value of the standard deviation from range, we calculate the following:

Sample No	Total	Mean	Range
1	4267	711.7	219
2	3518	586.33	167
3	3909	651.50	134
4	3847	641.17	171
5	4080	680.00	489
6	3836	639.33	200
7	3993	665.50	239

Now $\overline{\overline{x}} = 653.57$, $\overline{R} = 231.29$ and $A_2 = 0.483$

Therefore, LCL = $x - A_2\overline{R} = 653.57 - 0.483 \times 231.29$

$= 653.57 - 111.71 = 541.86$

CL = $\overline{\overline{x}} = 653.57$

and UCL = $\overline{\overline{x}} + A_2\overline{R} = 653.57 + 111.71 = 765.28$

All the mean values are plotted on the chart as follows :

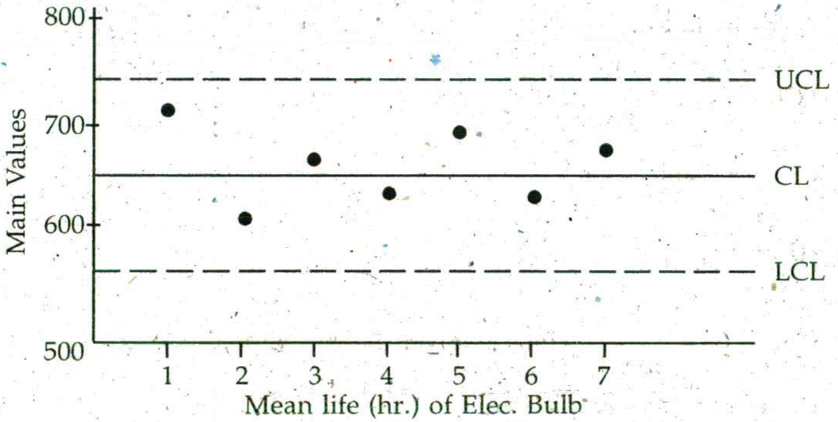


Fig 15.4 Control chart for mean

Comment : From the conrol chart, it is evident that all the points are seen within the control limits. Therefore, the process seems to be within control.

15.8 Control Charts for Standand Deviation

Case 1 (Standard Given) : For normally distributed random variable x, we know $E(s) = C_2\sigma$ and

$\sigma_s = \sigma\sqrt{\frac{n-1}{n} - C_2^2}$ where C_2 is defined earlier. Let us suppose that the value given for σ is σ′, then the control limits will be based on-

$$LCL = C_2\sigma' - 3\sigma'\sqrt{\frac{n-1}{n} - C_2^2} = B_1\sigma'$$

$$CL = C_2\sigma' \text{ and}$$

$$UCL = C_2\sigma' + 3\sigma'\sqrt{\frac{n-1}{n} - C_2^2} = B_2 s'$$

where $B_1 = C_2 - 3\sqrt{\frac{n-1}{n} - C_2^2}$ and $B_2 = C_2 + 3\sqrt{\frac{n-1}{n} - C_2^2}$.

The values of B_1, B_2 and C_2 can be obtained in Appendix-1 for different values of n, the sample size.

Example 15.4 : Draw a control chart for standard deviation with the data given in Example 15.1 when the standard value of standard deviation is given by 81.75.

Solution : In this example n = 6 and therefore, for n = 6 the values of $C_2 = 0.8686$, $B_1 = 0.026$ and $B_2 = 1.711$ are abtained from the table given in Appendix-1. Here the given value of the standard deviation is σ′ = 81.75 therfore,

$LCL = B_1\sigma' = 0.026 \times 81.75 = 21.3$

$CL = C_2\sigma' = 0.8686 \times 81.75 = 71.01$

$UCL = B_2\sigma' = 1.711 \times 81.75 = 139.87$

All the standard deviation values are plotted on the chart as follows:

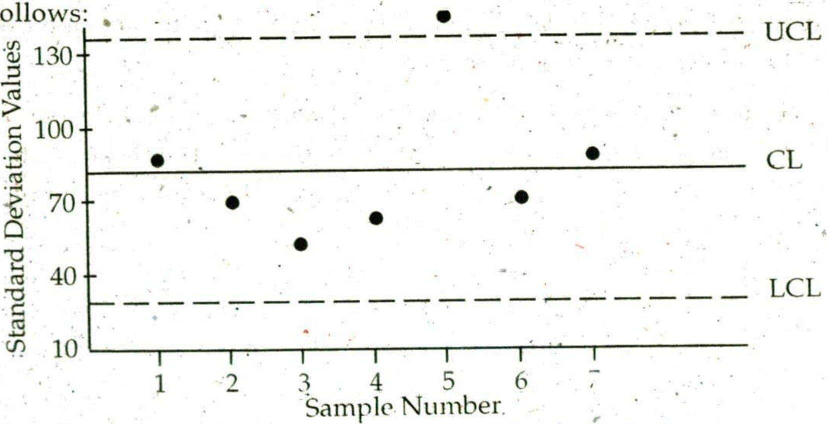


Fig 15.5 Control chart for standard deviation

Comment : All the points represented by the sample standard deviations are seen to be within the control limits except the 5th one. Therefore, the process seems to be out of control.

Case 2 (Standard Not Given) : In this case also, we assume that the random samples provided are drawn from normal distributions. Here, since the standard deviation value is not given, it is to be estimated from the samples, we have to estimate σ by $\overline{s}/C_2$ where $\overline{s}$ is the mean of the sample standard deviations and C_2 is constant dependent on n, the sample size, available is Appendix-1.

Now the control limits for the standard deviation are as follows:

$$LCL = \overline{s} - 3\frac{\overline{s}}{C_2}\sqrt{\frac{n-1}{n} - C_2^2} = B_3\overline{s}$$

$CL = \overline{s}$ and

$$UCL = \overline{s} + 3\frac{\overline{s}}{C_2}\sqrt{\frac{n-1}{n} - C_2^2} = B_4\overline{s}$$

where, $B_3 = 1 - \frac{3}{C_2}\sqrt{\frac{n-1}{n} - C_2^2}$ and $B_4 = 1 + \frac{3}{C_2}\sqrt{\frac{n-1}{n} - C_2^2}$

The values of B_3 and B_4 are available for differnt values of n, the sample size in Appendix-1.

Example 15.5 Draw a control chart for standard deviation of the data given in Example 15.1 when the standard value of standard deviation is not given and comment.

Solution : We calculate mean and standard deviation of the sample values as given is the solution of Example 15.2. The mean of the standard deviation values is $\overline{s} = 74.04$ and for n = 6, the sample size we have $B_3 = 0.03$ and $B_4 = 1.97$ Therefore,

$LCL = B_3\overline{s} = 0.03 \times 74.04 = 2.22$

$CL = \overline{s} = 74.04$ and

$UCL = B_4\overline{s} = 1.97 \times 74.04 = 145.86$

Now all the standard deviation values are plotted on the chart as follows:

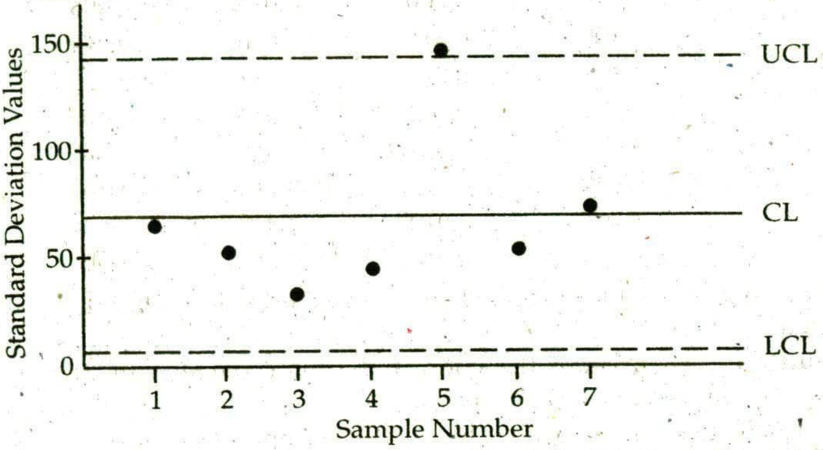


Fig 15.6 Control chart for standard deviation

Comment : All the sample standard deviation values except the 5th sample point are seen within the control limits. Therefore, the process seems to be out of control.

15.9 Control charts for Range

Case 1 (Standard Given) : For a normally distributed random variable x, We have $E(R) = d_2\sigma$ and $\sigma_R = D\sigma$ where d_2 and D are functions of n, the sample size. If the standard value of σ is given to be σ´ then the control limits for range can be constructed as follows :

$LCL = d_2\sigma' - 3D\sigma' = D_1\sigma'$

$CL = d_2\sigma'$ and

$UCL = d_2\sigma' + 3D\sigma' = D_2\sigma'$

where, $D_1 = d_2 - 3D$ and $D_2 = d_2 + 3D$. The values of D_1, D_2 and d_2 are available in the Appendix-1 for various values of n, the sample size.

Example 15.6 A machine is manufacturing mica discs with specified thickness between 0.008" and 0.015". Samples of size 4 are drawn every hour and their thickness in inches are recorded as follows:

Sample No	**Thickness of mica discs**			
1	0.014	0.008	0.012	0.012
2	0.011	0.010	0.013	0.008
3	0.011	0.012	0.016	0.014
4	0.017	0.012	0.017	0.016
5	0.015	0.012	0.014	0.010
6	0.013	0.008	0.015	0.015
7	0.014	0.012	0.013	0.010
8	0.011	0.010	0.008	0.016
9	0.014	0.010	0.012	0.009

Draw control chart for range when the standard value of $\sigma' = 0.01$ is given.

Solution : To prepare the control chart for range with given standard value σ, we proceed as follows:

We calculate the range values for all the 9 samples each of size 4.

Sample No:	1	2	3	4	5	6	7	8	9
Range :	0.006	0.005	0.005	0.005	0.005	0.007	0.004	0.008	0.005

For n = 4 we have from Appendix-1, $D_1 = 0$, $D_2 = 4.918$ and $d_2 = 2.059$.

For the standard value given, $\sigma' = 0.01$, we have

$LCL = D_1 \sigma' = 0 \times 0.01 = 0$

$CL = d_2\sigma' = 2.059 \times 0.01 = 0.02$ and

$UCL = D_2\sigma' = 4.918 \times 0.01 = 0,05$

Now all the values of range are plotted on the chart as follows:

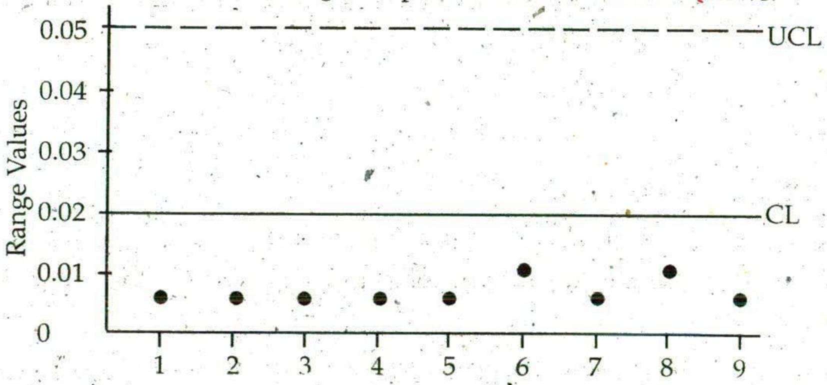


Fig 15.7 Control Chart for range

Comment : Since all the points are within the control limits. The process seems to be under control. But it can be easily noticed that all the points are lying below the central line, this indicates that there may be some systematic causes of variations which are to be checked.

Case 2 (Standard Not Given) : When the standard value of σ is not specified. it is to be estimated by the sample observations. It is usually done by the following method :

We Know, $E(R)=d_2\sigma \therefore \hat{\sigma}= \overline{R} /d_2$

where $\overline{R}$ is the mean of the range values obtained from the sample observations. Finally the control limits can be formed as follows:

$$LCL = \overline{R} -3\frac{D}{d_2}\overline{R} = D_3 \overline{R}$$

$CL = \overline{R}$ and

$$UCL = \overline{R} +3\frac{D}{d_2}\overline{R} = D_4 \overline{R}$$

where $D_3 = 1 - 3D/d_2$ and $D_4 = 1 + 3D/d_2$. The values of D_3 and D_4 for various values of n, the sample size are given in Appendix-1.

Note: 1) In case of control chart for range in both the above cases, if the lower control limit comes out to be negative, it is to be considered as zero because the value of range by its very nature can never be negative quantity.

2) Since the value of R can be calculated with the shortest possible time in calculating differnt measures of dispersion, the control chart for range is very much used in different production process where the decision is to be taken in short time. Therefore, it is a largely used technique of drawing control chart.

Example 15.7 Find out the control chart for range from the data given in Example 15.6 when the standard value of σ is not specified.

Solution : Since the value of σ is not specified, it is to be estimated from the values of ranges obtained from sample observations. First we calculate the sample ranges and thier mean is calculated. For calculation of mean of the ranges, we prepare the following table:

Sample No:	1	2	3	4	5	6	7	8	9
Ranges:	0.006	0.005	0.005	0.005	0.005	0.007	0.004	0.008	0.005

$\overline{R} = \Sigma R_i/9 = 0.005$ and for $n = 4$, $D_3 = 0$ and $D_4 = 2.282$; the control limits can be obtained as follows :

$LCL = D_3 \overline{R} = 0 \times 0.005 = 0$

$CL = \overline{R} = 0.005$ and

$UCL = D_4 \overline{R} = 2.282 \times 0.005 = 0.013$

Now all the range values are plotted on the chart as follows :

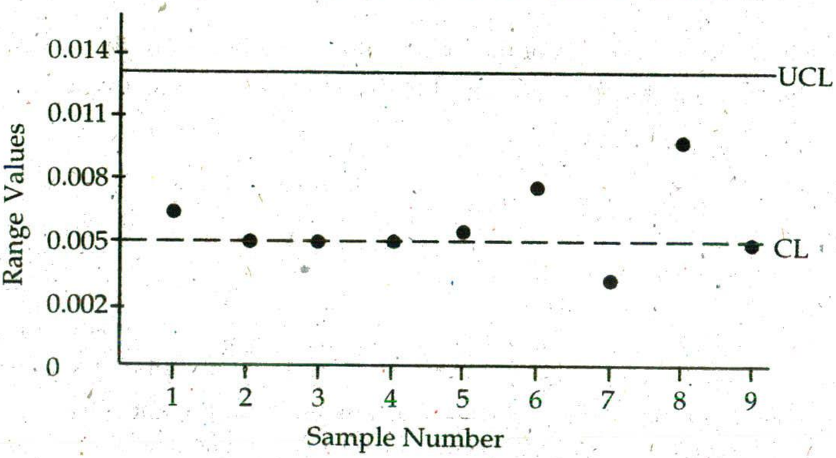


Fig 15.8 Control chart for range.

Comment : Since all the range values are lying within the control limits the process seems to be within control.

15.10 Control Charts for Number of Defectives

For qualitative characteristic of an attribute, each item is recorded either defective or non-defictive. In that case, to judge whether the production process is in control or not, one has to ascertain whether the population number of defectives are same from one sub-group to other i.e., the number of defectives say d in the sample or the fraction defectives $p = d/n$ in the sample where n in the number of items inspected per sub-sample are same from one sub-group to other. Here also we consider two cases one by one.

Case 1 (Standard Given) : In this case, we assume that each random sample is taken with replacement or even if taken without replacement, is taken from practically infinite population. We may suppose that d=np is distributed as Binomial form with E(np) = nP and $\sigma_{np} = \sqrt{nP(1-P)}$ where P is same in all sub-groups if and only if the process is in control.

Let the given standard value of P be p´ then the control limits will be constructed on the basis of

$LCL = np' - 3\sqrt{np'(1-p')}$

$CL = np'$ and

$UCL = np' + 3\sqrt{np'(1-p')}$

Note : Since p´ can never be negative, if LCL value calculated is negative then it is considered to be zero.

Example 15.8 Following are the number of defectives of 16 samples each containing 1500 certain types of transitors in a manufacturing process-

280, 306, 337, 305, 356, 402, 216, 264, 126, 409, 193, 326, 280, 389, 415, 315

Draw control chart for number of defectives and comment, when the standard value of P is given by 0.27.

Solution : To draw control chart for number of defectives for given value of P = 0.27, we proceed as follows:

Here, n = 1500 p = 0.27 ∴ 1 – p = 1 - 0.27 = 0.73.

Thus the

$LCL = 1500 \times 0.27 - 3\sqrt{1500 \times 0.27 \times 0.73} = 405 - 51.58 = 353.42$

$CL = 405$ and

$UCL = 405 + 51.58 = 456.58$

Now all the number of defectives are plotted on the chart as follows:

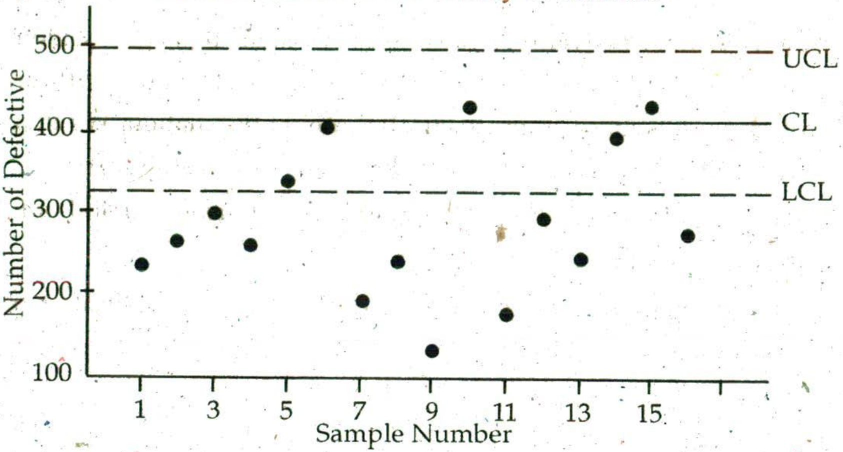


Fig 15.9 Control chart for number of defectives.

Comment : Since a number of points are seen to be outside the control limits, therefore, the process seems to be out of control.

Case 2 (Standard Not Given) : If the standard value of P is not given or specified then it is to be estimated from the sample values. The appropriate estimate of P is nothing but the average of the fraction defectives. Therefore, out of the total items inspected n, the number of defectives are usually given or detected, hence the fraction defective for each sub-sample can be obtained by the formula given by $p_i = d_i/n$ where i indicates the ith sub-sample.

Thus $\overline{p} = \sum_{i=1}^{m} p_i/m$ where m is the number of sub-samples. Finally the control limits can be formed as follows:

$$LCL = n\overline{p} - 3\sqrt{n\overline{p}\ (1 - \overline{p})}$$

$$CL = n\overline{p} \text{ and}$$

$$UCL = n\overline{p} + 3\sqrt{n\overline{p}\ (1 - \overline{p})}$$

Note : The number of defectives, np can never be negative, hence if LCL comes out to be negative is considered as zero.

Example 15.9 Draw control chart for number of defectives from the data given in Example 15.8 when no standard value is given and comment.

Solution : In this case, the standard value of P, the fraction defective is not given and therefore, it is to be estimated from the sub-sample values. From the given data, the estimation procedure can be done as follows:

Sample No.	No of Defectives (di)	Fraction Difectives (di/n)
1	280	0.187
2	306	0.204
3	337	0.22
4	305	0.203
5	356	0.237
6	402	0.268
7	216	0.144
8	264	0.176
9	126	0.084
10	409	0.273
11	193	0.129
12	326	0.217
13	280	0.187
14	389	0.259
15	415	0.277
16	315	0.210

Here n = 1500 and $\overline{p} = \Sigma p_i/16 = 3.28/16 = 0{:}205$, the estimated value of the fraction defective.

Now n = 1500, $\overline{p} = 0.205$ and $(1 - \overline{p}) = 0.795$. Therefore,

LCL = $1500 \times 0.205 - 3\sqrt{1500 \times 0.205 \times 0.795} = 307.50 - 46.91 = 260.59$

CL = 307.50 and

UCL = $307.50 + 46{:}91 = 354.41$

Now all the number of defectives are plotted on the control chart as follows :

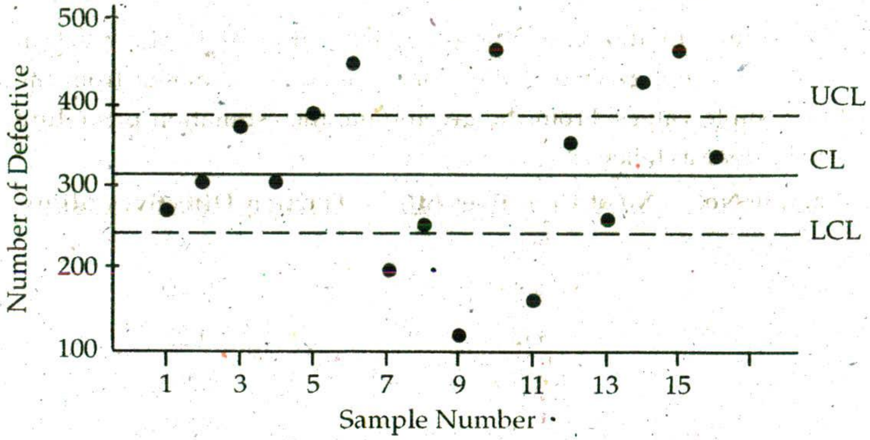


Fig 15.10 Control chart for number of defectives

Comment : Since a number of points are seen to be outside the control limits, the process seems to be out of control.

15.11 Control Charts for Fraction Defectives

In some cases, we are interested to find out the control chart for fraction defectives instead of the number of defectives. Here d is taken as the number of defectives and p is fraction defectives defiened by d/n where n is the number of items inspected per sub-sample.

Case 1 (Standard Given) : In this case, we construct control chart for p, the fraction defectives instead of np, the number of defectives. We know that

$E(p) = P$ and $\sigma_p = \sqrt{P(1-P)/n}$, since we know that $p = d/n$

$$\therefore\ V(p) = V(d/n) = \frac{1}{n^2}V(d) = nPQ/n^2 = PQ/n = P(1-P)/n$$

Let us consider that the standard value of the fraction defective, P is given by p´. Then the control limits will consist of

$$LCL = p' - 3\sqrt{p'(1-p')/n} = p' - A\sqrt{p'(1-p')}$$

$$CL = p' \text{ and}$$

$$UCL = p' + 3\sqrt{p'(1-p')/n} = p' + A\sqrt{p'(1-p')}$$

where $A = 3/\sqrt{n}$ which can be obtained for various values of n. the sample size in Appendix-1.

Note : Since the fraction defective can never be negative therefore, LCL if comes out to be negative will be considered as zero.

Example 15.10 Draw control chart for fraction defectives from the data given in Example 15.8 when the standard value of P is given by 0.2 and comment.

Solution : Here in the example n=1500, $p' = 0.2$ and therefore,

$$LCL = p' - 3\sqrt{p'(1-p')/n} = 0.2 - 3\sqrt{(0.2 \times 0.8)/1500} = 0.2 - 0.03 = 0.17$$

CL = 0.2 and

UCL = 0.2 + 0.03 = 0.23

Now we calculate the fraction defectives from the data given in Example 15.8 which are displayed in Example 15.9 and those values are plotted on the control chart as follows:

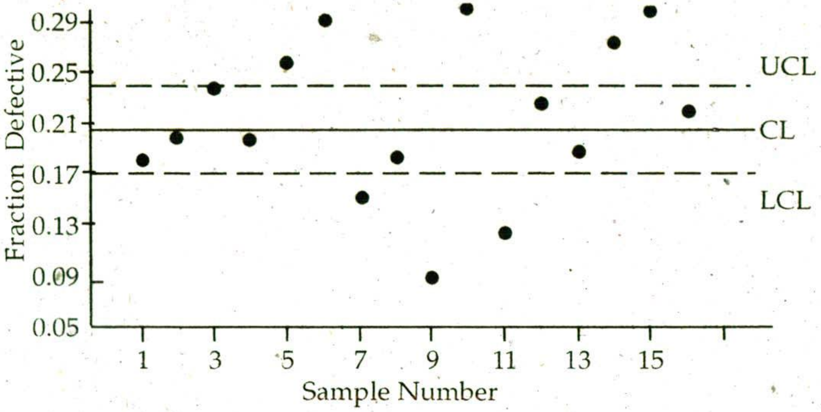


Fig 15.11 Control Chart for fraction defrctives.

Comment : Since so many points are seen to be outside the control limits, the process seems to be out of control.

Case 2 (Standard Not Given) : In the case, the value of fraction defective, p is not given and therefore, it is to be estimated from the sample values. We can calculate p_i, the fraction defective corresponding to ith sub-sample. Thus the estimated value of p is .

$$\hat{p} = \bar{p} = \sum_{i=1}^{m} p_i/m$$ where m is the number of sub-samples.

Therfore, the control limits for fraction defectives are as follows:

$$LCL = \bar{p} - 3\sqrt{\bar{p}(1-\bar{p})/n} = \bar{p} - A\sqrt{\bar{p}(1-\bar{p})}$$

CL = $\bar{p}$ and

$$UCL = \bar{p} + 3\sqrt{\bar{p}(1-\bar{p})/n} = \bar{p} + A\sqrt{\bar{p}(1-\bar{p})}$$

where $A = 3/\sqrt{n}$ which is given in Appendix - 1 for various values of n, the sample size.

Note : In the above case also, as $\overline{p}$ can never be negative the LCL if comes out to be negative according to the above formula, will be considered as zero.

Example 15.11 Draw control chart for fraction defectives from the data given in Example 15.8 when no standard value is given and comment.

Solution : Here the standard value of p is not given. It is to be estimated from the sample values. First we calculate the fraction defectives for each sub-sample having sample size 1500 with the formula di/n where di is the number of defectives in the ith sub-sample. The calculated values are seen in the solution of Example 15.9. Here $\hat{p} = \overline{p} = \sum p_i/16 = 3.28 / 16 = 0.205$ and $n = 1500$. Therefore, the control limits are.

$$LCL = \overline{p} - 3\sqrt{\overline{p}\ (1 - \overline{p})/n} = 0.205 - 3\ (0.205 \times 0795)/1500$$

$$= 0.205 - 0.031 = 0.174$$

$CL = 0.205$ and

$UCL = 0.205 + 0.031 = 0.236$

Now all the fraction defective values are plotted on the control chart as follows:

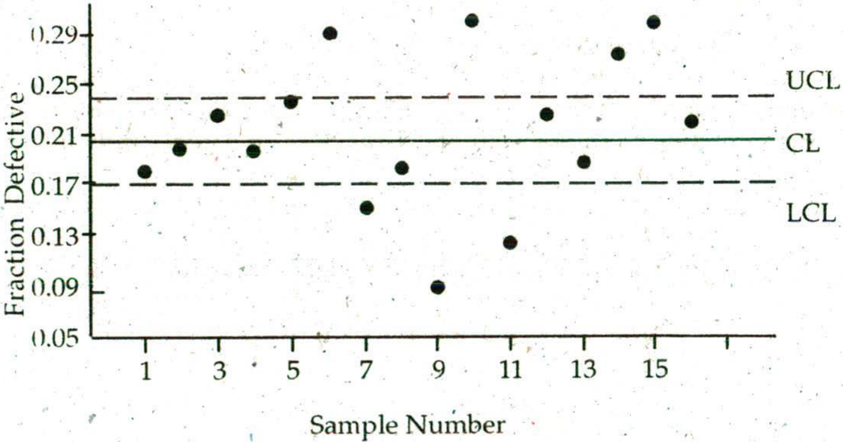


Fig 15.12 Control chart for fraction defectives

Comment : Since a number of points are seen outside the control limits, the process seems to be out of control.

15.12 Control Charts for Percent Defectives

In this case, we construct control chart for 100 p instead of p, the fraction defectives. In some cases, the values of p may be very small and in these cases, we usually use the values of percent defectives, 100 p and construct control chart for percent defectives.

Case 1 (Standard Given) : Here the value of the fraction defective p is given by p´ and the control chart for percent defectives is drawn depending on the following control limits :

$LCL = 100p' - 3 \times 100 \sqrt{p'(1-p')/n} = 100p' - A \times 100 \sqrt{p'(1-p')}$

$CL = 100p'$ and

$UCL = 100p' + 3 \times 100\sqrt{p'(1-p')/n} = 100p' + A \times 100 \sqrt{p'(1-p')}$

where $A = 3/\sqrt{n}$ which is given in Appendix-1 for various values of n, the sample size.

Note : Since the fraction defective can never be negative, the LCL if comes out to be negative will be considered as zero.

Example 15.12 Draw control chart for percent defectives from the data given in Example 15.8 for given value of p´= 0.2

Solutiion : Here n=1500 and the given value of p´ = 0.2. Therefore, the control limits are as follows:

$LCL = 100 \times 0.2 - 100 \times 3 \sqrt{(0.2 \times 0.8)/1500} = 20 - 3.1 = 16.9$

$CL = 20$

$UCL = 20 + 3.1 = 23.1$

Now before drawing the control chart we calculate the percent defectives for each of 16 sub-samples as follows:

Sample No.	Fraction Defective (p_i)	Percent defective ($100p_i$)
1	0.187	18.7
2	0.204	20.4
3	0.225	22.5
4	0.203	20.3
5	0.237	23.7
6	0.268	26.8
7	0.144	14.4
8	0.176	17.6
9	0.084	8.4
10	0.273	27.3
11	0.129	12.9
12	0.217	21.7
13	0.187	18.7
14	0.259	25.9
15	0.277	27.7
16	0.210	21.0

Now we plot all the percent defective values on the control chart as follows:

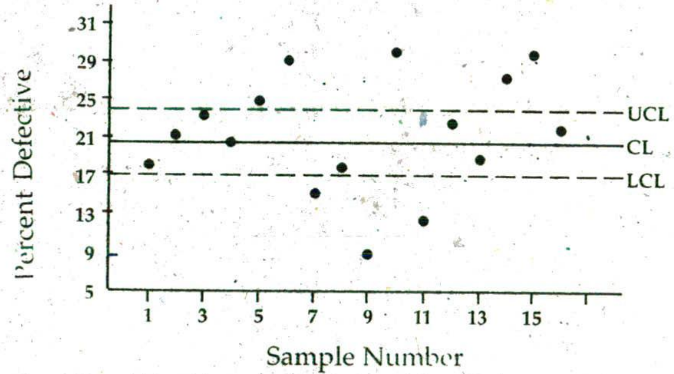


Fig 15.13 Control chart for percnt Defectives.

Comment : Since a member of points are seen outsides the control limits. The process seems to be out of control.

Case 2 (Standard Not Given) : When the value of the fraction defective, p is not given, we usmally estimate it from the sample values by $\overline{p} = \sum^{m} p_i / m$ where m is the number of sub-samples.

In this case, the control limits are as follows:

$LCL = 100\ \overline{p} - 3 \times 100 \sqrt{\overline{p}\ (1 - \overline{p})/n} = 100\ \overline{p} - 100\ A \sqrt{\overline{p}\ (1 - \overline{p})}$

$CL = 100\ \overline{p}$ and

$UCL = 100\ \overline{p} + 3 \times 100 \sqrt{\overline{p}\ (1 - \overline{p})/n} = 100\ \overline{p} + 100\ A \sqrt{\overline{p}\ (1 - \overline{p})}$

where $A = 3/\sqrt{n}$ which is given in Appendix-1 for different values of n,. the sample size.

Note : Since $\overline{p}$ can never be negative if LCL comes out to be negative will be considered as zero.

Example 15.13 Draw control chart for percent defectives from the data given in Example 15.8 when no standard value of p, the fraction defective is given and comment.

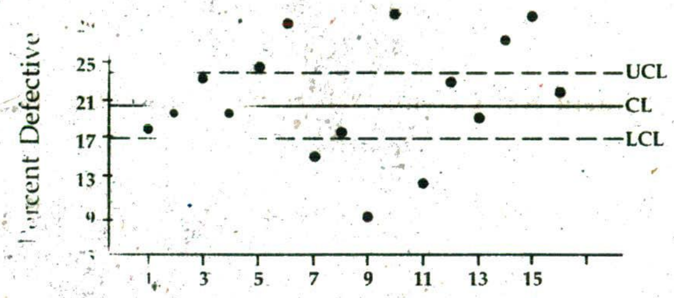
Solution : Here p, the fraction defective is not given. It is estimated from the sample values given and thus we get $\overline{p} = \sum_{i=1}^{m} p_i / m$ $= 3.28 / 16 = 0.205$. Therefore, for $n = 1500$ and $\overline{p} = 0.205$ the control limits are as follows:

$LCL = 100 \times 0.205 - 100 \times 3 \sqrt{(0.205 \times 0.795)/1500} = 20.5 - 3.13 = 17.37$

$CL = 20.5$ and

$UCL = 20.5 + 3.13 = 23.63$

Now we plot all the percent defective values as calculated in the solution of example 15.12 on the control chart as follows:



Sample Number

Fig 15.14 Control chart for percent defectives.

Comment : Since a large number of points are seen outside the control limits, the process seems to be out of control.

15.13 Control Charts for Number of Defects

A defective is an item that fails to fullfil one or more of the given specifications and a defect is any instance of the item's lack of conformity to specifications. Every defective items thus contains one or more defects. For example, the number of holes or tears in say 5 metres piece of cloth, the number of black spots on a roll of white paper or the photographic film, loose screws or exposed wires in a television set, the number of loose connections in differant parts of a finished car and so on.

In a continuous process of production system, the number of defects may be large in number, even though the probability of a defect to occur in any one spot is very negligible and may assume to follow a Poisson distribution with parameter λ. A control chart for member of defects, c will then aim at detecting any differences among the λ values in the sub-sample.

Case 1 (Standard Given) : We know that for a Poisson variable c with parameter λ, $E(c) = \lambda$ and $\sigma_c = \sqrt{\lambda}$. Hence if a standard value for λ say λ' is given, then the control chart for c will be based on

$LCL = \lambda' - 3\sqrt{\lambda'}$

$CL = \lambda'$ and

$UCL = \lambda' + 3\sqrt{\lambda'}$

Note : Since the number of defects cannot be negative, if LCL comes out to be negative will be considered as zero.

Example 15.14 Following are the number of defects found in 1000 items of cotton piece goods inspected everyday in a certain month:

Day:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
No of Defects:	1	1	3	7	8	1	2	6	1	1	10	5	0	19	16
Day:	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
No of Defects:	20	1	6	12	4	5	1	8	7	9	2	3	14	6	8

Draw control chart for the number of defects when the value of the expected number of defects, $\lambda = 8$ is given and comment.

Solution : Here $\lambda = 8$, therefore, the control limits can be constructed as fallows:

$LCL = 8 - 3\sqrt{8} = 8 - 8.49 = -0.49 = 0$

$CL = 8$ and

$UCL = 8 + 8.49 = 16.49$

Now we draw all the points representing the number of defects on the control chart as follows:

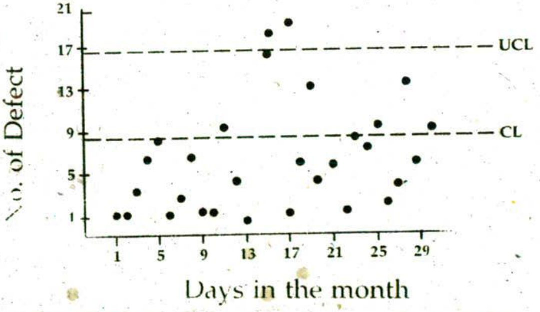


Fig 15.15 Control chart for number of defects.

Comment : Since the number of defects on the 14th and 16th days of the month are outside the control limits, the process seems to be out of control.

Case 2 (Standard Not Given) : When no standard is specified, λ is to be estimated from the observed c, the number of defects. Let us suppose that c_i be the number of defects for the ith sub-sample ($i = 1, 2, \ldots\ldots, m$), the appropriate estimate of λ will be-

$$\hat{\lambda} = \bar{c} = \sum_{i=1}^{m} c_i/m.$$

Now the control limits for the number of defects would be-

$$LCL = \bar{c} - 3\sqrt{\bar{c}}$$

$CL = \bar{c}$ and

$$UCL = \bar{c} + 3\sqrt{\bar{c}}$$

Note : It is easily understood that the number of defects cannot be negative and therefore, if LCL comes out to be negative will be considered as zero.

Example 15.15 Draw control chart for the number of defects of the data given in Example 15.14 and comment when the standard value is not given.

Solution : Since the standard value of λ is not given, it is to be estimated from the sample values.

Here $\hat{c} = \Sigma c_i/30 = 187/30 = 6.23$.

Therefore, the control limits are as follows:

$LCL = 6.23 - 3\sqrt{6.23} = 6.23 - 7.49 = - 1.26 = 0$

$CL = 6.23$ and

$UCL = 6.23 + 7.49 = 13.72$

Now we draw all the number of defects on the control chart as follows:

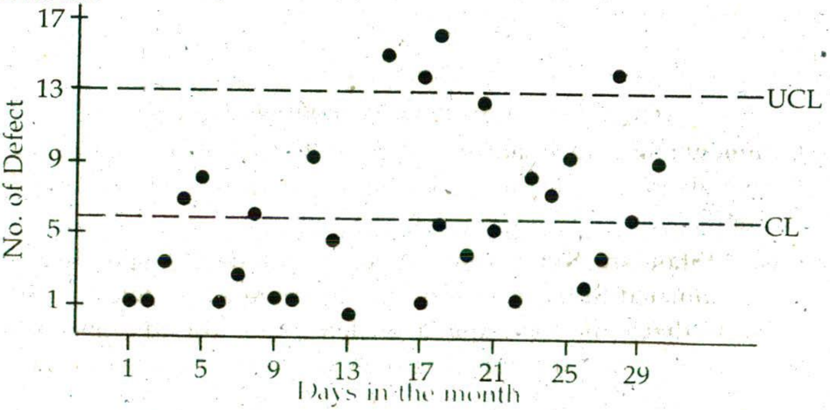


Fig 15.16 Control chart for number of defects.

Comment : The points indicated by the no of defects in different values of sub-samples namely 14, 15, 16 and 28 give evidence of lack of control since they are outside the control limits. Therefore, the process seems to be out of control.

15.14 Control Charts for Defects per Unit

Here we construct control chart for defects per unit which is almost exactly equal to the control chart for number of defects c if the area of opportunity or sample size varies from one sub-sample to another. Moreover, by means of this approach the overall picture can be given without too much detail.

Let us consider an example where the utility of this type of chart can be explained very clearly. A plant making trucks began to analyse the number of defects per truck in the final inspection point at the end of the assembly line. In a certain day, let there be 95 trucks ready for shipment in which in all 190 defects were detected. So the number of defects per truck is now can be defined by $u_1 = c/n_1 = 190/95 = 2.00$ In the next day, let there be 105 trucks ready for shipment in which a total of 215 defects were detected. So that $u_2 = c/n_2 = 215/105 = 2.05$ and so on. With all these u values a control chart can be drawn. If the number of trucks for shipment per day are considered to be same then the control chart can be constructed more easily.

The whole procedure can be displayed as follows:

Let us define some notations.

n = number of units of products per sub-sample,

c = total number of defects in the sample of n units,

u = c/n = number of defects per unit

$\overline{u} = \sum_{i=1}^{m} u_i/m$ = average of u's for m sub-samples.

Case 1.(Standard Given) : Let the standard value of the number of defects per unit be u´, then the 3σ control chart for defects per unit would be $u' \pm 3\sigma_u'$ where $\sigma_u' = \sigma_c/n = \sigma_c/n$. Since we know that

$u' = \frac{c}{n}$ or $c = nu'$ and $\sigma_c = \sqrt{c} = \sqrt{nu'}$. So that $\sigma_u' = \sqrt{nu'}/n = \sqrt{u'/n}$.

Therefore, the control limits for defects per unit are

$LCL = u' - 3\sqrt{u'/n}$

$CL = u'$ and

$UCL = u' + 3\sqrt{u'/n}$

Note : The number of defects per unit cannot be negative and therefore, if the LCL comes out to be negative will be considered as zero.

Example 15.16 The number of defects per unit truck in an assembly section of a truck producing company were recorded from November 2 to November 29 in the working days. Draw control chart for number of defects per unit and comment when the standard value is given, u = 19 and the number of trucks per day produced is considered to be same and n=95.

November :	4	5	6	7	8	11	12	13	14	15
No of Defects/Unit :	12	15	15	27	19	24	30	28	17	20
November :	18	19	20	21	22	25	26	27	29	
No of Defects/Unit :	21	19	18	12	16	22	25	20	14	

Solution : Here n = 95 and u´ = 19. Therefore, the control limits are

$LCL = 19 - 3\sqrt{19/95} = 19 - 1.34 = 17.66$

$CL = 19$ and

$UCL = 19 + 1.34 = 20.34$

Now we draw all the points of defects per unit on the control chart as follows:

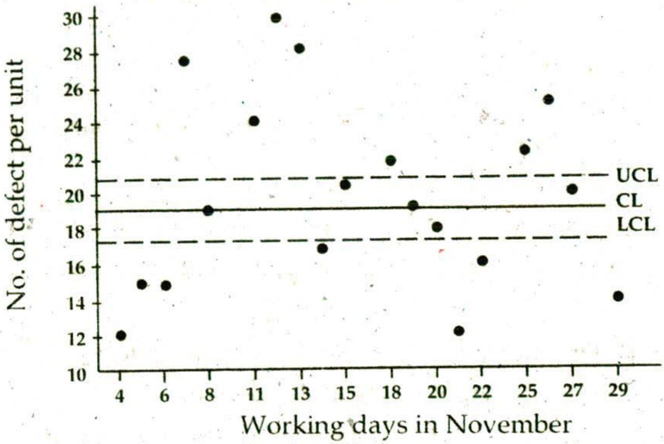


Fig 15.17 Control chart for defects per unit

Comment : A number of points are seen out side the control limits and therefore, the process seems to be out of control.

Case 2 (Standard Not Given) : If the standard value of u, the number of defects per unit is not given, it is to be estimated from the sample values which is $\overline{u} = \sum_{i=1}^{m} u_i/m$ = average of u's for m sub-sample in which the total number of units inspected in every sub-sample remaining same from one sub-sample to another. Now the control limits for number of defects per unit are as follows:

$$LCL = \overline{u} - 3\sqrt{\overline{u}/n}$$

$$CL = \overline{u} \text{ and}$$

$$UCL = \overline{u} + 3\sqrt{\overline{u}/n}$$

If the sub-sample sizes are not same then,

$\overline{u} = \sum_{i=1}^{m} n_i u_i/\Sigma n_i$ where u_i and n_i are respectively the number of defects per unit and sample size for the ith sub-sample. In this case, the control limits for number of defects per unit are as follows:

$$LCL = \overline{u} - 3\sqrt{\overline{u}/n}$$

$$CL = \overline{u} \text{ and}$$

$$UCL = \bar{u} + 3\sqrt{\bar{u}/n}$$

where $n = \dfrac{n_1 + n_2 + \ldots\ldots\ldots + n_m}{m} = \Sigma n_i/m$, the average size of the samples.

Note : The number of defects per unit cannot be negative and therefore, if the LCL comes out to be negative will be considered as zero.

Example 15.17 Draw control chart for number of defects per unit from the data given in Example 15.16 and comment.

Solution: Here n = 95 and the standard value of number of defects per unit is not given, it is to be estimated from the sub-sample values provided here in the example. Therefore, $\bar{u} = \Sigma u_i/19 = 374/19 = 19.68$

Now, the control limits are

$$LCL = \bar{u} - 3\sqrt{\bar{u}/n} = 19.68 - 1.37 = 17.63$$

$$CL = \bar{u} = 19.68 \text{ and}$$

$$UCL = \bar{u} + 3\sqrt{\bar{u}/n} = 19.68 + 1.37 = 21.05$$

Now we draw all the points of defects per unit on the control chart as follows:

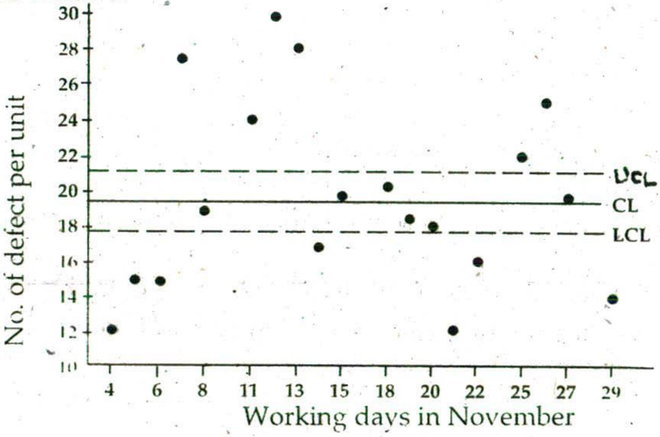


Fig 15.18 Control chart for defects per unit.

Comment : Since a number of points are seen out-side the control limits, the process seems to be out of control.

15.15 Sampling Inspection by Attributes

We know that the process control is achieved through the technique of control chart. Process control ensures that the quality of a manufacturing process is satisfactory or not. It may mean a very great saving as well as the reputation of the industry which always try to produce uniformly good products. If a trouble starts in the process, it is of great economic importance to detect it immediately.

Lot control is achieved through the sampling inspection. It is not practicable to inspect the lot fully. For simplicity, we here consider the sampling inspection for attribute i.e. the items are judged good or defective by inspection and the quality of the lot would be adjusted from the sample fraction defective.

A sampling plan may be of either the accceptance, rejection or the acceptance-rectification type. In the former, the lot is either accepted or rejected in the light of the sampling. In the latter, it is subjected to cent percent inspection and in either case, all defective items are replaced by good items. In this case, we accept the lot after replacing all defective items in the sample.

Here we are considering some of the important terms used in different sampling plans.

Producer's Risk : A producer produces goods. By producer we mean any person, firm or a department which produces goods and supplies to another person, firm of other department. In sampling we face two kinds of risks, one, a good quality of products may be rejected through sampling on the other hand a bad quality may be accepted as good one. Suppose a producer claims that he has standardised the quality at a level of fraction defective $\overline{P}$, which may be termed as producer's process average. The probabilities of rejecting a lot under the sampling inspection plan when the fraction defective is actually $\bar{p}$ is called the producer's risk and is usually denoted by Pp.

Consumer's Risk : By consumer, we mean a person, firm or department which receives goods from the producers. The consumers have to face a risk of accepting bad quality of goods as good quality. If p_i be the lot tolerance fraction defective i.e. maximum fraction defective in the lot that he will tolerate, then the probability of accepting a lot with fraction defective p_i under the sampling inspection plan is called the consumer's risk and is denoted by p_c.

Average Outgoing Quality Limit (AOQL) : Average outgoing quality is defined as the expected fraction defectives remaining in the lot after the application of the sampling plan, Let P be the actual fraction defective in the lot. The maximum value of average out going quality with respect to P is known as average out-going quality limit (AOQL).

Average Sample Number (ASN) : Average sample number (ASN) is defined as the expected value of the sample size requiered for coming to a decision for acceptance or rejection of a lot under the sampling inspection plan. ASN is a function of p, the actual fraction defective of the lot, A curve obtained by putting ASN against p is called the ASN curve.

Operating Characteristics (O. C) Curve : Let p be the fraction defective of the lot. The operating characteristics (OC) is the mathematical expression L(p) stating the probability of accepting a lot as a function of p. The curve obtained by plotting L(p) against p is called OC curve. Naturally the steeper the OC curve, the greater is the protection to the consumer. An ideal plan would reject all lots which are of worse quality and accepts all lots which are equal to or better than that quality. Such an ideal plan can never be attained.

16. DEMOGRAPHY

16.1 Introduction

At the very outset, let us introduce ourselves with vital statistics. Vital statistics may be defined as that branch of Biometry which deals with data and the law of human mortality, morbidity and demography. The term vital statistics refers to numerical data or the tecniques used to analyse the data related to vital events of a section of population. Here population refers to the aggregate of persons represented by certain types of statistics. By vital events we usually mean those events of human life such as fertility, mortality, marraige, divorce, seperation, adoption etc.

Vital statistics are extensively used in almost all the spheres of human activity. Some important applications are as follows :

1. **Study of Population Trend :** Vital statistics usually reflect the changing pattern of population in any region, community or country in terms of births, deaths and marriages. Comparison of birth rate and death rate of different populations help us to form some idea about the population trend of those region or countries and also the general standard of living and fertility of the races.

2. **Use in Public Administration :** The study of population movement, population projections and other related studies together with birth and death statistics according to age, sex distribution provides administrative and overall planning and evaluation of economic and social development programmes.

3. **Use in Medical Science :** Mortality and nautality statistics also provide guide spots for researchers in medical science, health department and pharmaceutical profession.

4. **Use in Insurance Schemes :** The whole of acturial science including life insurance is based on the mortality of life tables. The vital records concerning all possible factors contributing to deaths in various ages are important tool in all life insurance schemes.

16.2 Meaning of Demography

Demography may be defined as the numerical picture of a human population. It usually deals with the behaviour of the aggregate and not with the behaviour of the individuals. In the process of replacement, the numbers of population are always changing. Some people die each year and others are born. In addition, there may be some net gain or loss by migration. This is in short, called vital process. The structure of a population is dependent on the kind of balance of these vital process. There are usually two main aspects of the behavior of populations, the composition of the aggregate and changes that occur during some period of observation. Since a population is subject to constant change, We have a continuous process of having data of so many events of human life.

The data thus obtained are to be analysed. The statistical operations of demography can be stated shortly as (i) The calculation of conventional measurements (ii) Comparison of these measurements and (iii) Estimation of certain figures those are not fully available.

The basic techniques of population study are not straight forward and cannot by applied every where in the same manner. The common strategy for treating the statistical material of demography can be stated by asking three distinct questions. Those are as follows :

(1) What is the level of performances in some form of behavior in the population, for example, the birth or death rates, the percentages who have some regular work or the proportion married?

(2) Has the level of performance changed during some period of observations? This question is answered by comparing observations at the begining and at the end of the period.

(3) Are there patterns of variation within the population? The answer of this question is found by comparing observation calculated for different groups, age, sex, occupation, education groups and or the combination of these.

16.3 Sources of Demographic Data

There are two main sources of data viz. (1) Census and (2) Registration, which we deal in demography. By census we mean counting all the persons present in an entire population on or about the same date. The other form of statistics is a record of vital events in a calender year. This is provided by some scheme of registration designed to record all of certain events such as births, deaths, migrations, marriage and divorce as they occur. The main distinction of these two systems are that a person may be recorded several times during a year in registration process but in a census he should be counted only once. In census we have records of persons and in registration we have records of events. A census can ascertain the number of vital events indirectly by asking each persons or household of the past years but in registration we have data given by some persons who are involved during a year. A census is conducted at regular interval of time, usually 10 years but in registration we record the vital events in continuous process of time. Census gives almost all types of informations of a population but registration of birth provides informations on the place of birth, sex, age, religion of the parents, number of previous issues and their sexes, father's occupation etc. whereas the registration of death gives the place of death, sex, age, marital status, number of issues, birth place, occupation and cause of death.

What ever the sources of data of demographic analysis we must be sure of one thing that the data should be reliable, organised and consistent as much as possible to speak in general.

In Bangladesh, secondary demographic data are available mainly from-Bangladesh Bureau of Statistics, Monthly Statistical Bulletin, Village Population Statistics, Union Population Statistics, District Census Report, Bangladesh Population Census Report, Statistical Pocket Book, Statistical Year Book of Bangladesh etc.

16.4 Topics of the Population Study

Now-a-days population study is of great interest to a number of statistician and applied researchers specially demographers. Some common notations of statistical measurements describing conventional rates and ratios and how to compute them are important aspects to cover in advance as much of this ground as possible. Before looking very far into the facts, we must make some inquiry about the nature of the statistics – the method of collection, standard of classification, errors and biases. Life table which illustrates the basic model of most analysis must be well understood in order to make the best use of the existing literature. Measurements of different mortality characteristics are sometimes necessary seperately when the actual life table is not provided. Like that fertility measures are also very useful. Their net effect, the growth or decline in the number of people are sometimes useful to policy makers. The migration of people from one place to another affects the population in both places, since it is associated with very large and rapid changes of population and population structure.It generally receives a prominent share of attention. All these topics mentioned above are to be studied with great attention.

16.5 Concepts on Rates and Ratios

Sometimes absolute figures such as number of births occured last year, how many children of age 5-14 were counted in the last census etc. are considered as the raw material of demographers. More often, they may wish to measure these facts in relation to some other numbers such as the total population at that time. This brings us to rates and ratios which are relative numbers. They provide the measurements of demography by which the behavior in one population can be compared with that of another. Ratios are very easy to calculate. They also provide quick and concise comparisons between many corresponding sets of numbers. The ratios are generally multiplied by 100 or 1000 to make them easier to read and compare with others. This modification does not really change the meaning of the ratios. The term rate is applied to several different things in demography and this sometimes leads to confusion. Some ordinary percentage figures are called rates though they have nothing to do with changes or with vital events.

16.6 Sex Ratio

The ratio between the two numbers is called sex ratio and is calculated as follows :

$$\text{Sex Ratio} = \frac{M}{F} \times K$$

where M is the number of males recorded in some statistical universe of persons, F is the number of females in the same universe and K is arbitrary number usually 100 or 1000.

For certain portion of population the sex ratio is

$$\text{Sex Ratio} = \frac{M_i}{F_i} \text{ x K}$$

where i indicates the ith category or area in the population.

Example 16.1 In a census of a certain country counted 210326 males and 208859 females. Find the sex ratio of that country.

Solution : We know, sex ratio $= \frac{M}{F} \text{ x } 100 = \frac{210326}{208859} \text{ x } 100 = 100.7 \approx 101$

which indicates that according to the census data of that country, there are 101 males per 100 females.

16.7 Child women Ratio

To measure the incidence of childbearing in the population of adult women we consider the number of children under 5 years of age per 1000 women of childbearing age. We compute the following ratio as child women ratio

$$\frac{P_{0-4}}{f_{15-49}} \text{ x K}$$

where P_{0-4} is the number of children of both sexes under 5 years of age, f_{15-49} is the number of females between ages 15 – 49 and K= 1000.

Child-women ratio is useful where there is no adequate registration of birth. It is also used for all groups in the population whose fertility is to be studied.

Example 16.2 In a certain country, the number of children of both sexes under age 5 years is 30,557 and the numbers of females between ages 15-49 is 1057396 as recorded in a registration in the year 1993. Find out the child-women ratio and comment.

Solution : We know that child-women ratio is $\frac{30557}{1057396} \text{ x } 1000 = 28.9$

Comment : Child-women ratio gives us some idea on ordinary birth rate. Therefore, in that country, the ordinary birth rate is about 29 per thousand.

16.8 Population Growth Rate

Changes of total population size over a period of time is an important factor to be studied seriously. The change may increase or decrease depending on the population figures are larger, p_2 and p_1 where p_1 is the total size of population at the first census or date of first counting and p_2 is that of at the second census or latter date of counting. The relative change is measured by the ratio-

$$r = \left(\frac{P_2 - P_1}{P_1}\right) x\ 100 = \left(\frac{P_2}{P_1} - 1\right) x\ 100$$

where r is the rate of growth of population per year and p_1 and p_2 are the population sizes in two consequetive years.

When a quantity such as p_1 is subjected to regular annual increase at a constant increasing or decreasing rate, the relation between the ratio $\frac{P_2}{P_1}$ and the annual rate, $r = \frac{P_2}{P_1} = (1 + r)^n$

where p_1 is the number of population at the initial date, p_2 is the number of population at the latter date, n is the exact number of years between p_1 and p_2 and r is the annual rate of growth.

Note : This rate is also expressed in terms of percentages i.e. multiplying by 100.

Example 16.3 In a country, the total populaion according to the census in 1981 was 4347704 and 5848910 according to the registration in 1997. Find the growth rate and comment.

Solution : Considering the date and month of census period 1981 and 1997 to be same, the exact number of years become (1997-1981) = 16 years.

From the formula be have,

$\frac{P_2}{P_1} = (1 + r)^n$ Taking log of both sides we have,

$$n \log (1 + r) = \log p_2 - \log p_1$$

or $\log (1 + r) = (\log p_2 - \log p_1)/n$

$\therefore \log (1 + r) = (\log 5848910 - \log 4347704)/16 = (6.76707\text{-}6.63826)/16$

$= 0.008051$ Taking anti log, we have $1 + r = 1.01871$

or, $r = 0.01871$

In percentage, $r = 0.01871 \times 100 = 1.87$

Comment : The growth rate of that country is about 2 percent per year.

16.9 Estimation of Mid-year Population

Mid-year population size is usually required in many vital rates such as birth rates and death rates. Actually in all these rates "number of person years" is to be used which is difficult to calculate. As an approximation, the only practical alternative is to take the number of persons present at some moment, generally the middle of the year. If a continuous population registrar may be provided then the mid-year population P is defined as follows :

$$P = p_1 + \frac{1}{2} (p_2 - p_1)$$

where p_1 is the number of population at the begining of the year and p_2 is that of at the end or the begining of the next year.

Example 16.4 From the records of registration of a country, it is found that 12639032 is the population at the begining of the year 1995 and 12974326 is the population at the end of the year. Estimate the mid-year population of 1995.

Solution : Mid-year population of 1995 means the population in the month of July, 1995.

We know, that the estimate of mid-year population

$$P = p_1 + \frac{1}{2} (p_2 - p_1)$$

From the problem ; $p_1 = 12639032$ and $p_2 = 12974326$

therefore, $P = 12639032 + \frac{1}{2}(12974326\text{-}12639032) = 12639032 + 167647$ $= 12806679$

Hence the mid-year population of that country in the year 1995 is 12806679.

Note : More accurately, counting on the number of months between the periods of consideration, the mid-year population of a year can be defined as bellow :

$$P = p_1 + \frac{n}{N}(p_2 - p_1)$$

where P is the mid-year population, p_1 is the size of the population as determined in the first census, p_2 is the size of the population as determined in the second census. N is the number of months between the two censuses and n is the number of months of between the date of p_1 and the date of estimate.

Thus from the above formula we can estimate the mid-year population when the number of months between the periods are more than 12 months.

Example 16.5 : In a certain locality the total population according to census 1941 was 4537407 and 5892602 according to census 1951. Estimate the mid-year population of 1950. (Assuming the date of census in 1941 was April, 1 and for 1951 was November, 30.

Solution : Here $p_1 = 4537407$; $p_2 = 5892602$, $N = 9 + (12 \times 9) + 11 = 128$ months, $n = 9 + (12 \times 8) + 6 = 111$ months, number of months from the date of census of 1941 to July 1, 1950.

$$\text{Now } P = 4537407 + \frac{111}{128}(5892602 - 4537407)$$

$$= 4537407 + 1175208 = 5712615.$$

The mid-year population of 1950 of the locality is 5712615.

16.10 Crude Birth Rate (CBR)

Crude birth rate is a ratio of total registered livebriths to the total population in some specified year and also multiplied by 1000. Symbolically it can be written as follows :

$$\text{CBR} = \frac{B}{P} \text{ x K}$$

where B is the total livebirths registered during the calender year (Jan.1 to Dec. 31) P is the mid-year population of that calender year (July.1) and K is 1000.

Note : The notation of "livebirth" is to some extent difficult to define exactly. In 1950 international agencies defined livebirth as follows :

Livebirth is the complete expulsion or extraction from its mother of a product of conception irrespective of the duration of pregnancy, which after such seperation breaths and shows different evidence of life.

16.11 Crude Death Rate (CDR)

Crude death rate is a ratio of the total registered deaths of some specified year to the total population multiplied by 1000. Symbolically it can be written as follows :

$$\text{CDR} = \frac{D}{P} \text{ x K}$$

where D is the total number of registered deaths during the calender year (Jan. 1 to Dec. 31), P is the mid-year population of that calender year (July, 1) and K is 1000.

Crude Rate of Natural Increase : Crude rate of natural increase may be defined as the gap between the crude birth rate and crude death rate, Symbolically it can be written as ;

$$\text{CRNI} = \frac{B}{P} \text{ x K} - \frac{D}{P} \text{ x K} = \left(\frac{B-D}{P}\right) \text{x K}$$

where B is the total livebirths registered during the calender year (Jan1. to Dec 31.), D is the total number of registered deaths during the calender year (Jan 1 o Dec 31), P is the mid-year population of that calender year and K is 1000.

By crude rate of natural increase, we can measure the change in population size per 1000 due to the differences in the numbers of births and deaths in a given year. The change may be due to the loss or gain. If the number of births are greater than the number of deaths then we have a gain change otherwise we have loss change. This rate represents only the change due to the balance of births and deaths and does not include the effect of migration.

Example 16.6 The yearly total number of live birth is 150791, total number of registered deaths is 125803 and the mid-year population is 6657339 are recorded in a country in 1981. Calculate crude birth rate, crude death rate and crude rate of natural increase.

Solution : Here we have, B = 150791, D = 125803 and P = 6657339

$$\text{therefore, CBR} = \frac{150791}{6657339} \times 1000 = 22.7\,; \quad \text{CDR} = \frac{125803}{6657339} \times 1000 = 18.9$$

$$\text{CRNI} = \left(\frac{B-D}{P}\right) x K = \left(\frac{150791-125803}{6657339}\right) x\ 1000 = 3.8$$

Note : When we have CBR and CDR, the CRNI can be obtained by taking the simple difference of the two values. Actually in that case, CRNI = CBR - CDR. In the above problem, CRNI = 22.7 - 18.9 = 3.8 which is same as we have calculated independently.

16.12 Average Crude Birth Rate

Vital rates are most often calculated based on the data of census year because census provides a convenient and reliable figure for the total population. A rate based on just one year may be affected by fluctuations in the numbers of births and deaths. This drawback may be avoided by taking average of a number of yearly data centred on the census year. Therefore, crude birth rate taking three years may be calculated as follows :

$$\text{CBR (average)} = \frac{1}{3}\left(\frac{B_1}{P_1} + \frac{B_2}{P_2} + \frac{B_3}{P_3}\right) x\ 1000$$

where B_1 is the total number of livebirths registered during the previous year of census year, B_2 is the total number of livebirths registered during the census year. B_3 is the total number of livebirths registered during one year after the census year. P_2, P_1 and P_3 are the mid-year population corresponding to the census year, one year previous and one year after the census year.

Average Crude Death Rate : Considering the above reasons stated in case of average crude birth rate, the average crude death rates taking three consequitive years may be calculated as follows :

$$\text{CDR (average)} = \frac{1}{3}\left(\frac{D_1}{P_1} + \frac{D_2}{P_2} + \frac{D_3}{P_3}\right) \times 1000$$

where D_2 , D_1 and D_3 are the total number of deaths corresponding to census year, one previous year and one year after the census year. P_2, P_1 and P_3 are the mid-year population corresponding to census year, one previous year and one year after the census year.

Example 16.7 The total number of livebirths and death as well as the mid year population of a certain locality in three consequitive years taking census year at the middle are given below :

Year	Mid-year Population	Total No. of livebirths	Total No. of deaths
1950	34734133	4714001	4614500
1951	35891305	4303263	4228156
1952	37019126	3876774	3808473

Calculate average crude birth rate and average crude death rate.

Solution : According to the formula-

$$\text{CBR (average)} = \frac{1}{3}\left(\frac{4714001}{34734133} + \frac{4303263}{35891305} + \frac{3876774}{37019126}\right) \times 1000$$

$$= \frac{1}{3}(0.13572 + 0.11990 + 0.10472) \times 1000 = 120.11$$

Again,

$$\text{CDR (average)} = \frac{1}{3}\left(\frac{4614500}{34734133} + \frac{4228156}{35891305} + \frac{3808473}{37019126}\right) \times 1000$$

$$= \frac{1}{3}(0.13285 + 0.11780 + 0.10288) \times 1000 = 117.84$$

16.13 Monthly Crude Birth Rate, Monthly Crude Death Rate

Through registration we have monthly data of births and deaths of the year, therefore, monthly crude birth rate or monthly crude death rate can be calculated to show the monthly or seasonal pattern of variation

Monthly crude birth rate may be calculated as follows :

$$\text{CBR (monthly)} = \frac{365}{n} \times \frac{B}{P} \times K$$

where B is the total number of livebirths in the month, P is the mid-month population, n is the number of days in the month, 365 is the total number of days in a year and 366 is the total number of days in a leap year and K is 1000.

To show the monthly pattern of variation among death rates the monthly crude death rate may be calculated. The formula for the calculation of monthly crude death rate is as follows :

$$\text{CDR (monthly)} = \frac{365}{n} \times \frac{D}{P} \times K$$

where D is the total number of deaths in the month, P is the mid-month population, n is the number of days in the month, 365 is the total number of days in a year and 366 is the total number of days in a leap year and K is 1000.

16.14 Age Specific Birth Rate (ASBR)

It can be observed that the frequency of child birth varies markedly with age. The age which is to be specified is not that of the child being born but the age of each parents at the time of birth. To form a better idea as to the fertility situation obtaining in a community, it is necessary to compute a fertility rate for each group of either mother or father. Age specific birth rates are usually calculated for women rather than men.

They are computed as follows :

$$\text{ASBR} = \frac{b_i}{p_i} \text{ x K}$$

where b_i is the number of livebirths registered during the year to the women in the age interval i (usually an interval of 5 years); p_i is the mid-year population of women having ith age group and K is 1000.

For calculation of the above rate we require the population of women bearing the age groups 15-49. Some adjustments are necessary if there are some births registered for women below age 15 and over age 49. These two cases are transfered to the categories 15-19 and 45-49 respectively.

Example 16.8 Average yearly livebirths at different age groups along with female population in census data of 1950 are given below of a certain country, calculate the age specific birth rates.

Age groups	Average yearly livebirth	Female Population
15-19	56558	4229005
20--24	626240	3870468
25-29	809727	3341590
30-34	515268	2825769
35-39	291728	2657741
40-44	86238	1273441
45-49	4848	1078362

Solution : According to the formula the age specific birth rates are furnished in the following table :

Table-16.1

Age Group (1)	Average Yearly livebirths (b_i) (2)	Female Population (p_i) (3)	Age Specifice birth rates (2)/(3) x 1000
15-19	56558	4229005	13.4
20-24	626240	3870468	161.8
25-29	809727	3341590	242.3
30-34	515268	2825769	182.3
35-39	291728	2657741	109.8
40-44	86238	1273441	67.7
45-49	4848	1078362	4.5

16.15 General Fertility Rate (GFR)

This is another ratio of just two numbers, each from different universes. It is usually called general fertility rate which is the ratio of all livebirth to the number of women of child bearing age i.e. age between 15 and 49. Thus general fertility rate can be defined as below-

$$GFR = \frac{B}{P^F_{15-49}} x K$$

where B is the number of livebirths registered during the year, P^F_{15-49} is the mid-year female population between age 15 and 49 and K is 1000.

It is refered to as a ratio of births per 1000 women of child bearing age. This places it somewhere in between a crude birth rate and an age specific birth rate because it attributes all birth to all women within these age limits without further distinction among them. The general fertility ratio is the number of births per 1000 person-years lived by the population of adult women those who form the universe of potential mothers. The chief advantages of the general fertility ratio is that it can be computed in situations in which the registration of births and the enumeration of population are satisfactory, but where direct evidence of births by age of parents are lacking.

16.16 Total Fertility Rate (TFR)

The general fertility ratio is so affected independently of actual changes or differences in birth rates at each age. But a whole series of seven birth rates corresponding to 15-19, 20-24, 25-29, 30-34, 35-39, 40-44 and 45-49 are sometimes too long and cumbersome to be useful. Therefore, they can be condensed by combinding them into a single rate which is usually called the total fertility rate defined by -

$$\text{TFR} = \sum_{i=15}^{49} \frac{b_i(1)}{p_i} \times k$$

where b_i (1) is the number of livebirths registered during the year to mothers of age i where i is an interval of one year, p_i is the mid-year population of women on the same age and K is 1000 or sometimes 1.

In practice, TFR is calculated in some different way because birth rates in single year intervals of age are not usually calculated. The birth rate for each 5 years age group is an average for the groups. This is not quite the same as the sum of the rates for all five single years. It is a close approximation to multiply the birth rates for each groups of women by 5. Thus TFR can also be calculated as follows :

$$\text{TFR} = 5 \sum_{i=1}^{7} \frac{b_i(5)}{p_i} \times k$$

where $b_i(5)$ is the number of livebirths registered during the year to mothers of age i where i is an interval of 5 years started from 15-19, 20-24 and so on up to 44-49, p_i is the mid-year population of women of the same age group and K is sometimes 1000 or sometimes 1.

Example 16.9 Calculate the total fertility rate (TFR) of the data given in Example 16.8.

Solution : We calculate the Age Specific Birthrate (ASBR) as given in the solution of Example 16.8. The age groups and the corresponding ASBRs are shown in the following table

Table 16.2

Age Group	ASBR (Births/1000 women)
15-19	13.4
20-24	161.8
25-29	242.3
30-34	182.3
35-39	109.8
40-44	67.7
45-49	4.5

Therefore, TFR = 5 (13.4 + 161.8 + ... + 67.7 + 4.5)
= 5 x 781.8 = 3909

Note : The TFR is also the same as the total number of children that would ever be born to a group of women, if the groups passed through the reproductive span of life with these birth rates at each year of age. Here the reproductive span really means the age groups in which an women can bear child i.e. from age 15 to age 49.

16.17 Gross Reproduction Rate (GRR)

The gross reproduction rate is a slight modification of the total fertility rate, It is computed as follows :

$$\text{GRR} = 5 \sum_{i=1}^{7} \frac{b^{f}_{i(5)}}{p_i} \times k$$

where $b_{i(5)}^{f}$ is the number of livefemalebirths registered during the years to mothers of age i where i is an interval of 5 years starting from 15-19 and so on, p_i is the mid-year population of women of the same age-group and K is sometimes 1000 or sometimes 1.

Example 16.10 In a certain country the following data are registered in a census year 1971. Calculate the GRR and Comment.

Age gr.	Average yearly Female livebirth in census 1971	Female Population in the census yr. 1971
15-19	28279	5338115
20-24	313120	4880568
25-29	404864	4451602
30-34	257634	3836870
35-39	145864	3668752
40-44	43119	2684552
45-49	2424	1089463

Solution : For calculating GRR the following table is usually prepared.

Table 16.3

Age gr.	Ave. Yrly. Fem. births $b_{i(5)}^{f}$	Fem. Pop^n P_i	$\frac{b_{i(5)}^{f} \times 1000}{P_i}$
15-19	28279	5338115	5.30
20-24	313120	4880568	64.16
25-29	404864	4451602	90.95
30-34	257634	3836870	67.15
35-39	145864	3668752	39.76
40-44	43119	2684552	16.06
45-49	2424	1089463	2.22
Total			285.6

Therefore, GRR = 5 x 285.6 = 1428

Comment : GRR = 1428 means the number of daughters per 1000 women having reproductive capacity.

Note : If GRR is greater than unity then the population would increase no matter how high the death rates may be. Male gross reproduction rate can also be calculated on the basis of the male births rather than the female births and also the population of male having reproductive age period.

16.18 Net Reproduction Rate (NRR)

The main disadvantage of the GRR is that it does not consider the fact that some of the females who are assumed to begin life together may die before reaching age 15, some may die between ages 15 and 16 and so on. In simple words, GRR takes into account of current fertility only but ignores current mortality.

Therefore, the net reproduction rate consists of a hypothetical cohort of women, their deaths and all their female births during the childbearing period. Thus net reproduction rate (NRR) may be defined in symbols as follows :

$$NRR = \sum_{x=15}^{49} b_x \left(\frac{Lx}{l_o}\right) x\ k$$

where b_x is the rate of female birth per person year at each age x, Lx is the number of years lived by the persons who have reached each age before they reach the next higher age and l_o is the number of females in the cohort started at age 0, K is sometimes 1000 or sometimes 1.

If the age groups are made in 5 years, then

$$NRR = 5 \times \sum_{x=1}^{7} b_x \left(\frac{Lx}{l_o}\right) x\ k$$

where b_x, L_x, l_o and K are defined as above considering the age interval of 5 years.

16.19 Age Specific Death Rate (ASDR)

In crude vital rates, there is always some diversity hidden in it because they apply to everyone in the population. But if the population is grouped into groups whose members are to some extent homogeneous in nature, rates computed for each group will be less affected by diversity among individuals.

Thus crude death rate is modified by age specific death rate and can be defined in symbols as follows :

$$\text{ASDR} = \frac{d_i}{p_i} \times K$$

where d_i is the number of deaths during the year in the ith age group, p_i is the mid-year population of the same age group and K is 1000.

If the age groups are made of 5 years of age then

$$\text{ASDR} = \frac{d_i(5)}{p_i\ (5)} \times K$$

where d_i (5) is the number of deaths during the year in the ith age group of 5 years, p_i (5) is the mid-year population of the same age group of 5 years and K is 1000

Age specific death rates can be calculated for both male and female seperately and thus the formula of ASDR for men is as follows :

$$\text{ASDR (Men)} = \frac{d_i(m)}{p_i(m)} \times K$$

where $d_i(m)$ is the number of deaths of men during the year in the ith age group, $p_i(m)$ is the mid-year population of men of the same age group and K is 1000.

Similarly ASDR for women can be defined as follows :

$$\text{ASDR (Women)} = \frac{d_i(w)}{p_i(w)} \times K$$

where $d_i(w)$ is the number of deaths of women during the year in the ith age group, p_i (w) is the mid-year population of women in the same age group and K is 1000.

Example 16.11 Calculate the age specific death rates from the following data of both sexes of the population of a certain country during the year 1945-1947.

Age	Mid-year Popn. of 1946	Average Yrly death
1-4	700762	20683
5-9	811363	5451
10-14	805642	2589
15-19	680614	3345
20-24	641571	5104
25-34	1027405	9305
35-44	790514	8775
45-54	515695	8209
55-64	293598	8075
65 and above	229498	21958
All ages	6657339	125803

Solution : For calculation of Age Specific Death Rates, following table is presented below :

Table 16.4

Age Groups	Mid-year Popn. of 1946 (p_i)	Average yearly Deaths (1945-47) (d_i)**	Death Rates Death per 1000 (d_i/p_i) x 1000
1-4	700762	20683	29.5
5-9	811363	5451	6.7
10-14	805642	2589	3.2
15-19	680614	3345	4.9
20-24	641571	5104	8.0
25-34	1027405	9305	9.1
35-44	790514	8775	11.1
45-54	515695	8209	15.9
55-64	293598	8075	27.5
65 and above	229498	21958	97.7
* All Ages	6657339 (p)	125803 (d)	18.9

* Gives us the Crude Death Rate = $\frac{d}{p} \times 1000 = \frac{125803}{6657339} \times 1000 = 18.9$

** Average yearly deaths = $\frac{\text{Registered Deaths of 1945.1946. 1947}}{3}$

16.20 Standardised Death Rate (S_tDR)

Using the formulas of crude death rates and age specific death rates for two region say A and B we have the following relations :

$$CDR\ (A) = m\ (A) = \frac{d(A)}{p(A)} \times 1000 = \frac{\Sigma\ mi\ (A)\ P_i(A)}{\Sigma p_i(A)} \text{ and}$$

$$CDR\ (B) = m\ (B) = \frac{d(B)}{p(B)} \times 1000 = \frac{\Sigma m_i\ (B)\ p_i(B)}{\Sigma p_i\ (B)}$$

The above two expressions are nothing but the weighted arithmetic mean of the ASDRs, the weights being the corresponding mid-year population in the age groups. It is observed that even if ASDR s are same i. e. $m_i(A) = m_i(B)$ for all age groups i, $m_{(A)} \neq m_{(B)}$.

$$\text{Since in general, } \frac{p_i\ (A)}{\Sigma p_i\ (A)} \neq \frac{p_i\ (B)}{\Sigma p_i(B)}$$

Since the age distribution of the populations in the two regions A and B are not identical. This drawback can be removed if we use same set of weights. This is usually done by standardised death rate. (S_tDR) which can be written in symbols as follows :

Standardised death rate for the region A is

$$\therefore\ (S_tDR)A = \frac{\Sigma m_i(A)\ p_{i(S)}}{\Sigma p_{i(S)}}.$$

where, m_i (A) is the death rate per 1000 for ith age group of region A, $p_{i(S)}$ is the mid-year population of the same age-group of a standard region S which is usually given and the standardised death rate for the region B is

$$(S_t\ DR)_B = \frac{\Sigma m_i\ (B)\ p_{i(S)}}{\Sigma p_{i(S)}}.$$

where $m_{i(B)}$ is the death rate per 1000 for the ith age group of region B and $p_{i(S)}$ is defind as above.

These standardised death rates for regions A and B respectively are nothing but the crude death rates that would be observed in the standard population if it were subjected to the age specific death rates of the regions A and B.

Example 16.12 Compute the crude and standardised death rates of the two populations A and B, regarding A as standard population from the following data and comment.

Age group (years)	A		B	
	Population	Deaths	Population	Deaths
Under age 10	18000	500	10000	272
10-20	10000	140	28000	560
20-40	48000	1150	60000	1512
40-60	28000	950	13000	425
60 and above	8000	400	1000	80

Solution : For computation of crude death rate and standardised death rate, the following table is prepared :

Table-16.5

	Population-A			**Population-B**			
Age Groups (Years)	Mid-year Population $p_i(A)$	Deaths $d_i(A)$	Death Rate Per 1000 $m_i(A)$	Mid-year Population $p_i(B)$	Deaths $d_i(B)$	Death Rate per 1000 $m_i(B)$	Pi(A) x pi(B)
Under age 10	18000	500	27.78	10000	272	27.20	489600
10-20	10000	140	14.00	28000	560	20.00	200000
20-40	48000	1150	23.96	60000	1512	25.20	1209600
40-60	28000	950	33.93	13000	425	32.69	915320
60 &above	8000	400	50.00	1000	80	80.00	640000
Total	112000	3140		112000	2849		3454520

Crude Death Rates :

$$\text{CDR of Population A} = \frac{\Sigma d_i(A)}{\Sigma p_i(A)} \times 1000 = \frac{3140}{112000} \times 1000 = 28.04$$

$$\text{CDR of Population B} = \frac{\Sigma d_i(B)}{\Sigma p_i(B)} \times 1000 = \frac{2849}{112000} \times 1000 = 25.44$$

Standardised Death Rates :

Since the population A is considered as standard population

S_t DR (A) = CDR for the population A = 28.04

$$S_t\,DR(B) = \frac{\Sigma p_{i(A)} \times m_{i(B)}}{\Sigma p_{i(B)}} = \frac{3454520}{112000} = 30.84$$

Comment : We may therefore, conclude that the crude death rate of A is greater than that of population B. But the standardised death rate of population B is greater than the crude death rate of population A.

16.21 Infant Death Rate (IDR)

In demography, infants are usually defined as an exact age groups namely age "Zero" or those children in the first year of life who have not yet reached age one.

The infant death rate (IDR) is the ratio of registered deaths of infants during a year to livebirths registered during the same year. Symbolically IDR can be expressed as

$$\text{IDR} = \frac{d_o}{B} \times K$$

where d_o is the number of deaths below age one registered during the year, B is the number of livebirths registered during the same year and K is 1000.

Note : The infant death rate computed for both sexes together is often used for a brief and overall view of conditions or trends. A more detailed study of infant mortality requires seperate rates for both males and females, since infant mortality differs markedly by sex.

Example 16.13 The report on vital statistics of a certain country stated that in 1990 a total of livebirths were 238494 of which the number of deaths among infants were 33309. Find out the infant death rate and comment.

Solution : Here it is stated that in the country total number of livebirths registered during the year 1990 is B = 238494 and the number of deaths below age one registered during the year is d_o = 33309

Therefore, $\text{IDR} = \frac{d_o}{B} \times 1000 = \frac{33309}{238494} \times 1000 = 139.7$

Comment : The infant mortality rate of that country is about 140 per 1000 births.

16.22 Life Table

The life table is a life history of a hypothetical group or cohort of people. It is usually diminished gradually by death. The record starts at the birth of each numbers of the cohort and continue untill all have died. The cohort loses a predetermined proportion at each age. Mortality is traditionally represented by the life table.

Assumptions of the life tables : A few simplifying assumptions of the life table are as follows :

(i) The cohort is closed against migration in or out. Hence there are no changes in membership except the losses due to death.

(ii) People die at each age according to a schedule that is fixed in advance and does not change.

(iii) The cohort starts from some standard number of births (figure like 1,000 or 10,000) usually called the radix of the life table.

(iv) At each age deaths are evenly distributed between one birthday and the next.

(v) The cohort normally contains number of only one sex. Thus life table for male or female is usually constructed according to the necessity.

Description of a life Table : A typical life table has generally the following 8 columns :

1	2	3	4	5	6	7	8
x	l_x	d_x	q_x	p_x	L_x	T_x	e_x

Under the above assumptions, the descriptions and relationship between the columns are as follows:

1. In the first column, x indicates the age which usually starts from 0, 1, 2 .... and so on.

2. In the second column, l_x indicates the survivors at exact age x. This column values are not usually given but can be obtained from the knowledge of q_x & d_x etc.

3. In the third column, d_x indicates the number of deaths between age x and age x+1 i. e.

$dx = l_x - l_{x+1} \Rightarrow l_x = d_x + l_{x+1}$ also $l_{x+1} = l_x - dx$

4. In the fourth column, q_x indicates the probability of dying between age x and age x+1 i.e.

$$q_x = \frac{d_x}{l_x} = \frac{l_x - l_{x+1}}{l_x} = 1 - \frac{l_{x+1}}{l_x} = 1 - p_x$$

where in the fifth column, p_x indicates the probability of surviving between age x and age x+1 i.e.

$$p_x = 1 - q_x = 1 - \frac{d_x}{l_x} = \frac{l_x - d_x}{l_x} = \frac{l_{x+1}}{l_x}$$

5. In the sixth column, L_x indicates the years lived between age x and x + 1.

$$L_x = l_x - \frac{1}{2} d_x = l_x - \frac{1}{2} (l_x - l_{x+1}) = \frac{1}{2} (l_x + l_{x+1}).$$

For the begining of life i.e. at age o or 1 the estimate of L_x at the mid point between l_x and l_{x+1} would be too high. The following is often a good approximation:

$L_0 = 0.3l_o + 0.7l_1$

$L_1 = 0.4l_1 + 0.6l_2$

$L_2 = 0.5\ l_2 + 0.5\, l_3 = \frac{1}{2}(l_2 + l_3)$ and so on.

6. T_x indicates the total years lived after exact age x i.e.

$T_x = L_x + L_{x+1} + L_{x+2}$ .... $\Rightarrow T_{x+1} = T_x - L_x$

7. e_x indicates the expectation of life or the average number of years lived after exact age x.

$$e_x = \frac{T_x}{l_x}$$

Construction of life table : On the basis of above relationship and the relevant data provided, we can construct a life table. A sample of a life table construction is presented as below:

Here a complete life table is presented in which the age interval is one year through out the table of data of a certain area of a country of males-

Table—16.6

Age x (1)	Survivors at age x l_x (2)	No. of deaths between age x and x + 1 d_x (3)	Prob. of dying at age x q_x (4)	Prob. of survivors at age x p_x (5)	Years lived between age x and x +1 L_x (6)	Total years lived after exact age x T_x (7)	Expectation of life e_x (8)
0	10,000	1885	.1885	.8115	8680.5	399955.9	39.99
1	8115	491	.0605	.9395	7820.4	391275.4	48.22
2	7624	197	.0258	.9742	7525.5	383655.0	50.32
3	7427	91	.0123	.9877	7381.5	376129.5	50.64
4	7336	55	.0075	.9925	7308.5	368748.0	50.27
5	7281	37	.0051	.9949	7262.5	361439.5	49.64
6	7255	30	.0041	.9965	7229.0	354177.0	48.89
7	7214	25	.0035	.9965	7201.5	346948.0	48.09
8	7189	17	.0024	.9976	7180.5	339746.5	47.26
9	7172	14	.0020	.9980	7165.0	332566.0	46.37

Age	Survivors at age x	No. of deaths between age x and x + 1	Prob. of dying at age x	Prob. of survivors at age x	Years lived between age x and x +1	Total years lived after exact age x	Expectation of life
x	l_x	d_x	q_x	p_x	L_x	T_x	e_x
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
10	7158	15	.0021	.9979	7150.5	325401.0	45.46
11	7143	18	.0025	.9975	7134.0	318350.5	44.57
12	7125	22	.0031	.9969	7114.0	311216.5	43.68
13	7103	28	.0039	.9961	7089.0	304102.5	42.81
14	7075	33	.0047	.9953	7058.5	297013.5	41.98
15	7042	39	.0055	.9945	7022.5	289955.0	41.18
16	7003	42	.0060	.9940	6982.0	282932.5	40.40
17	6961	47	.0068	.9932	6937.5	275950.5	39.64
18	6914	51	.0074	.9926	6888.5	269013.0	38.91
19	6863	54	.0079	.9921	6836.0	262124.5	38.19
20	6809	57	.0084	.9916	6780.5	255288.5	37.49
21	6752	59	.0087	.9913	6722.5	248508.0	36.81
22	6693	60	.0090	.9910	6663.0	241785.5	36.13
23	6633	60	.0091	.9909	6603.0	235122.5	35.48
24	6573	61	.0093	.9907	6542.5	228519.5	34.77
25	6512	60	.0092	.9908	6482.0	221977.0	34.09
26	6452	60	.0093	.9907	6422.0	215495.0	33.40
27	6392	60	.0094	.9906	6362.0	209073.0	32.71
28	6332	59	.0093	.9907	6302.5	202711.0	32.01
29	6273	59	.0094	.9906	6243.5	196408.5	31.13
30	6214	59	.0095	.9905	6184.5	190165.0	30.60
31	6155	58	.0094	.9906	6126.0	183980.5	29.89
32	6097	59	.0097	.9903	6067.5	177854.5	29.17

Age	Survivors at age x	No. of deaths between age x and x + 1	Prob. of dying at age x	Prob. of survivors at age x	Years lived between age x and x +1	Total years lived after exact age x	Expectation of life
x	l_x	d_x	q_x	p_x	L_x	T_x	e_x
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
33	6038	59	.0098	.9902	6008.5	171787.0	28.45
34	5979	58	.0097	.9903	5950.0	165778.5	27.73
35	5921	60	.0101	.9899	5891.0	159828.5	26.99
36	5861	59	.0101	.9899	5831.5	153937.5	26.26
37	5802	60	.0103	.9897	5772.0	148106.0	25.53
38	5742	60	.0104	.9896	5712.0	142334.0	24.79
39	5682	61	.0107	.9893	5651.5	136622.0	24.04
40	5621	62	.0113	.9887	5590.0	130970.5	23.30
41	5559	62	.0112	.9888	5528.0	125380.5	22.55
42	5497	62	.0113	.9887	5466.0	119852.5	21.80
43	5435	63	.0116	.9884	5403.5	114386.5	21.05
44	5372	64	.0119	.9881	5340.0	108983.0	20.29
45	5308	64	.0121	.9879	5276.0	103643.0	19.53
46	5244	64	.0122	.9878	5212.0	98367.0	18.76
47	5372	66	.0127	.9873	5147	93155.0	14.98
48	5114	67	.0131	.9869	5080.5	88008.0	17.21
49.	5047	69	.0137	.9863	5012.5	82927.5	16.43
50	4978	71	.0143	.9857	4942.5	77915.0	15.65
51	4907	74	.0151	.9849	4870.0	72972.5	14.87
52	4883	78	.0161	.9839	4794.0	68102.5	14.09
53	4755	81	.0170	.9830	4714.5	63308.5	13.31
54	4674	85	.0182	.9818	4631.5	58594.0	12.54
55	4589	89	.0194	.9806	4544.5	53962.5	11.76

Age	Survivors at age x	No. of deaths between age x and x + 1	Prob. of dying at age x	Prob. of survivors at age x	Years lived between age x and x +1	Total years lived after exact age x	Expectation of life
x	l_x	d_x	q_x	p_x	L_x	T_x	e_x
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
56	4500	94	.0204	.9791	4453.0	49418.0	10.98
57	4406	96	.0218	.9782	4358.0	44965.0	10.20
58	4310	100	.0232	.9768	4260.0	40607.0	9.42
59	4210	103	.0245	.9755	4158.5	36347.0	8.63
60	4107	106	.0258	.9742	4054.0	32188.5	7.84
61	4001	110	.0275	.9725	3946.0	28134.5	7.03
62	3891	116	.0298	.9702	3833.0	24188.5	6.21
63	3775	122	.0323	.9677	3714.0	20355.5	5.39
64	3653	128	.0350	.9650	3589.0	16641.5	4.56
65	3525	136	.0386	.9614	3457.0	13052.5	3.70
66	3389	141	.0416	.9584	3318.5	9595.5	2.83
67	3248	146	.0450	.9550	3175.0	6277.0	1.93
68	3102	–	–	–	3102.0	–	–
:							
:							
:							

Abridged Life Table : The life table constructed from the data in which the age interval is one year throughout the table is called a complete life table. On the otherhand, in abridged life table the age interval is usually 5 years or 10 years.

A typical abridged life table consist of the following columns :

1) x to (x+n) exact age interval; n may be 5 or 10 years.

2) l_x, the number of persons out of a cohort of l_o persons, living at the begining of the interval x to (x+n)

3) ${}_nq_x$, the probability of the persons dying in the age interval x to (x + n) and is given by-

$$ {}_nq_x = \frac{l_x - l_{x+n}}{l_x} = 1 - \frac{l_{x+n}}{l_x} = 1 - {}_np_x \quad \text{where } {}_np_x = \frac{l_{x+n}}{l_x} $$

4) ${}_nd_x$, the number of deaths in the interval x to (x+n) and in given by-

$$ {}_nq_x = \frac{{}_nd_x}{l_x} \Rightarrow {}_nd_x = {}_nq_x \times l_x $$

5) ${}_nL_x$; the number of members of the life table of stationary population in the age group (x, x + n) and is given by-

$$ {}_nL_x = \int_0^n l_{x+t}\, dt $$

6) $Tx = \int_0^{\infty} L_{x+t}\, dt$ is the number of persons lived after age x.

7) $e_x = \frac{T_x}{l_x}$, complete expectation of life at age x.

Uses of Life Table : The basic objective of life table is to give a clear picture of the age distribution of mortality, in a given group of population. In the words of William Farr life table is a Berometer of the populations.

Some of the most important uses and applications of life table are mentioned as below :

1) Life tables provide the actuarial science with a sound foundation, converting the insurance business from a mare gambling in human lives to the ability to offer well calculated safe guard in the event of death. It also helps to fix up the reasonable rate of primium at the time of opening an insurance.

2) Life tables are used by demographers to devise measures such as Net Reproduction Rate to study the rate of growth of population.

3) Life tables for two or more different groups of population may be used for the relative comparison of various measures of mortality such as death rate, expection of life at various ages etc.

4) Life tables are as well used-

i) by the government and the private organisation for determining the rates of retirement benifits to be given to its employees,

ii) for predicting the school going population in connection with school building programmes,

iii) for estimating the probable number of future widows and orphans in a community,

iv) for computing the approximate size of future labour forces in different levels.

16.23 Stationary Population

Stationary population is a hypothetical cohort of population determined by each age by the fixed mortality rates. This population is closed against migration and therefore, is maintained solely by a balance of births and deaths. This balance is fixed and so the size of the population is stationary.

Finally, stationary population may be defined as the hypothetical model of a population based on unchanging condition of fertility, mortality and total size.

16.24 Stable Population

Stable population is a bit less restricted type of stationary population where independent schedule of fertility and mortality are permitted. That is why, some increament or decreament of population is possible in this case, Some of the characteritics of stationary population are present in this type of population but can be derived by a more complex procedure.

17. NON-PARAMETRIC TESTS

17.1 Introduction

In parametric statistical test, the null hypothesis usually specifies certain conditions about the parameters of the population from which the sample is drawn. For example, $H_0 : \mu = \mu_0$ is a null hypothesis in a parametric statistical test. The main assumption here is that the sample has been independently distributed as normal with mean μ and variance σ^2 (say).

But in non-parametric statistical test the null hypothesis does not specify conditions about the parameters of the population from which the sample is drawn, it usually specifies the form of the distribution. For example, H_0 : The sample data fit the Binomial distribution. Here the knowledge about the parameters of the distribution may or may not be known. The usual assumptions are (i) the obsurvations are independent (ii) the variable under study has continuity. These assumptions are weaker and fewer than the conditions required in a parametric test.

By convention, there are two main types of non-parametric statistical procedures (i) Truely non-parametric procedures which are valid under very general assumptions, (ii) Distribution free procedures whose validity usually does not depend on the functional form of the population.

Advantages of Non-parametric Statistical Test :

(i) Probability statement obtained from this type of test is exact. The accuracy of the probability statement does not depend on the shape of the population from which the sample is drawn.

(ii) In the type of test, the sample size is small.

(iii) Since most non-parametric tests depend on a minimum number of assumptions, so the chance of their improper use is minimum.

(iv) The calculation of these test procedures are quick and easy, so minimum knowledge of mathematics and statistics are required.

(v) These procedures are applied when the data are measured on weak measurement scale.

Disadvantages :

(i) If all assumptions in the data are met, then the use of non-parametric tests waste information.

(ii) There is no non-parametric test for any interaction effects in the analysis of variance.

(iii) The arithmetical calculations of these test procedures are simple but sometimes very labourious and tedious.

(iv) For conclusions, a large number of different types of tables are required (see Appendix-2)

17.2. Measurement Scale

The numerical data used in non-parametric statistical test are with the nature of scale. A statistician follows the views on measurement and measurement scale. Here we shall discuss four level of measurement scale namely (i) Nominal scale (ii) Ordinal Scale (iii) Interval Scale (iv) Ratio Scale.

(i) Nominal Scale : It is the weakest of the four measurement scales. It differentiates one object or event from the other on the basis of a name. For example, the item produced in a factory may be classified as defective or non-defective. A new born baby may be male or female. Frequently we use arbitrary numbers rather than the name to distinguish among the objects or events on the basis of some characteristics. For example-Defective as 0 and non-defective as 1 ; son as 1 and daughter as 0 etc.

In a nominal scale, the scaling operation is partitioning a given class into a set a mutually exclusive sub-classes. The only relation involved is that of equivalance. The equivalance relation is reflexive, symmetrical and transitive.

Reflexive : $x = x$ for all values of x.

Symmetrical : If $x = y$ then $y = x$.

Transitive : If $x = y$ and $y = z$ then $x = z$.

Nominal data are frequently refered to as count data or frequency data.

(ii) Ordinal Scale : It may happen that the objects or items in one category of a scale are not just different from the objects in other category of that scale, but that they stand in some kind of relation to them on the basis of the relative amounts of some characteristics they possess.

Ordinal scale makes it possible for objects to be ranked. For example, sales man can be ranked from the poorest, medium, the best on the basis of their personalities. Beauty contestants can be ranked from the least beautiful, 'medium beautiful', the most beautiful. Sometimes the ranks may have the numbers 1, 2, 3........, etc. For example, the runners in a competition may be ranked 1, 2, 3,.......according to the order in which they cross the finishing line.

The relation among the ranked values may be designated by > (greater than) to identify preferred to, is higher than, is more difficult than etc. For example, in social acceptibility, upper class members are higher than (>) all members of the lower middle class. The equivalence relation (=) holds among the members of the same class.

The greaters than (>) relation is irreflexive, asymmetrical and transitive.

Irreflexive : It is not true for any x that $x > x$.

Asymmerical : If $x > y$ then $y \ngtr x$.

Transitive : If $x > y$ and $y > z$ then $x > z$.

(iii) Interval Scale : When a scale has all the characteristics of an ordinal scale and when in addition the distance between any two members on the scale are of known size then the measurements are called interval scale. An interval scale is characterized by a comment and constant unit of measurement which assigns a real number to all pairs of objects in the ordered set.

Suppose for example, that four objects A, B, C & D are assigned scores of 20, 30, 60 and 70 respectively where measurements are on the interval scale. We can say that the difference between 20 and 30 is equal to the difference between 60 and 70. That is equal distance between the numbers of each two pairs of scores indicate equal differences in the amount of the trait being measured.

(iv) Ratio Scale : When a scale has all the characteristics of an interval scale and in addition has a true zero point at the origin, it is called a ratio scale. In a ratio scale, the ratio of any two scale points is independent of the units of measurements. The familiar measurements of height and weight are examples of measurement of ratio scale. We can easily say that a person who weighs 90 kg weighs 30 kg more than a person who weighs 60 kg. With a ratio scale we can also say that a 90 kg person weighs twice as much as a 45 kg person. The ratio scale represents the highest level of measurement.

General Comments on Measurement Scales : Among the four measurement scales—nominal, ordinal, interval and ratio scales ; nominal and ordinal measurement scales are the most common types achieved in behavioural science. Data measured by either nominal or ordinal scale should be analysed by the non-parametric tests. Data measured in interval or ratio scales may be analysed by parametric tests.

17.3 Some Non-parametric Tests that Use Data from a Single Sample

In this section, we present some important non-parametric tests that utilize data form a single sample. We consider estimation and hypothesis testing procedures that are appropriate for some important measures of central tendency, proportion and testing for randomness and the presence of trend.

17.3.1 The Sign Test

The sign test is the oldest of all non-parametric tests and it is so called as we usually convert the data for analysis to a series of plus and minus signs. The test statistic consists of either the number of plus signs or the number of minus signs.

Assumptions

(i) The observations may be considered as a random sample from a population with unknown median M.

(ii) The variable of interest is measured on ordinal scale.

(iii) The n sample values $x_1, x_2, \ldots\ldots, x_n$ may be considered as continuous.

Hypotheses

(a) Two sided $H_0 : M = M_0$, $H_1 : M \neq M_0$

(b) One sided $H_0 : M \leq M_0$, $H_1 : M > M_0$

(c) One sided $H_0 : M \geq M_0$, $H_1 : M < M_0$

We select α as the level of significance.

Test Statistic

We record the sign of the difference obtained by subtracting the median value M_0 stated in H_0 from each of the sample value. In short, we record the signs of the n differences

$$x_i - M_0, \ i = 1, 2, \ldots\ldots n$$

If the null hypothesis is true that the median is M_0 we expect a random sample from the population to have as many plus signs as minus signs.

If we observe a sufficiently small number of plus or minus signs, we reject H_0 in (a). So the test statistic in case (a) is the number of plus or minus signs whichever is minimum.

If we observe a sufficiently small number of minus signs, we reject H_0 in (b). So the test statistic in case (b) is the number of minus signs.

If we observe a sufficiently small number of plus signs we reject H_0 in (c). So the test statistic in case (c) is the number of plus signs.

Note : If $x_i = M_0$, that item is omitted from the sample size as well as analysis.

Conclusion can be made as follows :

(i) Reject H_0 in case (a) at α level of significance, if the probability of observing as few of the less frequently occuring signs when H_0 is true in a random sample of size n is less than or equal to $\alpha/2$.

(ii) Reject H_0 in case (b) at α level of significance if the probability of observing as few minus signs when H_0 is true in a random sample of size n is less than or equal to α.

(iii) Reject H_0 in case (c) at α level of significance if the probability of observing as few plus signs when H_0 is true in a random sample of size n is less than or equal to α.

Note : For each of the above 3 cases, we consult Table-1.

Example 17.1 The appearance transit times for 11 patients with significantly occluded right coronary arteries are as follows :

Patients' No.	:	1	2	3	4	5	6
Transit time (Sec)	:	1.80	3.30	5.65	2.25	2.50	3.50
Patients' No.	:	7	8	9	10	11	
Transit time (Sec)	:	2.75	3.25	3.10	2.70	3.00	

Can we conclude that at 5% level of significance, the median appearance transit time in the population from which the sample was drawn is 3.50 Seconds?

Solution : We set up $H_0 : M = 3.50$, $H_1 ; M \neq 3.50$.

Now we compute 11 differences $(x_i - 3.50)$ from the given observations in the sample as below :

Pat No.	Transit times (Sec)	Sign of the difference $(x_i - 3.50)$
1	1.80	—
2	3.30	—
3	5.65	+
4	2.25	—
5	2.50	—
6	3.50	0
7	2.75	—
8	3.25	—
9	3.10	—
10	2.70	—
11	3.00	—

Here we have seen 9 minus signs, 1 plus sign and 1 zero, so the size of the sample becomes, 11-1=10

The value of the test statistic, K=1

From Table-1, we have $P\{K \leq 1/10, 0.5\} = 0.0098$

Conclusion : Since 0.0098 is less than $\alpha/2 = 0.025$, we may reject H_0. That is, at $\alpha = 0.05$ level of significance we can conclude that the population from which the above sample is drawn has no median = 3.50.

17.3.2 The Wilcoxon Signed Rank Test

We have seen that in the sign test, we have considered only the signs of the differences between the observed values in the sample and the median value as given in the null hypothesis. The Wilcoxon signed rank test utilises both the signs of the differences as well as the magnitude of the differences, hence it provides more information than the sign test, it is often more powerful test.

In this test procedure, we, at first, find the rank of the absolute differences, we assign the original signs of the differences to the ranks and compute two sums : the sum of the ranks with negative signs and the sum of the ranks with positive signs. Since the population considered here is symmetric, the conclusion about the population median may also be applied to the population mean.

Assumptions

(i) The sample $x_1, x_2, \ldots\ldots x_n$ of size n is a random sample from a population with unknown median M.

(ii) The variable of interest is continuous.

(iii) The population from which the sample is drawn is symmetric.

(iv) The scale of measurement is at least interval.

(v) The observations are independent.

Hypotheses

(a) Two sided, $H_0 : M = M_0; H_1 : M \neq M_0$

(b) One sided $H_0 : M \geq M_0 ; H_1 : M < M_0$

(c) One sided $H_0 : M \leq M_0 ; H_1 : M > M_0$.

We select α as the level of significance.

Test Statistic

To obtain the test statistic, we follow the following steps :

(i) Subtract the median value given in null hypothesis from each observation i.e. find $D_i = x_i - M_0$. If $x_i = M_0$ eliminate these values from calculations and thus the sample size is reduced.

(ii) Find the absolute values of D_i i.e. $|D_i|$.

(iii) Rank the $|D_i|$ values from the smallest to the largest. If two or more $|D_i|$'s are equal, assign each tied value the mean of the rank positions occupied by the differences that are tied.

For example. if the three smallest differences are all equal, rank them, 1.2 and 3 but assign each a rank of (1+2+3)/3 = 2

(iv) The sign of the differences are attached to the corresponding ranks.

(v) Obtain the sum of the rank values with + ve signs and call it T+ and also obtain the sum of the rank values with - ve signs, call it T – . If any one of T + or T - is obtained the other can be calculated from T + = {n (n + 1) / 2] - T – .

Conclusion can be made as follows :

We consult Table-11 for various values of n from 3 to 25 at various levels of significance appear in column leveled by d.

(i) For case (a) ; the test statistic T which is the minimum value of either T+ or T – .

Reject H_0 at the α (two sided) level of significance if the calculated value of T is less than or equal to d for given n.

(ii) For case (b) ; the test statistic is T +

Reject H_0 at the α (one sided) level of significance if T + is less than or equal to d for given n.

(iii) For case (c). The test statistic is T – .

Reject H_0 at the α (one sided) level of significance if T – is less than or equal to d for given n.

Example 17.2 The 1.Q's of 15 persons in a test are as follows : 99, 100, 90, 94, 135, 108, 107, 111, 119, 104, 127, 109, 105, 125. Test whether the median 1.Q of the population from which the above sample has been drawn is 107.

Solution : We set up, $H_0 : M = 107; H_1 ; M \neq 107$.

The calculation of the test statistic can be carried out as follows :

1.Q.	$D_i = x_i - M_0$	$\lvert D_i \rvert$	Rank of $\lvert D_i \rvert$	Signed rank of $\lvert D_i \rvert$
99	- 8	8	7	- 7
100	- 7	7	6	- 6
90	- 17	17	11	- 11
94	- 13	13	10	- 10
135	+ 28	28	14	+ 14
108	+ 1	1	1	+ 1
107	0	Eliminate for analysis		
111	+ 4	4	5	+ 5
119	+ 12	12	9	+ 9
104	- 3	3	4	- 4
127	+ 20	20	13	+ 13
109	+ 2	2	2.5	+ 2.5
117	+ 10	10	8	+ 8
105	- 2	2	2.5	- 2.5
125	+ 18	18	12	+ 12

We calculate. T + = 64.5 and T - = 40.5 The test statistic in this case is T = 40.5

Conclusion : From Table-11 for n = 15 - 1 = 14 for two sided $\alpha = 0.05$ the d = 22. Since the calculated value of T is greated than 22 at α = 0.05, therefore, H_0 may be accepted. That is, the median value of the population from which the above sample is drawn is 107.

Large Sample Approximation

When the sample size, n is greater than 25 we cannot use Table-11. For n > 25 we can use the following large sample approximation and the test statistic is

$$T^* = \frac{T - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \qquad \text{........(17.1)}$$

which has approximately the standard normal distribution; here T is the value of T or T+ or T – according to the hypotheses and n is the sample size.

If there are 't' tied values obtained in calculating the rank of $|D_i|$, the correction factor is to be subtracted from the quantity under the radical in (17.1). So that T* reduced to the following value,

$$T^{**} = \frac{T - n(n+1)/4}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\Sigma t^3 - \Sigma t}{48}}}$$

T** has the critical value ±1.96 at $\alpha = 0.05$ and ± 2.58 at $\alpha = 0.01$. Therefore, the conclusion can be made as usual.

17.3.3 The χ^2-test

Frequently we may be interested in the subjects or responses which falls in various categories or classes. For example, children may be classified according to their most frequently modes of play.

To test the hypothesis that these modes will not differ in frequency, we usually use χ^2-test which is almost equivalent to the χ^2 for testing goodness of fit.

Assumptions

(i) The data available for analysis consist of a random sample of n independent observations.

(ii) The measurement scale may be nominal.

(iii) The observations are classified into mutually exclusive and exhustive classes.

Hypothesis

H_0 : The sample has been drawn from a population that follows a specified distribution.

H_1 : The sample has not been drawn from a population that follows the specified distribution.

Note that the alternative hypothesis does not indicate how the true distribution differs from the distribution mentioned in the null hypothesis.

We select α as the level of significance.

Test Statistic

Here we have given the observed frequencies and the expected frequencies are calculated in the light of H_0. The method is to compare observed frequencies with the corresponding expected frequencies. The null hypothesis states the proportion of objective falling in each of the classes in the presumed population.

The χ^2 test says whether the observed frequencies are sufficiently close to the expected ones to be likely to have occured under H_0.

The test statistic is

$$\chi^2 = \sum_{i=1}^{K} (O_i - E_i)^2/E_i = \sum_{i=1}^{K} O_i^2/E_i - n \qquad ....(17.2)$$

where $\sum_{i=1}^{K} Oi = \sum_{i=1}^{K} E_i = n$; O_i and E_i are observed and expected frequencies respectively corresponding to ith class, n is the size of the sample and K is the number of classes. Under H_0 ; for large n, the sample size (17.2) is distributed as χ^2 distribution with (K-1) d.f.

Conclusion can be made as follows :

If the calculated value of χ^2 is greater than or equal to the tabulated value of χ^2 with (K-1) d.f. at α level of significance, we may reject H_0 at α level of significance.

Note : Cochran recommends that in the above test, no expected frequency should be less than 1. When expected frequencies less than 1 occur, we usually combine the categories in which they occur with adjacent categories until the minimum frequency requirement has been met. We have to recompute the number of d.f. on the basis of the new number of categories. If we estimate p number of parameters from the data, the d.f. reduces by p also.

Example 17.3 Given below the data obtained in 8 different mutually exclusive classes where the observed frequencies represent the number of winners in a certain game.

Different Classes

	1	2	3	4	5	6	7	8
No. of Winners	29	19	18	25	17	10	15	11

Test whether the sample has been obtained from a rectangular population with $f_1 = f_2 = \ldots\ldots = f_8$.

Solution : We set up, $H_0 : f_1 = f_2 = f_1 = \ldots\ldots = f_8$

$H_1 : f_1 \neq f_2 \neq f_3 \neq \ldots\ldots\ldots \neq f_8$

Here $n = 144$, $K = 8$ under H_0, $E_i = \frac{n}{K} = \frac{144}{8} = 18$

We prepare a table of the following type-

Different Classes

	1	2	3	4	5	6	7	8
Obs No. of Winners (O) :	29	19	18	25	17	10	15	11
Expt. No. of Winners (E) :	18	18	18	18	18	18	18	18
$\frac{O^2}{E}$:	46.72	20.06	18.00	34.72	16.06	5.56	12.60	6.72

$\therefore \chi^2 = \Sigma O^2/E - 144 = 160.34 - 144 = 16.34$ with $8 - 1 = 7$ d.f.

Conclusion : The calculated value of χ^2 with 7 d.f. is greater than the tabulated value of χ^2 with same d.f. at 5% level of significance. Therefore, H_0 may be rejected at 0.05 level of significance.

17.3.4 The Kolmogorov-Smirnov Test

This test procedure is designed for testing the goodness of fit for data of continuous type and measured on ordinal scale. It has been developed by two Russiun Mathematicians-A. N. Kolmogorov and N.V. Smirnov.

In this type of test we focus on two cumulative distribution function (c.d.f) namely a hypothesized c.d.f and the observed c.d.f. Let us denote F(x) as the probability that the value of the random variable X is less than or equal to x. That is, $F(x) = P\{X \leq x\}$.

Suppose we draw a random sample from some unknown distribution function F(x). Under H_0 we have $F(x) = F_0(x)$ and S(x) is the c.d.f corresponding to sample observation. The objective of the Kolmogorov-Smirnov test is to show one sample goodness of fit test to determine whether the lack of aggrement between $F_0(x)$ and S(x) is sufficient to cast doubt on H_0 that $F(x) = F_0(x)$.

If the differences between S(x) and $F_0(x)$ are not too large then we can say that the sample data have been obtained from $F_0(x)$ otherwise we can cast doubt on the aggrement stated in H_0.

Assumptions

(i) The independent observations $x_1, x_2, \ldots\ldots, x_n$ is a random sample of size n from unknown distribution function F(x).

(ii) The data are of continuous type.

(iii) The measurement scale for this test is usually of ordinal scale type.

Hypotheses

a) Two sided, $H_0 : F(x) = F_0(x)$ for all values of x
$H_1 : F(x) \neq F_0(x)$ for at least one value of x

b) One sided, $H_0 : F(x) \geq F_0(x)$ for all values of x
$H_1 : F(x) < F_0(x)$ for at least one value of x

c) One sided, $H_0 : F(x) \leq F_0(x)$ for all values of x
$H_1 : F(x) > F_0(x)$ for at least one value of x

We select α as the level of significance

Test Statistic

Let S(x) be the cumulative distribution function computed from the sample data. It can be calculated as

$$S(x) = \frac{\text{the no of sample observation less than or equal to x}}{n}$$

where n = sample sige.

For hypothesis (a) The test statistic is D=Maximum $|S(x) - F_0(x)|$, that is maximum value of the absolute difference between S(x) and $F_0(x)$.

For hypothesis (b) the test statistic is D^+ = Maximum $[F_0(x)-S(x)]$

For hypothesis (c), the test statistic is D^- = Maximum $[S(x) - F_0(x)]$

Conclusion can be made as follows :

We consult Table-18 and reject H_0 at α level of significance, if the test statistic under consideration D or D+ or D–exceeds the (1- α) quantile shown in the Table-18. Otherwise we shall accept H_0.

Example 17.4 The following data give us the kidney weights in gram of 36 dogs. Test whether the data are from a normal distribution with population mean 85 gm and standard deviation 15 gm.

58, 78, 84, 90, 97, 70, 90, 86, 82, 59, 90, 70, 74, 83, 90, 76, 88, 84, 68, 93, 70, 94, 70, 110, 67, 68, 75, 80, 68, 82, 104, 92, 112, 84, 98, 80

Consider $\alpha = 0.05$ as the level of significance.

Solution : We set up, $H_0 : F(x) = F_0(x)$ for all values of x

$H_1 : F(x) \neq F_0(x)$ for at least one value of x.

where $F_0(x)$ is the c.d.f of a normally distributed variable with mean 85 gm and st. dev. 15 gm.

The required test statistic is D = Maximum $|S(x) - F_0(x)|$

We first obtain the values of S(x) which is

$$S(x) = \frac{\text{No. of sample observation} \leq x}{n} \quad \text{where n = sample size = 36.}$$

$F_0(x)$ can be obtained from Table-2

For calculation of the test statistic, we proceed systematically as follows :

Calculation of S(x)

x wt(gm)	Freq.	Cum. Freq.	S(x)
58	1	1	$\frac{1}{36} = 0.0278$
59	1	2	$\frac{2}{36} = 0.0556$
67	1	3	0.0833
68	3	6	0.1667
70	4	10	0.2778
74	1	11	0.3056
75	1	12	0.3333
76	1	13	0.3611
78	1	14	0.3889
80	2	16	0.4444
82	2	18	0.5000
83	1	19	0.5278
84	3	22	0.6111
86	1	23	0.6389
88	1	24	0.6667
90	4	28	0.7778
92	1	29	0.8056
93	1	30	0.8333
94	1	31	0.8611
97	1	32	0.8889
98	1	33	0.9167
104	1	34	0.9444
110	1	35	0.9722
112	1	36	1.0000

Calculation of F_0 (x)

x wt (gm)	$z = \frac{x - 85}{15}$	P $(0 \le z \le z)$	$F_0(x)$
58	- 1.80	0.4641	.5-.4641 = 0.0359
59	- 1.73	0.4582	0.0418
67	- 1.20	0.3849	0.1151
68	- 1.13	0.3708	0.1292
70	- 1.00	0,3413	0.1587
74	- 0.73	0.2673	0.2327
75	- 0.67	0.2486	0.2514
76	- 0.60	0.2257	0.2743
78	- 0.47	0.1808	0.3192
80	- 0.33	0.1293	0.3707
82	- 0.20	0.0793	0.4207
83	- 0.13	0.0517	0.4483
84	- 0.07	0.0279	0.4721
86	0.07	0.0279	.5+=0.5279
88	0.20	0.0793	0.5793
90	0.33	0.1293	0.6293
92	0.47	0.1808	0.6808
93	0.53	0.2019	0.7019
94	0.60	0.2257	0.7257
97	0.80	0.2881	0.7881
98	0.87	0.3078	0.8078
104	1.27	0.3980	0.8980
110	1.67	0.4525	0.9525
112	1.80	0.4641	0.9641

Calculation of $|S(x) - F_0(x)|$

x	$S(x)$	$F_0(x)$	$\lvert S(x) - F_0(x) \rvert$
58	0.0278	0.0359	0.0081
59	0.0556	0.0418	0.0138
67	0.0833	0.1151	0.0318
68	0.1667	0.1292	0.0375
70	0.2778	0.1587	0.1191
74	0.3056	0.2327	0.0729
75	0.3333	0.2514	0.0819
76	0.3611	0.2743	0.0868
78	0.3889	0.3192	0.0697
80	0.4444	0.3707	0.0737
82	0.5000	0.4207	0.0793
83	0.5278	0.4483	0.0795
84	0.6111	0.4721	0.1390
86	0.6389	0.5279	0.1110
88	0.6667	0.5793	0.0874
90	0.7778	0.6293	0.1485
92	0.8056	0.6808	0.1248
93	0.8333	0.7019	0.1314
94	0.8611	0.7257	0.1354
97	0.8889	0.7881	0.1008
98	0.9167	0.8078	0.1089
104	0.9444	0.8980	0.0464
110	0.9722	0.9525	0.0197
112	1.0000	0.9641	0.0359

From the above table, D = Maximum of $|S(x) - F_0(x)| = 0.1485$

Conclusion : From the Table-18 for n = 36 the (1-0.05) = 0.95 quantile value is 0.221. Since the calculated value of the test statistic is 0.149 (app). which does not exceed 0.221, therefore, H_0 may be accepted. That is, the sample has been obtained from a normal distribution with mean 85 gm and st. dev. 15 gm.

17.3.5 The Binomial Test.

There are populations which have only two classes, let us consider human population with two classes-male and female ; literate and illiterate ; member, non-member etc. Trial has two classes-success and failure. A number of such examples can be cited easily.

For any population of two classes, if the proportion in one class is p, then the other class must have proportion equals to (1-p) or q, such that p+q = 1. For a population p is fixed but it varies from population to population. Again for a fixed population with proportion p, if we draw random samples, we cannot expect that the proportion of one of the two classes to be exactly p and q for the other because of random effect. Such differences between the observed and the population values arise because of chance. If we draw a random sample of size n from a population, the binomial formula enables us to compute the probability, that the sample contains a specified number of elements in one class. For this the test is usually called the binomial test.

Assumptions

(i) The data consist of the outcomes of n repeated Bernoulli trials. The proportion of success is $\hat{p} = \frac{s}{n}$ where s is the number of trials having success.

(ii) Then trials are independent.

(iii) The probability of success p remains constant from trial to trial. We use p to designate the proportion of the population having characteristic of interest.

Hypotheses

(a) Two sided, $H_0 : p = p_0$ $H_1 : p \neq p_0$

(b) One sided, $H_0 : p \leq p_0$ $H_1 : p > p_0$

(c) One sided, $H_0 : p \geq p_1$ $H_1 : p < p_0$

we select α as the level of significance.

Test Statistic

Since we are interested in the number of successes S, Our test statistic is S = number of success which is usually smaller number of outcomes.

Conclusion can be made as follows :

(a) Reject H_0 if S is either less than or equal to s_1 or larger than s_2. s_1 and s_2 are determined by $p\ (x \leq s_1) \approx \alpha/2$ and $p\ (x > s_2) \approx \alpha/2$.

(b) Reject H_0 if S is greater than s where $p\ (x > s) = \alpha$.

(c) Reject H_0 if S is greater than s where $p\ (x \leq s) = \alpha$.

Example 17.5 In an experiment with 18 trials, the failures and success are recorded as follows :

	Trial		
	Success	failure	Total
Frequency	2	16	18

Test whether the experiment follows binomial distribution with p, the proportion of success equal to 0.5 at $\alpha = 0.05$.

Solution : We set up $H_0 : p = 0.5$ $H_1 : p \neq 0.5$

Here n = 18, S = 2. From the Table-1, we have $P\ (x \leq 2) = 0.001$

Conclusion : Since the probability 0.001 is very much smaller in comparison to $\alpha/2 = 0.025$. Therefore, H_0 may be rejected.

Large Sample Approximation : We know that for large. n, the sample size, the binomial distribution tends to normal and this tendency is rapid when $p = q = \frac{1}{2}$. Therefore, for large n, ($n > 25$), the sampling distribution of S is approximately normal with mean np and variance npq. So for large n, H_0 may be tested by-

$$z = \frac{S - np}{\sqrt{npq}} \qquad \text{..........(17.3)}$$

where z is a N(0.1) variate. In this case, conclusion can be made as usual.

If the value of the test statistic, $S < 5$, correction for continuity arises and z becomes as follows :

$$z = \frac{(S \pm 0.5) - np}{\sqrt{npq}} \qquad \text{..........(17.4)}$$

where $(S + 0.5)$ is used when $S < np$ and $(S - 0.5)$ is used when $S > np$.

Example 17.6 A sample of size 30 students, were interviewed to give comment 'in farour of' or 'against' the present education system of a country. The opinion is presented below :

	In favour	Against	Total
Frequency	8	22	30

Test whether the probability of opinion in favour or against is same.

Solution : We set up, $H_0 : p = 0.5$, $H_1 : p \neq 0.5$

Here $n = 30 > 25$ and the test statistic, $S = 8$. np in greater than S i.e. $30 \times 0.5 = 15 > 8$, therefore, we use

$$z = \frac{(8 + 0.5) - 30 \times 0.5}{\sqrt{30 \times 0.5 \times 0.5}} = \frac{8.5 - 15}{7.5} = -2.37$$

Conclusion : The critical value of z at 0.05 level of significance is ± 1.96. Since our calculated value is greater than - 1.96, H_0 may be rejected at 5% level of significance.

17.3.6 The Run Test

In many situations we are interested to know whether the data collected for statistical analysis is random or not. In all most all statistical inference we assume that the data in a random sample.

If the randomness of a sample is suspected, we test its randomness before going to statistical analysis in detail. There are many situations where we may investigate the assumption of randomness. The data of statistical quality control analysis and regression analysis are the most important situations in which we must be sure about randomness.

Procedures for investigating the randomness are based on the number and nature of the runs present in a data of interest. A run is defiened as a sequence of like events or items or symbols of the sample that is preceeded and followed by an event, item or symbol of a different type. The number of items or events in a run is called the length of the run. We doubt on randomness of a sample if there are too many or too few runs.

Consider an example of a sample of size 10 patients who are subjected to some treatment. If the sexes of the patients are MFMFMFMFMF we would suspect about the randomness of the data having 10 runs.

But if the sexes of the patients are of order MMMMMFFFFF then also we doubt on randomness since there are only 2 runs. The following procedure will give us a guideline to decide whether the sequence of events or items in a sample is the result of a random process.

Assumption

(i) The data given for statistical analysis consist of a sequence of observations recorded in the order of their occurence which can be classified into two mutually exclusive types. Let us consider n = the sample size, n_1 = number of observation in one type and n_2 = number of observations in the other type.

Hypotheses

(a) : Two sided

H_0 : The pattern of occurence of the two types of observation is determined by a random process.

H_1 : The pattern of occurrence is not random.

(b) One sided.

H_0 : The pattern of occurence of the two types of observation is determined by a random process.

H_1 : The pattern is not random (because there are too few runs to be attributed to chance).

(c) One sided.

H_0 : The pattern of occurence of the two types of observation is determined by a random process.

H_1 : The pattern is not random (because there are too many runs to be attributed to chance).

We select α as the level of significance.

Test Statistic

The test statistic is r. the total number of runs.

Conclusion can be made as below :

(a) Reject H_0, if r is less than or equal to the lower critical value (consult Table-6) or greater than or equal to the upper critical value (consult Table-7)

Table-6 and Table-7 give lower and upper critical values respectively of the test statistic for the 0.05 level of significance and values of n_1 and n_2 through 20.

(b) Enter Table-6 with n_1 and n_2. If r is less than or equal to the tabulated value of the test statistic, reject H_0 at 0.025 level of significance.

(c) Enter Table-7 with n_1 and n_2. If r is greater than or equal to the tabulated value of the test statistic reject H_0 at 0.025 level of significance.

Note : If the exact value of n_1 and n_2 are not in the tables we use the closest value as an approximation.

Example 17.6 The data of blood flow in lung capillaries in 16 patients with neuromuscular weakness have the following sequence of sexes

FFFMFFMMMFFFFFFM

Test the null hypothesis that the sequence is random.

Solution : We set up,

H_0 : The pattern of occurnce of the two types of observations is determine by a random process.

H_1 : The pattern of occurence is not random.

Here in the sequence, first 3 F constitute a run whereas 1M constitute another run, In the similar ways, there are 6 runs, therefore, the test statistic, r = 6 and n_1 = 11 and n_2 = 5. From Table-6 for n_1 = 11 and n_2 = 5 lower critical value = 4 ; upper critical value is not given.

Conclusion : In the data, the number of runs = 6 which is greater than the lower critical value at 0.05 level of significance. Therefore, H_0 may be accepted.

Example 17.7 The deviation of daily temperature from normal of a certain city in January 1998 are given below :

Day	1	2	3	4	5	6	7	8	9	10	11	12				
Temp. (Dev. from normal)	5	13	12	11	12	2	-1	2	-1	2	3	-7				
	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
	-6	-7	-12	-9	6	7	10	6	1	1	-3	7	-2	-5	-6	-6
	29	30	31													
	-2	1	-1													

Test whether the pattern of departure above and below normal is the result of random process.

Solution : We set up,

H_0 : The pattern of occurence of negative and positive deviation from normal is due to random process.

H_1 : The pattern of occurence of negative and positive deviations from normal is not due to random process.

From the data, we have seen positive deviations for days 1 to 6 constitute a run, three runs each of length one follow this run. Days 10 and 11 contribute another run. Next 5 negative deviations together give us another run and similarly in all we get 11 runs in total. And $n_1 = 18$, $n_2 = 13$ give the number of items or observations with positive deviations and negative deviations respectively. The value of the test statistic, r=11.

Conclusion : Consulting Table-6 and Table-7 with $n_1 = 18$ and $n_2 = 13$ we find the critical values for this test are 10 and 22. Since the calculated value of test statistic is greater than the lower critical value but less than the upper critical value, therefore, the H_0 may be accepted and we conclude that the pattern of occurence of temperatures above and below normal is random.

Large Sample Approximation : When either n_1 or n_2 is greater than 20 we cannot use Table-6 and Table-7 to test the above hypothesis. However, for large samples,

$$z = \frac{r - [\{(2n_1n_2)/(n_1+n_2)\}+1]}{\sqrt{\dfrac{2n_1n_2\,(2n_1n_2 - n_1-n_2)}{(n_1+n_2)^2(n_1+n_2 - 1)}}}$$

which is distributed approximately as standard normal distribution when H_0 is true. Finally the conclusion can be made as usual by considering the critical value ± 1.96 at 0.05 level of significance and the critical value ± 2.58 at 0.01 level of significance.

17.4 Some Non-parametric Tests that use Data from Two or More Samples.

In this section, we shall discuss several non-parametric test procedures based on data obtained from two or more populations. In some of the cases, the two populations are independent and in that case, there is independence within samples as well as between samples. In case of two related populations, we obtain data in which measurements are taken "before and after" or "pre or post test" experiment on the same subjects.

17.4.1. The Sign Test for Two Related Samples

When we use measurement scale of two variates by ordinal scale, the sign test for two related samples is very useful. In this case, each pair of measurements has one with larger value than the others and if so which is larger. The hypothesis testing procedure and the idea are almost same as given in one sample case.

Assumptions

(i) The data consist of a random sample of n pairs of measurements (x_1, y_1) ; (x_2, y_2); ....; (x_n, y_n) where each pair of measurement is taken on same subject or subjects that have been paired with respect to one or more variables.

(ii) The n pair of measurements are independent.

(iii) The measurement scale is at least ordinal within each pair so that one can determine which of the two members is larger.

(iv) The variable under study is continuous.

Hypotheses

(a) Two sided

H_0 : The median of the population of differences $(x_i - y_i) = D_i$ is zero.

H_1 : The median of the population of differences is not zero.

(b) One sided

H_0: The median of the population of differences is less than or equal to zero.

H_1 : The median of the population of differences is greater than zero.

C. (One sided)

H_0 : The median of the population of differences is greater than or equal to zero.

H_1 : The median of the population of differences is less than zero.

We select α as the level of significance.

Test Statistic

For each pair of observations (x_i, y_i), we record the sign of the difference obtained by subtracting y_i from x_i. In short, we record the sign of $(x_i - y_i)$ i.e. we record + sign if $(x_i - y_i) > 0$ and - sign if $(x_i - y_i) < 0$. If ties occur i.e. if $x_i = y_i$ for any pair, eliminate those pairs from the analysis and reduce the sample size n according to the number of ties. If there are 't' ties in a series of n pairs of observations, the reduced number of sample size would be (n-t).

For case (a) If H_0 is true, we expect a sample of differences to include as many + signs and - sings. Either a sufficiently small number of + signs or a sufficiently small number of - signs causes us to reject H_0.

Therefore, the test statistic for case (a). is the number of + signs or the number of - signs whichever is smaller.

For case (b) Since a sufficiently small number of signs causes us to reject H_0. Therefore, the test statistic for case (b), is the number of - signs.

For case (c) Since a sufficiently small number of + signs causes us to reject H_0. Therefore, the test statistic for case (c), is the number of + signs.

Conclusion can be made as below :

We can state the hypotheses in terms of probabilities of + and - signs.

For (a), $H_0 : P(+) = P(-) = 0.5$ $H_1 : P(+) \neq p(-) \neq 0.5$

For (b) $H_0 : P(+) \leq P(-)$ $H1 : P(+) > P(-)$

For (c) $H_0 : P(+) \geq P(-)$ $H_1 : P(+) < P(-)$

From this we can see that the sign test for two related samples may be considered as a special case of binomial test.

Therefore, we can reject H_0 on the magnitude of $P (K \le k'/n, 0.5)$ which can be obtained from Table-1.

Finally the specific conclusions for the above 3 cases are as follows :

a. (Two sided) : Reject H_0 at the α level of significance if

$P (K \le k'/n, 0.5) \le \alpha/2$.

where k' is the number of + signs or - signs whichever is smaller.

(b) and (c) One sided : Reject H_0 at α level of significance if

$P (k \le k'/n, 0.5) = \alpha$ where k' is the number of - signs for case

(b) and k' is the number of + signs for case (c).

Example 17.8 The yields in kg of a certain paddy variety are recorded in 10 equal plots before and after using manure.

Plot No.	1	2	3	4	5	6	7	8	9	10
Yield before Using Manure :	463	462	462	456	450	426	418	415	409	402
Yield after Using Manure :	523	494	461	456	476	454	448	408	470	437

Test whether the median yields of paddy before and after using manure are same at 0.05 level of significance.

Solution : We set up,

H_0 : The median of the population of differences $(x_i - y_i) = D_i$ is zero.

H_1 : The median of the population of differences is not zero.

Otherwise we can state the above H_0 and H_1, as follows :

$H_0 : P (+) = P (-) = 0.5$ $\qquad$ $H_1 : P (+) \neq P (-) \neq 0.5$

The test statistic is the number of + signs or the - signs whichever is smaller. For this we try to find out the number of + and - signs as follows :

Yield of paddy Before using Manuare (x_i)	Yield of paddy After Using Manure (y_i)	Difference (x_i-y_i) D_i	Sign of D_i
463	523	-60	-
462	494	-32	-
462	461	+1	+
456	456	0	Eliminate
450	476	-26	-
426	454	-28	-
418	448	-30	-
415	408	+7	+
409	470	-61	-
402	437	-35	-

In this case, the modified or reduced sample size, $n = 10 - 1 = 9$ and the test statistic, $K = 2$. From Table-1 we have $P\ (K \leq 2/\ 9,\ 0.5) =$ 0.09

Conclusion : Since the probability 0.09 is greater than $\alpha/_2 = 0.025$, therefore, H_0 may be accepted.

17.4.2 The Tukey's Quick Test

This procedure is mainly used for making inference about the difference between two location parameters. The hypothesis usually considered in this case is that the two location parameters are equal. In parametric test, for this type of hypothesis, we use t-test for two independent as well as related samples where some strict assumptions are to be met. If any one of these assumptions fails to meet we may use Tukey's quick test which is based on the fact that when we compare two groups, the less over lap in the observations, the more likely we are to reach the conclusion that the groups differ.

Assumptions

(1) The data consist of observations $x_1, x_2, \ldots\ldots\ldots, x_n$ and $y_1, y_2, \ldots\ldots\ldots, y_m$ to form two independent random samples of sizes n and m respectively, such that $n > m$.

(2) The measurement scale employed is atleast ordinal.

(3) The random variables are of continuous type.

Hypotheses

a. (Two sided)

H_0 : The two samples come from identical populations.

H_1 : Either the x's tend to be larger than the y's or the y's tend to be larger than the x's.

b. (One sided)

H_0 : The two samples come from identical populations.

H_1 : The x's tend to be larger than y's.

c. (One sided)

H_0 : The two samples come from identical populations.

H_1 : The y's tend to be larger than x's.

We select α as the level of significance.

Test Statistic

The test statistic depends on which set of hypotheses we are considering. So for different sets of hypotheses we can obtain the test statistic as follows :

(a). Two sided

From n observations $x_1, x_2, \ldots\ldots\ldots, x_n$ and m observations $y_1, y_2, \ldots\ldots\ldots, y_m$, the numerical value of the test statistic T can be obtained from the following processes :

(1) Find out the smallest and the largest values from the two samples combined.

(2) If both the largest and the smallest values occur in the same sample i.e. either in x's or in y's. We cannot conclude that the location parameters of the two populations are different. In this case the value of the test statistic, T= 0.

But if one sample contains the largest value and the other the smallest value, the following steps [(3) to (7)] must be carried out.

(3) Determine the largest and the smallest values in Sample- 1.

(4) Determine the largest and the smallest values in Simple – 2.

(5) Consider the sample which contains the smallest value in the two samples combined (as in step-1).

Count the number of values in this sample that are smaller than the smallest value in the other sample.

(6) Consider the sample which contains the largest value in the two samples combined (as in step-1).

Count the number of values in this sample that are larger than the largest value in the other sample.

(7) The sum of the counts obtained in step-5 and step-6 gives the numerical value of the test statistic T.

b. (One sided)

If the largest value in both the samples combined is an x value, count the number of values that are larger than the largest in y values. The number of counts is the value of the test statistic, T_1.

c. (One sided)

If the smallest value in both the samples combined is an y value, count the number of values that are smaller than the smallest in the x values. The number of counts is the value of the test statistic, T_2.

The conclusion can be made as below :

a. **(Two sided)**

Reject H_0 at the 0.05, 0.01 or 0.001 level of significance if T is equal to or greater, than 7, 10 or 13 respectively.

b. c. **(One sided)** Reject H_0 at the 0.025, 0.005 or 0.0005 level of significance if T_1 or T_2 is greater than or equal to 7, 10 or 13. respectively.

Ties : In this case, ties occur in two ways.

(1) The first type occurs when the largest value in the two groups combined or the smallest value in both the groups combined occurs in both samples. In either of these cases, count the value of the test statistic as 0.

(2) The second type of tie occurs when the largest value of the combined samples lies in both the samples. Simillar type of tie may occur in case of the smallest value also. Count values of this type as $\frac{1}{2}$ when determining the value of the test statistic.

Example 17.9 In a research study differential personality inventory to male and female alcoholics were determined by the scores on cynicism scale as given below :

Male (x) :	5	6	5	4	3	2	3	4
Female (y) :	1	2	2	3	1	2	3	

using $\alpha = 0.05$, test whether these data provide sufficient evidence to indicate a difference in population.

Solution : We set up,

H_0 : The two sampled populations are identical.

H_1: Either the x's tend to be larger than the y's or the y's tend to be larger than the x's.

Test Statistic

To obtain the value of the test statistic, we proceed as follows :

(1) The largest value of the two samples combined is 6 which occurs is the Male group (x). The smallest value is 1, which occurs in the Female group (y).

(2) Since the two extreme values occur in different groups, we proceed with step 3 to 7 as below.

(3) The largest and the smallest values in the Male group are 6 and 2 respectively.

(4) The largest and the smallest values in the Female group are 3 and 1 respectively.

(5) The smallest value in the two groups combined is 1, which occurs in the Female group. This group contains 2 values (1, 1) that are smaller than the smallest value in Male group. There are 3 values (2, 2, 2) equal to 2, the smallest value of the Male group. Therefore, we count $\frac{1}{2}+\frac{1}{2}+\frac{1}{2}=1\frac{1}{2}$. So in this process the total count is $2+1\frac{1}{2}=3\frac{1}{2}$.

(6) The largest value in the two groups combined is 6, which occurs in the Male group. This group contains 5 values (5, 6, 5, 4, 4) that are larger than the largest value in the Female group. There are 2 values (3, 3) equal to 3 the largest value of the Female group. Therefore, we count $\frac{1}{2}+\frac{1}{2}=1$. So in this process the total count is 5 + 1= 6.

(7) The sum of the counts in step (5) and (6) is equal to $3\frac{1}{2}+6=9\frac{1}{2}$, the value of the test statistic i.e. $T=9\frac{1}{2}$.

Conclusion : Since the calculated value of T is greater than the tabulated value 7 at 0.05 level of significance. Therefore, the H_0

may be rejected, i.e. the data provide sufficient evidence to indicate a difference in populations.

17.4.3. The Median Test

The median test is one of the simplest and most widely used procedures for testing the null hypothesis that two independent samples have been drawn from the populations with equal median value.

Assumptions

(1) The data consist of two independent random samples $x_1, x_2,$ ---- --, x_{n_1} and $y_1, y_2,$---, y_{n_2} of sizes n_1 and n_2 respectively.

(2) The variable of interest is continuous.

(3) The measurement scale used is at least ordinal.

(4) If the two populations have the same median, then for each population, the probability p is the same that an observed value will exceed the grand median.

Hypothesis

H_0 : The two populations are identical.

H_1 : Either the X's tend to be larger than the Y's or the Y's tend to be larger than the X's

We select α as the level of significance.

Test Statistic

If the two populations have the same median, we would expect about half of observations in each of the two samples to be above the common median and about half to be below. Under the H_0 that the two population medians are equal, we may estimate this common parameter by calculating the median of the sample values of the combined two samples of size $(n_1 + n_2)$.

If H_0 is true, we expect about half the observations in each sample to fall above the combined sample median and about half to fall below. The median test allows us to conclude on the basis of sample data whether it is likely that the null hypothesis is false.

The usual test statistic is

$$T = \frac{\frac{A}{n_1} - \frac{B}{n_2}}{\sqrt{\hat{p}\left(1 - \hat{p}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \text{where } \hat{p} = \frac{A + B}{n_1 + n_2}$$

A=No: of observations from sample–1 falling above the median ; B=No. of observations from sample–2 falling above the median; n_1 = 1st sample size ; n_2= 2nd sample size.

Conclusion can be made as below :

If the combined sample size, $N = n_1 + n_2$ is such that Np and N(1 - p) are larger than 5 and if H_0 is true then T is approximately distributed as normal with mean zero and standard deviation one. Therefore, the conclusion can be made as usual by following the level of significance $\alpha = 0.05$, the critical value is ± 1.96 and for $\alpha = 0.01$, the crilical value is ± 2.58.

Example 17.10 The following data give the marks obtained in a test of certain subject by two groups A and B consisting of 32 and 16 students respectively. Test whether these data provide sufficient evidence to indicate that the medians of the two populations from which two groups of students are selected are same. Consider $\alpha = 0.05$

Marks obtained by Group-A students	Marks obtained by Group–B students
25, 25, 17, 26, 18, 30, 24, 21	31, 21, 38, 19
13, 30, 20, 23, 26, 12, 20, 37	38, 41, 68, 28
9, 17, 37, 20, 11, 32, 16, 31	43, 42, 30, 20
46, 20, 25, 17, 36, 54, 8, 26	29, 13, 32, 30.

Solution : We set up,

H_0 : The population median mark of group-A is equal to that of group-B.

H_1 : Either the marks of Group–A students are larger than that of Group–B students or the marks of Group–B students are larger than that of Group–A students.

To calculate the test statistic, we compute the median of the combined data by arranging them in order of magnitude as below :

8, 9, 11, 12, 13, 13, 16, 17, 17, 17, 18, 19, 20, 20, 20, 20, 20, 21, 21, 23, 24, 25, 25, 25, 26, 26, 26, 28, 29, 30, 30, 30, 30, 31, 31, 32, 32, 36, 37, 37, 38, 38, 41, 42, 43, 46, 54, 68.

Therefore, the median $= \dfrac{25 + 26}{2} = 25.5$

From Group A, the number of marks above the median mark is 12 i.e. A = 12. From Group-B, the number of marks above the median mark is 12 i.e B = 12

$$\therefore \hat{p} = \frac{12 + 12}{32 + 16} = 0.5 \,;\, N = 32 + 16 = 48$$

Since $N\hat{p} = 48 \times 0.5 = 24 = N(1 - \hat{p}) > 5$, the test statistic is

$$T = \frac{\frac{12}{32} - \frac{12}{16}}{\sqrt{0.5\,(1\text{-}0.5)\left(\frac{1}{32} + \frac{1}{16}\right)}} = -\,2.45$$

Conclusion : For $\alpha = 0.05$, the level of significance the critical value is ± 1.96. Since –2.45 is less than –1.96 we may reject H_0 at 5% level of significance and conclude that the data provide no evidence that the two samples have been obtained from the populations with same median value.

17.4.4. The Mann-Whitney Test

The procedure for testing the null hypothesis of equal population location parameters is given by Mann-Whitney. They proposed the test for two samples of unequal sizes. Wilcoxon gave this type of test for two samples of equal sizes and used the rank sum as the test statistic. This test is, therefore, called as the Mann-Whitney-Wilcoxon test also.

Assumptions

(1) The data consist of a random sample of observations $x_1, x_2, \ldots\ldots\ldots, x_{n_1}$ of size n_1 from population 1 and another random sample of observations $y_1, y_2, \ldots\ldots\ldots, y_{n_2}$ of size n_2 from population- 2.

(2) The two samples are independent.

(3) The variables are of continuous type.

(4) The distribution functions of the two populations differ only with respect to location parameters if they differ at all.

(5) The measurement scale employed is at least ordinal.

Hypotheses

The following hypotheses are appropriate if the assumption (4) is met.

(a) Two sided

H_0: The two populations have identical distribution.

H_1: The two populations differ with respect to location parameters.

(b) One sided

H_0: The two populations have identical distribution.

H_1: The X's tend to be smaller then Y's

c. (One sided)

H_0 : The two populations have identical distribution

H_1 : The X's tend to be larger than Y's.

We select α as the level of significance.

Test Statistic

To calculate the test statistic, we first combine the two samples and rank all observations from the smallest to the largest, We assign tied observations the mean of the rank positions they would have occupied if there is no tie. The sum of the rank values in sample-1 is obtainted and let it be called S. Then the test statistic is given by,

$$T = S - \frac{n_1 (n_1 + 1)}{2}$$

Conclusion can be made as below :

1. For **case (a),** From Table-9 we note down $W_{\alpha/2}$ and also calculate $W_{(1-\alpha/2)} = n_1 \times n_2 - W_{\alpha/2}$.

   Reject H_0 if the calculated value of T is less than $W_{\alpha/2}$ or greater than $W_{(1-\alpha/2)}$.

2. For **case (b),** reject H_0 if the calculated value of T is less than W_{α}.

3. For, **case (c),** we calculate the value of $W_{(1-\alpha)}$ where $W_{(1-\alpha)} = n_1 n_2 - W_{\alpha}$.

Reject H_0 if the calculated value of T is greater than $W_{(1-\alpha)}$

Example 17.11 The following data give the scores of two groups of individuals namely improved and unimproved after giving a certain type of therapy.

Improved Individuals		Unimproved Individuals	
Sample No.	Score (X)	Sample No.	Score (Y)
1	11.9	1	6.6
2	11.7	2	5.8
3	9.5	3	5.4
4	9.4	4	5.1
5	8.7	5	5.0
6	8.2	6	4.3
7	7.7	7	3.9
8	7.4	8	3.3
9	7.4	9	2.4
10	7.1	10	1.7
11	6.9		
12	6.8		
13	6.3		
14	5.0		
15	4.2		
16	4.1		
17	2.2		

Test whether the two represented populations from which these two samples drawn are different with respect to location parameter. Given $\alpha = 0.05$.

Solution : We set up, H_0 : The two populations are identical

H_1 : The two populations differ with respect to location parameters.

At first we arrange the combined scores in rank as follows :

X scores	Rank	Y scores	Rank
		1.7	1
2.2	2		
		2.4	3
		3.3	4
		3.9	5
4.1	6		
4.2	7		
		4.3	8
5.0	9.5	5.0	9.5
		5.1	11
		5.4	12
		5.8	13
6.3	14		
		6.6	15
6.8	16		
6.9	17		
7.1	18		
7.4	19.5		
7.4	19.5		
7.7	21		
8.2	22		
8.7	23		
9.4	24		
9.5	25		
11.7	26		
11.9	27		

Rank Total 296.5= S

The test statistic is $T = S - \frac{n_1 (n_1 + 1)}{2} = 296.5 - \frac{17 \text{ x } 18}{2} = 143.5$

Conclusion : From Table-9 we have, for $n_1 = 17$ $n_2 = 10$ and $\alpha/2 = 0.025$, $W\alpha/2 = 46$ and $W(1-\alpha/2) = 17 \times 10 - 46 = 124$. Since the calculated value of T is greater than 124 we may reject H_0 and we conclude that the two population location parameters are different.

Large Sample Approximation

When the sample sizes are greater than 20 i.e if n_1 and $n_2 > 20$. We cannot use Table-9 It is seen that for large sample sizes i.e n_1, $n_2 > 20$, the distribution of T is approximately normal with mean $= \frac{n_1 n_2}{2}$

and variance $= \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$

Therefore, from central limit theorem, we have,

$z = \frac{T - n_1 n_2/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}$ which is distributed as standardised normal variate i.e N (0.1). Therefore, the conclusion can be made as usual by considering the critical value of $z = \pm 1.96$ at 0.05 level of significance and $z = \pm 2.58$ at 0.01 level of significance.

Correction for tied sample values

Ties may occur within or between the group observations. We may use the correction factor in the z value to have more appropriate test statistic. Let 't' be the number of ties in a given rank

Then correction factor for ties is $\frac{n_1 n_2 (\Sigma t^3 - \Sigma t)}{12 (n_1 + n_2)(n_1 + n_2 - 1)}$ which is subtracted from the variance term in the z value. So the adjusted z value is as follows :

$$z_{Adj} = \frac{T - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 - 1)}{12} - \frac{n_1 n_2 (\Sigma t^3 - \Sigma t)}{12 (n_1 + n_2)(n_1 + n_2 - 1)}}}$$

which is also N (0.1) variate.

For a selected value of α, the level of significance, the conclusion can be made as usual.

17.4.5 The Mood Test

In parametric test, F test is usually used to test the null hypothesis that the two population dispersion parameters designated by σ_1^2 and σ_2^2 are equal. To support this test, there is a strict assumption that the populations from which the two samples drawn, are assumed to be normal. If this assumption is not satisfied, we usually go forward to test the above hypothesis by the Mood test. This test is distribution free and assumes only that the two populations' medians are equal.

Assumptions

(1) The data consist of two random samples $x_1, x_2, ---, x_{n_1}$ of size n_1 and $y_1, y_2, ----, y_{n_2}$ of size n_2 from population 1 and 2 respectively where $n_1 \leq n_2$.

(2) The population distributions are continuous.

(3) The two samples are independent.

(4) The data are measured on at least an ordinal scale.

(5) The two populations are identical with equal median except for a possible difference in dispersion.

Hypotheses

Let σ_1 be the dispersion parameter of population 1 and σ_2 be the dispersion parameter of population 2. We may consider the following type of hypotheses.

a. (Two sided) $H_0 : \sigma_1 = \sigma_2$ $H_1 : \sigma_1 \neq \sigma_2$

b. (One sided) $H_0 : \sigma_1 \leq \sigma_2$ $H_1 : \sigma_1 > \sigma_2$

c. (One sided) $H_0 : \sigma_1 \geq \sigma_2$ $H_1 : \sigma_1 < \sigma_2$

We select α as the level of significance.

Note : σ should be interpreted not as population standard deviation but as a general measure of dispersion .

Test Statistic

The test statistic is given by—

$$M = \sum_{i=1}^{n_1} \left(r_i - \frac{N+1}{2} \right)^2$$ where $N = n_1 + n_2$ and r_i is the rank of the ith observation of X in the combined rank of both the samples. $\frac{N+1}{2}$ is the mean of the combined rank. The quantity in the parentheses is nothing but the deviation of the ith rank of X observation from the mean rank of the combined sample. This has some similarity to the term $\Sigma(x_i - \overline{x})^2$ in the total sum of squares. For that we can deduce that if the Y's are dispersed relative to the X's (i.e. if $\sigma_1 < \sigma_2$) M tends to be small, on the other hand if the X', s are dispersed relative to the Y's (i.e. if $\sigma_1 > \sigma_2$), M tends to be large. For example, if we have observation Y Y X X X Y Y having rank 1, 2, 3, 4, 5, 6, 7 then $M = (3-4)^2 + (4-4)^2 + (5-4)^2 = 2$. Again if we have observations X Y Y Y Y X X having rank 1, 2, 3, 4, 5, 6, 7, then $M = (1-4)^2 + (6-4)^2 + (7-4)^2 = 22$. Therefore, we can say, that sufficiently small value of M causes us to reject H_0 in (a). and (c) and sufficiently large value of M causes us to reject H_0 in (a). and (b). Intermidiate values of M indicate that we may not reject H_0 in any case. Thus the objective of this test is to determine whether M is sufficiently small or large to cause us to reject H_0.

Conclusion can be made as follows :

We consult Table-10 where for different n_1 and n_2 and for α a chosen level of significance there are four values under a column leveled α, 1st one is M' 2nd one is α' 3rd one is M'' and 4th one is α''. The relationship $\alpha' \le \alpha \le \alpha''$ holds.

(a) For two sided test, reject H_0 at a level of significance α if

(1) the computed $M \le M'$ (or M'') for α' (or α'') closer to $\alpha/_2$ in a column headed by α.

(2) if the computed $M \geq M''$(or M') for α' (or α'') closer to $(1- \alpha/2)$ in a column headed by $(1- \alpha)$ value that is equal to $(1-\alpha/2)$.

(b) For one sided test, with $H_1 : \alpha_1 > \alpha_2$, reject H_0 at a level of significance α if the computed value of $M \geq M'$ (or M'') for α' (or α'') closer to α in a column leveled $(1-\alpha)$.

(c) For one sided test, with $H_1 :.\alpha_1 < \alpha_2$ reject H0 at a level of significance α if the computed value of $M \leq M'$ (or M'') for α' (or α'') closer to α in the column lev[illegible]ed α.

Example 17.12 Two groups of hens were given a certain improved feed and after 7 days thier increase in weights (gm) were recorded as below :

Group-1 (x) :	3.84	2.60	1.19	2.00	
Group-2 (y) :	3.97	2.50	2.70	3.36	2.30

Test whether the dispersion with respect to variable are same in the population. Use α =.05.

Solution : We set up, $H_0 : \sigma_1 = \sigma_2$, $H_1 : \alpha_1 \neq \alpha_2$.

To calculate the test statistic we at first find out the rank of the combined sample as follows :

Observation : in order	1.19	2.00	2.30	2.50	2.60	2.70	3.36	3.84	3.97
Group :	X	X	Y	Y	X	Y	Y	X	Y
Rank :	1	2	3	4	5	6	7	8	9

Mean of the rank $= \frac{N+1}{2} = \frac{9+1}{2} = 5$ and the test statistic is

$$M= (1-5)^2+ (2-5)^2 + (5-5)^2 + (8-5)^2 = 34$$

Conclusion : From Table-10 for $n_1 = 4$ and $n_2 = 5$ and $\alpha = 0.05$ (Two sided test) shows the critical value to be $M'= 9$ and $M''=42$. Since $9 < 34 < 42$, we may accept the H_0 and conclude that the two population dispersion parameters may be equal.

17.4.6. Wilcoxon Matched- pairs Signed-ranks Test :

The sign test for two related samples uses the only information whether x is larger, smaller or equal to y. But if the data provide more information regarding the magnitude as well as the direction of the differences between x and y, the sign test is not at all powerful, in this situation.Wilcoxon matched- pair signed-rank test fulfills the need for the above cases. It is appropriate when we determine the amount of any difference between pairs of observations x_i and y_i as well as the direction. When the magnitude of the differences are determined we can rank them and Wilcoxon test utilises this additional information.

Assumptions

(1) The data for analysis consist of n values of differences $D_i = (y_i - x_i)$, each pair measurements (x_i, y_i) is taken on the same subject.

(2) The sample of (x_i, y_i) pair is random.

(3) The differences represent observations on a continuous random variable.

(4) The distribution of the population of differences is normal.

(5) The differences are independent.

(6) The differences are measured on an interval scale.

Hypotheses

(a) Two sided

H_0 : The median of the population of differences is zero.

H_1 : The median of the population of differences is not zero.

(b). One sided

H_0 : The median of the population of differences is less than or equal to zero.

H_1 : The median of population of differences is greater than zero.

(c). (One sided)

H_0 : The median of the population of differences is greater than or equal to zero.

H_1 : The median of the population of differences is less than zero

We select α as the level of significance.

Test Statistic

The procedure of calculating the numerical value of the test statistic is as follows :

i) Obtain each of the signed differences, $D_i = (y_i - x_i)$.

ii) Rank the absolute value of D_i from the smallest to the largest value. That is, rank $|D_i| = |y_i - x_i|$.

iii) Assign the signs of the differences to the ranks.

iv) Calculate –

T+ = the sum of the ranks with + signs and T - = the sum of the ranks with – signs .

Finally T + or T- is the value of the test statistic depending on the hypotheses.

Ties : There are two types of ties, one or both may occur in a given problem.

a) The first type occurs when $D_i = y_i - x_i = 0$, and the number of such pairs are reduced from n, the sample sizes in case of analysis.

b) The other type occurs when two or more values of $|D_i|$ are equal. For ties of this type, the $|Di|$'s receive the average of the ranks that otherwise would be assigned to them.

The conclusion can be made as follows :

(a) **Two sided case :** A sufficiently small value of T + or T- will cause the rejection of H_0, that the median of the population of differences is equal to zero. Therefore, in this case, the value of the test statistic is T + or T- whichever is smaller. We consult Table-11 and reject H_0 at the α level of significance if T is smaller than or equal to d for n and tabulated α (two sided).

(b) **One sided :** Small value of T- will cause the rejection of H_0 that the median of the population of differences is less than or equal to zero. Reject H_0 at the α level of significance if T- is smaller than or equal to d for n and α (one sided) from Table-11.

(c) **One sided :** Small value of T + will cause the rejection of H_0 that the median of the population of differences is greater than or equal to zero. We reject H_0 at the α level of significance if T + is smaller than or equal to d for n and α (one sided) from Table-11

Example 17.13 The yields of a certain paddy vareity per plot are given at control period as well as after using fertiliser in plots of an agricultural farm

Plot No	:	1	2	3	4	5	6	7	8	9
Yield (using fertiliser x)	:	33	17	33	25	36	25	31	20	18
Yield (at control y)	:	21	17	22	13	33	20	19	13	9

Test whether the median of the population of differences is zero at 0.05 level of significance.

Solution : We set up,

H_0 : The median of the population of differences is zero.

H_1 : The median of the population of differences is not zero

To calculate the value of the test statistic we proceed to prepare a table of the following type :

Plot No.	Yield (x) (Using fertiliser)	Yield (y) (Control)	$D_i = y_i - x_i$	$\lvert Di \rvert$	Rank of $\lvert D_i \rvert$	Signed Rank
1	33	21	- 12	12	7	- 7
2	17	17	0	Omit from analysis		
3	30	22	- 8	8	4	- 4
4	25	13	- 12	12	7	- 7
5	36	33	- 3	3	1	- 1
6	25	20	- 5	5	2	- 2
7	31	19	- 12	12	7	- 7
8	20	13	- 7	7	3	- 3
9	18	9	- 9	9	5	- 5

Here the reduced sample size is 9 - 1= 8. The, test statistic, T = 0 (since no value of differences has + sign).

Conclusion : From Table-11 for n= 8 and at 0.05 level of significance (Two sided) the value of d = 5.

Since the value of T is smaller than d we may reject the null hypothesis that the median of the population differences is zero.

Large Sample Approximation :

When the sample size, n is greater than 25, we cannot use Table-11. for n > 20 ; we may calculate.

$$z = \frac{T - [n (n - 1)]/4}{\sqrt{n (n + 1) (2n + 1)/24}}$$ which is approximately N (0.1) variate.

Therefore, the conclusion can be made as usual by considering the critical value ± 1.96 at 0.05 level of significance and ± 2.58 at 0.01 level of significance.

17.4.7 The Kolmogorov- Smirnov Test

This test is used for testing the null hypothesis that the two independent samples come form the populations that are identical with respect of location parameters and dispersion. This test is more general since it is sensitive to the differences of all types that may exist between two distributions.

Assumptions

(1) The data for analysis consist of two independent random samples designated as $x_1, x_2, \ldots\ldots\ldots, x_{n_1}$ and $y_1, y_2, ---, y_{n_2}$ of sizes n_1 and n_2 respectively.

(2) The data are measured on at least an ordinal scale.

Hypotheses

We define $F_1(x)$ and $F_2(y)$ as the unknown population distribution functions of x's and y's respectively. We can test the following two sided or one sided hypothesis.

(a) Two sided

$H_0 : F_1(x) = F_2(y)$ for all x and y.

$H_1 : F_1(x) \neq F_2(y)$ for at least one value of x or y.

(b) One sided

$H_0 : F_1(x) \leq F_2(y)$ for all x and y.

$H_1 : F_1(x) > F_2(y)$ for at least one value of x or y.

(c). One sided

$H_0 : F_1(x) \geq F_2(y)$ for all x and y.

$H_1 : F_1(x) < F_2(y)$ for at least one value of x or y.

We select α as the level of significance.

Test Statistic

Let $S_1(x)$ and $S_2(y)$ be the sample distribution functions of the observed x´s and the observed y´s respectively. We define—

$$S_1(x) = \frac{\text{number of observed x´s} \leq x}{n_1} \quad \text{and}$$

$$S_2\,(x) \;=\; \frac{\text{number of observed } y\text{'s} \leq y}{n_2}$$

We first arrange the data of two samples in order of magnitude from the smallest to the largest. The test statistic for three sets of hypotheses are as follows :

(a). Two sided : $D = \text{Maximum} \mid S_1\,(x) - S_2\,(y) \mid$

(b). One sided : $D^+ = \text{Maximum}\,[S_1\,(x) - S_2\,(y)]$

(c). One sided : $D^- = \text{Maximum}\,[S_2\,(y) - S_1\,(x)]$

The conclusion can be made as follows :

We consult Table-13 for $n_1 = n_2$ and Table-14 for $n_1 \neq n_2$.

Reject H_0 at the α level of significance if the appropriate test statistic D or D^+ or D^- exceeds the value given in the quantile column (1-α).

Example 17.14 The following data give the basal metabolic rates (ml oxygen/mnt) on 5 non-athletic males (x) and 6 athletic males (y):

x	y
206	236
238	209
224	287
257	276
230	252
	251

Test whether the two populations from which two samples have been drawn have same distribution functions. Consider $\alpha = 0.05$ as the level of significance.

Solution : We set up, $H_0 : F_1\,(x) = F_2\,(y)$ $H_1 : F_1(x) \neq F_2\,(y)$

where F_1 and F_2 represent two population distribution functions of x and y respectively.

To calculate the test statistic, we arrange the data given in two samples in order of magnitude from the smallest to the largest, calculate the sample distribution function and finally the differences between the sample distribution functions—

x_i	$S_1(x)$	y_i	$S_2(y)$	$\lvert S_1(x) - S_2(y) \rvert$
206	$\frac{1}{5}$			$\lvert \frac{1}{5} - 0 \rvert = \frac{1}{5} = \frac{6}{30}$
		209	$\frac{1}{6}$	$\lvert \frac{1}{5} - \frac{1}{6} \rvert = 1/30$
224	$\frac{2}{5}$			$\lvert 2/5 - 1/6 \rvert = 7/30$
203	$\frac{3}{5}$			$\lvert \frac{3}{5} - \frac{1}{6} \rvert = 13/30$
		236	$\frac{2}{6}$	$\lvert \frac{3}{5} - \frac{2}{6} \rvert = 8/30$
238	$\frac{4}{5}$			$\lvert \frac{4}{5} - \frac{2}{6} \rvert = 14/30$
		251	$\frac{3}{6}$	$\lvert \frac{4}{5} - \frac{3}{6} \rvert = 9/30$
		252	$\frac{4}{6}$	$\lvert \frac{4}{5} - \frac{4}{6} \rvert = 4/30$
257	$\frac{5}{5}$			$\lvert \frac{5}{5} - \frac{4}{6} \rvert = 10/30$
		276	$\frac{5}{6}$	$\lvert \frac{5}{5} - \frac{5}{6} \rvert = 5/30$
		278	$\frac{6}{6}$	$\lvert \frac{5}{5} - \frac{6}{6} \rvert = 0$

From the column given by $\lvert s_1(x) - s_2(y) \rvert$ we have seen that the maximum of $\lvert s_1(x) - s_2(y) \rvert$ is $D = \frac{14}{30} = 0.467$.

Conclusion : From the Table-14 we have seen for $n_1 = 5$ and $n_2 = 6$ at (1-0.95) =0.05 level of significance, the tabulated value is 2/3 =

0.667. Since the calculated value of D is less than the tabulated values, H_0 may be accepted.

17.4.8. Fisher's Exact Probability Test

It may happen like this that the individual observation of a sample may come from any of the two mutually exclusive classes and there are only two independent samples of same type of classification. For example, we get such data when we compare two treatments and classify the observations as either responding or not responding. We may draw independent random samples from each of the two populations and thus classify the observations as above. We can arrange the data in a 2 x 2 contingency table as follows :

Sample No	Responding	Non-responding	Total
1	a	A-a	A
2	b	B-b	B
Total	(a + b)	A+B- (a+b)	A+B

We must arrange the data in such a way that $A \geq B$ and $\frac{a}{A} \geq \frac{b}{B}$.

The objective of the study is to test whether the two populations from which the two samples have been drawn differ with respect to the proportion of observations that fall into the two classifications. When the sample observations are very small that we cannot perform χ^2–test for testing independence of attributes in a 2 x 2 contingency table we use Fisher's exact probability test.

Assumptions

(1) The data consist of A number of sample observations from population–1 and B number fo sample observations from population–2.

(2) The samples are random and independent.

(3) Each observation can be classified as one of the two mutually exclusive types.

Hypotheses

(a). Two sided

H_0 : The proportion of the characteristic of interest is the same in both the populations.

H_1 : The proportion of the characteristic of interest is not the same in both the populations.

(b). One Sided

H_0 : The proportion of the characteristic of interest in populations 1 is less than or the same as the proportion in population–2.

H_1 : The proportion of the characteristic of interest in population-1 is greater in population–1 than in population–2

We select α as the level of significance.

Test Statistic

The test statistic is b, the number of observations in sample–2 with the characteristic of interest.

Conclusion can be made as follows :

Table-15 gives the critical value of b for A and B between 3 and 20 at different levels of significance 0.05, 0.025, 0.001, 0.005.

(a) Enter Table-15 with A, B and a, if the observed value of b is equal to or less than the integer in a given column (at different level of significance), reject H_0 at a level of significance equal to the twice the level of significance shown at the top at the column. For example, We have A=8, B=7, a=7 and observed value of b=1. From Table-15 we can reject H_0 at the 2(0.05)=0.1, the 2(0.025) = 0.05 and the 2(0.01)= 0.02 level of significance but not at the 2(0.005)=0.01 level of significance.

(b) Enter Table-15 with A, B and a, if the observed value of b is less than or equal to the integer in a given column, reject H_0 at the level of significance shown at the top of the column. For example, we have A=16, B=8, a=14 and the observed value of b=3, we can reject H_0 at 0.05 and 0.025 level of significance but not at the 0.01 or 0.005 level of significance.

Example 17.15. The following data give the residential location of two social groups and the electoral relationship in a certain period—

Residential location of	Electoral relationship High	Low	Total
Group-1	9	1	10
Group-2	2	3	5

Test the dependency of two attributes at $\alpha = 0.05$, level of significance.

Solution : We set up,

H_0 : The proportion of location with high electoral relationship within class groups is the lower in population-1 than in population-2.

H_1 : The proportion is higher in population–1 than in population-2.

The data in the table show that A=10, B=5, a=9 and b=2. Therefore, the test statistic is, b=2.

Conclusion : We enter Table-15 with A=10, B=5 and a=9 we find that the critical value of b=1 at 2 (0.025) = 0.05 level of significance, Since the value of the test statistic b=2 is greater than the tabulated value of b, we therefore, may accept H_0 at 0.05 level of significance.

Large Sample Approximation

For a sufficiently large samples, that is, when A and B are both greater than 20, we should not use the above test. In that case, for testing the null hypotheses of the equality of two population proportions we can proceed by using the normal approximation as follows :

$$z = \frac{\frac{a}{A} - \frac{b}{B}}{\sqrt{\hat{p}\left(1 - \hat{p}\right)\left(\frac{1}{A} + \frac{1}{B}\right)}} \quad \text{where } \hat{p} = \frac{a + b}{A + B}$$

z is usually N (0.1) variate and the conclusion can be made as usual by considering the critical value ± 1.96 at 0.05 level of significance and the critical value ± 2.58 at 0.01 level of significance.

The use of normal approximation is considered to be satisfactory if a, b, (A-a) and (B-b) are all greater than or equal to 5. Alternatively when all the cell frequencies i.e a, b, (A-a) and (B-b) are greater than 5, we can use the χ^2 test for 2 x 2 contingency table as we have done for testing the independence of attributes in **3.b of Section 10.6**

17.4.9 The Wald-Wolfowitz Runs Test :

In one sample run test, we defined run as a sequence of like events, items or symbols that is preceeded or followed by an event, item or symbol of a different type or by none at all. Here in the Wald-Wolfowitz run test, we use the number of runs present in the data of two samples to test the null hypothesis that the sample have come from identical population in any respect of location, dispersion or skewness.

Assumptions

(1) The data consist of observations $x_1, x_2, \ldots\ldots\ldots, x_{n_1}$ of size n_1 and $y_1, y_2, \ldots\ldots\ldots, y_{n_2}$ of size n_2 giving us random samples drawn from population-1 and population-2 respectively. n_1 may or may not be equal to n_2.

(2) The two samples are independent.

(3) The variable of interest is continuous.

Hypotheses

H_0 : The x´s and y´s have come from identically distributed population.

H_1 : The population of x´s and the population of y´s are not identically distributed.

Test Statistic

The definition of a run is already given, The test statistic is r, the number of runs in a complete set of data when both the samples combined are arranged in order of magnitude. In the ordering, we must keep track of the sample to which each observation belongs.

When there are ties between the observations of the samples i.e. one or more x values are equal to one or more y values, we must apply some methods of dealing with the problem. One simple approach is to prepare two ordered arrangements one resulting in the fewest number of runs r´ and the other resulting in the largest number of runs r´´, finally the test statistic, $r = \frac{r' + r''}{2}$. If there are ties within the samples x or y, we do not present any problem.

The rationale underlying the Wald-Wilfowitz run test is that if the samples of x's and y's have come from identically distributed populations, we expect them to be well mixed and thus the number of runs is relatively large. Therefore, large value of r will support the null hypothesis H_0, sufficiently small value of r will reject the H_0 concluding that the two populations from which the two samples have been obtained are not identically distributed. This indicates that the location parameters, dispersions or the degree of skewness are different in the two populations.

Conclusion can be made as below :

Reject H_0 at the 0.025 level of significance if the calculated value of r is less than or equal to the tabulated value of r for n_1 and n_2 as given in Table-6.

Example 17.16. Following are the two samples of marks in a subject out of 100 of two sections of a school in annual examination :

Group-1 (x) : 68, 67, 58, 62, 55, 60, 67.

Group-2 (y) : 60, 59, 72, 73, 56, 53, 43, 50, 65, 56, 56, 56, 57, 37

Test whether the two populations from which the above two samples have been obtained are identically distributed. Use α = 0.025 as the level of significance.

Solution : We set up,

H_0 : The two samples have come from identically distributed population.

H_1 : The two samples have not come from identically distributed population.

Here we have seen that there is only one tie between the marks of two samples. That is, the mark 60 is seen in both the sets. Therefore, we have two ordered arrangements which are as follows :

Ordered arrangement for r´

37,	43,	50,	53,	55,	56,	56,	56,	56,	57,	58,	59,	60
y	y	y	y	x	y	y	y	y	y	x	y	y

60,	62,	65,	67,	67,	68,	72,	73
x	x	y	x	x	x	y	y

Here r´ = 9.

Ordered arrangement for r´´

37,	43,	50,	53,	55,	56,	56,	56,	56,	57,	58,	59,	60,
y	y	y	y	x	y	y	y	y	y	x	y	x

60,	62,	65,	67,	68,	72,	73
y	x	y	x	x	y	y

Here r'' =11.

Therefore, the test statistic, $r = \frac{r_1 + r_2}{2} = \frac{9 + 11}{2} = 10.$

Conclusion : From Table-6 for $n_1 = 7$ and $n_2 = 14$ at 0.025 level of significance, the tabulated value of r = 5. Since the calculated value of r = 10, which is greater than the tabulated value, therefore, H_0 may be accepted.

17.4 10. χ^2 –test of Homogeneity

In parametric test we tackled the testing of null hypothesis

$H_0 : p_1 = p_2$ where p_1 and p_2 are two population proportion **(vide 1.d of Section 10.6)**. This H_0 may be alternatively stated as the two populations are homogeneous with respect to the proportion of subjects possessing some characteristic of interest.

Here a sample is drawn from each of the populations under study and the subjects are classified according to whether they possess the characteristic of interst. The result is usually displayed in a 2 x 2 contingency table which is almost same as given in Example 10.14 and the testing procedure is also same as χ^2-test given there. We calculate χ^2 for both the test of homogeneity and the test for independence by using the same formula. The test of homogeneity differs in two respects : (1) sampling procedure (2) the rationale underlying the calculation of expected frequencies. Here we test the null hypothesis that the two populations represented by the two samples are homogeneous by comparing the calculated value of χ^2 with the tabulated value of χ^2 with 1.d. f. This test is defined by the χ^2-test for homogeneity. The above testing procedure for the 2 x 2 contingency table is a special case of the general case of r x c contingency table where r indicates the number of rows and c indicates the number of columns.

Assumptions

(1) The samples are independent.

(2) The samples are random.

(3) Each subject in the population may be classified into one of the two mutually exclusive categories, according to whether it has or does not have the characteristic of interest.

Hypothesis

H_0 : The sampled populations are homogeneous.

H_1 : The sampled populations are not homogeneous.

We select α as the level of significance.

Test Statistic

The test statistic is $$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$= \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{O_{ij}^2}{E_{ij}} - n \qquad \text{............}(17.5)$$

where O_{ij} is the observed frequency corresponding to the ith row and the jth column of the contingency table and E_{ij} is the expected frequency corresponding to O_{ij} usually $E_{ij} = \frac{n_{i.} \times n_{.j}}{n}$; $n_{i.}$ is the marginal total of the ith row and $n_{.j}$ is the marginal total of the jth column and $n = \sum_i \sum_j O_{ij} = \sum_i \sum_j E_{ij}$. (17.5) is distributed as χ^2 with (r-1) (c-1) d. f.

For a 2 x 2. contingency table with cell frequencies $\frac{a}{c}\Big|\frac{b}{d}$

the calculated value of $\chi^2 = \frac{(ad - bc)^2 n}{(a + b)\ (c + d)\ (a + c)\ (b + d)}$

with 1 d.f where $n = a + b + c + d$.

The conclusion can be made as follows :

Reject H_0 if the calculated value of χ^2 with appropriate d. f. is greater than or equal to the tabulated value of χ^2 with same d. f at α level of significance.

Example 17.17. To determines the public awareness and concern of the family planning procedures a sample of 40 couples of each of three village areas of Dhaka district were interviewed. The responses are given below :

	Opinion				
Area of Residence	No.	Yes	Doubtful	Don't Know	Total
A	5	31	2	2	40
B	10	21	4	5	40
C	11	20	7	2	40
Total	26	72	13	9	120

Test at 0.05 level of significance whether the responses of three areas are homogeneous.

Solution : We set up,

H_0 : The three populations of residents are homogeneous with respect to knowledge of family planning.

H_1 : The three populations are not homogeneous.

We calculate the expected cell frequencies as follows :

	Opinion				
Area of Residence	No.	Yes	Doubtful	Don't Know	Total
A	8.67	24	4.33	3	40
B	8.67	24	4.33	3	40
C	8.67	24	4.33	3	40
Total	26.01	72	12.99	9	120

We then calculate O_{ij}^2 / E_{ij} corresponding to different cells as follows :

	Opinion				
Area of Residence	No.	Yes	Doubtful	Don't know	Total
A	2.88	40.04	0.92	0.44	44.28
B	11.53	18.38	3.70	8.33	41.94
C	13.96	16.67	11.32	0.44	42.39
Total					128.61

$$\therefore \chi^2 = \sum_{i}^{r}\sum_{j}^{c} O_{ij}^2/E_{ij} - n = 128.61 - 120 = 8.61.$$

Conclusion : The tabulated value of χ^2 with 6 d. f. at 0.05 level of significance is 12.592. Since the calculated value of χ^2 is smaller than the tabulated value of χ^2 with same d. f. we can not reject the H_0. Therefore, we may conclude that the populations are homogeneous with respect to the knowledge of family planning procedures.

The r x 2 Contingency table : When we proceed for χ^2- test of homogeneity to the case of three or more populations whose subjects can be classified into one of two mutually exclusive categories, the data can be displayed in a r x 2 contingency table where the r rows represent the r populations and the two columns represent the two classifications of the characteristic of interest. The testing procedures are same as given in Example 17.17. Here the d.f of χ^2 test statistic is (r-1). The conclusion can be made as usual.

17.4.11. The Kendall's Co-efficient of Concordance.

In Spearman rank correlation co-efficient, r and in Kendall's τ, we are interested with the extent to which two sets of ranking of n individuals agree or disagree. In many practical cases we may also be interested in the degree of agreement among several say m sets of rankings of n individuals. Such a situation may occur in two ways as follows :

(1) We may rank a group of n individuals on the basis of each of m characteristics. For example, n= 6 students may be ranked according to thier marks obtained in m= 4 subjects is a certain examination. The result can be displayed in the following table :

	Students					
Subjects	A	B	C	D	E	F
Bengali	2	5	1	4	3	6
English	1	3	2	4	6	5
Mathematics	4	2	3	5	1	6
Gn. Science	5	3	2	1	6	4

(2) A group of m Judges may rank a group of n individuals on the basis of the same charactristic. For example, a group of m=3 judges may rank n=5 employees in the basis of leadership ability. The result can be displayed in the following table :

	Employees				
Judge	A	B	C	D	E
1	2	1	3	5	4
2	3	2	4	1	5
3	5	1	3	2	4

For the above situations, we would like to have a measure of the strength of the agreement among the m sets of rankings. For this we are to test null hypothesis of no association among the rankings. We can achieve the above objective by using Kendall's co-efficient of concordance.

Assumptions

(1) The data consist of m complete sets of observations on n individuals.

(2) The measurement scale is at least ordinal

(3) The observations as collected may consist of ranks. If the original data are not rank, they must be capable of being converted to ranks.

Hypothesis

H_0 : The m sets of rankings are not associated.

H_1 : The m sets of rankings are associated.

We consider α as the level of significance.

Test Statistic

The most convenient form of the test statistic is

$$W = \frac{12 \sum_{j}^{n} R_j^2 - 3\, m^2 n\, (n + 1)^2}{m^2 n\, (n^2 - 1)} \qquad \text{..........(17.6)}$$

where m is the number of sets of ranking, n is the number of individuals and R_j is the sum of the ranks assigned to the jth individual.

The conclusion can be made as follows :

For values of m and n, we consult Table-16 and reject H_0 at α level of significance if the value of probabitity in Table-16 associated with the appropriate values of W, m and n is less than or equal to α.

Example 17.18 Hand writing specimen of 5 competitors were judged by 4 judges. The result is displayed as below :

	Competitors				
Judges	A	B	C	D	E
1	4	3	2	5	1
2	3	4	5	2	1
3	4	3	1	5	2
4.	4	3.	5	2	1

Test whether there is no agreement among the judges. use $\alpha = 0.05$.

Solution : We set up,

H_0 : The rankings of the judgement are not associated.

H_1 : The rankings of the judgement are associated.

To calculate the test statistic we proceed as follows :

The rank totals of different competitors are :

$R_A=15$, $R_B=13$, $R_C=13$, $R_D=14$, $R_E=5$.

We know, m=4 n=5, therefore,

$$W = \frac{12\,(15^2 + 13^2 + 13^2 + 14^2 + 5^2) - 3 \times 4^2 \times 5\,(5+1)^2}{4^2 \times 5\,(5^2 - 1)}$$

$$= \frac{9408 - 8640}{1920} = \frac{468}{1920} = 0.4$$

Conclusion : From the Table-16 for m=4 and n=5 the probability of W = 0.4 is approximately 0.119, which is greater than $\alpha = 0.05$, therefore, we can not reject the null hypothesis.

Large Sample Approximation

For $m > 5$ we cannot use Table-16. Also for $n > 7$ the large sample approximation to the formula given in (17.6) is $\chi^2 = m\,(n-1)\,W$ which is distributed as χ^2 with (n-1) d.f. So for $m > 5$ and $n > 7$ we can use the above approximation.

Ties : In a set of observations to be ranked, if two or more observations are equal, we assign each the mean of the rank positions for which it is tied. We adjust the test statistic for ties by replacing the demoninator of W in (17.6) by $m^2n\,(n^2-1) - m\,\Sigma(t^3-t)$, where t is the number of observations in any set of ranking tied for a given rank. This adjustment does not usually effect the large sample approximation.

17.4.12 The Kruskal-Wallis Test

It is the most widely used non-parametric test when we are to test null hypothesis that several samples have been drawn from identical populations is the Kruskal-Wallis one-way analysis of variance by ranks. In short, it is usually called Kruskal-Wallis test. When we deal with two samples only the Kruskal-Wallis test is equivalent to the Mann - Whitney test, discussed is 17.4.4.

The Kruskal-Wallis test uses more information than the Median test as a result it is usually more powerful and is preferred when the available data are measured on at least the ordinal scale.

Assumptions

(1) The data for analysis consist of k random samples of sizes n_1, n_2, ..., n_k. Some or all the sample sizes may also be equal.

(2) The observations are independent both within and between samples.

(3) The variable of interest is continuous.

(4) The measurement scale is at least ordinal.

(5) The populations are identical except for a possible difference in location for at least one population.

Hypothesis

H_0 : The k population distribution functions are identical.

H_1 : The k populations do not all have the same median.

We select α as the level of significance.

Test Statistic

The data for analysis may be displayed as below :

Samples

1	2	3------------------------k
x_{11}	x_{21}	x_{31}------------------x_{k1}
x_{12}	x_{22}	x_{32}------------------x_{k2}
-	-	- -
-	-	- -
x_{1n_1}	x_{2n_2}	x_{3n_3}------------------x_{kn_k}

Now we replace each original observation by its rank relative to all the observations in the k samples. Let us suppose that N= Σn_i be the total number of observations in the k samples, we assign the rank 1 to the smallest observation, the rank 2 to the next in size and so on to the largest which is given by rank N. In case of ties , we assign the tied observations the average of the ranks that would be assigned if there were no ties. Now let R_i be the sum of the ranks corresponding to the ith sample, then the test statistic is

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{1}{n_i} \left[R_i - \frac{n_i (N+1)}{2} \right]^2 \quad \text{.........(17.7)}$$

which is a weighted sum of squares of deviations of sums of ranks from the expected sum of ranks using reciprocals of sample sizes as the weights. A more convenient way of calculation of the above form (17.7) is given by---

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1) \quad \text{.........(17.8)}$$

Conclusion can be made as follows :

When we deal with three samples each with 5 or fewer observations we can compare the calculated value of H with that of given in Table-17 and reject H_0 if the calculated value of H is equal or smaller than the tabulated value of H at α level of significance.

For more than three samples with each more than 5 observations we usually compare the calculated value of H with tabulated value of χ^2 with (k-1) d.f where k is the no of samples considered in the analysis.

Example 17.19 : The following data give the yields is kg of three different varieties of paddy grown on an almost homogeneous plot of land.

Variety-1 : 262, 307, 211, 323, 454, 339, 304, 154, 287, 356.

Variety-2 : 465, 501, 455, 355, 468, 362.

Variety-3 : 343, 772, 207, 1048, 838, 687.

Test at 0.05 level of significance whether the above three samples have been obtained from three identical distributions.

Solutions : We set up,

H_0 : The three populations represented by the data are identical.

H_1 : The three populations do not have the same median.

To calculate the test statistic we assign ranks to the original data as follows:

											Rank sum
Var-1:	4	7	3	8	14	9	6	1	5	12	R_1=69
Var-2:	16	18	15	11	17	13					R_2=90
Var-3:	10	20	2	22	21	19					R_3=94

We know, N= 22 and the test statistic is

$$H = \frac{12}{22\,(22+1)}\left[\frac{69^2}{10} + \frac{90^2}{6} + \frac{94^2}{6}\right] - 3\,(22+1) = 9.232.$$

Since the sample sizes all exceeds 5, we must use χ^2 table for conclusion.

Conclusion : The calculated value of H is seen to be greater than the tabulated value of $\chi 2$ (5.991) with 2 d.f at 0.05 level of significance. Therefore, the null hypothesis may be rejected.

Correction for Ties

If there are a number of ties, we may want to adjust the test statistic. The adjustment factor is $1-\frac{\Sigma(t^3 - t)}{N^3 - N}$ where t is the number of tied observations in a group of tied scores. The adjusted test statistic is

$$H_{Adj} = \frac{H}{1 - \Sigma(t^3 - t) \,/\, (N^3 - N)} \qquad \text{............(17.9)}$$

where H is given in (17.8). From (17.9) we have seen that the adjustment factor inflates the value of the test statistic. Therefore, if H is significant at a desired level of significance α, there is no point of calculating H_{Adj}.

17.4.13 The Durbin Test

In design of experiment, when the number of plots (k) per block is smaller than the number of treatments (t), we usually proceed with incomplete block designs (IBD). When we can adopt a design where the number of plots per block are same (k, $k < t$), each pair of treatments is occuring a constant number of times (λ) and each treatment is replicated equal number of times say r in the experiment, we get a balanced incomplete block design (BIBD). In an experiment with such type of design we have performed the analysis of variance to test the parametric hypotheses that there is no difference among the treatment means subject to the condition that the population distribution of the treatments should obey some assumptions. When the distributions do not meet these assumptions we proceed by the Durbin test to test the above null hypothesis.

Assumptions

(1) The blocks are mutually independent of each other.

(2) The observations within each block may be ranked in order of magnitude.

Hypothesis

H_0 : The treatments have equal effects.

H_1 : The response to at least one treatment tend to be larger than the response to at least one other treatment.

We select α as the level of significance.

Test statistic

The data of the experiment can be systematically written in a (Block x Treatment) table and then rank the observations in the block from the smallest to the largest. We assign tied observations the mean of the rank positions for which they are tied. A moderate number of ties does not greatly effect the result.

The test statistic for the Durbin test is,

$$T = \frac{12\,(t-1)}{rt\,(k-1)\,(k+1)} \sum_{i=1}^{t} R_i^2 - \frac{3r\,(t-1)\,(k+1)}{k-1}$$

where t= the number of treatments under investigation,

k = the number of plots per block,

r = the number of replications of the treatment,

R_i = the sum of the ranks of the ith treatment.

The conclusion can be made as follows :

Reject H_0 at the selected level of significance α if the calculated value of T is greater than the tabulated value of χ^2 with (t-1) d.f. at α level of significance.

Note : The χ^2 approximation to T is good when r is large. However, since there is no other non-parametric test is available for balanced incomplete block design, one may use the Durbin test even when r is small. One should realize that the results are probably very crude where r is small.

Example 17.20 In an experiment of a road test conducted to assess the effect of four different compounding ingredients on the life of automobile tires, a balanced incomplete block design was performed with the following relative wear values :

Tire (Block)	Compound (Treatments)			
	A	B	C	D
1	238	238	279	
2	196	213		308
3	254		334	367
4		312	421	412

Test at 0.05 level of significance whether there is no difference among the compounds.

Solution : We set up,

H_0 : The treatments have equal effect.

H_1 : The treatments have no equal effect for at least one.

We have, t = 4, k= 3, r = 3.

The given data can be ranked in order of magnitude within block as follows :

	Compound (Treatments)			
Tire (Block)	A	B	C	D
1	1.5	1.5	3	
2	1	2		3
3	1		2	3
4		1	3	2
Total	3.5	4.5	8	8

The Durbin test statistic,

$$T = \frac{12\,(4-1)}{3 \times 4\,(3-1)\,(3+1)}\left[3.5^2 + 4.5^2 + 8^2 + 8^2\right] - \frac{3 \times 3\,(4-1)\,(3+1)}{(3-1)}$$

$= 6.19.$

Conclusion : The tabulated value of χ^2 with (4-1) = 3 d.f. at 0.05 level of significance is 7.815. Since the calculated value of T is smaller than the tabulated value of χ^2, we cannot reject the H_0, therefore, we conclude that all the treatments have equal effects.

Appendix-1

Factors Useful in the Construction of Control Charts

	Mean chart			Standard deviation chart					Range Chart				
Sampl size	Factors for control limits			Factor for central line	Factors for control limits				Factor for central line	Factors for contrl limits			
n	A	A_1	A_2	C_2	B_1	B_2	B_3	B_4	d_2	D_1	D_2	D_3	D_4
2	1.121	3.760	1.880	0.5642	0	1.843	0	3.267	1.128	0	3.686	0	3.267
3	1.732	2.394	1.023	0.7236	0	1.858	0	2.568	1.693	0	4.358	0	2.575
4	1.500	1.880	0.729	0.7979	0	1.808	0	2.266	2.059	0	4.698	0	2.282
5	1.342	1.596	0.577	0.8407	0	1.756	0	2.089	2.326	0	4.918	0	2.115
6	1.225	1.410	0.483	0.8686	0.026	1.711	0.030	1.970	2.534	0	5.078	0	2.004
7	1.134	1.277	0.419	0.8882	0.105	1.672	0.118	1.882	2.704	0.205	5.203	0.076	1.924
8	1.061	1.175	0.373	0.9027	0.167	1.638	0.185	1.815	2.847	0.387	5.307	0.136	1.864
9	1.000	1.094	0.337	0.9139	0.219	1.609	0.239	1.761	2.970	0.546	5.394	0.184	1.816
10	0.949	1.028	0.308	0.9227	0.262	1.584	0.284	1.716	3.078	0.687	5.469	0.223	1.777

Mean chart				Standard deviation chart					Range Chart				
Sampl size	Factors for control limits			Factor for central line	Factors for control limits				Factor for central line	Factors for contrl limits			
n	A	A_1	A_2	C_2	B_1	B_2	B_3	B_4	d_2	D_1	D_2	D_3	D_4
11	0.905	0.973	0.285	0.9300	0.299	1.561	0.321	1.679	3.173	0.812	5.534	0.256	1.744
12	0.866	0.925	0.266	0.9359	0.331	1.541	0.354	1.646	3.258	0.924	5.592	9.284	1.716
13	0.832	0.884	0.249	0.9410	0.359	1.523	0.382	1.618	3.336	1.026	5.646	0.308	1.692
14	0.802	0.848	0.235	0.9453	0.384	1.507	0.406	1.594	3.407	1.121	5.693	0.329	1.671
15	0.775	0.816	0.223	0.9490	0.406	1.492	0.428	1.572	3.472	1.207	5.737	0.348	1.652
16	0.750	0.788	0.212	0.9523	0.427	1.478	0.448	1.552	3.532	1.285	5.779	0.364	1.636
17	0.728	0.762	0.203	0.9551	0.445	1.465	0.466	1.534	3.583	1.359	5.817	0.379	1.621
18	0.707	0.738	0.194	0.9576	0.461	1.454	0.482	1.518	3.640	1.426	5.854	0.392	1.608
19	0.688	0.717	0.187	0.9599	0.477	1.443	0.497	1.503	3.689	1.490	5.888	0.404	1.596
20	0.671	0.697	0.180	0.9619	0.491	1.433	0.510	1.490	3.735	1.548	5.922	0.414	1.586
21	0.655	0.679	0.173	0.9638	0.504	1.424	0.523	1.477	3.778	1.606	5.950	0.425	1.575
22	0.640	0.662	0.167	0.9655	0.516	1.415	0.534	1.466	3.819	1.659	5.979	0.434	1.566
23	0.626	0.647	0.162	0.9670	0.527	1.407	0.545	1.455	3.858	1.710	6.006	0.443	1.557
24	0.612	0.632	0.157	0.9684	0.538	1.399	0.555	1.445	3.895	1.759	6.031	0.452	1.548
25	0.600	0.619	0.153	0.9696	0.548	1.392	0.565	1.435	3.931	1.804	6.058	0.459	1.541

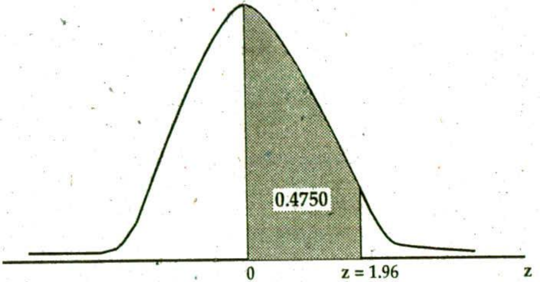
APPENDIX—II

TABLE-1. TABLES OF PROBABILITIES ASSOCIATED WITH VALUES AS SMALL AS OBERSERVED VALUES OF X IN THE BINOMIAL TEST

Given in the body of this table are one tailed probabilities under H_0 for the Bionmial test when $P = Q = \frac{1}{2}$. To save space , demical points are omitted in the p' s.

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	998
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	885

Table-2 Normal curve areas. (Entries in the body of the table give the area under the standard normal curve from o to z.)



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0395
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	.3289	.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	.4382	.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	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

TABLE—3 X^2 DISTRIBUTION*

Values of $X^2_{\alpha,\, v}$

α v	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
1	0.000	0.000	0.001	0.004	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.832	15.086	16.750
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.688	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672
40	20.706	22.164	24.433	26.509	55.759	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	67.505	71.420	76.154	79.490
60	35.535	37.485	40.482	43.188	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	124.342	129.561	135.807	140.169

TABLE—4. t DISTRIBUTION

Values of $t_{\alpha.v}$

α v	0.05	0.025	0.01	0.005
1	6.314	12.706	31.821	63.657
2	2.920	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
4	2.132	2.776	3.747	4.604
5	2.015	2.571	3.365	4.032
6	1.943	2.447	3.143	3.707
7	1.895	2.365	2.998	3.499
8	1.860	2.306	2.896	3.355
9	1.833	2.262	2.821	3.250
10	1.812	2.228	2.764	3.169
11	1.796	2.201	2.718	3.106
12	1.782	2.179	2.681	3.055
13	1.771	2.160	2.650	3.012
14	1.761	2.145	2.624	2.977
15	1.753	2.131	2.602	2.947
16	1.746	2.120	2.583	2.921
17	1.740	2.110	2.567	2.898
18	1.734	2.101	2.552	2.878
19	1.729	2.093	2.539	2.861
20	1.725	2.086	2.528	2.845
21	1.721	2.080	2.518	2.831
22	1.717	2.074	2.508	2.819
23	1.714	2.069	2.500	2.807
24	1.711	2.064	2.492	2.797
25	1.708	2.060	2.485	2.787
26	1.706	2.056	2.479	2.779
27	1.703	2.052	2.473	2.771
28	1.701	2.048	2.467	2.763
29	1.699	2.045	2.462	2.756
30	1.697	2.042	2.457	2.750
40	1.684	2.021	2.423	2.704
60	1.671	2.000	2.390	2.660
120	1.658	1.980	2.358	2.617
∞	1.645	1.960	2.326	2.576

TABLE-5. *F.* DISTRIBUTION*

Values of $F_{.0.5} : v_1, v_2$

v_1 v_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∝
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∝	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

For other values of v_1 and v_2 one may use linear interpolation , taking $1/v_1$ and $1/v_2$.

TABLE-5 (Contd.)

Values of $F_{.0.1} : v_1, v_2$

v_1 / v_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∝
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∝	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

For other values of v_1 and v_2 one may use linear interpolation ,taking $1/v_1$ and $1/v_2$ as the independent variables .

Table-6 Lower critical values of r in the runs test

n_2 n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

Note : For the one sample runs test , any value of r that is equal to or smaller than that shown in the body of this table in given value of n_1 and n_2 in significant of the 0.05 level.

Table-7 Upper critical values of r in the runs test.

n_2 n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8				11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9					13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10					13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11					13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12					13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13						15	16	17	18	19	19	20	20	21	21	22	22	23	23
14						15	16	17	18	19	20	20	21	22	22	23	23	23	24
15						15	16	18	18	19	20	21	22	22	23	23	24	24	25
16							17	18	19	20	21	21	22	23	23	24	25	25	25
17							17	18	19	20	21	22	23	23	24	25	25	26	26
18							17	18	19	20	21	22	23	24	25	25	26	26	27
19							17	18	20	21	22	23	23	24	25	26	26	27	27
20							17	18	20	21	22	23	24	25	25	26	27	27	28

Note **:** For the one sample runs test, any value of r that is equal to or smaller than that shown in the body of this table in given value of n_1 and n_2 in significant of the 0.05 level.

Table-8 Critical values of the Turkey quick test statistic

N-n	*n*	2.5%/0.5%/0.05% (one sided) 5%/1%/0.1% (Two sided)	*N-n*	*n*	2.5%/0.5%/0.05% (one sided) 5%/1%/0.1% (Two sided)
0	4–8	7/9/13	**9**	3	10/13/—
	9–21	7/10/13		4	10/13/16
	22–24	7/10/14		5–7	9/12/16
	25–	8/10/14		8	8/12/15
1	3–4	7/—/—		9–18	8/11/15
	5–6	7/9/—		19–31	8/11/14
	7	7/9/13		32–	8/10/14
	8–20	7/10/13	**10**	3	11/14/—
	21–23	7/10/14		4	10/13/17
	24	8/10/14		5	9/13/17
2	3–4	7/9/—		6–	9/13/16
	5	7/10/—		7–9	9/12/16
	6–18	7/10/13		10	8/12/15
	19–21	7/10/14		11–22	8/11/15
	22–	8/10/14		23–42	8/11/14
3	3–5	7/10/—		43–	8/10/14
	6–14	7/10/13	**11**	2	12/—/—
	15–17	7/10/14		3	11/15/—
	18–	8/10/4		4	10/14/18
4	3	8/—/—		5	10/13/17
	4–7	8/10/13		6–7	9/12/17
	8–	8/10/14		8–10	9/12/16
5	3–4	9/11/—		11–12	8/12/16
	5–6	8/11/14		13–19	8/11/15
	7–	8/10/14	**12**	2	12/—/—
6	3–4	9/11/—		3	12/15/—
	5–11	8/11/14		4	11/15/18
	12–	8/10/14		5	10/14/18
7	3–4	9/12/—		6	10/13/17
	5	9/12/15		7–8	9/12/17
	6–8	8/11/15		9–10	9/12/16
	9–17	8/11/14		11–13	8/12/16
	18–	8/10/4		14	8/11/16
8	3	10/13/—		15–18	8/11/15
	4	9/12/—	**13**	2	13/—/—
	5–6	9/12/15		3	12/16/—
	7–14	8/11/15		4	11/15/19
	15–24	8/11/14		5–6	10/14/18
	25–	8/10/14		7	9/13/18
13	8–9	9/13/17	**17**	2	16/—/—
	10	9/12/17		3	14/19/—
	11	9/12/16		4	12/18/—
	12–15	8/12/16		5	11/16/21
	16–17	8/11/16		6	11/16/20

14	2	13/-/-
	3	13/17/-
	4	11/16/19
	5	11/15/19
	6	10/14/19
	7	10/14/18
	8	9/13/18
	9–10	9/13/17
	11–12	9/12/17
	13	9/12/16
	14–16	8/12/16
15	2	14/-/-
	3	13/18/-
	4	12/16/20
	5	11/15/20
	6	10/15/19
	7	10/14/19
	8	10/14/18
	9	9/13/18
	10–11	9/13/17
	12–13	9/12/17
	14	9/12/16
	15	8/12/16
16	2	16/-/-
	3	13/18/-
	4	12/17/-
	5	11/16/20
	6	10/15/20
	7–8	10/14/19
	9	9/14/18
	10–11	9/13/18
	12	9/13/17
	13–14	9/12/17
	7	10/15/20
	8–9	10/14/19
	10–12	9/13/18
	13	9/13/17
18	2	17/—/—
	3	14/20
	4	13/18/—
	5	11/17/22
	6	11/16/21
	7–8	10/15/20
	9	10/14/19
	10	9/14/19
	11–12	9/13/18
19	2	17/—/—
	3	14/20/—
	4	13/19/23
	5	12/17/22
	6	11/16/22
	7	11/16/21
	8	10/15/20
	9	10/14/20
	10	10/14/19
	11	9/14/19
20	2	18/—/—
	3	15/21/-
	4	13/19/24
	5	12/18/23
	6	11/17/22
	7	11/16/24
	8	10/15/21
	9	10/15/20
	10	10/14/20

Entries in this table are critical values of the Tukey quick test statistics T_1 and T_2 for $\alpha = 0.025, 0.005$, and 0.0005 for one-sided tests, and T for $a = 0.05, 0.01$, and 0.001 for a two-sided test. Let n = size of the smaller of two independent samples and N = size of the larger, and reject H_0 if the computed value of the test statistics is greater than or equal to the critical value corresponding to n, N, and desired a. Because of the discrete nature of the test statistics, use of this table yields actual levels of significance that are smaller than or equal to stated levels.

$$P(T \geq h) \approx \frac{2\lambda}{\lambda^2 - 1}\left(\frac{\lambda}{\lambda + 1}\right)^h$$

where $\lambda = N/n$ for the two-sided test and half that probability applies for a one-sided test.

Table-9 Quantiles of the Mann-Whitney test statistic

n_1	p	$n_2 = 2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	.005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	.01	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2	2
	.025	0	0	0	0	0	0	1	1	1	1	2	2	2	2	2	3	3	3	3
	.05	0	0	0	1	1	1	2	2	2	2	3	3	4	4	4	4	5	5	5
	.10	0	1	1	2	2	2	3	3	4	4	5	5	5	6	6	7	7	8	8
3	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	.005	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	3	3	4	4
	.01	0	0	0	0	0	1	1	2	2	2	3	3	3	4	4	5	5	5	6
	.025	0	0	0	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9
	.05	0	1	1	2	3	3	4	5	5	6	6	7	8	8	9	10	10	11	12
	.10	1	2	2	3	4	5	6	6	7	8	9	10	11	11	12	13	14	15	16
4	.001	0	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4
	.005	0	0	0	0	1	1	2	2	3	3	4	4	5	6	6	7	7	8	9
	.01	0	0	0	1	2	2	3	4	4	5	6	6	7	9	8	9	10	10	11
	.025	0	0	1	2	3	4	5	5	6	7	8	9	10	11	12	12	13	14	15
	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19
	.10	1	2	4	5	6	7	8	10	11	12	13	14	16	17	18	19	21	22	23

n_1	p	$n_2 = 2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	.001	0	0	0	0	0	0	1	2	2	3	3	4	4	5	6	6	7	8	8
	.005	0	0	0	1	2	2	3	4	5	6	7	8	8	9	10	11	12	13	14
	.01	0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	.025	0	1	2	3	4	6	7	8	9	10	12	13	14	15	16	18	19	20	21
	.05	1	2	3	5	6	7	9	10	12	13	14	16	17	19	20	21	23	24	26
	.10	2	3	5	6	8	9	11	13	14	16	18	19	21	23	24	26	28	29	31
6	.001	0	0	0	0	0	0	2	3	4	5	5	6	7	8	9	10	11	12	13
	.005	0	0	1	2	3	4	5	6	7	8	10	11	12	13	14	16	17	18	19
	.01	0	0	2	3	4	5	7	8	9	10	12	13	14	16	17	19	20	21	23
	.025	0	2	3	4	6	7	9	11	12	14	15	17	18	20	22	23	25	26	28
	.05	1	3	4	6	8	9	11	13	15	17	18	20	22	24	26	27	29	31	33
	.10	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	35	37	39
7	.001	0	0	0	0	1	2	3	4	6	7	8	9	10	11	12	14	15	16	17
	.005	0	0	1	2	4	5	7	8	10	11	13	14	16	17	19	20	22	23	25
	.01	0	1	2	4	5	7	8	10	12	13	15	17	18	20	22	24	25	27	29
	.025	0	2	4	6	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
	.05	1	3	5	7	9	12	14	16	18	20	22	25	27	29	31	34	36	38	40
	.10	2	5	7	9	12	14	17	19	22	24	27	29	32	34	37	39	42	44	47
8	.001	0	0	0	1	2	3	5	6	7	9	10	12	13	15	16	18	19	21	22
	.005	0	0	2	3	5	7	8	10	12	14	16	18	19	21	23	25	27	29	31
	.01	0	1	3	5	7	8	10	12	14	16	18	21	23	25	27	29	31	33	35
	.025	1	3	5	7	9	11	14	16	18	20	23	25	27	30	32	35	37	39	42
	.05	2	4	6	9	11	14	16	19	21	24	27	29	32	34	37	40	42	45	48
	.10	3	6	8	11	14	17	20	23	25	28	31	34	37	40	43	46	49	52	55

n_1	p	$n_2 = 2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	0	2	3	4	6	8	9	11	13	15	16	18	20	22	24	26	27
	.005	0	1	2	4	6	8	10	12	14	17	19	21	23	25	28	30	32	34	37
9	.01	0	2	4	6	8	10	12	15	17	19	22	24	27	29	32	34	37	39	41
	.025	1	3	5	8	11	13	16	18	21	24	27	29	32	35	38	40	43	46	49
	.05	2	5	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55
	.10	3	6	10	13	16	19	23	26	29	32	36	39	42	46	49	53	56	59	63
	.001	0	0	1	2	4	6	7	9	11	13	15	18	20	22	24	26	28	30	33
	.005	0	1	3	5	7	10	12	14	17	19	22	25	27	30	32	35	38	40	43
10	.01	0	2	4	7	9	12	14	17	20	23	25	28	31	34	37	39	42	45	48
	.025	1	4	6	9	12	15	18	21	24	27	30	34	37	40	43	46	49	53	56
	.05	2	5	8	12	15	18	21	25	28	32	35	38	42	45	49	52	56	59	63
	.10	4	7	11	14	18	22	25	29	33	37	40	44	48	52	55	59	63	67	71
	.001	0	0	1	3	5	7	9	11	13	16	18	21	23	25	28	30	33	35	38
	.005	0	1	3	6	8	11	14	17	19	22	25	28	31	34	37	40	43	46	49
11	.01	0	2	5	8	10	13	16	19	23	26	29	32	35	38	42	45	48	51	54
	.025	1	4	7	10	14	17	20	24	27	31	34	38	41	45	48	52	56	59	63
	.05	2	6	9	13	17	20	24	28	32	35	39	43	47	51	55	58	62	66	70
	.10	4	8	12	16	20	24	28	32	37	41	45	49	53	58	62	66	70	74	79
	.001	0	0	1	3	5	8	10	13	15	18	21	24	26	29	32	35	38	41	43
	.005	0	2	4	7	10	13	16	19	22	25	28	32	35	38	42	45	48	52	55
12	.01	0	3	6	9	12	15	18	22	25	29	32	36	39	43	47	50	54	57	61
	.025	2	5	8	12	15	19	23	27	30	34	38	42	46	50	54	58	62	66	70
	.05	3	6	10	14	18	22	27	31	35	39	43	48	52	56	61	65	69	73	78
	.10	5	9	13	18	22	27	31	36	40	45	50	54	59	64	68	73	78	82	87

n_1	p	$n_2 = 2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
13	.001	0	0	2	4	6	9	12	15	18	21	24	27	30	33	36	39	43	46	49
	.005	0	2	4	8	11	14	18	21	25	28	32	35	39	43	46	50	54	58	61
	.01	1	3	6	10	13	17	21	24	28	32	36	40	44	48	52	56	60	64	68
	.025	2	5	9	13	17	21	25	29	34	38	42	46	51	55	60	64	68	73	77
	.05	3	7	11	16	20	25	29	34	38	43	48	52	57	62	66	71	76	81	85
	.10	5	10	14	19	24	29	34	39	44	49	54	59	64	69	75	80	85	90	95
14	.001	0	0	2	4	7	10	13	16	20	23	26	30	33	37	40	44	47	51	55
	.005	0	2	5	8	12	16	19	23	27	31	35	39	43	47	51	55	59	64	68
	.01	1	3	7	11	14	18	23	27	31	35	39	44	48	52	57	61	66	70	74
	.025	2	6	10	14	18	23	27	32	37	41	46	51	56	60	65	70	75	79	84
	.05	4	8	12	17	22	27	32	37	42	47	52	57	62	67	72	78	83	88	93
	.10	5	11	16	21	26	32	37	42	48	53	59	64	70	75	81	86	92	98	103
15	.001	0	0	2	5	8	11	15	18	22	25	29	33	37	41	44	48	52	56	60
	.005	0	3	6	9	13	17	21	25	30	34	38	43	47	52	56	61	65	70	74
	.01	1	4	8	12	16	20	25	29	34	38	43	48	52	57	62	67	71	76	81
	.025	2	6	11	15	20	25	30	35	40	45	50	55	60	65	71	76	81	86	91
	.05	4	8	13	19	24	29	34	40	45	51	56	62	67	73	78	84	89	95	101
	.10	6	11	17	23	28	34	40	46	52	58	64	69	75	81	87	93	99	105	111
16	.001	0	0	3	6	9	12	16	20	24	28	32	36	40	44	49	53	57	61	66
	.005	0	3	6	10	14	19	23	28	32	37	42	46	51	56	61	66	71	75	80
	.01	1	4	8	13	17	22	27	32	37	42	47	52	57	62	67	72	77	83	88
	.025	2	7	12	16	22	27	32	38	43	48	54	60	65	71	76	82	87	93	99
	.05	4	9	15	20	26	31	37	43	49	55	61	66	72	78	84	90	96	102	108
	.10	6	12	18	24	30	37	43	49	55	62	68	75	81	87	94	100	107	113	120

n_1	p	$n_2 = 2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	1	3	6	10	14	18	22	26	30	35	39	44	48	53	58	62	67	71
	.005	0	3	7	11	16	20	25	30	35	40	45	50	55	61	66	71	76	82	87
17	.01	1	5	9	14	19	24	29	34	39	45	50	56	61	67	72	78	83	89	94
	.025	3	7	12	18	23	29	35	40	46	52	58	64	70	76	82	88	94	100	106
	.05	4	10	16	21	27	34	40	46	52	58	65	71	78	84	90	97	103	110	116
	.10	7	13	19	26	32	39	46	53	59	66	73	80	86	93	100	107	114	121	128
	.001	0	1	4	7	11	15	19	24	28	33	38	43	47	52	57	62	67	72	71
	.005	0	3	7	12	17	22	27	32	38	43	48	54	59	65	71	76	82	88	93
18	.01	1	5	10	15	20	25	31	37	42	48	54	60	66	71	77	83	89	95	101
	.025	3	8	13	19	25	31	37	43	49	56	62	68	75	81	87	94	100	107	113
	.05	5	10	17	23	29	36	42	49	56	62	69	76	83	89	96	103	110	117	124
	.10	7	14	21	28	35	42	49	56	63	70	78	85	92	99	107	114	121	129	136
	.001	0	1	4	8	12	16	21	26	30	35	41	46	51	56	61	67	72	78	83
	.005	1	4	8	13	18	23	29	34	40	46	52	58	64	70	75	82	88	94	100
19	.01	2	5	10	16	21	27	33	39	45	51	57	64	70	76	83	89	95	102	108
	.025	3	8	14	20	26	33	39	46	53	59	66	73	79	86	93	100	107	114	120
	.05	5	11	18	24	31	38	45	52	59	66	73	81	88	95	102	110	117	124	131
	.10	8	15	22	29	37	44	52	59	67	74	82	90	98	105	113	121	129	136	144
	.001	0	1	4	8	13	17	22	27	33	38	43	49	55	60	66	71	77	83	89
	.005	1	4	9	14	19	25	31	37	43	49	55	61	68	74	80	87	93	100	106
20	.01	2	6	11	17	23	29	35	41	48	54	61	68	74	81	88	94	101	108	115
	.025	3	9	15	21	28	35	42	49	56	63	70	77	84	91	99	106	113	120	128
	.05	5	12	19	26	33	40	48	55	63	70	78	85	93	101	108	116	124	131	139
	.01	8	16	23	31	39	47	55	63	71	79	87	95	103	111	120	128	136	144	152

Table-10 Critical value of Mood's test statistic ($n_1 = m$, $n_2 = n$). The table gives M′. P(M ≤ M′), M″ , P(M ≤ M″) such that P(M ≤ M′) ≤ α < (P(M ≤ M″).]

Sample size		Normal significiance level α									
m	n	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
2	2						2.50	2.50	2.50	2.50	2.50
							0.8333	0.8333	0.8333	0.8333	0.8333
		0.50	0.50	0.50	0.50	0.50	4.50	4.50	4.50	4.50	4.50
		0.1667	0.1667	0.1667	0.1667	0.1667	1.0000	1.0000	1.0000	1.0000	1.0000
2	3						4.00	5.00	5.00	5.00	5.00
							0.5000	0.9000	0.9000	0.9000	0.9000
		1.00	1.00	1.00	1.00	1.00	5.00	8.00	8.00	8.00	8.00
		0.2000	0.2000	0.2000	0.2000	0.2000	0.9000	1.0000	1.0000	1.0000	1.0000
2	4					0.50	6.50	8.50	8.50	8.50	8.50
						0.0667	0.6667	0.9333	0.9333	0.9333	0.9333
		0.50	0.50	0.50	0.50	2.50	8.50	12.50	12.50	12.50	12.50
		0.0667	0.0667	0.0667	0.0667	0.3333	0.9333	1.0000	1.0000	1.0000	1.0000
2	5					1.00	10.00	10.00	13.00	13.00	13.00
						0.0952	0.7619	0.7619	0.9524	0.9524	0.9524
		1.00	1.00	1.00	1.00	2.00	13.00	13.00	18.00	18.00	18.00
		.00952	0.0952	0.0952	0.0952	0.1429	0.9524	0.9524	1.0000	1.0000	1.0000
2	6				0.50	0.50	14.50	14.50	18.50	18.50	18.50
					0.0357	0.0357	0.8214	0.8214	0.9643	0.9643	0.9643
		0.50	0.50	0.50	2.50	2.50	18.50	18.50	24.50	24.50	24.50
		0.0357	0.0357	0.0357	0.1786	0.1786	0.9643	0.9643	1.0000	1.0000	1.0000
2	7					2.00	20.00	20.00	25.00	25.00	25.00
						0.0833	0.8611	0.8611	0.9722	0.9722	0.9722
		1.00	1.00	1.00	1.00	4.00	25.00	25.00	32.00	32.00	32.00
		0.0556	0.0556	0.0556	0.0556	0.1389	0.9722	0.9722	1.0000	1.0000	1.0000

Sample size		Normal significiance level α									
m	n	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
3	7			2.75	6.75	6.75	34.75	40.75	44.75	46.75	46.75
				0.0167	0.5000	0.0500	0.8500	0.9333	0.9667	0.9833	0.9833
		2.75	2.75	4.75	8.75	8.75	38.75	42.75	46.75	52.75	52.75
		0.0167	0.0167	0.0333	0.1167	0.1167	0.9167	0.9500	0.9833	1.0000	1.0000
3	8		2.00	2.00	8.00	11.00	45.00	50.00	54.00	59.00	59.00
			0.0061	0.0061	0.0485	0.0970	0.8848	0.9394	0.9636	0.9879	0.9879
		2.00	5.00	5.00	9.00	13.00	50.00	51.00	57.00	66.00	66.00
		0.0061	0.0303	0.0303	0.0606	0.1212	0.9394	0.9515	0.9758	1.0000	1.0000
3	9		2.75	4.75	6.75	12.75	54.75	60.75	66.75	70.75	72.75
			0.0091	0.0182	0.0273	0.0909	0.8727	0.9182	0.9727	0.9818	0.9909
		2.75	4.75	6.75	8.75	14.75	56.75	62.75	70.75	72.75	80.75
		0.0091	0.0182	0.0273	0.0636	0.1364	0.9091	0.9636	0.9818	0.9909	1.0000
3	10	2.00	2.00	6.00	10.00	14.00	68.00	76.00	77.00	86.00	88.00
		0.0035	0.0035	0.0245	0.0490	0.0979	0.8986	0.9441	0.9720	0.9860	0.9930
		5.00	5.00	8.00	11.00	17.00	70.00	77.00	81.00	88.00	97.00
		0.0175	0.0175	0.0280	0.0559	0.1189	0.9266	0.9720	0.9790	0.9930	1.0000
3	11		2.75	6.75	10.75	16.75	74.75	84.75	90.75	102.75	104.75
			0.0055	0.0165	0.0440	0.0879	0.8846	0.9451	0.9560	0.9890	0.9945
		2.75	4.75	8.75	12.75	18.75	78.75	86.75	92.75	104.75	114.75
		0.0055	0.0110	0.0385	0.0549	0.1099	0.9066	0.9505	0.9780	0.9945	1.0000
3	12	2.00	2.00	9.00	13.00	20.00	89.00	99.00	107.00	114.00	121.00
		0.0022	0.0022	0.0220	0.0440	0.0945	0.8879	0.9385	0.9648	0.9868	0.9912
		5.00	5.00	10.00	14.00	21.00	90.00	101.00	110.00	121.00	123.00
		0.0110	0.0110	0.0308	0.0615	0.1121	0.9055	0.9560	0.9824	0.9912	0.9956
3	13	2.75	4.75	8.75	12.75	20.75	102.75	114.75	124.75	132.75	140.75
		0.0036	0.0071	0.0250	0.0357	0.0893	0.8893	0.9464	0.9714	0.9893	0.9929
		4.75	6.75	10.75	14.75	22.75	104.75	116.75	128.75	140.75	142.75
		0.0071	0.0107	0.0286	0.0536	0.1036	0.9071	0.9500	0.9857	0.9929	0.9964

Sample size		Normal significiance level α									
m	n	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
3	14	2.00	5.00	11.00	17.00	25.00	116.00	128.00	138.00	149.00	162.00
		0.0015	0.0074	0.0235	0.0500	0.0868	0.8926	0.9353	0.9735	0.9882	0.9941
		5.00	6.00	13.00	18.00	26.00	117.00	129.00	144.00	153.00	164.00
		0.0074	0.0103	0.0294	0.0544	0.1044	0.9044	0.9500	0.9765	0.9912	0.9971
3	15	4.75	6.75	12.75	18.75	26.75	132.75	146.75	156.75	164.75	174.75
		0.0049	0.0074	0.0245	0.0490	0.0907	0.8995	0.9485	0.9681	0.9804	0.9926
		6.75	8.75	14.75	20.75	28.75	134.75	148.75	158.75	170.75	184.75
		0.0074	0.0172	0.0368	0.0613	0.1005	0.9191	0.9583	0.9779	0.9902	0.9951
3	16	2.80	8.00	13.00	20.00	32.00	146.00	164.00	179.00	187.00	198.00
		0.0010	0.0083	0.0206	0.0444	0.0970	0.8937	0.9463	0.9732	0.9835	0.9938
		5.00	9.00	14.00	21.00	33.00	149.00	166.00	181.00	194.00	209.00
		0.0052	0.0103	0.0289	0.0526	0.1011	0.9102	0.9567	0.9814	0.9917	0.9959
3	17	4.75	6.75	12.75	20.75	34.75	162.75	180.75	192.75	210.75	222.75
		0.0035	0.0053	0.0175	0.0439	0.1000	0.8930	0.9421	0.9719	0.9860	0.9947
		6.75	8.75	14.75	22.75	36.75	164.75	182.75	200.75	218.75	234.75
		0.0053	0.0123	0.0263	0.0509	0.1070	0.9018	0.9509	0.9754	0.9930	0.9965
4	4			5.00	5.00	9.00	29.00	31.00	31.00	33.00	33.00
				0.0143	0.0143	0.0714	0.8714	0.9286	0.9286	0.9857	0.9857
		5.00	5.00	9.00	9.00	11.00	31.00	33.00	33.00	37.00	37.00
		0.0143	0.0143	0.0714	0.0714	0.1286	0.9286	0.9857	0.9857	1.0000	1.0000
4	5			6.00	10.00	11.00	37.00	41.00	42.00	42.00	45.00
				0.0159	0.0397	0.0556	0.8730	0.9286	0.9603	0.9603	0.9921
		6.00	6.00	9.00	11.00	14.00	38.00	42.00	45.00	45.00	50.00
		0.0159	0.0159	0.0317	0.0556	0.1190	0.9048	0.9603	0.9921	0.9921	1.0000
4	6	5.00	5.00	9.00	13.00	15.00	47.00	51.00	53.00	55.00	55.00
		0.0048	0.0048	0.0238	0.0476	0.0857	0.8952	0.9333	0.9571	0.9762	0.9762
		9.00	9.00	11.00	15.00	17.00	49.00	53.00	55.00	59.00	59.00
		0.0238	0.0238	0.0429	0.0857	0.1095	0.9143	0.9571	0.9762	0.9952	0.9952
4	7		6.00	11.00	14.00	20.00	58.00	63.00	68.00	70.00	70.00
			0.0061	0.0212	0.0455	0.0909	0.8848	0.9394	0.9727	0.9848	0.9848
		6.00	9.00	14.00	15.00	21.00	59.00	66.00	70.00	75.00	75.00
		0.0061	0.0121	0.0455	0.0576	0.1152	0.9030	0.9576	0.9848	0.9970	0.9970

Sample size		Normal significiance level α									
m	n	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
4	8	5.00	5.00	13.00	17.00	21.00	69.00	77.00	81.00	87.00	87.00
		0.0020	0.0020	0.0202	0.0465	0.0869	0.8970	0.9475	0.9636	0.9899	0.9899
		9.00	9.00	15.00	19.00	23.00	71.00	79.00	83.00	93.00	93.00
		0.0101	0.0101	0.0364	0.0545	0.1030	0.9051	0.9556	0.9798	0.9980	0.9980
4	9	6.00	11.00	14.00	20.00	27.00	85.00	92.00	98.00	104.00	106.00
		0.0028	0.0098	0.0210	0.0420	0.0965	0.8979	0.9497	0.9748	0.9874	0.9930
		9.00	14.00	15.00	21.00	29.00	86.00	93.00	101.00	106.00	113.00
		0.0056	0.0210	0.0266	0.0531	0.1077	0.9231	0.9552	0.9804	0.9930	0.9986
4	10	9.00	13.00	17.00	21.00	31.00	97.00	105.00	115.00	121.00	125.00
		0.0050	0.0100	0.0230	0.0430	0.0969	0.8961	0.9491	0.9740	0.9860	0.9910
		11.00	15.00	19.00	23.00	33.00	99.00	107.00	117.00	123.00	127.00
		0.0090	0.0180	0.0270	0.0509	0.1129	0.9161	0.9530	0.9820	0.9900	0.9950
4	11	10.00	11.00	20.00	26.00	35.00	113.00	125.00	134.00	143.00	148.00
		0.0037	0.0051	0.0220	0.0462	0.0967	0.8967	0.9495	0.9722	0.9897	0.9934
		11.000	14.00	21.00	27.00	36.00	114.00	126.00	135.00	146.00	150.00
		0.0051	0.0110	0.0278	0.0505	0.1011	0.9099	0.9612	0.9780	0.9927	0.9963
4	12	11.00	15.00	21.00	29.00	39.00	129.00	141.00	153.00	161.00	171.00
		0.0049	0.0099	0.0236	0.0489	0.0978	0.8962	0.9495	0.9747	0.9879	0.9945
		13.00	17.00	23.00	31.00	41.00	131.00	143.00	155.00	163.00	173.00
		0.0055	0.0126	0.0280	0.0533	0.1033	0.9159	0.9538	0.9791	0.9901	0.9951
4	13	11.00	17.00	25.00	33.00	45.00	146.00	162.00	173.00	186.00	193.00
		0.0029	0.0088	0.0227	0.0475	0.0971	0.8933	0.9496	0.9710	0.9891	0.9941
		14.00	18.00	26.00	34.00	46.00	147.00	163.00	174.00	187.00	198.00
		0.0063	0.0113	0.0265	0.0504	0.1071	0.9000	0.9529	0.9777	0.9908	0.9958
4	14	13.00	19.00	27.00	37.00	49.00	163.00	181.00	195.00	207.00	217.00
		0.0033	0.0088	0.0235	0.0477	0.0928	0.8931	0.9487	0.9739	0.9889	0.9941
		15.00	21.00	29.00	39.00	51.00	165.00	183.00	197.00	213.00	221.00
		0.0059	0.0141	0.0291	0.0582	0.1059	0.9049	0.9539	0.9755	0.9915	0.9954

Sample size		Normal significiance level α									
m	n	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
4	15	15.00 0.0049	21.00 0.0098	29.00 0.0199	41.00 0.0472	56.00 0.0993	183.00 0.8965	202.00 0.9466	218.00 0.9727	234.00 0.9892	245.00 0.9943
		17.00 0.0054	22.00 0.0114	30.00 0.0261	42.00 0.0524	57.00 0.1045	185.00 0.9017	203.00 0.9518	219.00 0.9768	235.00 0.9902	247.00 0.9954
4	16	17.00 0.0047	21.00 0.0089	33.00 0.0233	43.00 0.0436	61.00 0.0962	203.00 0.8933	223.00 0.9451	241.00 0.9728	259.00 0.9870	275.00 0.9946
		19.00 0.0056	23.00 0.0105	35.00 0.0283	45.00 0.0504	63.00 0.1061	205.00 0.9028	225.00 0.9525	243.00 0.9752	261.00 0.9903	277.00 0.9955
5	5		11.25 0.0079	15.25 0.0159	17.25 0.0317	23.25 0.0952	55.25 0.8889	59.25 0.9365	61.25 0.9683	65.25 0.9841	67.25 0.9921
		11.25 0.0079	25.25 0.0159	17.25 0.0317	21.25 0.0635	25.25 0.1111	57.25 0.9048	61.25 0.9683	65.25 0.9841	67.25 0.9921	71.25 1.0000
5	6	10.00 0.0022	10.00 0.0022	19.00 0.0238	24.00 0.0476	27.00 0.0758	69.00 0.8810	75.00 0.9459	76.00 0.9632	83.00 0.9870	84.00 0.9913
		15.00 0.0108	15.00 0.0108	20.00 0.0260	25.00 0.0563	30.00 0.1104	70.00 0.9069	76.00 0.9632	79.00 0.9805	84.00 0.9913	86.00 0.9957
5	7	11.25 0.0025	15.25 0.0051	21.25 0.0202	27.25 0.0480	33.25 0.0884	83.25 0.8990	89.25 0.9495	93.25 0.9646	101.25 0.9899	105.25 0.9949
		15.25 0.0051	17.25 0.0101	23.25 0.0303	29.25 0.0631	35.25 0.1136	85.25 0.9167	91.25 0.9520	95.25 0.9773	103.25 0.9924	107.25 0.9975
5	8	15.00 0.0039	20.00 0.0093	26.00 0.0225	31.00 0.0490	39.00 0.0979	99.00 0.8974	106.00 0.9448	113.00 0.9697	118.00 0.9852	123.00 0.9938
		18.00 0.0070	22.00 0.0124	27.00 0.0272	33.00 0.0521	40.00 0.1049	101.00 0.9068	107.00 0.9510	114.00 0.9759	122.00 0.9922	126.00 0.9953
5	9	17.25 0.0040	21.25 0.0080	29.25 0.0250	35.25 0.0450	45.25 0.0999	115.25 0.8951	123.25 0.9411	133.25 0.9710	141.25 0.9890	145.25 0.9910
		21.25 0.0080	23.25 0.0120	31.25 0.0300	37.25 0.0509	47.25 0.1149	117.25 0.9121	125.25 0.9500	135.25 0.9790	143.25 0.9900	147.25 0.9960
5	10	20.00 0.0040	26.00 0.0097	33.00 0.0223	41.00 0.0456	52.00 0.0989	134.00 0.8934	146.00 0.9494	154.00 0.9724	166.00 0.9897	174.00 0.9947
		22.00 0.0053	27.00 0.0117	34.00 0.0266	42.00 0.0503	53.00 0.1002	135.00 0.9068	147.00 0.9547	155.00 0.9757	168.00 0.9923	175.00 0.9973

Sample size		Normal significiance level α									
m	n	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
5	11	21.25	27.25	37.25	45.25	57.25	153.25	165.25	177.25	187.25	197.25
		0.0037	0.0087	0.0234	0.0458	0.0934	0.8997	0.9473	0.9748	0.9881	0.9950
		23.25	29.25	39.25	47.25	59.25	155.25	167.25	179.25	191.25	199.25
		0.0055	0.0114	0.0275	0.0527	0.1053	0.9125	0.9519	0.9776	0.9918	0.9954
5	12	26.00	30.00	42.00	53.00	65.00	174.00	189.00	202.00	216.00	226.00
		0.0047	0.0082	0.0244	0.0486	0.0931	0.8993	0.9473	0.9746	0.9888	0.9945
		27.00	31.00	43.00	54.00	66.00	175.00	190.00	203.00	217.00	227.00
		0.0057	0.0102	0.0267	0.0535	.01021	0.9071	0.9551	0.9772	0.9901	0.9952
5	13	27.25	33.25	45.25	57.25	73.25	195.25	211.25	227.25	243.25	255.25
		0.0044	0.0082	0.0233	0.0476	0.0997	0.8985	0.9444	0.9741	0.9893	0.9946
		29.25	35.25	47.25	59.25	75.25	197.25	213.25	229.25	245.25	257.25
		0.0058	0.0105	0.0268	0.0537	0.1076	0.9059	0.9512	0.9762	0.9904	0.9958
5	14	30.00	38.00	51.00	65.00	81.00	219.00	238.00	254.00	275.00	285.00
		0.0044	0.0088	0.0248	0.0495	0.0978	0.8999	0.9479	0.9720	0.9896	0.9946
		31.00	39.00	52.00	66.00	82.00	220.00	239.00	255.00	276.00	287.00
		0.0054	0.0108	0.0255	0.0544	0.1034	0.9037	0.9520	0.9754	0.9906	0.9953
5	15	33.25	39.25	55.25	69.25	89.25	241.25	265.25	283.25	305.25	319.25
		0.0045	0.0077	0.0235	0.0470	0.0988	0.8951	0.9494	0.9739	0.9896	0.9946
		35.25	41.25	57.25	71.25	91.25	243.25	267.25	285.25	307.25	321.25
		0.0058	0.0103	0.0263	0.0526	0.1053	0.9005	0.9542	0.9763	0.9906	0.9957
6	6	17.50	27.50	33.50	39.50	45.50	93.50	99.50	105.50	111.50	115.50
		0.0011	0.0097	0.0238	0.0465	0.0963	0.8734	0.9307	0.9675	0.9848	0.9946
		23.50	29.50	35.50	41.50	47.50	95.50	101.50	107.50	113.50	119.50
		0.0054	0.0152	0.0325	0.0693	0.1266	0.9037	0.9535	0.9762	0.9903	0.9989
6	7	27.00	31.00	38.00	45.00	54.00	114.00	122.00	129.00	135.00	140.00
		0.0047	0.0099	0.0204	0.0466	0.0973	0.8980	0.9476	0.9749	0.9883	0.9948
		28.00	34.00	39.00	46.00	55.00	115.00	123.00	130.00	138.00	142.00
		0.0052	0.0146	0.0251	0.0524	0.1206	0.9108	0.9580	0.9979	0.9918	0.9971

Sample size		Normal significiance level α									
m	n	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
6	8	29.50	35.50	41.50	49.50	59.50	131.50	141.50	149.50	157.50	165.50
		0.0047	0.0100	0.0213	0.0430	0.0942	0.8924	0.9461	0.9737	0.9873	0.9940
		31.50	37.50	4350	51.50	61.500	133.50	143.50	151.50	159.[illegible]	167.50
		0.0060	0.0130	0.0266	0.0509	0.1062	0.9004	0.9540	0.9750	[illegible]	0.9[illegible]
6	9	34.00	39.00	49.00	58.00	69.00	154.00	165.00	175.00	186.[illegible]	193.[illegible]
		0.0050	0.0086	0.0232	0.0488	0.0969	0.8973	0.9467	0.9734	0.9894	0.99[illegible]
		35.00	40.500	50.00	59.00	70.00	155.00	166.00	176.00	187.00	195.00
		0.0062	0.0110	0.0256	0.0547	0.1039	0.9065	0.9504	0.9766	0.9910	0.9956
6	10	37.50	43.50	53.50	63.50	75.50	175.50	189.50	201.50	213.50	221.50
		0.0049	0.0100	0.0237	0.0448	0.0888	0.8976	0.9476	0.9734	0.9891	0.9948
		39.50	45.50	55.50	65.50	77.50	177.50	191.50	203.50	215.50	223.50
		0.0054	0.0111	0.0262	0.0521	0.1010	0.9063	0.9540	0.9784	0.9901	0.9953
6	11	42.00	49.00	61.00	73.00	87.00	200.00	216.00	229.00	244.00	253.00
		0.0048	0.0094	0.0243	0.0490	0.0977	0.8998	0.9491	0.9737	0.9898	0.9941
		43.50	50.00	62.00	74.00	88.00	201.00	217.00	230.00	245.00	254.00
		0.0060	0.0103	0.0255	0.0512	0.1037	0.9009	0.9504	0.9758	0.9901	0.9954
6	12	45.50	51.50	67.50	79.50	95.50	223.50	243.50	257.50	273.50	285.50
		0.0048	0.0082	0.0248	0.0470	0.0950	0.8954	0.9494	0.9733	0.9879	0.9944
		47.50	53.50	69.50	81.50	97.50	225.50	245.50	259.50	275.50	287.50
		0.0063	0.0102	0.0273	0.0513	0.1033	0.9004	0.9542	0.9757	0.9900	0.9950
6	13	50.00	58.00	74.00	89.00	107.00	252.00	273.00	290.00	310.00	323.00
		0.0047	0.0090	0.0234	0.0483	0.0985	0.8979	0.9499	0.9736	0.9898	0.9949
		51.00	59.00	75.00	90.00	108.00	253.00	274.00	291.00	311.0	324.00
		0.0053	0.0101	0.0256	0.0503	0.1008	0.9001	0.9510	0.9751	0.9902	0.9951
6	14	53.50	63.50	81.50	97.50	117.50	279.50	301.50	321.50	343.50	357.50
		0.0049	0.0093	0.0246	0.0495	0.0974	0.8972	0.9459	0.9730	0.9888	0.9944
		55.50	65.50	83.50	99.50	119.50	281.50	303.50	323.50	345.50	359.50
		0.0054	0.0108	0.0281	0.0527	0.1043	0.9040	0.9501	0.9754	0.9901	0.9950
7	7	41.75	47.75	57.75	65.75	75.75	147.75	157.75	165.75	175.75	179.75
		0.0029	0.0082	0.0233	0.0466	0.0950	0.8869	0.9452	0.9709	0.9889	0.9948
		43.75	49.75	59.75	67.75	77.75	149.75	159.75	167.75	177.75	183.75
		0.0052	0.0111	0.0291	0.0548	0.1131	0.9050	0.9534	0.9767	0.9918	0.9971

Sample size		Normal significiance level α									
m	n	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
7	8	50.00	55.00	66.00	75.00	87.00	173.00	184.00	195.00	204.00	211.00
		0.0050	0.0082	0.0238	0.0479	0.0977	0.8988	0.9455	0.9745	0.9890	0.9939
		51.00	56.00	67.00	76.00	88.00	174.00	185.00	196.00	205.00	212.00
		0.0059	0.0110	0.0272	0.0533	0.1052	0.9004	0.9510	0.9778	0.9002	0.9552
7	9	53.75	59.75	71.75	83.75	95.75	197.75	211.75	221.75	235.75	245.75
		0.0049	0.0087	0.0224	0.0495	0.0920	0.8970	0.9495	0.9706	0.9895	0.9949
		55.75	61.75	73.75	85.75	97.75	199.75	213.75	223.75	237.75	247.75
		0.0058	0.0103	0.0267	0.0556	0.1016	0.9073	0.9549	0.9764	0.9911	0.9963
7	10	59.00	67.00	82.00	94.00	109.00	226.00	242.00	254.00	270.00	279.00
		0.0046	0.0090	0.0243	0.0478	0.0975	0.8978	0.9499	0.9726	0.9896	0.9949
		60.00	68.00	83.00	95.00	110.00	227.00	243.00	255.00	271.00	280.00
		0.0053	0.0100	0.0268	0.0521	0.1009	0.9051	0.9544	0.9753	0.9902	0.9951
7	11	63.75	73.75	89.75	103.75	119.75	253.75	271.75	287.75	303.75	315.75
		0.0042	0.0096	0.0246	0.0495	0.0946	0.8991	0.9483	0.9742	0.9882	0.9943
		65.75	75.75	91.75	105.75	121.75	255.75	273.75	289.75	305.75	317.75
		0.0050	0.0103	0.0272	0.0526	0.1012	0.9053	0.9506	0.9767	0.9904	0.9952
7	12	71.00	82.00	99.00	115.00	135.00	285.00	306.00	323.00	343.00	357.00
		0.0048	0.0094	0.0241	0.0489	0.0996	0.8997	0.9491	0.9738	0.9893	0.9950
		72.00	83.00	100.00	116.00	136.00	286.00	307.00	324.00	344.00	358.00
		0.0051	0.0104	0.0258	0.0519	0.1044	0.9020	0.9515	0.9754	0.9900	0.9952
7	13	75.75	87.75	107.75	125.75	147.75	315.75	339.75	359.75	381.75	397.75
		0.0042	0.0089	0.0239	0.0487	0.0983	0.8972	0.9487	0.9745	0.9889	0.9949
		77.75	89.75	109.75	127.75	149.75	317.75	341.75	361.75	383.75	399.75
		0.0050	0.0101	0.0261	0.0528	0.1054	0.9039	0.9523	0.9758	0.9905	0.9953
8	8	72.00	78.00	92.00	104.00	118.00	218.00	232.00	244.00	258.00	264.00
		0.0043	0.0078	0.0239	0.0496	0.0984	0.8908	0.9457	0.9740	0.9900	0.9942
		74.00	80.00	94.00	106.00	120.00	220.00	234.00	246.00	260.00	266.00
		0.0058	0.0100	0.0260	0.0543	0.1092	0.9016	0.9504	0.9761	0.9922	0.9957

Sample size		Normal significiance level α									
m	n	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
8	9	79.00	90.00	103.00	116.00	132.00	250.00	266.00	279.00	294.00	303.00
		0.0042	0.0096	0.0229	0.0487	0.0988	0.8959	0.9477	0.9742	0.9896	0.9945
		80.00	91.00	104.00	117.00	133.00	251.00	267.00	280.00	295.00	304.00
		0.0050	0.0102	0.0253	0.0510	0.1016	0.9005	0.9520	0.9760	0.9901	0.9952
8	10	88.00	98.00	114.00	128.00	146.00	280.00	300.00	316.00	332.00	344.00
		0.0050	0.100	0.0245	0.0481	0.0980	0.8917	0.9487	0.9744	0.9891	0.9948
		90.00	100.00	116.00	130.00	148.00	282.00	302.00	318.00	334.00	346.00
		0.0059	0.0112	0.0280	0.0525	0.1033	0.9001	0.9532	0.9768	0.9900	0.9950
8	11	95.00	107.00	126.00	143.00	163.00	316.00	337.00	355.00	376.00	388.00
		0.0047	0.0095	0.0247	0.500	0.0988	0.8984	0.9489	0.9739	0.9900	0.9948
		96.00	108.00	127.00	144.00	164.00	317.00	338.00	356.00	377.00	389.00
		0.0051	0.0105	0.0256	0.0530	0.1039	0.9021	0.9501	0.9759	0.9909	0.9953
8	12	102.00	116.00	136.00	156.00	178.00	352.00	376.00	396.00	418.00	434.00
		0.0044	0.0097	0.0234	0.0496	0.0970	0.8995	0.9497	0.9749	0.9894	0.9949
		104.00	118.00	138.00	158.00	180.00	354.00	378.00	398.00	420.00	436.00
		0.0051	0.0103	0.0252	0.0533	0.1031	0.9056	0.9531	0.9763	0.9903	0.9953
9	9	110.25	120.25	138.25	154.25	172.25	308.25	326.25	342.25	360.25	370.25
		0.0045	0.0085	0.0230	0.0481	0.0973	0.8975	0.9476	0.9742	0.9899	0.9949
		112.25	122.25	140.25	156.25	174.25	310.25	328.25	344.25	362.25	372.25
		0.0051	0.0101	0.0258	0.0524	0.1025	0.9027	0.9519	0.9770	0.9915	0.9955
9	10	122.00	134.00	154.00	171.00	191.00	347.00	368.00	385.00	404.00	419.00
		0.0049	0.0096	0.0250	0.0492	0.0963	0.8987	0.9489	0.9738	0.9890	0.9950
		123.00	135.00	155.00	172.00	192.00	348.00	369.00	386.00	405.00	420.00
		0.0050	0.0101	0.0256	0.0514	0.1003	0.9021	0.9515	0.9751	0.9900	0.9955
9	11	132.25	144.25	166.25	186.25	210.25	384.25	408.25	430.25	452.25	468.25
		0.0049	0.0089	0.0235	0.0484	0.0984	0.8942	0.9465	0.9744	0.9896	0.9950
		134.25	146.25	168.25	188.25	212.25	386.25	410.25	432.25	454.25	470.25
		0.0056	0.0102	0.0251	0.0519	0.1049	0.9005	0.9500	0.9765	0.9900	0.9955
10	10	162.50	176.50	198.50	218.50	242.50	418.50	442.50	462.50	484.50	498.50
		0.0050	0.0098	0.0241	0.0489	0.0982	0.8966	0.9479	0.9740	0.9891	0.9944
		164.50	178.50	200.50	220.50	244.50	420.50	444.50	464.50	486.50	500.50
		0.0056	0.0109	0.0260	0.0521	0.1034	0.9018	0.9511	0.9759	0.9902	0.9950

Table-11 *d*– factors for Wilcoxon signed-rank test and confidence intervals for the median (a'= one side significance level, a''=two sided significance level)

n	d	Confidence coefficient	α''	α'
3	1	.750	.250	.125
4	1	.875	.125	.063
5	1	.938	.062	.031
	2	.875	.125	.063
6	1	.969	.031	.016
	2	.937	.063	.031
	3	.906	.094	.047
	4	.884	.156	.078
7	1	.984	.016	.008
	2	.969	.031	.016
	4	.922	.078	.039
	5	.891	.109	.055
8	1	.922	.008	.004
	2	.984	.016	.008
	4	.961	.039	.020
	5	.945	.055	.027
	6	.922	.078	.039
	7	.891	.109	.055
9	2	.992	.008	.004
	3	.988	.012	.006
	6	.961	.039	.020
	7	.945	.055	.027
	9	.902	.098	.049
	10	.871	.129	.065
10	4	.990	.010	.005
	5	.986	.014	.007
	9	.951	.049	.024
	10	.936	.064	.032
	11	.916	.084	.042
14	13	.991	.009	.004
	14	.989	.011	.005
	22	.951	.049	.025
	23	.942	.058	.029
	26	.909	.091	.045
	27	.896	.104	.052
15	16	.992	.008	.004
	17	.990	.010	.005
	26	.952	.048	.024
	27	.945	.055	.028
	31	.905	.095	.047
	32	.893	.107	.054
16	20	.991	.009	.005
	21	.989	.011	.006
	30	.956	.44	.022
	31	.949	.051	.025
	36	.907	.093	.047
	37	.895	.105	.052
17	24	.991	.009	.005
	25	.989	.011	.006
	35	.955	.045	.022
	36	.949	.051	.025
	42	.902	.098	.049
	43	.891	.109	.054
18	28	.991	.009	.005
	29	.990	.010	.005
	41	.952	.048	.024
	42	.946	.054	.027
	48	.901	.099	.049

TABLE-12. TABLE OF CRITICAL VALUES OF D IN THE KOLMOGOROV-SMIRNOW

One –Samples Test

Sample size (N)	Level of significances for D = Maximum $\Omega\ F_0(X) - S_N(X)\ \Omega$				
	.20	.15	.10	.05	.01
1	.900	.925	.950	.975	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.828
4	.494	.525	.564	.624	.733
5	.446	.474	.510	.565	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.410	.490
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.392
17	.250	.266	.286	.318	.381
18	.244	.259	.278	.309	.371
19	.237	.252	.272	.301	.363
20	.231	.246	.264	.294	.356
25	.21	.22	.24	.27	.32
30	.19	.20	.22	.24	.29
35	.18	.19	.21	.23	.27
Over 35	$\frac{1.07}{\sqrt{N}}$	$\frac{1.14}{\sqrt{N}}$	$\frac{1.22}{\sqrt{N}}$	$\frac{1.36}{\sqrt{N}}$	$\frac{1.63}{\sqrt{N}}$

	12	.895	.105	.053		49	.892	.108	.054
11	6	.990	.010	.005	19	33	.991	.009	.005
	7	.986	.014	.007		34	.989	.011	.005
	11	.958	.042	.021		47	.951	.049	.025
	12	.946	.054	.027		48	.945	.055	.027
	14	.917	.083	.042		54	.904	.096	.048
	15	.898	.102	.051		55	.896	.104	.052
12	8	.991	.009	.005	20	38	.991	.009	.005
	9	.988	.012	.006		39	.989	.011	.005
	14	.958	.042	.021		53	.952	.048	.024
	15	.948	.052	.026		54	.947	.053	.027
	18	.908	.092	.046		61	.903	.097	.049
	19	.890	.110	.055		62	.895	.105	.053
13	10	.992	.008	.004	21	43	.991	.009	.005
	11	.990	.010	.005		44	.990	.010	.005
	18	.952	.048	.024		59	.954	.046	.023
	19	.943	.057	.029		60	.950	.050	.025
	22	.906	.094	.047		68	.904	.096	.048
	23	.890	.110	.055		69	.897	.103	.052
22"	49	.991	.009	.005	24	62	.990	.010	.005
	50	.990	.010	.005		63	.989	.011	.005
	66	.954	.046	.023		82	.951	.049	.025
	67	.950	.050	.025		83	.947	.053	.026
	76	.902	.098	.049		92	.905	.095	.048
	77	.895	.105	.053		93	.899	.101	.051
23	55	.991	.009	.005	25	69	.990	.010	.005
	56	.990	.010	.005		70	.989	.011	.005
	74	.952	.048	.024		90	.952	.048	.024
	75	.948	.052	.026		91	.948	.052	.026
	84	.902	.098	.049		101	.904	.096	.048
	85	.895	.105	.052		102	.899	.101	.051

Note : For $n > 25$ use $d \approx \frac{1}{2}\left[\frac{1}{2}n(n+1) + 1 - z\sqrt{n\ (n+1)\ (2n+1)/6}\right]$, where z is read from Table-2.

Table-13 Quantilies of the Smirnov test Statistic for two samples of equal size n

One sided test	$p = 0.90$	0.95	0.975	0.99	0.995
Two sided test	$p = 0.80$	0.90	0.95	0.98	0.99
$n = 3$	2/3	2/3			
4	3/4	3/4	3/4		
5	3/5	3/5	4/5	4/5	4/5
6	3/6	4/6	4/6	5/6	5/6
7	4/7	4/7	5/7	5/7	5/7
8	4/8	4/8	5/8	5/8	6/8
9	4/9	5/9	5/9	6/9	6/9
10	4/10	5/10	6/10	6/10	7/10
11	5/11	5/11	6/11	7/11	7/11
12	5/12	5/12	6/12	7/12	7/12
13	5/13	6/13	6/13	7/13	8/13
14	5/14	6/14	7/14	7/14	8/14
15	5/15	6/15	7/15	8/15	8/15
16	6/16	6/16	7/16	8/16	9/16
17	6/17	7/17	7/17	8/17	9/17
18	6/18	7/18	8/18	9/18	9/18
19	6/19	7/19	8/19	9/19	9/19
20	6/20	7/20	8/20	9/20	10/20
21	6/21	7/21	8/21	9/21	10/21
22	7/22	8/22	8/22	10/22	10/22
23	7/23	8/23	9/23	10/23	10/23
24	7/24	8/24	9/24	10/24	11/24
25	7/25	8/25	9/25	10/25	11/25
26	7/26	8/26	9/26	10/26	11/26
27	7/27	8/27	9/27	11/27	11/27
28	8/28	9/28	10/28	11/28	12/28
29	8/29	9/29	10/29	11/29	12/29
30	8/30	9/30	10/30	11/30	12/30
31	8/31	9/31	10/31	11/31	12/31
32	8/32	9/32	10/32	12/32	12/32
33	8/33	9/33	11/33	12/33	13/33
34	8/34	10/34	11/34	12/34	13/34
35	8/35	10/35	11/35	12/35	13/35
36	9/36	10/36	11/36	12/36	13/36
37	9/37	10/37	11/37	13/37	13/37
38	9/38	10/38	11/38	13/38	14/38
39	9/39	10/39	11/39	13/39	14/39
40	9/40	10/40	12/40	13/40	14/40
Approximation for n > 40 :	$\frac{1.52}{\sqrt{n}}$	$\frac{1.73}{\sqrt{n}}$	$\frac{1.92}{\sqrt{n}}$	$\frac{2.15}{\sqrt{n}}$	$\frac{2.30}{\sqrt{n}}$

Table-14 Quantiles of the Smirnov test statistic for two sample of different size.

One sided test		$p = 0.90$	0.95	0.975	0.99	0.995
Two-sided test		$p = 0.80$	0.90	0.95	0.98	0.99
$N_1 = 1$	$N_2 = 9$	17/18				
	10	9/10				
$N_1 = 2$	$N_2 = 3$	5/6				
	4	3/4				
	5	4/5	4/5			
	6	5/6	5/6			
	7	5/7	6/7			
	8	3/4	7/8	7/8		
	9	7/9	8/9	8/9		
	10	7/10	4/5	9/10		
$N_1 = 3$	$N_2 = 4$	3/4	3/4			
	5	2/3	4/5	4/5		
	6	2/3	2/3	5/6		
	7	2/3	5/7	6/7	6/7	
	8	5/8	3/4	3/4	7/8	
	9	2/3	2/3	7/9	8/9	8/9
	10	3/5	7/10	4/5	9/10	9/10
	12	7/12	2/3	3/4	5/6	11/12
$N_1 = 4$	$N_2 = 5$	3/5	3/4	4/5	4/5	
	6	7/12	2/3	3/4	5/6	5/6
	7	17/28	5/7	3/4	6/7	6/7
	8	5/8	5/8	3/4	7/8	7/8
	9	5/9	2/3	3/4	7/9	8/9
	10	11/20	13/20	7/10	4/5	4/5
	12	7/12	2/3	2/3	3/4	5/6
	16	9/16	5/8	11/16	3/4	13/16
$N_1 = 5$	$N_2 = 6$	3/5	2/3	2/3	5/6	5/6
	7	4/7	23/35	5/7	29/35	6/7
	8	11/20	5/8	27/40	4/5	4/5
	9	5/9	3/5	31/45	7/9	4/5
	10	1/2	3/5	7/10	7/10	4/5
	15	8/15	3/5	2/3	11/15	11/15
	20	1/2	11/20	3/5	7/10	3/4

$N_1 = 6$	$N_2 = 7$	23/42	4/7	29/42	5/7	5/6
	8	1/2	7/12	2/3	3/4	3/4
	9	1/2	5/9	2/3	13/18	7/9
	10	1/2	17/30	19/30	7/10	11/15
	12	1/2	7/12	7/12	2/3	3/4
	18	4/9	5/9	11/18	2/3	13/18
	24	11/24	1/2	7/12	5/8	2/3
$N_1 = 7$	$N_2 = 8$	27/56	33/56	5/8	41/56	3/4
	9	31/63	5/9	40/63	5/7	47/63
	10	33/70	39/70	43/70	7/10	5/7
	14	3/7	1/2	4/7	9/14	5/7
	28	3/7	13/28	15/28	17/28	9/14
$N_1 = 8$	$N_2 = 9$	4/9	13/24	5/8	2/3	3/4
	10	19/40	21/40	23/40	27/40	7/10
	12	11/24	1/2	7/12	5/8	2/3
	16	7/16	1/2	9/16	5/8	5/8
	32	13/32	7/16	1/2	9/16	19/32
$N_1 = 9$	$N_2 = 10$	7/15	1/2	26/45	2/3	31/45
	12	4/9	1/2	5/9	11/18	2/3
	15	19/45	22/45	8/15	3/5	29/45
	18	7/18	4/9	1/2	5/9	11/18
	36	13/36	5/12	17/36	19/36	5/9
$N_1 = 10$	$N_2 = 15$	2/5	7/15	1/2	17/30	19/30
	20	2/5	9/20	1/2	11/20	3/5
	40	7/20	2/5	9/20	1/2	
$N_1 = 12$	$N_2 = 15$	23/60	9/20	1/2	11/20	7/12
	16	3/8	7/16	23/48	13/24	7/12
	18	13/36	5/12	17/36	19/36	5/9
	20	11/30	5/12	7/15	31/60	17/30
$N_1 = 15$	$N_2 = 20$	7/20	2/5	13/30	29/60	31/60
$N_1 = 16$	$N_2 = 20$	27/80	31/80	17/40	19/40	41/80

Large sample approximation :

$$1.07\sqrt{\frac{m+n}{mn}} \quad 1.22\sqrt{\frac{m+n}{mn}} \quad 1.36\sqrt{\frac{m+n}{mn}} \quad 1.52\sqrt{\frac{m+n}{mn}} \quad 1.63\sqrt{\frac{m+n}{mn}}$$

Table-15 Significance tests in a 2 × 2 Contingency table

			Probaility			
		a	0.05	0.025	0.01	0.005
A = 3	*B* = 3	3	0.050	—	—	—
A = 4	*B* = 4	4	0.014	0.014	—	—
	3	4	0.029	—	—	—
A = 5	*B* = 5	5	1.024	1.024	0.004	0.004
		4	0.024	0.024	—	—
	4	5	1.048	0.008	0.008	—
		4	0.040	—	—	—
	3	5	0.018	0.018	—	—
	2	5	0.048	—	—	—
A = 6	*B* = 6	6	2.030	1.008	1.008	0.001
		5	1.040	0.008	0.008	—
		4	0.030	—	—	—
	5	6	1.015+	1.015+	0.002	0.002
		5	0.013	0.013	—	—
		4	0.045+	—	—	—
	4	6	1.033	0.005-	0.005-	0.005-
		5	0.024	0.024	—	—
	3	6	0.012	0.012	—	—
		5	0.048	—	—	—
	2	6	0.036	—	—	—
A = 7	*B* = 7	7	3.035-	2.010+	1.002	1.002
		6	1.015-	1.015-	0.002	0.002
		5	0.010+	0.010+	—	—
		4	0.035-	—	—	—
	6	7	2.021	2.021	1.005-	1.005-
		6	1.025+	0.004	0.004	0.004
		5	0.016	0.016	—	—
		4	0.049	—	—	—
	5	7	2.045+	1.010+	0.001	0.001
		6	1.045+	0.008	0.008	—
		5	0.027	—	—	—
	4	7	1.024	1.024	0.003	0.003
		6	0.015+	0.015+	—	—
		5	0.045+	—	—	—
	3	7	0.008	0.008	0.008	—
		6	0.033	—	—	—
	2	7	0.028	—	—	—
A = 8	*B* = 8	8	4.038	3.013	2.003	2.003
		7	2.020	2.020	1.005+	0.001
		6	1.020	0.020	0.003	0.003
		5	0.013	0.013	—	—

		4	0.038	—	—	—
A = 8	*B* = 7	8	3.026	2.007	2.007	1.001
		7	2.035-	1.009	1.009	0.001
		6	1.032	0.006	0.006	—
		5	0.019	0.019	—	—
	6	8	2.015-	2.015-	1.003	1.003
		7	1.016	1.016	0.002	0.002
		6	0.009	0.009	0.009	—
		5	0.028	—	—	—
	5	8	2.035-	1.007	1.007	0.001
		7	1.032	0.005-	0.005-	0.005-
		6	0.016	0.016	—	—
		5	0.044	—	—	—
	4	8	1.018	1.018	0.002	0.002
		7	0.010+	0.010+	—	—
		6	0.030	—	—	—
	3	8	0.006	0.006	0.006	—
		7	0.024	0.024	—	—
	2	8	0.022	0.022	—	—
A = 9	*B* = 9	9	5.041	4.015-	3.005-	3.005-
		8	3. 025-	3.025-	2.008	1.002
		7	2.028	1.008	1.008	0.001
		6	1.025-	1.025-	0.005-	0.005-
		5	0.015-	0.015-	—	—
		4	0.041	—		
	8	9	4.029	3.009	3.009	2.002
		8	3.043	2.013	1.003	0.003
		7	2.044	1.012	0.002	0.002
		6	1.034	0.007	0.007	—
		5	0.020	0.020	—	—
	7	9	3.019	3.019	2.005-	2.005-
		8	2.024	2.024	1.006	0.001
		7	1.020	1.020	0.003	0.003
		6	0.010+	0.010+	—	—
		5	0.029	—	—	—
	6	9	3.044	2.011	1.002	1.002
		8	2.047	1.011	0.001	0.001
		7	1.035-	0.006	0.006	—
		6	0.017	0.017	—	—
		5	0.42	—	—	—
A = 9	**B = 5**	9	2.027	1.005-	1.005-	1.005-
		8	1.023	1.023	0.003	0.003
		7	0.010+	0.010+	—	—
		6	0.028	—	—	—
	4	9	1.014	1.014	0.001	0.001
		8	0.007	0.007	0.007	—
		7	0.021	0.021	—	—
		6	0.049	—	—	—

	3	9	1.045+	0.005-	0.005-	0.005-
		8	0.018	0.018	—	—
		7	0.045+	—	—	—
	2	9	0.018	0.018	—	—
A = 10	**B = 10**	10	6.043	5.016	4.005+	3.002
		9	4.029	3.010-	3.010-	2.003
		8	3.035-	2.012	1.003	1.003
		7	2.035-	1.010-	1.010-	0.002
		6	1.029	0.005+	0.005+	—
		5	0.016	0.016	—	—
		4	0.043	—	—	—
A = 10	*B* = 9	10	5.033	4.011	3.003	3.003
		9	4.050-	3.017	2.005-	2.005-
		8	2.019	2.019	1.004	1.004
		7	1.015-	1.015-	0.002	0.002
		6	1.040	0.008	0.008	—
		5	0.022	0.022	—	—
	8	10	4.023	4.023	3.007	2.002
		9	3.032	2.009	2.009	1.002
		8	2.031	1.008	1.008	0.001
		7	1.023	1.023	0.004	0.004
		6	0.011	0.011	—	—
		5	0.029	—	—	—
	7	10	3.015-	3.015-	2.003	2.003
		9	2.018	2.018	1.004	1.004
		8	1.013	1.013	0.002	0.002
		7	1.036	0.006	0.006	—
		6	0.017	0.017	—	—
		5	0.041	—	—	—
	6	10	3.036	2.008	2.008	1.001
		9	2.036	1.008	1.008	0.001
		8	1.024	1.024	0.003	0.003
		7	0.010+	0.010+	—	—
		6	0.026	—	—	—
	5	10	2.022	2.022	1.004	1.004
		9	1.017	1.017	0.002	0.002
		8	1.047	0.007	0.007	—
		7	0.019	0.019	—	—
		6	0.042	—	—	—
	4	10	1.011	1.011	0.001	0.001
		9	1.041	0.005-	0.005-	0.005-
		8	0.015-	0.015-	—	—
		7	0.035-	—	—	—
	3	10	1.038	0.003	0.003	0.003
		9	0.014	0.014	—	—
		8	0.035-	—	—	—
	2	10	0.015+	0.015+	—	—
		9	0.045+	—	—	—

$A = 11$	$B = 11$	11	7.045	6.018	5.006	4.002
		10	5.032	4.012	3.004	3.004
		9	4.040	3.015-	2.004	2.004
		8	3.043	2.015-	1.004	1.004
		7	2.040	1.012	0.002	0.002
		6	1.032	0.006	0.006	—
		5	0.018	0.018	—	—
		4	0.045+	—	—	—
	10	11	6.035+	5.012	4.004	4.004
		10	4.021	4.021	3.007	2.002
		9	3.024	3.024	2.007	1.002
		8	2.023	2.023	1.006	0.001
		7	1.017	1.017	0.003	0.003
		6	1.043	0.009	0.009	—
		5	0.023	0.023	—	—
	9	11	5.026	4.008	4.008	3.002
		10	4.038	3.012	2.003	2.003
		9	3.040	2.012	1.003	1.003
		8	2.035-	1.009	1.009	0.001
		7	1.025-	1.025-	0.004	0.004
		6	0.012	0.012	—	—
		5	0.030	—	—	—
$A = 11$	$B = 8$	11	4.018	4.018	3.005-	3.005-
		10	3.024	3.024	2.006	1.001
		9	2.022	2.022	1.005-	1.005-
		8	1.015-	1.015-	0.002	0.002
		7	1.037	0.007	0.007	—
		6	0.017	0.017	—	—
		5	0.040	—	—	—
	7	11	4.043	3.011	2.002	2.002
		10	3.047	2.013	1.002	1.002
		9	2.039	1.009	1.009	0.001
		8	1.025-	1.025-	0.004	0.004
		7	0.010+	0.010+	—	—
		6	0.025-	0.025-	—	—
	6	11	3.029	2.006	2.006	1.001
		10	2.028	1.005+	1.005+	0.001
		9	1.018	1.018	0.002	0.002
		8	1.043	0.007	0.007	—
		7	0.017	0.017	—	—
		6	0.037	—	—	—
	5	11	2.018	2.018	1.003	1.003
		10	1.013	1.013	0.001	0.001
		9	1.035	0.005-	0.005-	0.005-
		8	0.013	0.013	—	—
		7	0.029	—	—	—
	4	11	1.009	1.009	1.009	0.001
		10	1.033	0.004	0.004	0.004

		9	0.011	0.011	—	—
		8	0.026	—	—	—
	3	11	1.033	0.003	0.003	0.003
		10	0.011	0.011	—	—
		9	0.027	—	—	—
	2	11	0.013	0.013	—	—
		10	0.038	—	—	—
A = 12	*B* = 12	12	8.047	7.019	6.007	5.002
		11	6.034	5.014	4.005-	4.005-
		10	5.045-	4.018	3.006	2.002
		9	4.050-	3.020	2.006	1.001
		8	3.050-	2.018	1.005-	1.005-
		7	2.045-	1.014	0.002	0.002
		6	1.034	0.007	0.007	—
		5	0.019	0.019	—	—
		4	0.047	—	—	—
	11	12	7.037	6.014	5.005-	5.005-
		11	5.024	5.024	4.008	3.002
		10	4.029	3.010+	2.003	2.003
		9	3.030	2.009	2.009	1.002
		8	2.026	1.007	1.007	0.001
		7	1.019	1.019	0.003	0.003
		6	1.045-	0.009	0.009	—
		5	0.024	0.024	—	—
	10	12	6.029	5.010-	5.010-	4.003
		11	5.043	4.015+	3.005-	3.005-
		10	4.048	3.017	2.005-	2.005-
		9	3.046	2.015-	1.004	1.004
		8	2.038	1.010+	0.002	0.002
		7	1.026	0.005-	0.005-	0.005-
		6	0.012	0.012	—	—
		5	0.030	—	—	—
A = 12	*B* = 9	12	5.021	5.021	4.006	3.002
		11	4.029	3.009	3.009	2.002
		10	3.029	2.008	2.008	1.002
		9	2.024	2.024	1.006	0.001
		8	1.016	1.016	0.002	0.002
		7	1.037	0.007	0.007	—
		6	0.017	0.017	—	—
		5	0.039	—	—	—
	8	12	5.049	4.014	3.004	3.004
		11	3.018	3.018	2.004	2.004
		10	2.015+	2.015+	1.003	1.003
		9	2.040	1.010-	1.010-	0.001
		8	1.025-	1.025-	0.004	0.004
		7	0.010+	0.010+	—	—
		6	0.024	0.024	—	—

	7	12	4.036	3.009	3.009	2.002
		11	3.038	2.010-	2.010-	1.002
		10	2.029	1.006	1.006	0.001
		9	1.017	1.017	0.002	0.002
		8	1.040	0.007	0.007	—
		7	0.016	0.016	—	—
		6	0.034	—	—	—
	6	12	3.025-	3.025-	2.005-	2.005-
		11	2.022	2.022	1.004	1.004
		10	1.013	1.013	0.002	0.002
		9	1.032	0.005-	0.005-	0.005-
		8	0.011	0.011	—	—
		7	0.025-	0.025-	—	—
		6	0.050	—	—	—
	5	12	2.015-	2.015-	1.002	1.002
		11	1.010-	1.010-	1.010-	0.001
		10	1.028	0.003	0.003	0.003
		9	0.009	0.009	0.009	—
		8	0.020	0.020	—	—
		7	0.041	—	—	—
	4	12	2.050	1.007	1.007	0.001
		11	1.027	0.003	0.003	0.003
		10	0.008	0.008	0.008	—
		9	0.019	0.019	—	—
		8	0.038	—	—	—
	3	12	1.029	0.002	0.002	0.002
		11	0.009	0.009	0.009	—
		10	0.022	0.022	—	—
		9	0.044	—	—	—
	2	12	0.011	0.011	—	—
		11	0.033	—	—	—
A = 13	*B* = 13	13	9.048	8.020	7.007	6.003
		12	7.037	6.015+	5.006	4.002
		11	6.048	5.021	4.008	3.002
		10	4.024	4.024	3.008	2.002
		9	3.024	3.024	2.008	1.002
		8	2.021	2.021	1.006	0.001
		7	2.048	1.015+	0.003	0.003
		6	1.037	0.007	0.007	—
		5	0.020	0.020	—	—
		4	0.048	—	—	—
	12	13	8.039	7.015-	6.005+	5.002
		12	6.027	5.010-	5.010-	4.003
		11	5.033	4.013	3.004	3.004
		10	4.036	3.013	2.004	2.004
A = 13	*B* = 12	9	3.034	2.011	1.003	1.003
		8	2.029	1.008	1.008	0.001
		7	1.020	1.020	0.004	0.004
		6	1.046	0.010-	0.010-	—

	5	0.024-	0.024-	—	—
11	13	7.031	6.011	5.003	5.003
	12	6.048	5.018	4.006	3.002
	11	4.021	4.021	3.007	2.002
	10	3.021	3.021	2.006	1.001
	9	3.050-	2.017	1.004	1.004
	8	2.040	1.011	0.002	0.002
	7	1.027	0.005-	0.005-	0.005-
	6	0.013	0.013	—	—
	5	0.030	—	—	—
10	13	6.024	6.024	5.007	4.002
	12	5.035-	4.012	3.003	3.003
	11	4.037	3.012	2.003	2.003
	10	3.033	2.010+	1.002	1.002
	9	2.026	1.006	1.006	0.001
	8	1.017	1.017	0.003	0.003
	7	1.038	0.007	0.007	—
	6	0.017	0.017	—	—
	5	0.038	—	—	—
9	13	5.017	5.017	4.005-	4.005-
	12	4.023	4.023	3.007	2.001
	11	3.022	3.022	2.006	1.001
	10	2.017	2.017	1.004	1.004
	9	2.040	1.010+	0.001	0.001
	8	1.025-	1.025-	0.004	0.004
	7	0.010+	0.010+	—	—
	6	0.023	0.023	—	—
	5	0.049	—	—	—
8	13	5.042	4.012	3.003	3.003
	12	4.047	3.014	2.003	2.003
	11	3.041	2.011	1.002	1.002
	10	2.029	1.007	1.007	0.001
	9	1.017	1.017	0.002	0.002
	8	1.037	0.006	0.006	—
	7	0.015-	0.015-	—	—
	6	0.032	—	—	—
7	13	4.031	3.007	3.007	2.001
	12	3.031	2.007	2.007	1.001
	11	2.022	2.022	1.004	1.004
	10	1.012	1.012	0.002	0.002
	9	1.029	0.004	0.004	0.004
	8	0.010+	0.010+	—	—
	7	0.022	0.022	—	—
	6	0.044	—	—	—
6	13	3.021	3.021	2.004	2.004
	12	2.017	2.017	1.003	1.003
	11	2.046	1.010-	1.010-	0.001
	10	1.024	1.024	0.003	0.003
	9	1.050-	0.008	0.008	—
	8	0.017	0.017	—	—

		7	0.037	—	—	—
	5	13	2.012	2.012	1.002	1.002
		12	2.004	1.008	1.008	0.001
		11	1.022	1.022	0.002	0.002
		10	1.047	0.007	0.007	—
		9	0.015-	0.015-	—	—
		8	0.029	—	—	—
A = 13	*B* = 4	13	2.044	1.006	1.006	0.000
		12	1.022	1.022	0.002	0.002
		11	0.006	0.006	0.006	—
		10	0.015-	0.015-	—	—
		9	0.029	—	—	—
	3	13	1.025	1.025	0.002	0.002
		12	0.007	0.007	0.007	—
		11	0.018	0.018	—	—
		10	0.036	—	—	—
	2	13	0.010-	0.010-	0.010-	—
		12	0.029	—	—	—
A = 14	*B* = 14	14	10.049	9.020	8.008	7.003
		13	8.038	7.016	6.006	5.002
		12	6.023	6.023	5.009	4.003
		11	5.027	4.011	3.004	3.004
		10	4.028	3.011	2.003	2.003
		9	3.027	2.009	2.009	1.002
		8	2.023	2.023	1.006	0.001
		7	1.016	1.016	0.003	0.003
		6	1.038	0.008	0.008	—
		5	0.020	0.020	—	—
		4	0.049	—	—	—
	13	14	9.041	8.016	7.006	6.002
		13	7.029	6.011	5.004	5.004
		12	6.037	5.015+	4.005-	3.002
		11	5.041	4.017	3.006	2.001
		10	4.041	3.016	2.005-	2.005-
		9	3.038	2.013	1.003	1.003
		8	2.031	1.009	1.009	0.001
		7	1.021	1.021	0.004	0.004
		6	1.048	0.010+	—	—
		5	0.025-	0.025-	—	—
	12	14	8.033	7.012	6.004	6.004
		13	6.021	6.021	5.007	4.002
		12	5.025+	4.009	4.009	3.003
		11	4.026	3.009	3.009	2.002
		10	3.024	3.024	2.007	1.002
		9	2.019	2.019	1.005-	1.005-
		8	2.042	1.012	0.002	0.002
		7	1.028	0.005+	0.005+	—
		6	0.013	0.013	—	—
		5	0.030	—	—	—

	11	14	7.026	6.009	6.009	5.003
		13	6.039	5.014	4.004	4.004
		12	5.043	4.016	3.005-	3.005-
		11	4.042	3.015-	2.004	2.004
		10	3.036	2.011	1.003	1.003
		9	2.027	1.007	1.007	0.001
		8	1.017	1.017	0.003	0.003
		7	1.038	0.007	0.007	—
		6	0.017	0.017	—	—
		5	0.038	—	—	—
	10	14	6.020	6.020	5.006	4.002
		13	5.028	4.009	4.009	3.002
		12	4.028	3.009	3.009	2.002
		11	3.024	3.024	2.007	1.001
		10	2.018	2.018	1.004	1.004
		9	2.040	1.011	0.002	0.002
		8	1.024	1.024	0.004	0.004
		7	0.010-	0.010-	0.010-	—
$A = 14$	$B = 10$	6	0.022	0.022	—	—
		5	0.047	—	—	—
	9	14	6.047	5.014	4.004	4.004
		13	4.018	4.018	3.005-	3.005-
		12	3.017	3.017	2.004	2.004
		11	3.042	2.012	1.002	1.002
		10	2.029	1.007	1.007	0.001
		9	1.017	1.017	0.002	0.002
		8	1.036	0.006	0.006	—
		7	0.014	0.014	—	—
		6	0.030	—	—	—
	8	14	5.036	4.010-	4.010-	3.002
		13	4.039	3.011	2.002	2.002
		12	3.032	2.008	2.008	1.001
		11	2.022	2.022	1.005-	1.005-
		10	2.048	1.012	0.002	0.002
		9	1.026	0.004	0.004	0.004
		8	0.009	0.009	0.009	—
		7	0.020	0.020	—	—
		6	0.040	—	—	—
	7	14	4.026	3.006	3.006	2.001
		13	3.025	2.006	2.006	1.001
		12	2.017	2.017	1.003	1.003
		11	2.041	1.009	1.009	0.001
		10	1.021	1.021	0.003	0.003
		9	1.043	0.007	0.007	—
		8	0.015-	0.015-	—	—
		7	0.030	—	—	—
	6	14	3.018	3.018	2.003	2.003
		13	2.014	2.014	1.002	1.002
		12	2.037	1.007	1.007	0.001

		11	1.018	1.018	0.002	0.002
		10	1.038	0.005+	0.005+	—
		9	0.012	0.012	—	—
		8	0.024	0.024	—	—
		7	0.044	—	—	—
	5	14	2.010+	2.010+	1.001	1.001
		13	2.037	1.006	1.006	0.001
		12	1.017	1.017	0.002	0.002
		11	1.038	0.005-	0.005-	0.005-
		10	0.011	0.011	—	—
		9	0.022	0.022	—	—
		8	0.040	—	—	—
	4	14	2.039	1.005-	1.005-	1.005-
		13	1.019	1.019	0.002	0.002
		12	1.044	0.005-	0.005-	0.005-
		11	0.011	0.011	—	—
		10	0.023	0.023	—	—
		9	0.041	—	—	—
	3	14	1.022	1.022	0.001	0.001
		13	0.006	0.006	0.006	—
		12	0.015-	0.015-	—	—
		11	0.029	—	—	—
	2	14	0.008	0.008	0.008	—
		13	0.025	0.025	—	—
		12	0.050	—	—	—
$A = 15$	**$B = 15$**	15	11.050-	10.021	9.008	8.003
		14	9.040	8.018	7.007	6.003
		13	7.025+	6.010+	5.004	5.004
		12	6.030	5.013	4.005-	4.005-
$A = 15$	**$B = 15$**	11	5.033	4.013	3.005-	3.005-
		10	4.033	3.013	2.004	2.004
		9	3.030	2.010+	1.003	1.003
		8	2.025+	1.007	1.007	0.001
		7	1.018	1.018	0.003	0.003
		6	1.040	0.008	0.008	—
		5	0.021	0.021	—	—
		4	0.050-	—	—	—
	14	15	10.042	9.017	8.006	7.002
		14	8.031	7.013	6.005-	6.005-
		13	7.041	6.017	5.007	4.002
		12	6.046	5.020	4.007	3.002
		11	5.048	4.020	3.007	2.002
		10	4.046	3.018	2.006	1.001
		9	3.041	2.014	1.004	1.004
		8	2.033	1.009	1.009	0.001
		7	1.022	1.022	0.004	0.004
		6	1.049	0.011	—	—
		5	0.025+	—	—	—
	13	15	9.035-	8.013	7.005-	7.005-

	14	7.023	7.023	6.009	5.003
	13	6.029	5.011	4.004	4.004
	12	5.031	4.012	3.004	3.004
	11	4.030	3.011	2.003	2.003
	10	3.026	2.008	2.008	1.002
	9	2.020	2.020	1.005+	0.001
	8	2.043	1.013	0.002	0.002
	7	1.029	0.005+	0.005+	—
	6	0.013	0.013	—	—
	5	0.031	—	—	—
12	15	8.028	7.010-	7.010-	6.003
	14	7.043	6.016	5.006	4.002
	13	6.049	5.019	4.007	3.002
	12	5.049	4.019	3.006	2.002
	11	4.045+	3.017	2.005-	2.005-
	10	3.038	2.012	1.003	1.003
	9	2.028	1.007	1.007	0.001
	8	1.018	1.018	0.003	0.003
	7	1.038	0.007	0.007	—
	6	0.017	0.017	—	—
	5	0.037	—	—	—
11	15	7.022	7.022	6.007	5.002
	14	6.032	5.011	4.003	4.003
	13	5.034	4.012	3.003	3.003
	12	4.032	3.010+	2.003	2.003
	11	3.026	2.008	2.008	1.002
	10	2.019	2.019	1.004	1.004
	9	2.040	1.011	0.002	0.002
	8	1.024	1.024	0.004	0.004
	7	1.049	0.010-	0.010-	—
	6	0.022	0.022	—	—
	5	0.046	—	—	—
10	15	6.017	6.017	5.005-	5.005-
	14	5.023	5.023	4.007	3.002
	13	4.022	4.022	3.007	2.001
	12	3.018	3.018	2.005-	2.005-
	11	3.042	2.013	1.003	1.003
	10	2.029	1.007	1.007	0.001
	9	1.016	1.016	0.002	0.002
	8	1.034	0.006	0.006	—
$A = 15$ $B = 10$	0.013	0.013	—	—	
	6	0.028	—	—	—
9	15	6.042	5.012	4.003	4.003
	14	5.047	4.015-	3.004	3.004
	13	4.042	3.013	2.003	2.003
	12	3.032	2.009	2.009	1.002
	11	2.021	2.021	1.005-	1.005-
	10	2.045	1.011	0.002	0.002
	9	1.024	1.024	0.004	0.004

	8	1.048	0.009	0.009	—
	7	0.019	0.019	—	—
	6	0.037	—	—	—
8	15	5.032	4.008	4.008	3.002
	14	4.033	3.009	3.009	2.002
	13	3.026	2.006	2.006	1.001
	12	2.017	2.017	1.003	1.003
	11	2.037	1.008	1.008	0.001
	10	1.019	1.019	0.003	0.003
	9	1.038	0.006	0.006	—
	8	0.013	0.013	—	—
	7	0.026	—	—	—
	6	0.050	—	—	—
7	15	4.023	4.023	3.005-	3.005-
	14	3.021	3.021	2.004	2.004
	13	2.014	2.014	1.002	1.002
	12	2.032	1.007	1.007	0.001
	11	1.015+	1.015+	0.002	0.002
	10	1.032	0.005-	0.005-	0.005-
	9	0.010+	0.010+	—	—
	8	0.020	0.020	—	—
	7	0.038	—	—	—
6	15	3.015+	3.015+	2.003	2.003
	14	2.011	2.011	1.002	1.002
	13	2.031	1.006	1.006	0.001
	12	1.014	1.014	0.002	0.002
	11	1.029	0.004	0.004	0.004
	10	0.009	0.009	0.009	—
	9	0.017	0.017	—	—
	8	0.032	—	—	—
5	15	2.009	2.009	2.009	1.001
	14	2.032	1.005-	1.005-	1.005-
	13	1.014	1.014	0.001	0.001
	12	1.031	0.004	0.004	0.004
	11	0.008	0.008	0.008	—
	10	0.016	0.016	—	—
	9	0.030	—	—	—
4	15	2.035+	1.004	1.004	1.004
	14	1.016	1.016	0.001	0.001
	13	1.037	0.004	0.004	0.004
	12	0.009	0.009	0.009	—
	11	0.018	0.018	—	—
	10	0.033	—	—	—
3	15	1.020	1.020	0.001	0.001
	14	0.005-	0.005-	0.005-	0.005-
	13	0.012	0.012	—	—
	12	0.025-	0.025-	—	—
	11	0.043	—	—	—
2	15	0.007	0.007	0.007	—
	14	0.022	0.022	—	—

		13	0.044	—	—	—
A = 16	*B* = 16	16	11.022	11.022	10.009	9.003
		15	10.041	9.019	8.008	7.003
		14	8.027	7.012	6.005-	6.005-
		13	7.033	6.015-	5.006	4.002
		12	6.037	5.016	4.006	3.002
		11	5.038	4.016	3.006	2.002
		10	4.037	3.015-	2.005-	2.005-
		9	3.033	2.012	1.003	1.003
		8	2.027	1.008	1.008	0.001
		7	1.019	1.019	0.003	0.003
		6	1.041	0.009	0.009	—
		5	0.022	0.022	—	—
	15	16	11.043	10.018	9.007	8.002
		15	9.033	8.014	7.005+	6.002
		14	8.044	7.019	6.008	5.003
		13	6.023	6.023	5.009	4.003
		12	5.024	5.024	4.009	3.003
		11	4.023	4.023	3.008	2.002
		10	4.049	3.020	2.006	1.001
		9	3.043	2.016	1.004	1.004
		8	2.035-	1.010+	0.002	0.002
		7	1.023	1.023	0.004	0.004
		6	0.011	0.011	—	—
		5	0.026	—	—	—
	14	16	10.037	9.014	8.005+	7.002
		15	8.025+	7.010-	7.010-	6.003
		14	7.032	6.013	5.005-	5.005-
		13	6.035+	5.014	4.005+	3.001
		12	5.035+	4.014	3.005-	3.005-
		11	4.033	3.012	2.004	2.004
		10	3.028	2.009	2.009	1.002
		9	2.021	2.021	1.006	0.001
		8	2.045-	1.013	0.002	0.002
		7	1.030	0.006	0.006	—
		6	0.013	0.013	—	—
		5	0.031	—	—	—
	13	16	9.030	8.011	7.004	7.004
		15	8.047	7.019	6.007	5.002
		14	6.023	6.023	5.008	4.003
		13	5.023	5.023	4.008	3.003
		12	4.022	4.022	3.007	2.002
		11	4.048	3.018	2.005+	1.001
		10	3.039	2.013	1.003	1.003
		9	2.029	1.008	1.008	0.001
		8	1.018	1.018	0.003	0.003
		7	1.038	0.007	0.007	—
		6	0.017	0.017	—	—
		5	0.037	—	—	—

	12	16	8.024	8.024	7.008	6.002
		15	7.036	6.013	5.004	5.004
		14	6.040	5.015-	4.005-	4.005-
		13	5.039	4.014	3.004	3.004
		12	4.034	3.012	2.003	2.003
		11	3.027	2.008	2.008	1.002
		10	2.019	2.019	1.005-	1.005-
		9	2.040	1.011	0.002	0.002
		8	1.024	1.024	0.004	0.004
		7	1.048	0.010-	0.010-	—
		6	0.021	0.021	—	—
		5	0.044	—	—	—
$A = 16$	$B = 11$	16	7.019	7.019	6.006	5.002
		15	6.027	5.009	5.009	4.002
		14	5.027	4.009	4.009	3.002
		13	4.024	4.024	3.008	2.002
		12	3.019	3.019	2.005+	1.001
		11	3.041	2.013	1.003	1.003
		10	2.028	1.007	1.007	0.001
		9	1.016	1.016	0.002	0.002
		8	1.033	0.006	0.006	—
		7	0.013	0.013	—	—
		6	0.027	—	—	—
	10	16	7.046	6.014	5.004	5.004
		15	5.018	5.018	4.005+	3.001
		14	4.018	4.018	3.005-	3.005-
		13	4.042	3.014	2.003	2.003
		12	3.032	2.009	2.009	1.002
		11	2.021	2.021	1.005-	1.005-
		10	2.042	1.011	0.002	0.002
		9	1.023	1.023	0.004	0.004
		8	1.045-	0.008	0.008	—
		7	0.017	0.017	—	—
		6	0.035	—	—	—
	9	16	6.037	5.010-	5.010-	4.002
		15	5.040	4.012	3.003	3.003
		14	4.034	3.010-	3.010-	2.002
		13	3.025+	2.007	2.007	1.001
		12	2.016	2.016	1.003	1.003
		11	2.033	1.008	1.008	0.001
		10	1.017	1.017	0.002	0.002
		9	1.034	0.006	0.006	—
		8	0.012	0.012	—	—
		7	0.024	0.024	—	—
		6	0.045+	—	—	—
	8	16	5.028	4.007	4.007	3.001
		15	4.028	3.007	3.007	2.001
		14	3.021	3.021	2.005-	2.005-
		13	3.047	2.013	1.002	1.002

		12	2.028	1.006	1.006	0.001
		11	1.014	1.014	0.002	0.002
		10	1.027	0.004	0.004	0.004
		9	0.009	0.009	0.009	—
		8	0.017	0.017	—	—
		7	0.033	—	—	—
	7	16	4.020	4.020	3.004	3.004
		15	3.017	3.017	2.003	2.003
		14	3.045+	2.011	1.002	1.002
		13	2.026	1.005-	1.005-	1.005-
		12	1.012	1.012	0.001	0.001
		11	1.024	1.024	0.003	0.003
		10	1.045-	0.007	0.007	—
		9	0.014	0.014	—	—
		8	0.026	—	—	—
		7	0.047	—	—	—
	6	16	3.013	3.013	2.002	2.002
		15	3.046	2.009	2.009	1.001
		14	2.025+	1.004	1.004	1.004
		13	1.011	1.011	0.001	0.001
		12	1.023	1.023	0.003	0.003
		11	1.043	0.006	0.006	—
		10	0.012	0.012	—	—
$A = 16$	$B = 6$	9	0.023	0.023	—	—
		8	0.040	—	—	—
	5	16	3.048	2.008	2.008	1.001
		15	2.028	1.004	1.004	1.004
		14	1.011	1.011	0.001	0.001
		13	1.025+	0.003	0.003	0.003
		12	1.047	0.006	0.006	—
		11	0.012	0.012	—	—
		10	0.023	0.023	—	—
		9	0.039	—	—	—
	4	16	2.032	1.004	1.004	1.004
		15	1.013	1.013	0.001	0.001
		14	1.032	0.003	0.003	0.003
		13	0.007	0.007	0.007	—
		12	0.014	0.014	—	—
		11	0.026	—	—	—
		10	0.043	—	—	—
	3	16	1.018	1.018	0.001	0.001
		15	0.004	0.004	0.004	0.004
		14	0.010+	0.010+	—	—
		13	0.021	0.021	—	—
		12	0.036	—	—	—
	2	16	0.007	0.007	0.007	—
		15	0.020	0.020	—	—
		14	0.039	—	—	—
$A = 17$	$B = 17$	17	12.022	12.022	11.009	10.004

	16	11.043	10.020	9.008	8.003
	15	9.029	8.013	7.005+	6.002
	14	8.035+	7.016	6.007	5.002
	13	7.040	6.018	5.007	4.003
	12	6.042	5.019	4.007	3.002
	11	5.042	4.018	3.007	2.002
	10	4.040	3.016	2.005	1.001
	9	3.035+	2.013	1.003	1.003
	8	2.029	1.008	1.008	0.001
	7	1.020	1.020	0.004	0.004
	6	1.043	0.009	0.009	—
	5	0.022	0.022	—	—
16	17	12.044	11.018	10.007	9.003
	16	10.035-	9.015-	8.006	7.002
	15	9.046	8.021	7.009	6.003
	14	7.025+	6.011	5.004	5.004
	13	6.027	5.011	4.004	4.004
	12	5.027	4.011	3.004	3.004
	11	4.025+	3.009	3.009	2.003
	10	3.022	3.022	2.007	1.002
	9	3.046	2.017	1.004	1.004
	8	2.036	1.011	0.002	0.002
	7	1.024	1.024	0.005-	0.005-
	6	0.011	0.011	—	—
	5	0.026	—	—	—
15	17	11.038	10.015-	9.006	8.002
	16	9.027	8.011	7.004	7.004
	15	8.035+	7.015-	6.006	5.002
	14	7.040	6.017	5.006	4.002
	13	6.041	5.017	4.006	3.002
	12	5.039	4.016	3.005+	2.001
	11	4.035+	3.013	2.004	2.004
	10	3.029	2.010-	2.010-	1.002
	9	2.022	2.022	1.008	0.001
A = 17 *B* = 15	8	2.046	1.014	0.002	0.002
	7	1.030	0.006	0.006	—
	6	0.014	0.014	—	—
	5	0.031	—	—	—
14	17	10.032	9.012	8.004	8.004
	16	8.021	8.021	7.008	6.003
	15	7.026	6.010-	6.010-	5.003
	14	6.028	5.011	4.004	4.004
	13	5.027	4.010-	4.010-	3.003
	12	4.024	4.024	3.008	2.002
	11	4.049	3.019	2.006	1.001
	10	3.040	2.014	1.003	1.003
	9	2.029	1.008	1.008	0.001
	8	1.018	1.018	0.003	0.003
	7	1.038	0.007	0.007	—

		6	0.017	0.017	—	—
		5	0.036	—	—	—
	13	17	9.026	8.009	8.009	7.003
		16	8.040	7.015+	6.005+	5.002
		15	7.045+	6.018	5.006	4.002
		14	6.045+	5.018	4.006	3.002
		13	5.042	4.016	3.005+	2.001
		12	4.035+	3.013	2.004	2.004
		11	3.028	2.009	2.009	1.002
		10	2.019	2.019	1.005-	1.005-
		9	2.040	1.011	0.002	0.002
		8	1.024	1.024	0.004	0.004
		7	1.047	0.010-	0.010-	—
		6	0.021	0.021	—	—
		5	0.043	—	—	—
	12	17	8.021	8.021	7.007	6.002
		16	7.030	6.011	5.003	5.003
		15	6.033	5.012	4.004	4.004
		14	5.030	4.011	3.003	3.003
		13	4.026	3.008	3.008	2.002
		12	3.020	3.020	2.006	1.001
		11	3.041	2.013	1.003	1.003
		10	2.028	1.007	1.007	0.001
		9	1.016	1.016	0.002	0.002
		8	1.032	0.006	0.006	—
		7	0.012	0.012	—	—
		6	0.026	—	—	—
	11	17	7.016	7.016	6.005-	6.005-
		16	6.022	6.022	5.007	4.002
		15	5.022	5.022	4.007	3.002
		14	4.019	4.019	3.006	2.001
		13	4.042	3.014	2.004	2.004
		12	3.031	2.009	2.009	1.002
		11	2.020	2.020	1.005-	1.005-
		10	2.040	1.011	0.001	0.001
		9	1.022	1.022	0.004	0.004
		8	1.042	0.008	0.008	—
		7	0.016	0.016	—	—
		6	0.033	—	—	—
	10	17	7.041	6.012	5.003	5.003
		16	6.047	5.015+	4.004	4.004
		15	5.043	4.014	3.004	3.004
		14	4.034	3.010+	2.002	2.002
		13	3.024	3.024	2.007	1.001
		12	3.049	2.015+	1.003	1.003
A = 17	*B* = 10	11	2.031	1.007	1.007	0.001
		10	1.016	1.016	0.002	0.002
		9	1.031	0.005+	0.005+	—
		8	0.011	0.011	—	—

	7	0.022	0.022	—	—
	6	0.042	—	—	—
9	17	6.032	5.008	5.008	4.002
	16	5.034	4.010-	4.010-	3.002
	15	4.028	3.008	3.008	2.002
	14	3.020	3.020	2.005-	2.005-
	13	3.042	2.012	1.002	1.002
	12	2.025+	1.006	1.006	0.001
	11	2.048	1.012	0.002	0.002
	10	1.024	1.024	0.004	0.004
	9	1.045-	0.008	0.008	—
	8	0.016	0.016	—	—
	7	0.030	—	—	—
8	17	5.024	5.024	4.006	3.001
	16	4.023	4.023	3.006	2.001
	15	3.017	3.017	2.004	2.004
	14	3.039	2.010-	2.010-	1.002
	13	2.022	2.022	1.004	1.004
	12	2.043	1.010-	1.010-	0.001
	11	1.020	1.020	0.003	0.003
	10	1.038	0.006	0.006	—
	9	0.012	0.012	—	—
	8	0.022	0.022	—	—
	7	0.040	—	—	—
7	17	4.017	4.017	3.003	3.003
	16	3.014	3.014	2.003	2.003
	15	3.038	2.009	2.009	1.001
	14	2.021	2.021	1.004	1.004
	13	2.042	1.009	1.009	0.001
	12	1.018	1.018	0.002	0.002
	11	1.034	0.005-	0.005-	0.005-
	10	0.010-	0.010-	0.010-	—
	9	0.019	0.019	—	—
	8	0.033	—	—	—
6	17	3.011	3.011	2.002	2.002
	16	3.040	2.008	2.008	1.001
	15	2.021	2.021	1.003	1.003
	14	2.045+	1.009	1.009	0.001
	13	1.018	1.018	0.002	0.002
	12	1.035-	0.005-	0.005-	0.005-
	11	0.009	0.009	0.009	—
	10	0.017	0.017	—	—
	9	0.030	—	—	—
	8	0.050-	—	—	—
5	17	3.043	2.006	2.006	1.001
	16	2.024	2.024	1.003	1.003
	15	1.009	1.009	1.009	0.001
	14	1.021	1.021	0.002	0.002
	13	1.039	0.005-	0.005-	0.005-
	12	0.010-	0.010-	0.010-	—

		11	**0**.018	**0**.018	—	—
		10	**0**.030	—	—	—
		9	**0**.049	—	—	—
	4	17	**2**.029	**1**.003	**1**.003	**1**.003
		16	**1**.011	**1**.011	**0**.001	**0**.001
		15	**1**.028	**0**.003	**0**.003	**0**.003
		14	**0**.006	**0**.006	**0**.006	—
A = 17	*B* = 4	13	**0**.012	**0**.012	—	—
		12	**0**.021	**0**.021	—	—
		11	**0**.035+	—	—	—
	3	17	**1**.016	**1**.016	**0**.001	**0**.001
		16	**1**.046	**0**.004	**0**.004	**0**.004
		15	**0**.009	**0**.009	**0**.009	—
		14	**0**.018	**0**.018	—	—
		13	**0**.031	—	—	—
		12	**0**.049	—	—	—
	2	17	**0**.006	**0**.006	**0**.006	—
		16	**0**.018	**0**.018	—	—
		15	**0**.035+	—	—	—
A = 18	*B* = 18	18	**13**.032	**13**.032	**12**.010-	**11**.004
		17	**12**.044	**11**.020	**10**.009	**9**.004
		16	**10**.030	**9**.014	**8**.006	**7**.002
		15	**9**.038	**8**.018	**7**.008	**6**.003
		14	**8**.043	**7**.020	**6**.009	**5**.003
		13	**7**.046	**6**.022	**5**.009	**4**.003
		12	**6**.047	**5**.022	**4**.009	**3**.003
		11	**5**.046	**4**.020	**3**.008	**2**.002
		10	**4**.043	**3**.018	**2**.006	**1**.001
		9	**3**.038	**2**.014	**1**.004	**1**.004
		8	**2**.030	**1**.009	**1**.009	**0**.001
		7	**1**.020	**1**.020	**0**.004	**0**.004
		6	**1**.044	**0**.010-	**0**.010-	—
		5	**0**.023	**0**.023	—	—
	17	18	**13**.045+	**12**.019	**11**.008	**10**.003
		17	**11**.036	**10**.016	**9**.007	**8**.002
		16	**10**.049	**9**.023	**8**.010-	**7**.004
		15	**8**.028	**7**.012	**6**.005-	**6**.005-
		14	**7**.030	**6**.013	**5**.005+	**4**.002
		13	**6**.031	**5**.013	**4**.005-	**4**.005-
		12	**5**.030	**4**.012	**3**.004	**3**.004
		11	**4**.028	**3**.010+	**2**.003	**2**.003
		10	**3**.023	**3**.023	**2**.008	**1**.002
		9	**3**.047	**2**.018	**1**.005-	**1**.005-
		8	**2**.037	**1**.011	**0**.002	**0**.002
		7	**1**.025-	**1**.025-	**0**.005-	**0**.005-

1. Bold type, for given a, *A* and *B*, shows the vale of *b* (< *a*) which is just significant at the probability level quoted (single-tails test).

2. Small type, for given *A*, *B* and $r = a + b$, shows the exact probability (if there is independence) that *b* is equal to or less than the integer shown in bold type.

Table A-16 Kendall's confficient of concordance

$m = 3$

$n = 2$

W	p
.000	1.000
.250	.833
.750	.500
1.000	.167

$n = 3$

W	p
.000	1.000
.111	.944
.333	.528
.444	.361
.778	.194
1.000	0.28

$n = 4$

W	p
.000	1.000
.062	.931
.188	.653
.250	.431
.438	.273
.562	.125
.750	.069
.812	.042
1.000	.005

$n = 5$

W	p
.000	1.000
.040	.954
.120	.691
.160	.522
.280	.367
.360	.182
.480	1.24
.520	.093
.640	.039
.760	.024
.840	.008
1.000	.001

$n = 6$

$n = 6$ (cont.)

W	p
.250	.252
.333	.184
.361	.142
.444	.072
.528	.052
583	.029
.694	.012
.750	.008
.778	.006
.861	.002
1.000	.000

$n = 7$

W	p
.000	1.000
.020	.964
.061	.768
.082	.620
.143	.486
.184	.305
.245	.237
.265	.192
.326	.112
.388	.085
.429	.051
.510	.027
.551	.021
.571	.016
.633	.008
.735	.004
.755	.003
.796	.001
.878	.000
1.000	.000

$n = 8$

W	p
.000	1.000
.016	.967
.047	.794
.062	.654
.391	.047
.422	.038
.438	.030
.484	.018
.562	.010
.578	.008
.609	.005
.672	.002
.750	.001
.766	.001
.812	.000
.891	.000
1.000	.000

$n = 9$

W	p
.000	1.000
.012	.971
.037	.814
.049	.685
.086	.569
.111	.398
.148	.328
.160	.278
.198	.187
.235	.154
.259	.107
.309	.069
.333	.057
.346	.048
.383	.031
.444	.019
.457	.016
.482	.010
.531	.006
.593	.004
.605	.003
.642	.001
.704	.001
.753	.000
-	-

$n = 10$ (cont.)

W	p
.010	.947
.030	.830
.040	.710
.070	.601
.090	.436
.120	.368
.130	.316
.160	222
.190	.187
.210	.135
.250	.092
.270	.078
.280	.066
.310	.046
.360	.030
.370	.026
.390	.018
.430	.012
.480	.007
.490	.006
.520	.003
.570	.002
.610	.001
.630	.001
.640	.001
.670	.000
-	-
-	-
-	-
1.000	.000

$n = 11$

W	p
.000	1.000
.008	.976
.025	.844
.033	.732
.058	.629
.074	.470
.099	.403
.107	.351
.298	.043
.306	.037
.322	.027
.355	.019
.397	.013
.405	.011
.430	.007
.471	.005
.504	.003
.521	.002
.529	.002
.554	.001
.603	.001
.620	.000
-	-
-	-
-	-
1.000	.000

$n = 12$

w	p
.000	1.000
.007	.978
.021	.856
.028	.751
.049	.654
.062	.500
.083	.434
.090	.383
.111	.287
.132	.249
.146	.191
.174	.141
.188	1.23
.194	.108
.215	.080
.250	.058
.257	.050
.271	.038
.299	.028
.333	.019

W	p
.000	1.000
.028	.956
.083	.740
.111	.570
.194	.430

.109	.531
.141	.355
.188	.285
.203	.236
.250	.149
.297	.120
.328	.079

-	-
-	-
1.000	.000

$n = 10$

W	p
.000	1.000

.132	.256
.157	.219
.174	.163
.207	.116
.223	.100
.231	.087
.256	.062

.340	.017
.361	.011
.396	.008
.424	.005
.438	.004
.444	.004
.465	.002

$m = 3$

$n = 12$ (cont.)

W	p
.507	.002
.521	.001
.528	.001
.549	.001
.562	.001
.583	.000
-	-
-	-
-	-
1.000	.000

$n = 13$

W	p
.000	1.000
.006	.980
.018	.866
.024	.767
.041	.657
.053	.527
.071	.463
.077	.412
.095	.316
.112	.278
.124	.217
.148	.165
.160	.145
.166	.129
.183	.098
.213	.073

$n = 13$ (cont.)

W	p
.219	.064
.231	.050
.254	.038
.284	.027
.290	.025
.308	.016
.337	.012
.361	.008
.373	.007
.379	.006
.396	.004
.432	.003
.444	.002
.450	.002
.467	.001
.479	.001
.497	.001
.538	.001
.550	.000
-	-
-	-
-	-
1.000	.000

$n = 14$

W	p
.000	1.000
.005	.981
.015	.874

$n = 14$ (cont.)

W	p
.020	.781
.036	.694
.046	.551
.061	.489
.066	.438
.082	.344
.097	.305
.107	.242
.128	.188
.138	.167
.143	.150
.158	.117
.184	.089
.189	.079
.199	.063
.219	.049
.245	.036
.250	.033
.265	.023
.291	.018
.311	.011
.321	.010
.327	.009
.342	.007
.372	.005
.383	.003
.388	.003
.403	.003
.413	.002

$n = 14$ (cont.)

W	p
.429	.002
.464	.001
.474	.001
.495	.000
-	-
-	-
-	-
1.000	.000

$n = 15$

W	p
.000	1.000
.004	.982
.013	.882
.018	.794
.031	.711
.040	.573
.053	.513
.058	.463
.071	.369
.084	.330
.093	.267
.111	.211
.120	.189
.124	.170
.138	.136
.160	.106
.164	.096
.173	.077

$n = 15$ (cont.)

W	p
.191	.059
.213	.047
.218	.043
.231	.030
.253	.022
.271	.018
.280	.015
.284	.011
.298	.010
.324	.007
.333	.005
.338	.005
.351	.004
.360	.004
.373	.003
.404	.002
.413	.001
.431	.001
.444	.001
.458	.001
.480	.000
-	-
-	-
-	-
1.000	.000

$m = 4$

$n = 2$

W	p
.001	1.000
.100	.958
.200	.833

$n = 4$ (cont.)

W	p
.675	.035
.700	.020
.725	.013

$n = 6$

W	p
.000	1.000
.011	.996
.022	.952

$n = 6$ (cont.)

W	p
.733	.001
.744	.001
.756	.000

$n = 7$ (cont.)

W	p
.543	.004
.559	.004
.567	.003

.300	.792
.400	.625
.500	.542
.600	.458
.700	.375
.800	.208
.900	.167
1.000	.042

n = 3

w	*p*
.022	1.000
.067	.958
.111	.910
.200	.727
.244	.615
.289	.524
.378	.446
.422	.328
.467	.293
.556	.207
.600	.182
.644	.161
.733	.075
.778	.054
.822	.026
.911	.017
1.000	.002

n = 4

W	*p*
.000	1.000
.025	.992
.050	.930
.075	.898
.100	.794
.125	.753
.150	.680
.175	.651
.200	.528
.225	.513
.250	.432
.275	.390
.300	.352
.325	.321
.775	.011
.800	.006
.825	.005
.850	.002
.900	.002
.925	.001
1.000	.000

n = 5

W	*p*
.008	1.000
.024	.974
.040	.944
.072	.857
.088	.769
.104	.710
.136	.652
.152	.563
.168	.520
.200	.443
.216	.406
.232	.368
.264	.301
.280	.266
.296	.232
.328	.213
.344	.162
.360	.151
.392	.119
.408	.102
.424	.089
.456	.071
.472	.067
.488	.057
.520	.049
.536	.033
.552	.032
.584	.024
.600	.021
.616	.015
.648	.011
.664	.009
.680	.008
.712	.006
.728	.003

.033	.938
.044	.878
.056	.843
.067	.797
.078	.779
.089	.676
.100	.666
.111	.608
.122	.566
.133	.541
1.44	.517
.167	.427
.178	.385
.189	.374
.200	.337
.211	.321
.222	.274
.233	.259
.244	.232
.256	.221
.267	.193
.278	.190
.289	.162
.300	.154
.311	.127
.322	.113
.344	.109
.356	.088
.367	.087
.378	.073
.389	.067
.400	.063
.411	.058
.422	.043
.433	.041
.444	.036
.456	.033
.467	.031
.478	.027
.489	.021
.500	.021
.522	.017
.533	.015
.544	.015
.556	.011
-	-
-	-
-	-
1.000	.000

n = 7

W	*p*
.004	1.000
.012	.984
.020	.964
.037	.905
.045	.846
.053	.795
.069	.754
.078	.678
.086	.652
.102	.596
.110	.564
.118	.533
.135	.460
.143	.420
.151	.378
.167	.358
.176	.306
.184	.300
.200	.264
.208	.239
.216	.216
.233	.188
.241	.182
.249	.163
.265	.150
.273	.122
.282	.118
.298	.101
.306	.093
.314	.081
.331	.073
.339	.062
.347	.058
.363	.051
.371	.040
.380	.037
.396	.034
.404	.032
.576	.003
.592	.003
.600	.002
.608	.002
.624	.001
.633	.001
.641	.001
.657	.001
.665	.001
.673	.001
.690	.000
-	-
-	-
-	-
1.000	.000
-	-

n = 8

W	*p*
.000	1.000
.006	.998
.012	.967
.019	.957
.025	.914
.031	.890
.038	.853
.044	.842
.050	.764
.056	.754
.062	.709
.069	.677
.075	.660
.081	.637
.094	.557
.100	.509
.106	.500
.112	.471
.119	.453
.125	.404
.131	.390
.137	.364
.144	.348
.156	.325
.162	.297
.169	.283
.175	.247

.375	.237	.744	.002	.567	.010	.412	.030	.181	.231
.400	.199	.776	.002	.578	.009	.429	.024	.194	.217
.425	.188	.792	.001	.589	.008	.437	.021	.200	.185
.450	.159	.808	.001	.600	.006	.445	.018	.206	.182
.475	.141	.840	.000	.611	.006	.461	.016	.212	.162
.500	.106	-	-	.633	.004	.469	.014	.219	.155
.525	.093	-	-	.644	.003	.478	.013	.225	.153
.550	.077	-	-	.656	.003	.494	.009	.231	.144
.575	.069	1.000	.000	.667	.002	.502	.008	.238	.122
.600	.058	-	-	.678	.002	.510	.008	.244	.120
.625	.054	-	-	.700	.001	.527	.007	.250	.112
.650	.036	-	-	.711	.001	.535	.006	.256	.106

$m = 4$				$m = 5$					
$n = 8$ (cont.)		$n = 8$ (cont.)		$n = 3$		$n = 3$ (cont.)		$n = 3$ (cont.)	
W	*p*	*W*	*p*	*W*	*p*	*W*	*p*	*W*	*p*
.262	.098	.456	.008	.000	1.000	.333	.475	.667	.063
.269	.091	.462	.007	.022	1.000	.356	.432	.689	.056
.281	.077	.469	.007	.044	.988	.378	.406	.711	.045
.294	.067	.475	.006	.067	.972	.400	.347	.733	.038
.300	.062	.481	.005	.089	.941	.422	.326	.756	.028
.306	.061	.494	.004	.111	.914	.444	.291	.778	.026
.312	.052	.500	.004	.133	.845	.467	.253	.800	.017
.319	.049	.506	.004	.156	.831	.489	.236	.822	.015
.325	.046	.512	.003	.178	.768	.511	.213	.844	.008
.331	.043	.519	.003	.200	.720	.533	.172	.867	.005
.338	.038	.525	.002	.222	.682	.556	.163	.889	.004
.344	.037	.531	.002	.244	.649	.578	.127	.911	.003
.356	.031	.538	.002	.267	.595	.600	.117	.956	.001
.362	.028	.544	.002	.289	.559	.622	.096	1.000	.000
.369	.026	.550	.002	.311	.493	.644	.080		
.375	.023	.556	.002						
.381	.021	.562	.001						
.394	.019	.569	.001						
.400	.015	.575	.001						
.406	.015	.581	.001						
.412	.013	.594	.001						
.419	.013	.606	.001						
.425	.011	.612	.000						
.431	.010	-	-						
.438	.009	-	-						
.444	.008	-	-						
.450	.008	1.000	.000						

Note : *P* = the probability of a computed value of W greater than or equal to the tabulated value.

Table-17 Critical values of the Kruskal-Wallis test statistic

Sample sizes n_1	n_2	n_3	Critical value	a
2	1	1	2.7000	.0500
2	2	1	3.6000	0.200
2	2	2	4.5714	0.067
			3.7143	0.200
3	1	1	3.2000	0.300
3	2	1	4.2857	0.100
			3.8571	0.133
3	2	2	5.3572	0.029
			4.7143	0.048
			4.5000	0.067
			4.4643	0.105
3	3	1	5.1429	0.043
			4.5714	0.100
			4.0000	0.129
3	3	2	6.2500	0.011
			5.3611	0.032
			5.1389	0.061
			4.5556	0.100
			4.2500	0.121
3	3	3	7.2000	0.004
			6.4889	0.011
			5.6889	0.029
			5.6000	0.050
			5.0667	0.086
			4.6222	0.100
4	1	1	3.5714	0.200
4	2	1	4.8214	0.057
			4.5000	0.076
			4.0179	0.114
4	2	2	6.0000	0.014
			5.3333	0.033
			5.1250	0.052
			4.4583	0.100
			4.1667	0.105
4	3	1	5.8333	0.021
			5.2083	0.050
			5.0000	0.057
			4.7000	0.101
4	4	1	6.6667	0.010
			6.1667	0.022
			4.9667	0.048
			4.8667	0.054
			4.1667	0.082
			4.0667	0.102
4	4	2	7.0364	0.006
			6.8727	0.011
			5.4545	0.046
			5.2364	0.052
			4.5545	0.098
			4.4455	0.103
4	4	3	7.1439	0.010
			7.1364	0.011
			5.5985	0.049
			5.5758	0.051
			4.5455	0.099
			4.4773	0.102
4	4	4	7.6538	0.008
			7.5385	0.011
			5.6923	0.049
			5.6538	0.054
			4.6539	0.097
			4.5001	0.104
5	1	1	3.8571	0.143
5	2	1	5.2500	0.036
			5.0000	0.048
			4.4500	0.071
			4.2000	0.095
			4.0500	0.119
5	2	2	6.5333	0.008
			6.1333	0.013
			5.1600	0.034
			5.0400	0.056
			4.3733	0.090
			4.2933	0.122

			4.0556	0.093
			3.8889	0.129
4	3	2	6.4444	0.008
			6.3000	0.011
			5.4444	0.046
			5.4000	0.051
			4.5111	0.098
			4.4444	0.102
4	3	3	6.7455	0.010
			6.7091	0.013
			5.7909	0.046
			5.7273	0.050
			4.7091	0.092
5	3	3	5.6485	0.049
			5.5152	0.051
			4.5333	0.097
			4.4121	0.109
5	4	1	6.9545	0.008
			6.8400	0.011
			4.9855	0.044
			4.8600	0.056
			3.9873	0.098
			3.9600	0.102
5	4	2	7.2045	0.009
			7.1182	0.010
			5.2727	0.049
			5.2682	0.050
			4.5409	0.098
			4.5182	0.101
5	4	3	7.4449	0.010
			7.3949	0.011
			5.6564	0.049
			5.6308	0.050
			4.5487	0.099
			4.5231	0.103
5	4	4	7.7604	0.009
			7.7440	0.011
			5.6571	0.049
			5.6176	0.050
			4.6187	0.100
			4.5527	0.102
5	5	1	7.3091	0.009
5	3	1	6.4000	0.012
			4.9600	0.048
			4.8711	0.052
			4.0178	0.095
			3.8400	0.123
5	3	2	6.9091	0.009
			6.8218	0.010
			5.2509	0.049
			5.1055	0.052
			4.6509	0.091
			4.4945	0.101
5	3	3	7.0788	0.009
			6.9818	0.011
5	5	1	6.8364	0.011
			5.1273	0.046
			4.9091	0.053
			4.1091	0.086
			4.0364	0.105
5	5	2	7.3385	0.010
			7.2692	0.010
			5.3385	0.047
			5.2462	0.051
			4.6231	0.097
			4.5077	0.100
5	5	3	7.5780	0.010
			7.5429	0.010
			5.7055	0.046
			5.6264	0.051
			4.5451	0.100
			4.5363	0.102
5	5	4	7.8229	0.010
			7.7914	0.010
			5.6657	0.049
			5.6429	0.050
			4.5229	0.099
			4.5200	0.101
5	5	5	8.0000	0.009
			7.9800	0.010
			5.7800	0.049
			5.6600	0.051
			4.5600	0.100
			4.5000	0.102

Table-18 Quantilies of the Kolmogorov test statistic

One sided test	$p = 0.90$	0.95	0.975	0.99	0.995
Two sided test	$p = 0.80$	0.90	0.95	0.98	0.99
$n = 1$	.900	.950	.975	.990	.995
2	.684	.776	.842	.900	.929
3	.565	.636	.708	.785	.829
4	.493	.565	.624	.689	.734
5	.447	.509	.563	.627	.669
6	.410	.468	.519	.557	.617
7	.381	.436	.483	.538	.576
8	.358	.410	.454	.507	.542
9	.339	.387	.430	.480	.513
10	.323	.369	.409	.457	.489
11	.308	.352	.391	.437	.468
12	.296	.338	.375	.419	.449
13	.285	.325	.361	.404	.432
14	.275	.314	.349	.390	.418
15	.266	.304	.338	.377	.404
16	.258	.295	.327	.366	.392
17	.250	.286	.318	.355	.381
18	.244	.279	.309	.346	.317
19	.237	.271	.301	.337	.361
20	.232	.265	.294	.329	.352
21	.226	.259	.287	.321	.344
22	.221	.253	.281	.314	.337
23	.216	.247	.275	.307	.330
24	.212	.242	.269	.301	.323
25	.208	.238	.264	.295	.317
26	.204	.233	.259	.290	.311
27	.200	.229	.254	.284	.305
28	.197	.225	.250	.279	.300
29	.193	.221	.246	.275	.295
30	.190	.218	.242	.270	.290
31	.187	.214	.238	.266	.285
32	.184	.211	.234	.262	.281
33	.182	.208	.231	.258	.277
34	.179	.205	.227	.254	.273
35	.177	.202	.224	.251	.269
36	.174	.199	.221	.247	.265
37	.172	.196	.218	.244	.262
38	.170	.194	.215	.241	.258
39	.168	.191	.213	.238	.255
40	.165	.189	.210	.235	.252
Approximation for $n > 40$:	$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$