Statistical Methods for Environmental Pollution Monitoring

Richard O. Gilbert
Pacific Northwest Laboratory

Dedicated to my parents, Mary Margaret and Donald I. Gilbert

This book is printed on acid-free paper. ⊖

Library of Congress Cataloging-in-Publication Data:

Gilbert, Richard O.
Statistical methods for environmental pollution monitoring.
Bibliography: p.
Includes index.
1. Pollution—Environmental aspects—Statistical

Contents

16 Detecting and Estimating Trends

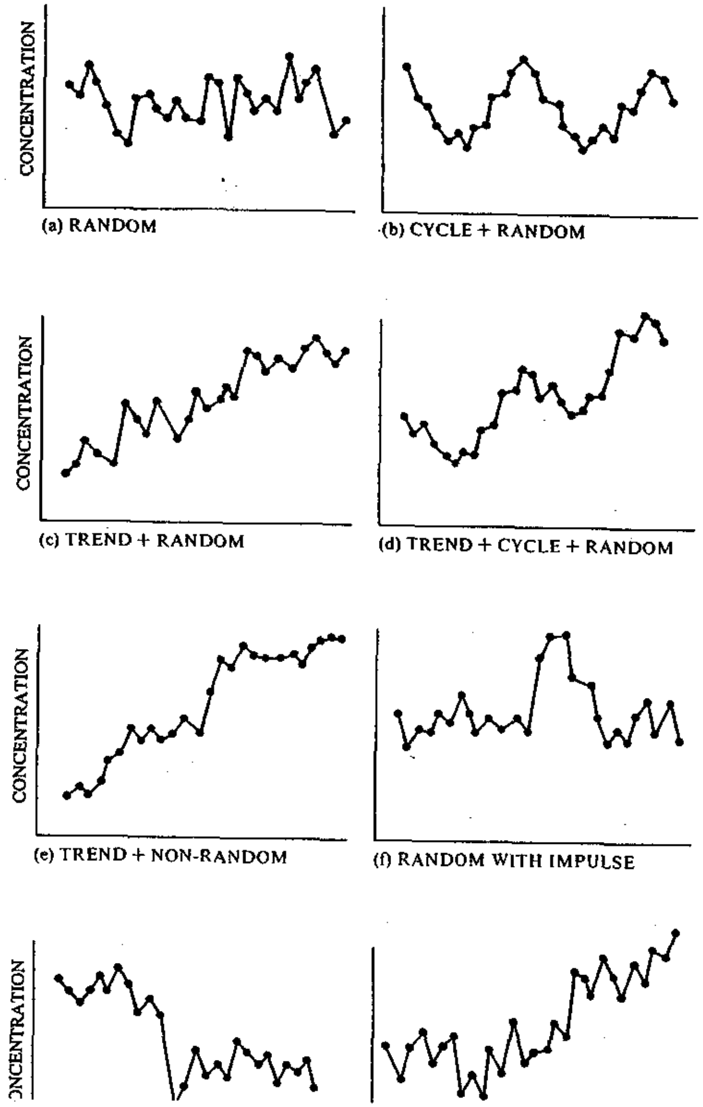
An important objective of many environmental monitoring programs is to detect changes or trends in pollution levels over time. The purpose may be to look for increased environmental pollution resulting from changing land use practices such as the growth of cities, increased erosion from farmland into rivers, or the startup of a hazardous waste storage facility. Or the purpose may be to determine if pollution levels have declined following the initiation of pollution control programs.

The first sections of this chapter discuss types of trends, statistical complexities in trend detection, graphical and regression methods for detecting and estimating trends, and Box-Jenkins time series methods for modeling pollution processes. The remainder of the chapter describes the Mann-Kendall test for detecting monotonic trends at single or multiple stations and Sen's (1968b) nonparametric estimator of trend (slope). Extensions of the techniques in this chapter to handle seasonal effects are given in Chapter 17. Appendix B lists a computer code that computes the tests and trend estimates discussed in Chapters 16 and 17.

16.1 TYPES OF TRENDS

Figure 16.1 shows some common types of trends. A sequence of measurements with no trend is shown in Figure 16.1(a). The fluctuations along the sequence are due to random (unassignable) causes. Figure 16.1(b) illustrates a cyclical pattern wih no long-term trend, and Figure 16.1(c) shows random fluctuations about a rising linear trend line. Cycles may be caused by many factors including seasonal climatic changes, tides, changes in vehicle traffic patterns during the day, production schedules of industry, and so on. Such cycles are not "trends" because they do not indicate long-term change. Figure 16.1(d) shows a cycle with a rising long-term trend with random fluctuation about the cycle.

Frequently, pollution measurements taken close together in time or space are positively correlated, that is, high (low) values are likely to be followed by



treatment plant. Finally, a sequence of random measurements fluctuating about a constant level may be followed by a trend as shown in Figure 16.1(h). We concentrate here on tests for detecting monotonic increasing or decreasing trends as in (c), (d), (e), and (h).

16.2 STATISTICAL COMPLEXITIES

The detection and estimation of trends is complicated by problems associated with characteristics of pollution data. In this section we review these problems, suggest approaches for their alleviation, and reference pertinent literature for additional information. Harned et al. (1981) review the literature dealing with statistical design and analysis aspects of detecting trends in water quality. Munn (1981) reviews methods for detecting trends in air quality data.

16.2.1 Changes in Procedures

A change of analytical laboratories or of sampling and/or analytical procedures may occur during a long-term study. Unfortunately, this may cause a shift in the mean or in the variance of the measured values. Such shifts could be incorrectly attributed to changes in the underlying natural or man-induced processes generating the pollution.

When changes in procedures or laboratories occur abruptly, there may not be time to conduct comparative studies to estimate the magnitude of shifts due to these changes. This problem can sometimes be avoided by preparing duplicate samples at the time of sampling: one is analyzed and the other is stored to be analyzed if a change in laboratories or procedures is introduced later. The paired, old-new data on duplicate samples can then be compared for shifts or other inconsistencies. This method assumes that the pollutants in the sample do not change while in storage, an unrealistic assumption in many cases.

16.2.2 Seasonality

The variation added by seasonal or other cycles makes it more difficult to detect long-term trends. This problem can be alleviated by removing the cycle before applying tests or by using tests unaffected by cycles. A simple nonparametric test for trend using the first approach was developed by Sen (1968a). The seasonal Kendall test, discussed in Chapter 17, uses the latter approach.

16.2.3 Correlated Data

Pollution measurements taken in close proximity over time are likely to be positively correlated, but most statistical tests require uncorrelated data. One approach is to use test statistics developed by Sen (1963, 1965) for dependent data. However, Lettenmaier (1975) reports that perhaps several hundred mea-

and provide tables of adjusted critical values for the Wilcoxon rank sum and Spearman tests. Their paper summarizes the latest statistical techniques for trend detection.

16.2.4 Corrections for Flow

The detection of trends in stream water quality is more difficult when concentrations are related to stream flow, the usual situation. Smith, Hirsch, and Slack (1982) obtain flow-adjusted concentrations by fitting a regression equation to the concentration-flow relationship. Then the residuals from regression are tested for trend by the seasonal Kendall test discussed in Chapter 17. Harned, Daniel, and Crawford (1981) illustrate two alternative methods, discharge compensation and discharge-frequency weighting. Methods for adjusting ambient air quality levels for meteorological effects are discussed by Zeldin and Meisel (1978).

16.3 METHODS

16.3.1 Graphical

Graphical methods are very useful aids to formal tests for trends. The first step is to plot the data against time of collection. Velleman and Hoaglin (1981) provide a computer code for this purpose, which is designed for interactive use on a computer terminal. They also provide a computer code for "smoothing" time series to point out cycles and/or long-term trends that may otherwise be obscured by variability in the data.

Cumulative sum (CUSUM) charts are also an effective graphical tool. With this method changes in the mean are detected by keeping a cumulative total of deviations from a reference value or of residuals from a realistic stochastic model of the process. Page (1961, 1963), Ewan (1963), Gibra (1975), Wetherill (1977), Berthouex, Hunter, and Pallesen (1978), and Vardeman and David (1984) provide details on the method and additional references.

16.3.2 Regression

If plots of data versus time suggest a simple linear increase or decrease over time, a linear regression of the variable against time may be fit to the data. A t test may be used to test that the true slope is not different from zero; see, for example, Snedecor and Cochran (1980, p. 155). This t test can be misleading if seasonal cycles are present, the data are not normally distributed, and/or the data are serially correlated. Hirsch, Slack, and Smith (1982) show that in these situations, the t test may indicate a significant slope when the true slope actually is zero. They also examine the performance of linear regression applied to deseasonalized data. This procedure (called *seasonal regression*) gave a t test

16.3.3 Intervention Analysis and Box-Jenkins Models

If a long time sequence of equally spaced data is available, intervention analysis may be used to detect changes in average level resulting from a natural or man-induced intervention in the process. This approach, developed by Box and Tiao (1975), is a generalization of the autoregressive integrated moving-average (ARIMA) time series models described by Box and Jenkins (1976). Lettenmaier and Murray (1977) and Lettenmaier (1978) study the power of the method to detect trends. They emphasize the design of sampling plans to detect impacts from polluting facilities. Examples of its use are in Hipel et al. (1975) and Roy and Pellerin (1982).

Box-Jenkins modeling techniques are powerful tools for the analysis of time series data. McMichael and Hunter (1972) give a good introduction to Box-Jenkins modeling of environmental data, using both deterministic and stochastic components to forecast temperature flow in the Ohio River. Fuller and Tsokos (1971) develop models to forecast dissolved oxygen in a stream. Carlson, MacConnick, and Watts (1970) and McKerchar and Delleur (1974) fit Box-Jenkins models to monthly river flows. Hsu and Hunter (1976) analyze annual series of air pollution SO_2 concentrations. McCollister and Wilson (1975) forecast daily maximum and hourly average total oxidant and carbon monoxide concentrations in the Los Angeles Basin. Hipel, McLeod, and Lennox (1977*a*, 1977*b*) illustrate improved Box-Jenkins techniques to simplify model construction. Reinsel et al. (1981*a*, 1981*b*) use Box-Jenkins models to detect trends in stratospheric ozone data. Two introductory textbooks are McCleary and Hay (1980) and Chatfield (1984). Box and Jenkins (1976) is recommended reading for all users of the method.

Disadvantages of Box-Jenkins methods are discussed by Montgomery and Johnson (1976). At least 50 and preferably 100 or more data collected at equal (or approximately equal) time intervals are needed. When the purpose is forecasting, we must assume the developed model applies to the future. Missing data or data reported as trace or less-than values can prevent the use of Box-Jenkins methods. Finally, the modeling process is often nontrivial, with a considerable investment in time and resources required to build a satisfactory model. Fortunately, there are several packages of statistical programs that contain codes for developing time series models, including Minitab (Ryan, Joiner, and Ryan 1982), SPSS (1985), BMDP (1983), and SAS (1985). Codes for personal computers are also becoming available.

16.4 MANN-KENDALL TEST

In this section we discuss the nonparametric Mann-Kendall test for trend (Mann, 1945; Kendall, 1975). This procedure is particularly useful since missing values

than their measured values. We note that the Mann-Kendall test can be viewed as a nonparametric test for zero slope of the linear regression of time-ordered data versus time, as illustrated by Hollander and Wolfe (1973, p. 201).

16.4.1 Number of Data 40 or Less

If n is 40 or less, the procedure in this section may be used. When n exceeds 40, use the normal approximation test in Section 16.4.2. We begin by considering the case where only one datum per time period is taken, where a time period may be a day, week, month, and so on. The case of multiple data values per time period is discussed in Section 16.4.3.

The first step is to list the data in the order in which they were collected over time: $x_1, x_2, \ldots, x_n$, where x_i is the datum at time i. Then determine the sign of all $n(n-1)/2$ possible differences $x_j - x_k$, where $j > k$. These differences are $x_2 - x_1, x_3 - x_1, \ldots, x_n - x_1, x_3 - x_2, x_4 - x_2, \ldots, x_n - x_{n-2}, x_n - x_{n-1}$. A convenient way of arranging the calculations is shown in Table 16.1.

Let $\mathrm{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $x_j - x_k$:

$$\begin{aligned} \mathrm{sgn}(x_j - x_k) &= 1 \quad \text{if } x_j - x_k > 0 \\ &= 0 \quad \text{if } x_j - x_k = 0 \\ &= -1 \quad \text{if } x_j - x_k < 0 \end{aligned} \qquad 16.1$$

Then compute the Mann-Kendall statistic

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \mathrm{sgn}(x_j - x_k) \qquad 16.2$$

which is the number of positive differences minus the number of negative differences. These differences are easily obtained from the last two columns of Table 16.1. If S is a large positive number, measurements taken later in time tend to be larger than those taken earlier. Similarly, if S is a large negative number, measurements taken later in time tend to be smaller. If n is large, the computer code in Appendix B may be used to compute S. This code also computes the tests for trend discussed in Chapter 17.

Suppose we want to test the null hypothesis, H_0, of no trend against the alternative hypothesis, H_A, of an upward trend. Then H_0 is rejected in favor of H_A if S is positive and if the probability value in Table A18 corresponding to the computed S is less than the a priori specified α significance level of the test. Similarly, to test H_0 against the alternative hypothesis H_A of a downward trend, reject H_0 and accept H_A if S is negative and if the probability value in the table corresponding to the absolute value of S is less than the a priori specified α value. If a two-tailed test is desired, that is, if we want to detect either an upward or downward trend, the tabled probability level corresponding to the absolute value of S is doubled and H_0 is rejected if that doubled value

Table 16.1 Differences in Data Values Needed for Computing the Mann-Kendall Statistic S to Test for Trend

Data Values Listed in the Order Collected Over Time							No. of + Signs	No. of − Signs
x_1	x_2	x_3	x_4	...	x_{n-1}	x_n		
	$x_2 - x_1$	$x_3 - x_1$	$x_4 - x_1$	...	$x_{n-1} - x_1$	$x_n - x_1$		
		$x_3 - x_2$	$x_4 - x_2$	...	$x_{n-1} - x_2$	$x_n - x_2$		
			$x_4 - x_3$	...	$x_{n-1} - x_3$	$x_n - x_3$		
				⋮	⋮	⋮		
					$x_{n-1} - x_{n-2}$	$x_n - x_{n-2}$		
						$x_n - x_{n-1}$		
						$S =$	(sum of + signs)	+ (sum of − signs)

Table 16.2 Computation of the Mann-Kendall Trend Statistic S for the Time Ordered Data Sequence 10, 15, 14, 20

Time / *Data*	*1* / *10*	*2* / *15*	*3* / *14*	*4* / *20*		*No. of + Signs*		*No. of − Signs*
		15 − 10	14 − 10	20 − 10		3		0
			14 − 15	20 − 15		1		1
				20 − 14		1		0
				S	=	5	−	1 = 4

significance level. For ease of illustration suppose only 4 measurements are collected in the following order over time or along a line in space: 10, 15, 14, and 20. There are 6 differences to consider: 15 − 10, 14 − 10, 20 − 10, 14 − 15, 20 − 15, and 20 − 14. Using Eqs. 16.1 and 16.2, we obtain $S = +1 + 1 + 1 - 1 + 1 + 1 = +4$, as illustrated in Table 16.2. (Note that the sign, not the magnitude of the difference is used.) From Table A18 we find for $n = 4$ that the tabled probability for $S = +4$ is 0.167. This number is the probability of obtaining a value of S equal to +4 or larger when $n = 4$ and when no upward trend is present. Since this value is greater than 0.10, we cannot reject H_0.

If the data sequence had been 18, 20, 23, 35, then $S = +6$, and the tabled probability is 0.042. Since this value is less than 0.10, we reject H_0 and accept the alternative hypothesis of an upward trend.

Table A18 gives probability values only for $n \leq 10$. An extension of this table up to $n = 40$ is given in Table A.21 in Hollander and Wolfe (1973).

16.4.2 Number of Data Greater Than 40

When n is greater than 40, the normal approximation test described in this section is used. Actually, Kendall (1975, p. 55) indicates that this method may be used for n as small as 10 unless there are many tied data values. The test procedure is to first compute S using Eq. 16.2 as described before. Then compute the variance of S by the following equation, which takes into account that ties may be present:

$$\mathrm{VAR}(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p - 1)(2t_p + 5)\right] \qquad 16.3$$

where g is the number of tied groups and t_p is the number of data in the pth group. For example, in the sequence {23, 24, trace, 6, trace, 24, 24, trace, 23} we have $g = 3$, $t_1 = 2$ for the tied value 23, $t_2 = 3$ for the tied value 24, and $t_3 = 3$ for the three trace values (considered to be of equal but unknown value less than 6).

Then S and VAR(S) are used to compute the test statistic Z as follows:

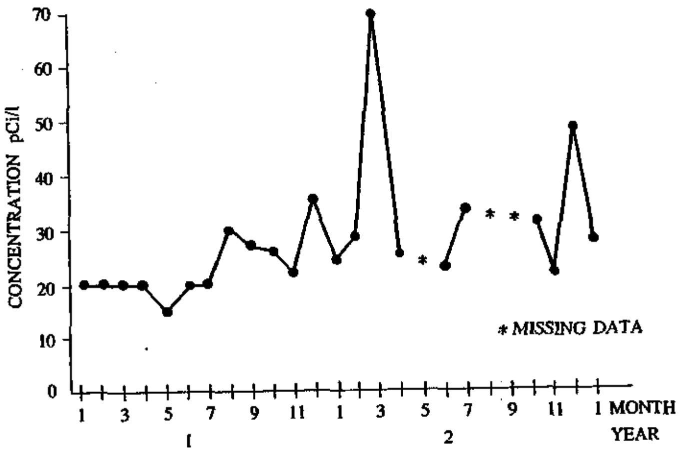


Figure 16.2 Concentrations of ^{238}U in ground water in well E at the former St. Louis Airport storage site for January 1981 through January 1983 (after Clark and Berven, 1984).

A positive (negative) value of Z indicates an upward (downward) trend. If the null hypothesis, H_0, of no trend is true, the statistic Z has a standard normal distribution, and hence we use Table A1 to decide whether to reject H_0. To test for either upward or downward trend (a two-tailed test) at the α level of significance, H_0 is rejected if the absolute value of Z is greater than $Z_{1-\alpha/2}$, where $Z_{1-\alpha/2}$ is obtained from Table A1. If the alternative hypothesis is for an upward trend (a one-tailed test), H_0 is rejected if Z (Eq. 16.4) is greater than $Z_{1-\alpha}$. We reject H_0 in favor of the alternative hypothesis of a downward trend if Z is negative and the absolute value of Z is greater than $Z_{1-\alpha/2}$. Kendall (1975) indicates that using the standard normal tables (Table A1) to judge the statistical significance of the Z test will probably introduce little error as long as $n \geq 10$ unless there are many groups of ties and many ties within groups.

EXAMPLE 16.2

Figure 16.2 is a plot of $n = 22$ monthly ^{238}U concentrations $x_1, x_2, x_3, \ldots, x_{22}$ obtained from a groundwater monitoring well from January 1981 through January 1983 (reported in Clark and Berven, 1984). We use the Mann-Kendall procedure to test the null hypothesis at the $\alpha = 0.05$ level that there is no trend in ^{238}U groundwater concentrations at this well over this 2-year period. The alternative hypothesis is that an upward trend is present.

$$\begin{aligned}\mathrm{VAR}(S) &= \tfrac{1}{18}[22(21)(44+5) \\ &\quad - 6(5)(12+5) - 2(1)(4+5) - 2(1)(4+5)] \\ &= 1227.33\end{aligned}$$

or $[\mathrm{VAR}(S)]^{1/2} = 35.0$. Therefore, since $S > 0$, Eq. 16.4 gives $Z = (108 - 1)/35.0 = 3.1$. From Table A1 we find $Z_{0.95} = 1.645$. Since Z exceeds 1.645, we reject H_0 and accept the alternative hypothesis of an upward trend. We note that the three missing values in Figure 16.2 do not enter into the calculations in any way. They are simply ignored and constitute a regrettable loss of information for evaluating the presence of trend.

16.4.3 Multiple Observations per Time Period

When there are multiple observations per time period, there are two ways to proceed. First, we could compute a summary statistic, such as the median, for each time period and apply the Mann-Kendall test to the medians. An alternative approach is to consider the $n_i \geq 1$ multiple observations at time i (or time period i) as ties in the time index. For this latter case the statistic S is still computed by Eq. 16.2, where n is now the sum of the n_i, that is, the total *number of observations rather than the number of time periods*. The differences between data obtained at the same time are given the score 0 no matter what the data values may be, since they are tied in the time index.

When there are multiple observations per time period, the variance of S is computed by the following equation, which accounts for ties in the time index:

$$\begin{aligned}\mathrm{VAR}(S) &= \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p-1)(2t_p+5) - \sum_{q=1}^{h} u_q(u_q-1)(2u_q+5)\right] \\ &\quad + \frac{\sum_{p=1}^{g} t_p(t_p-1)(t_p-2)\sum_{q=1}^{h} u_q(u_q-1)(u_q-2)}{9n(n-1)(n-2)} \\ &\quad + \frac{\sum_{p=1}^{g} t_p(t_p-1)\sum_{q=1}^{h} u_q(u_q-1)}{2n(n-1)} \qquad 16.5\end{aligned}$$

where g and t_p are as defined following Eq. 16.3, h is the number of time periods that contain multiple data, and u_q is the number of multiple data in the qth time period. Equation 16.5 reduces to Eq. 16.3 when there is one observation

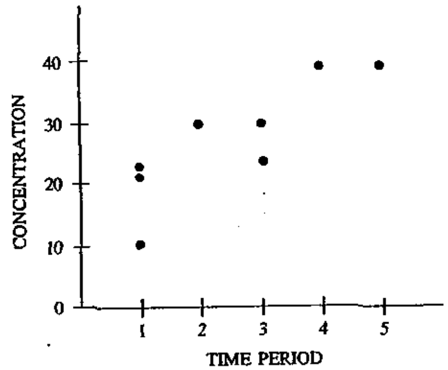


Figure 16.3 An artificial data set to illustrate the Mann-Kendall test for trend when ties in both the data and time are present.

EXAMPLE 16.3

To illustrate the computation of S and VAR(S), consider the following artificial data set:

(concentration, time period)

$$= (10, 1), (22, 1), (21, 1), (30, 2), (22, 3), (30, 3), (40, 4), (40, 5)$$

as plotted in Figure 16.3. There are 5 time periods and $n = 8$ data. To illustrate computing S, we lay out the data as follows:

Time Period :	1	1	1	2	3	3	4	5
Data :	10	22	21	30	22	30	40	40

We shall test at the $\alpha = 0.05$ level the null hypothesis, H_0, of no trend versus the alternative hypothesis, H_A, of an upward trend, a one-tailed test.

Now, look at all $8(7)/2 = 28$ possible data pairs, remembering to give a score of 0 to the 4 pairs within the same time index. The differences are shown in Table 16.3. Ignore the magnitudes of the differences, and sum the number of positive and negative signs to obtain $S = 19$. It is clear from Figure 16.3 that there are $g = 3$ tied data groups (22, 30, and 40) with $t_1 = t_2 = t_3 = 2$. Also, there are $h = 2$ time index ties (times 1 and 3) with $u_1 = 3$ and $u_2 = 2$. Hence, Eq. 16.5 gives

$$\text{VAR}(S) = \frac{1}{18}[8(7)(16 + 5) - 3(2)(1)(4 + 5) - 3(2)(6 + 5)$$

Table 16.3 Illustration of Computing S for Example 16.3

Time Period / *Data*	1 / 10	1 / 22	1 / 21	2 / 30	3 / 22	3 / 30	4 / 40	5 / 40	*Sum of + Signs*	*Sum of − Signs*
		NC	NC	+20	+12	+20	+30	+30	5	0
			NC	+8	0	+8	+18	+18	4	0
				+9	+1	+9	+19	+19	5	0
					−8	0	+10	+10	2	1
						NC	+18	+18	2	0
							+10	+10	2	0
								0	0	0
								S	= 20	−1
									= 19	

NC = Not computed since both data values are within the same time period.

= 2.4. Referring to Table A1, we find $Z_{0.95} = 1.645$. Since $Z >$ 1.645, reject H_0 and accept the alternative hypothesis of an upward trend.

16.4.4 Homogeneity of Stations

Thus far only one station has been considered. If data over time have been collected at $M > 1$ stations, we have data as displayed in Table 16.4 (assuming one datum per sampling period). The Mann-Kendall test may be computed for each station. Also, an estimate of the magnitude of the trend at each station can be obtained using Sen's (1968b) procedure, as described in Section 16.5.

When data are collected at several stations within a region or basin, there may be interest in making a basin-wide statement about trends. A general statement about the presence or absence of monotonic trends will be meaningful if the trends at all stations are in the same direction—that is, all upward or all downward. Time plots of the data at each station, preferably on the same graph to make visual comparison easier, may indicate when basin-wide statements are possible. In many situations an objective testing method will be needed to help make this decision. In this section we discuss a method for doing this that

Table 16.4 Data Collected over Time at Multiple Stations

		Station 1				. . .		*Station M*			
		Sampling Time				. . .		*Sampling Time*			
		1	2		K			1	2		K
	1	x_{111}	x_{211}	. . .	x_{K11}	. . .	1	x_{11M}	x_{21M}	. . .	x_{K1M}
	2	x_{121}	x_{221}	. . .	x_{K21}	. . .	2	x_{12M}	x_{22M}	. . .	x_{K2M}
Year	⋮						⋮				
	L	x_{1L1}	x_{2L1}	. . .	x_{KL1}	. . .	L	x_{1LM}	x_{2LM}	. . .	x_{KLM}

makes use of the Mann-Kendall statistic computed for each station. This procedure was originally proposed by van Belle and Hughes (1984) to test for homogeneity of trends between seasons (a test discussed in Chapter 17).

To test for homogeneity of trend direction at multiple stations, compute the homogeneity chi-square statistic, χ^2_{homog}, where

$$\chi^2_{\text{homog}} = \chi^2_{\text{total}} - \chi^2_{\text{trend}} = \sum_{j=1}^{M} Z_j^2 - M\bar{Z}^2 \qquad \textbf{16.6}$$

$$Z_j = \frac{S_j}{[\text{VAR}(S_j)]^{1/2}} \qquad \textbf{16.7}$$

S_j is the Mann-Kendall trend statistic for the jth station,

and $\quad \bar{Z} = \frac{1}{M}\sum_{j=1}^{M} Z_j$

If the trend at each station is in the same direction, then χ^2_{homog} has a chi-square distribution with $M - 1$ degrees of freedom (df). This distribution is given in Table A19. To test for trend homogeneity between stations at the α significance level, we refer our calculated value of χ^2_{homog} to the α critical value in Table A19 in the row with $M - 1$ df. If χ^2_{homog} exceeds this critical value, we reject the H_0 of homogeneous station trends. *In that case no regional-wide statements should be made about trend direction.* However, a Mann-Kendall test for trend at each station may be used. If χ^2_{homog} does not exceed the α critical level in Table A19, then the statistic $\chi^2_{\text{trend}} = M\bar{Z}^2$ is referred to the chi-square distribution with 1 df to test the null hypothesis H_0 that the (common) trend direction is significantly different from zero.

The validity of these chi-square tests depends on each of the Z_j values (Eq. 16.7) *having a standard normal* distribution. Based on results in Kendall (1975), this implies that the number of data (over time) for each station should exceed 10. Also, the validity of the tests requires that the Z_j be independent. This requirement means that the data from different stations must be uncorrelated. We note that the Mann-Kendall test and the chi-square tests given in this section may be computed even when the number of sampling times, K, varies from year to year and when there are multiple data collected per sampling time at one or more times.

EXAMPLE 16.4

We consider a simple case to illustrate computations. Suppose the following data are obtained:

Time				
1	*2*	*3*	*4*	*5*

8 and $S_2 = -1 + 0 - 1 - 1 + 1 - 1 + 0 - 1 - 1 + 1 = 2 - 6 = -4$. Equation 16.3 gives

$$\text{VAR}(S_1) = \frac{5(4)(15)}{18} = 16.667 \quad \text{and} \quad \text{VAR}(S_2)$$

$$= \frac{[5(4)(15) - 2(1)(9) - 2(1)(9)]}{18} = 14.667$$

Therefore Eq. 16.4 gives

$$Z_1 = \frac{7}{(16.667)^{1/2}} = 1.71 \quad \text{and} \quad Z_2 = \frac{-3}{(14.667)^{1/2}} = -0.783$$

Thus

$$\chi^2_{\text{homog}} = 1.71^2 + (-0.783)^2 - 2\left(\frac{1.71 - 0.783}{2}\right)^2 = 3.1$$

Referring to the chi-square tables with $M - 1 = 1$ df, we find the $\alpha = 0.05$ level critical value is 3.84. Since $\chi^2_{\text{homog}} < 3.84$, we cannot reject the null hypothesis of homogeneous trend direction over time at the 2 stations. Hence, an overall test of trend using the statistic χ^2_{trend} can be made. [Note that the critical value 3.84 *is only* approximate (somewhat too small), since the number of data at both stations is less than 10.] $\chi^2_{\text{trend}} = M\bar{Z}^2 = 2(0.2148) = 0.43$. Since $0.43 < 3.84$, we cannot reject the null hypothesis of no trend at the 2 stations.

We may test for trend at each station using the Mann-Kendall test by referring $S_1 = 8$ and $S_2 = -4$ to Table A18. The tabled value for $S_1 = 8$ when $n = 5$ is 0.042. Doubling this value to give a two-tailed test gives 0.084, which is greater than our prespecified $\alpha = 0.05$. Hence, we cannot reject H_0 of no trend for station 1 at the $\alpha = 0.05$ level. The tabled value for $S_2 = -4$ when $n = 5$ is 0.242. Since $0.484 > 0.05$, we cannot reject H_0 of no trend for station 2. These results are consistent with the χ^2_{trend} test before. Note, however, that station 1 still appears to be increasing over time, and the reader may confirm it is significant at the $\alpha = 0.10$ level. *This result suggests that this station be carefully watched in the future.*

16.5 SEN'S NONPARAMETRIC ESTIMATOR OF SLOPE

As noted in Section 16.3.2, if a linear trend is present, the true slope (change per unit time) may be estimated by computing the least squares estimate of the slope, b, by linear regression methods. However, b computed in this way ...

gross data errors or outliers, and it can be computed when data are missing. Sen's estimator is closely related to the Mann-Kendall test, as illustrated in the following paragraphs. The computer code in Appendix B computes Sen's estimator.

First, compute the N' slope estimates, Q, for each station:

$$Q = \frac{x_{i'} - x_i}{i' - i} \qquad 16.8$$

where $x_{i'}$ and x_i are data values at times (or during time periods) i' and i, respectively, and where $i' > i$; N' is the number of data pairs for which $i' > i$. The median of these N' values of Q is Sen's estimator of slope. If there is only one datum in each time period, then $N' = n(n - 1)/2$, where n is the number of time periods. If there are multiple observations in one or more time periods, then $N' < n(n - 1)/2$, where n is now the total number of observations, not time periods, since Eq. 16.8 cannot be computed with two data from the same time period, that is, when $i' = i$. If an x_i is below the detection limit, one half the detection limit may be used for x_i.

The median of the N' slope estimates is obtained in the usual way, as discussed in Section 13.3.1. That is, the N' values of Q are ranked from smallest to largest (denote the ranked values by $Q_{[1]} \le Q_{[2]} \le \cdots \le Q_{[N'-1]} \le Q_{[N']}$) and we compute

$$\begin{aligned} \text{Sen's estimator} &= \text{median slope} \\ &= Q_{[(N'+1)/2]} && \text{if } N' \text{ is odd} \\ &= \tfrac{1}{2}\left(Q_{[N'/2]} + Q_{[(N'+2)/2]}\right) && \text{if } N' \text{ is even} \end{aligned} \qquad 16.9$$

A $100(1 - \alpha)\%$ two-sided confidence interval about the true slope may be obtained by the nonparametric technique given by Sen (1968b). We give here a simpler procedure, based on the normal distribution, that is valid for n as small as 10 unless there are many ties. This procedure is a generalization of that given by Hollander and Wolfe (1973, p. 207) when ties and/or multiple observations per time period are present.

1. Choose the desired confidence coefficient α and find $Z_{1-\alpha/2}$ in Table A1.
2. Compute $C_\alpha = Z_{1-\alpha/2}[\mathrm{VAR}(S)]^{1/2}$, where VAR($S$) is computed from Eqs. 16.3 or 16.5. The latter equation is used if there are multiple observations per time period.
3. Compute $M_1 = (N' - C_\alpha)/2$ and $M_2 = (N' + C_\alpha)/2$.
4. The lower and upper limits of the confidence interval are the M_1th largest and $(M_2 + 1)$th largest of the N' ordered slope estimates, respectively.

EXAMPLE 16.5

We use the data set in Example 16.3 to illustrate Sen's procedure.

Table 16.5 Illustration of Computing an Estimate of Trend Slope Using Sen's (1968*b*) Nonparametric Procedure (for Example 16.5). Tabled Values Are Individual Slope Estimates, Q

Time Period	*1*	*1*	*1*	*2*	*3*	*3*	*4*	*5*
Data	*10*	*22*	*21*	*30*	*22*	*30*	*40*	*40*
		NC	NC	+20	+6	+10	+10	+7.5
			NC	+8	0	+4	+6	+4.5
				+9	+0.5	+4.5	+6.33	+4.75
					−8	0	+5	+3.33
						NC	+18	+9
							+10	+5
								0

NC = Cannot be computed since both data values are within the same time period.

slope estimates Q for these pairs are obtained by dividing the differences in Table 16.3 by $i' - i$. The 24 Q values are given in Table 16.5.

Ranking these Q values from smallest to largest gives

−8, 0, 0, 0, 0.5, 3.33, 4, 4.5, 4.5, 4.75, 5, 5, 6, 6, 6.33, 7.5, 8, 9, 9, 10, 10, 10, 18, 20

Since $N' = 24$ is even, the median of these Q values is the average of the 12th and 13th largest values (by Eq. 16.8), which is 5.5, the *Sen estimate of the true slope. That is, the average (median) change* is estimated to be 5.5 units per time period.

A 90% confidence interval about the true slope is obtaied as follows. From Table A1 we find $Z_{0.95} = 1.645$. Hence,

$$C_\alpha = 1.645[\mathrm{VAR}(S)]^{1/2} = 1.645[58.1]^{1/2} = 12.54$$

where the value for VAR(S) was obtained from Example 16.3. Since $N' = 24$, we have $M_1 = (24 - 12.54)/2 = 5.73$ and $M_2 + 1 = (24 + 12.54)/2 + 1 = 19.27$. From the list of 24 ordered slopes given earlier, the lower limit is found to be 2.6 by interpolating between the 5th and 6th largest values. The upper limit is similarly found to be 9.3 by interpolating between the 19th and 20th largest values.

16.6 CASE STUDY

This section illustrates the procedures presented in this chapter for evaluating trends. The computer program in Appendix B is used on the hypothetical data listed in Table 16.6 and plotted in Figure 16.4. These data, generated on a

Table 16.6 Simulated Monthly Data at Two Stations over a Four-Year Period

NUMBER OF YEARS = 4
NUMBER OF STATIONS = 2

STATION 1 YEAR	MONTH	NUMBER OF DATA POINTS 48 STATION 1	STATION 2 YEAR	MONTH	NUMBER OF DATA POINTS 48 STATION 2
1	1	6.00	1	1	5.09
1	2	5.41	1	2	5.07
1	3	4.58	1	3	4.93
1	4	4.34	1	4	4.94
1	5	4.77	1	5	5.15
1	6	4.54	1	6	11.82
1	7	4.50	1	7	5.48
1	8	5.02	1	8	5.18
1	9	4.38	1	9	5.79
1	10	4.27	1	10	5.11
1	11	4.33	1	11	5.10
1	12	4.33	1	12	5.94
2	13	5.00	2	13	6.91
2	14	5.02	2	14	7.11
2	15	4.14	2	15	5.40
2	16	5.16	2	16	6.77
2	17	6.33	2	17	5.35
2	18	5.49	2	18	6.04
2	19	4.54	2	19	5.45
2	20	6.62	2	20	6.95
2	21	4.64	2	21	5.54
2	22	4.45	2	22	5.71
2	23	4.57	2	23	6.14
2	24	4.09	2	24	7.13
3	25	5.06	3	25	5.80
3	26	4.83	3	26	5.91
3	27	4.92	3	27	5.88
3	28	6.02	3	28	7.21
3	29	4.77	3	29	8.29
3	30	5.03	3	30	6.00
3	31	7.15	3	31	6.28
3	32	4.30	3	32	5.69
3	33	4.15	3	33	6.52
3	34	5.13	3	34	6.27
3	35	5.28	3	35	6.46
3	36	4.31	3	36	6.94
4	37	6.53	4	37	6.28
4	38	5.11	4	38	6.74
4	39	4.31	4	39	6.91
4	40	4.64	4	40	7.81
4	41	4.87	4	41	6.53
4	42	4.89	4	42	6.26
4	43	4.92	4	43	7.01
[illegible]	[illegible]	[illegible]	4	44	7.42

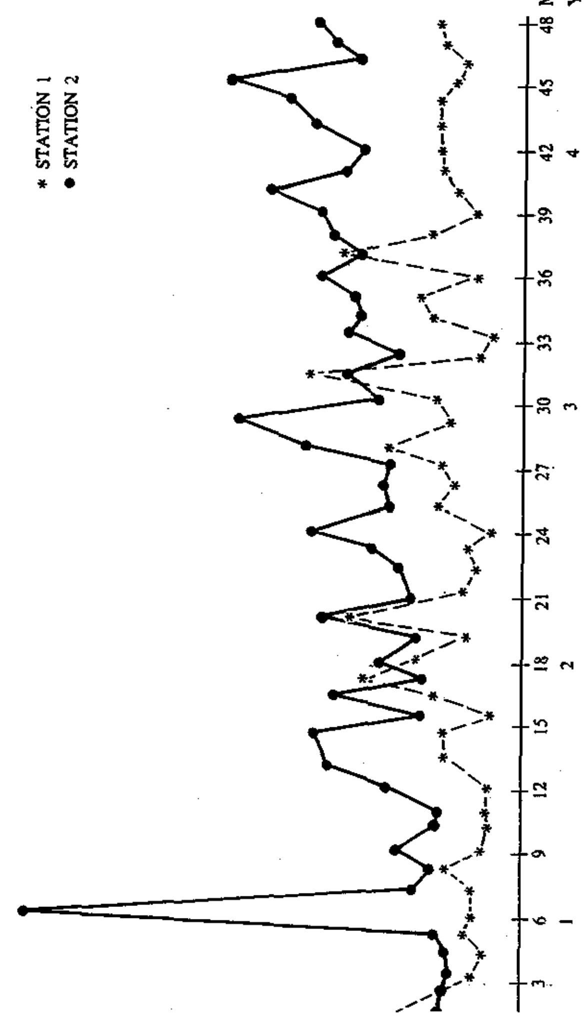


Figure 16.4 Data at two stations each month for four years. Data were simulated using the lognormal independent model given by Hirsch, Slack, and Smith (1982, Eq. 14b). Simulated data were obtained by D. W. Engel.

and the data for station 2 are lognormal with a trend of 0.4 units per year or 0.0333 units per month. These models were among those used by Hirsch, Slack, and Smith (1982) to evaluate the power of the seasonal Kendall test for trend, a test we discuss in Chapter 17.

The results obtained from the computer code in Appendix B are shown in Table 16.7. The first step is to decide whether the two stations have trends in the same direction. In this example we know it is not so, since one station has a trend and the other does not. But in practice this a priori information will not be available.

Table 16.7 shows that the chi-square test of homogeneity (Eq. 16.6) is highly significant ($\chi^2_{\text{homog}} = 10.0$; computed significance level of 0.002). Hence, we ignore the chi-square test for trend that is automatically computed by the program and turn instead to the Mann-Kendall test results for each station. This test for station 1 is nonsignificant (P value of 0.70), indicating no strong evidence for trends, but that for station 2 is highly significant. All of these test results agree with the true situation. Sen's estimates of slope are 0.002 and 0.041 per month for stations 1 and 2, whereas the true values are 0.0 and 0.0333, respectively. The computer code computes $100(1 - \alpha)\%$ confidence limits for the true slope for $\alpha = 0.20, 0.10, 0.05$, and 0.01. For this example the 95% confidence limits are −0.009 and 0.012 for station 1, and 0.030 and 0.050 for station 2.

The computer code allows one to split up the 48 observations at each station into meaningful groups that contain multiple observations. For instance, suppose

Table 16.7 Chi-Square Tests for Homogeneity of Trends at the Two Stations, and Mann-Kendall Tests for Each Station

HOMOGENEITY TEST RESULTS

CHI-SQUARE STATISTICS		df	PROB. OF A LARGER VALUE	
TOTAL	23.97558	2	0.000	Trend not equal at the 2 stations
HOMOGENEITY	10.03524	1	0.002	←
TREND	13.94034	1	0.000	← Not meaningful

STATION	SEASON	MANN-KENDALL S STATISTIC	Z STATISTIC	n	PROB. OF EXCEEDING THE ABSOLUTE VALUE OF THE Z STATISTIC (TWO-TAILED TEST) IF n > 10
1	1	45.00	0.39121	48	0.696
2	1	549.00	4.87122	48	0.000

SEN SLOPE CONFIDENCE INTERVALS

STATION	SEASON	ALPHA	LOWER LIMIT	SLOPE	UPPER LIMIT
1	1	0.010	-0.013	0.002	0.016
		0.050	-0.009	0.002	0.012

Table 16.8 Analyses of the Data in Table 16.6 Considering the Data as Twelve Multiple Observations in Each of Four Years

NUMBER OF YEARS = 4
NUMBER OF SEASONS = 1
NUMBER OF STATIONS = 2

HOMOGENEITY TEST RESULTS

SOURCE	CHI-SQUARE	df	PROB. OF A LARGER VALUE
TOTAL	21.45468	2	0.00
HOMOGENEITY	5.79732	1	0.016
TREND	15.65736	1	0.000

STATION	SEASON	MANN-KENDALL S STATISTIC	Z STATISTIC	n	PROB. OF EXCEEDING THE ABSOLUTE VALUE OF THE Z STATISTIC (TWO-TAILED TEST) IF n > 10
1	1	119.00	1.08623	48	0.277
2	1	489.00	4.49132	48	0.000

SEN SLOPE CONFIDENCE INTERVALS

STATION	SEASON	ALPHA	LOWER LIMIT	SLOPE	UPPER LIMIT
1	1	0.010	-0.120	0.080	0.225
		0.050	-0.065	0.080	0.190
		0.100	-0.037	0.080	0.176
		0.200	-0.014	0.080	0.153
2	1	0.010	0.290	0.467	0.670
		0.050	0.353	0.467	0.620
		0.100	0.370	0.467	0.600
		0.200	0.390	0.467	0.575

we regard the data in this example as 12 multiple data points in each of four years. Applying the code using this interpretation gives the results in Table 16.8.

The conclusions of the tests are the same as obtained in Table 16.7 when the data were considered as one time series of 48 single observations. However, this may not be the case with other data sets or groupings of multiple observations. Indeed, the Mann-Kendall test statistic Z for station 1 is larger in Table 16.8 than in Table 16.7, so that the test is closer to (falsely) indicating a significant trend when the data are grouped into years. For station 2 the Mann-Kendall test in Table 16.8 is smaller than in Table 16.7, indicating the test has less power to detect the trend actually present. The best strategy appears to be to not group data unnecessarily. The estimates of slope are now 0.080 and 0.467 per year, whereas the true values are 0.0 and 0.40, respectively.

and estimating trends, intervention analysis, and problems that arise when using regression methods to detect and estimate trends.

Next, the Mann-Kendall test for trend was described and illustrated in detail, including how to handle multiple observations per sampling time (or period). A chi-square test to test for homogenous trends at different stations within a basin was also illustrated. Finally, methods for estimating and placing confidence limits on the slope of a linear trend by Sen's nonparameter procedure were given and the Mann-Kendall test on a *simulated* data set was illustrated.

EXERCISES

16.1 Use the Mann-Kendall test to test for a rising trend over time, using the following data obtained sequentially over time.

Time	1	2	3	4	5	6	7
Data	ND	1	ND	3	1.5	1.2	4

Use $\alpha = 0.05$. What problem is encountered in using Table A18? Use the normal approximate test statistic Z.

16.2 Use the data in Exercise 16.1 to estimate the magnitude of the trend in the population. Handle NDs in two ways: (a) treat them as missing values, and (b) set them equal to one half the detection limit. Assume the detection limit is 0.5. What method do you prefer? Why?

16.3 Compute a 90% confidence interval about the true slope, using the data in part (b) of Exercise 16.2.

ANSWERS

16.1 $n = 7$. The 2 NDs are treated as tied at a value less than 1.1. $S = 16 - 4 = 12$. Since there is a tie, there is no probability value in Table A18 for $S = +12$, but the probability lies between 0.035 and 0.068. Using the large sample approximation gives $\text{Var}(S) = 43.3$ and $Z = 1.67$. Since $1.67 > 1.645$, we reject H_0 of no trend.

16.2 (a) The median of the 10 estimates of slope is 0.23. (b) The median of the 21 estimates of slope is 0.33.

Pros and Cons: Using one half of the detection limit assumes the actual measurements of ND values are equally likely to fall anywhere between zero and the detection limit. One half of the detection limit is the mean of that distribution. This method, though approximate, is preferred to treating NDs as missing values.

16.3 From Eq. 16.3, $\text{VAR}(S) = 44.3$. (The correction for ties in Eq. 16.3

17 Trends and Seasonality

Chapter 16 discussed trend detection and estimation methods that may be used when there are no cycles or seasonal effects in the data. Hirsch, Slack, and Smith (1982) proposed the seasonal Kendall test when seasonality is present. This chapter describes the seasonal Kendall test as well as the extention to multiple stations developed by van Belle and Hughes (1984). It also shows how to estimate the magnitude of a trend by using the nonparametric seasonal Kendall slope estimator, which is appropriate when seasonality is present. All these techniques are included in the computer code listed in Appendix B. A computer code that computes only the seasonal Kendall test and slope estimator is given in Smith, Hirsch, and Slack (1982).

17.1 SEASONAL KENDALL TEST

If seasonal cycles are present in the data, tests for trend that remove these cycles or are not affected by them should be used. This section discusses such a test: the seasonal Kendall test developed by Hirsch, Slack, and Smith (1982) and discussed further by Smith, Hirsch, and Slack (1982) and by van Belle and Hughes (1984). This test may be used even though there are missing, tied, or ND values. Furthermore, the validity of the test does not depend on the data being normally distributed.

The seasonal Kendall test is a generalization of the Mann-Kendall test. It was proposed by Hirsch and colleagues for use with 12 seasons (months). In brief, the test consists of computing the Mann-Kendall test statistic S and its variance, $\text{VAR}(S)$, separately for each month (season) with data collected over years. These seasonal statistics are then summed, and a Z statistic is computed. If the number of seasons and years is sufficiently large, this Z value may be referred to the standard normal tables (Table A1) to test for a statistically significant trend. If there are 12 seasons (e.g., 12 months of data per year), Hirsch, Slack, and Smith (1982) show that Table A1 may be used as long as

Table 17.1 Data for the Seasonal Kendall Test at One Sampling Station

		Season			
		1	2		K
Year	1	x_{11}	x_{21}	...	x_{K1}
	2	x_{12}	x_{22}	...	x_{K2}
	⋮				
	L	x_{1L}	x_{2L}	...	x_{KL}
		S_1	S_2	...	S_K
		$S' = \sum_{i=1}^{K} S_i$	$\text{Var}(S') = \sum_{i=1}^{K} \text{Var}(S_i)$		

exact test is important, the exact distribution of the seasonal Kendall test statistic can be obtained on a computer for any combination of seasons and years by the technique discussed by Hirsch, Slack, and Smith (1982).

Let x_{il} be the datum for the ith season of the lth year, K the number of seasons, and L the number of years. The data for a given site (sampling station) are shown in Table 17.1. The null hypothesis, H_0, is that the x_{il} are independent of the time (season and year) they were collected. The hypothesis H_0 is tested against the alternative hypothesis, H_A, that for one or more seasons the data are not independent of time.

For each season we use data collected over years to compute the Mann-Kendall statistic S. Let S_i be this statistic computed for season i, that is,

$$S_i = \sum_{k=1}^{n_i-1} \sum_{l=k+1}^{n_i} \text{sgn}(x_{il} - x_{ik}) \qquad 17.1$$

where $l > k$, n_i is the number of data (over years) for season i, and

$$\begin{aligned} \text{sgn}(x_{il} - x_{ik}) &= 1 && \text{if } x_{il} - x_{ik} > 0 \\ &= 0 && \text{if } x_{il} - x_{ik} = 0 \\ &= -1 && \text{if } x_{il} - x_{ik} < 0 \end{aligned}$$

$\text{VAR}(S_i)$ is computed as follows:

$$\text{VAR}(S_i) = \frac{1}{18}\left[n_i(n_i - 1)(2n_i + 5) - \sum_{p=1}^{g_i} t_{ip}(t_{ip} - 1)(2t_{ip} + 5) - \sum_{q=1}^{h_i} u_{iq}(u_{iq} - 1)(2u_{iq} + 5)\right] + \frac{\sum_{p=1}^{g_i} t_{ip}(t_{ip} - 1)(t_{ip} - 2) \sum_{q=1}^{h_i} u_{iq}(u_{iq} - 1)(u_{iq} - 2)}{9n_i(n_i - 1)(n_i - 2)}$$

where g_i is the number of groups of tied (equal-valued) data in season i, t_{ip} is the number of tied data in the pth group for season i, h_i is the number of sampling times (or time periods) in season i that contain multiple data, and u_{iq} is the number of multiple data in the qth time period in season i. These quantities are illustrated in Example 17.1.

After the S_i and $\text{Var}(S_i)$ are computed, we pool across the K seasons:

$$S' = \sum_{i=1}^{K} S_i \qquad 17.3$$

and

$$\text{VAR}(S') = \sum_{i=1}^{K} \text{VAR}(S_i) \qquad 17.4$$

Next, compute

$$\begin{aligned} Z &= \frac{(S' - 1)}{[\text{VAR}(S')]^{1/2}} && \text{if } S' > 0 \\ &= 0 && \text{if } S' = 0 \\ &= \frac{(S' + 1)}{[\text{VAR}(S')]^{1/2}} && \text{if } S < 0 \end{aligned} \qquad 17.5$$

To test the null hypothesis, H_0, of no trend versus the alternative hypothesis, H_A, of either an upward or downward trend (a two-tailed test), we reject H_0 if the absolute value of Z is greater than $Z_{1-\alpha/2}$, where $Z_{1-\alpha/2}$ is from Table A1. If the alternative hypothesis is for an upward trend at the α level (a one-tailed test), we reject H_0 if Z (Eq. 17.5) is greater than $Z_{1-\alpha}$. Reject H_0 in favor of a downward trend (one-tailed test) if Z is negative and the absolute value of Z is greater than $Z_{1-\alpha}$. The computer code in Appendix B computes the seasonal Kendall test for multiple or single observations per time period. Example 17.1 in the next section illustrates this test. The +1 added to the S' in Eq. 17.5 is a correction factor that makes Table A1 more exact for testing the null hypothesis. This correction is not necessary if there are ten or more data for each season ($n_i \geq 10$).

17.2 SEASONAL KENDALL SLOPE ESTIMATOR

The seasonal Kendall slope estimator is a generalization of Sen's estimator of slope discussed in Section 16.5. First, compute the individual N_i slope estimates for the ith season:

$$Q_i = \frac{x_{il} - x_{ik}}{\ldots}$$

estimates and find their median. This median is the seasonal Kendall slope estimator.

A 100(1 − α)% confidence interval about the true slope is obtained in the same manner as in Section 16.5:

1. Choose the desired confidence level α and find $Z_{1-\alpha/2}$ in Table A1.
2. Compute $C_\alpha = Z_{1-\alpha/2}[\text{VAR}(S')]^{1/2}$.
3. Computer $M_1 = (N' - C_\alpha)/2$ and $M_2 = (N' + C_\alpha)/2$.
4. The lower and upper confidence limits are the M_1th largest and the $(M_2 + 1)$th largest of the N' ordered slope estimates, respectively.

EXAMPLE 17.1

We use a simple data set to illustrate the seasonal Kendall test and slope estimator. Since the number of data are small, the tests and confidence limits are only approximations. All computations are given in Table 17.2. Suppose data are collected twice a year (e.g., December and June) for 3 years at a given location. The data are listed below and plotted in Figure 17.1.

	Year 1			Year 2			Year 3	
Season	1	1	2	1	2	2	1	2
Data	8	10	15	12	20	18	15	20

Note that two observations were made in season 1 of year 1 and in season 2 of year 2. Also, there is 1 tied data value, 20, in season 2.

Table 17.2, Part A, gives the $N'_1 + N'_2 = 5 + 5 = 10$ individual slope estimates for the 2 seasons and their ranking from smallest to largest. The seasonal Kendall slope estimate, 2.75, is the median of these 10 values. In Table 17.2, Part B, the seasonal Kendall Z statistic is calculated to be 2.1 by Eqs. 17.3–17.5. To test for an upward trend (one-tailed test) at the $\alpha = 0.05$ level, we reject the null hypothesis, H_0, of no trend if $Z > Z_{0.95}$, that is, if $Z > 1.645$. Since $Z = 2.10$, we reject H_0 and accept that an upward trend is present.

A 90% confidence interval on the true slope is obtained by computing $C_\alpha = 1.645[\text{VAR}(S')]^{1/2} = 1.645(3.808) = 6.264$, $M_1 = (10 - 6.264)/2 = 1.868$, and $M_2 + 1 = (10 + 6.264)/2 + 1 = 9.132$. Hence, the lower limit is found by interpolating between the first and second largest values to obtain 1.7. The upper limit is similarly found to be 4.1.

Table 17.2 Illustration of the Seasonal Kendall Test and Slope Estimator. Tabled Values Are Individual Slope Estimates Obtained from Eq. 17.6

Part A. Computing the Seasonal Kendall Slope Estimate

	Season 1							Season 2						
Year	1	1	2	3	Sum of	Sum of		1	2	2	3	Sum of	Sum of	
Data	8	10	12	15	+ Signs	− Signs		15	20	18	20	+ Signs	− Signs	
		a	+4	+3.5	2	0			+5	+3	+2.5	3	0	
			+2	+2.5	2	0				a	0	0	0	
				+3	1	0					+2	1	0	
				S_1 =	5	+ 0	= 5				S_2 =	4	+ 0	= 4

Ordered values of individual slope estimates:
0, 2, 2, 2.5, 2.5, 3, 3, 3.5, 4, 5
Median: Seasonal Kendall slope estimate = 2.75
80% Limits: 0.936 and 4.53

Part B. Computing the Seasonal Kendall Test

$n_1 = 4$ $\quad n_2 = 4$
$g_1 = 0$ $\quad g_2 = 1,\ t_{21} = 2$
$h_1 = 1,\ u_{11} = 2$ $\quad h_2 = 1,\ u_{21} = 2$
$N'_1 = 5$ $\quad N'_2 = 5$

$\text{Var}(S_1) = \frac{1}{18}[4(3)(13) - 2(1)(9)] + 0 + 0 = 7.667$
$\text{Var}(S_2) = \frac{1}{18}[4(3)(13) - 2(1)(9) - 2(1)(9)] + 0 + [2(1)][2(1)]/8(3)$
$= 6.667 + 0.1667 = 6.834$
$[\text{Var}(S_1)]^{1/2} = 2.8 \quad [\text{VAR}(S_2)]^{1/2} = 2.6$
$S' = S_1 + S_2 = 5 + 4 = 9$
$\text{VAR}(S') = \text{VAR}(S_1) + \text{VAR}(S_2) = 7.667 + 6.834 = 14.5$
$[\text{VAR}(S')]^{1/2} = 3.808 \quad Z = \frac{(9-1)}{3.808} = 2.1^b$

[a] Cannot be computed since both data values are within the same time period.
[b] Referring this value to Table A1 is only an approximate test for this example, since n_1 and n_2 are small and there are only two seasons.

procedure developed by van Belle and Hughes (1984) to test for homogeneity of trend direction in different seasons at a given station. This latter test is important, since if the trend is upward in one season and downward in another, the seasonal Kendall test and slope estimator will be misleading.

The procedure is to compute

$$\chi^2_{\text{homog}} = \chi^2_{\text{total}} - \chi^2_{\text{trend}} = \sum_{i=1}^{K} Z_i^2 - K\bar{Z}^2$$

where

$$Z_i = \frac{S_i}{\cdots}$$

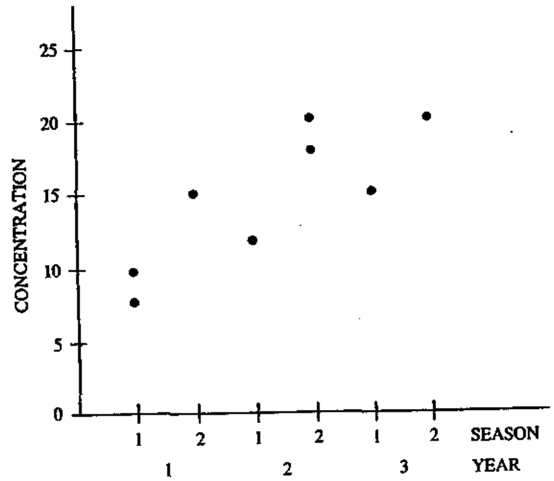


Figure 17.1 Artificial data set to illustrate computation of the seasonal Kendall slope estimator.

If χ^2_{homog} exceeds the α critical value for the chi-square distribution with $K - 1$ df, we reject the null hypothesis, H_0, of homogeneous seasonal trends over time (trends in the same direction and of the same magnitude). In that case the seasonal Kendall test and slope estimate are not meaningful, and it is best to compute the Mann-Kendall test and Sen's slope estimator for each individual season. If χ^2_{homog} does not exceed the critical value in the chi-square tables (Table A19), our calculated value of $\chi^2_{\text{trend}} = K\bar{Z}^2$ is referred to the chi-square distribution with 1 df to test for a common trend in all seasons.

The critical value obtained from the chi-square tables will tend to be too small unless (1) the number of data used to compute each Z_i is 10 or more, and (2) the data are spaced far enough apart in time so that the data in different seasons are not correlated. For some water quality variables Lettenmaier (1978) found that this implies that sampling should be at least two weeks apart.

Van Belle and Hughes (1984) show how to test whether there is a pattern to the trend heterogeneity when χ^2_{homog} is significantly large. They illustrate by showing how to test whether trends in summer and winter months are significantly different.

17.4 SEN'S TEST FOR TREND

when there are missing values, and the test is inexact in that case. Given these facts, Sen's test is preferred to the seasonal Kendall test when no data are missing. The computer code in Appendix B also computes Sen's test. Computational procedures are given in van Belle and Hughes (1984).

17.5 TESTING FOR GLOBAL TRENDS

In Section 17.3 the χ^2_{homog} statistic was used to test for homogeneity of trend direction in different seasons at a given sampling station. This test is a special case of that developed by van Belle and Hughes (1984) for $M > 1$ stations. Their procedures allow one to test for homogeneity of trend direction at different stations when seasonality is present. The test for homogeneity given in Section 16.4.4 is a special case of this test. Van Belle and Hughes illustrate the tests, using temperature and biological oxygen demand data at two stations on the Willamette River.

The required data are illustrated in Table 17.3. The first step is to compute the Mann-Kendall statistic for each season at each station by Eq. 17.1. Let S_{im} denote this statistic for the ith season at the mth station. Then compute

$$Z_{im} = \frac{S_{im}}{[\text{VAR}(S_{im})]^{1/2}}, \qquad i = 1, 2, \cdots, K, \quad m = 1, 2, \cdots, M \qquad \textbf{17.6}$$

where $\text{VAR}(S_{im})$ is obtained by using Eq. 17.2. (For this application all quantities in Eq. 17.2 relate to the data set for the ith season and mth station.) Note that missing values, NDs, or multiple observations per time period are allowed, as discussed in Section 17.1. Also, note that the correction for continuity (± 1 added to S in Eq. 16.5 and S' in Eq. 17.5) is not used in Eq. 17.6 for reasons discussed by van Belle and Hughes (1984).

Next, compute

$$\bar{Z}_{i\cdot} = \frac{1}{M}\sum_{m=1}^{M} Z_{im}, \qquad i = 1, 2, \cdots, K$$

$$= \text{mean over } M \text{ stations for the } i\text{th season}$$

Table 17.3 Data to Test for Trends Using the Procedure of van Belle and Hughes (1984)

		Station 1				...		Station M			
Season		1	2	...	K	...		1	2	...	K
Year	1	x_{111}	x_{211}	...	x_{K11}	...	1	x_{11M}	x_{21M}	...	x_{K1M}
	2	x_{121}	x_{221}	...	x_{K21}	...	2	x_{12M}	x_{22M}	...	x_{K2M}
	⋮						⋮				
	L	x_{1L1}	x_{2L1}	...	x_{KL1}	...	L	x_{1LM}	x_{2LM}	...	x_{KLM}

$$\bar{Z}_{\cdot m} = \frac{1}{K} \sum_{i=1}^{K} Z_{im}, \qquad m = 1, 2, \cdots, M$$

= mean over K seasons for the mth station

$$\bar{Z}_{\cdot\cdot} = \frac{1}{KM} \sum_{i=1}^{K} \sum_{m=1}^{M} Z_{im}$$

= grand mean over all KM stations and seasons

Now, compute the chi-square statistics in Table 17.4 in the following order: χ^2_{total}, χ^2_{trend}, χ^2_{station}, and χ^2_{season}. Then compute

$$\chi^2_{\text{homog}} = \chi^2_{\text{total}} - \chi^2_{\text{trend}}$$

and

$$\chi^2_{\text{station-season}} = \chi^2_{\text{homog}} - \chi^2_{\text{station}} - \chi^2_{\text{season}}$$

Refer χ^2_{station}, χ^2_{season}, and $\chi^2_{\text{station-season}}$ to the α level critical values in the chi-square tables with $M - 1$, $K - 1$, and $(M - 1)(K - 1)$ df, respectively.

If all three tests are nonsignificant, refer χ^2_{trend} to the chi-square distribution with 1 df to test for global trend. If χ^2_{season} is significant, but χ^2_{station} is not, that is, if trends have significantly different directions in different seasons but not at different stations, then test for a different trend direction in each season by computing the K seasonal statistics

$$M\bar{Z}^2_{i\cdot} \qquad i = 1, 2, \ldots, K \text{ seasons} \qquad 17.7$$

and referring each to the α-level critical value of the chi-square distribution with 1 df.

If χ^2_{station} is significant, but χ^2_{season} is not, that is, if trends have significantly different directions at different stations but not in different seasons, then test for a significant trend at each station by computing the M station statistics

Table 17.4. Testing for Trends Using the Procedure of van Belle and Hughes (1984)

Chi-Square Statistics	*Degrees of Freedom*	*Remarks*
$\chi^2_{\text{total}} = \sum_{i=1}^{K} \sum_{m=1}^{M} Z^2_{im}$	KM	
$\chi^2_{\text{homog}} = \sum_{i=1}^{K} \sum_{m=1}^{M} Z^2_{im} - KM\bar{Z}^2_{\cdot\cdot}$	$KM - 1$	Obtained by subtraction
$\chi^2_{\text{season}} = M \sum_{i=1}^{K} \bar{Z}^2_{i\cdot} - KM\bar{Z}^2_{\cdot\cdot}$	$K - 1$	Test for seasonal heterogeneity
$\chi^2_{\text{station}} = K \sum_{m=1}^{M} \bar{Z}^2_{\cdot m} - KM\bar{Z}^2_{\cdot\cdot}$	$M - 1$	Test for station heterogeneity
$\chi^2_{\text{station-season}} = \sum^{K} \sum^{M} Z^2_{im} - M \sum^{K} \bar{Z}^2_{i\cdot} -$	$(M - 1)(K - 1)$	Test for interaction Obtained by subtraction

$$K\bar{Z}^2_{\cdot m} \qquad m = 1, 2, \ldots, M \text{ stations}$$

and refer to the α-level critical value of the chi-square distribution with 1 df.

If both χ^2_{station} and χ^2_{season} are significant or if $\chi^2_{\text{station-season}}$ is significant, then the χ^2 trend test should not be done. The only meaningful trend tests in that case are those for individual station-seasons. These tests are made by referring each Z_{im} statistic (see Table 17.3) to the α-level critical value of the standard normal table (Table A1), as discussed in Section 16.4.2 (or Section 16.4.3 if multiple observations per season have been collected). For these individual Mann-Kendall tests, the Z_{im} should be recomputed so as to include the correction for continuity (± 1) as given in Eq. 16.4.

The computer code listed in Appendix B computes all the tests we have described as well as Sen's estimator of slope for each station-season combination. In addition, it computes the seasonal Kendall test, Sen's aligned test for trends, the seasonal Kendall slope estimator for each station, the equivalent slope estimator (the "station Kendall slope estimator") for each season, and confidence limits on the slope.

The code will compute and print the K seasonal statistics (Eq. 17.7) to test for equal trends at different sites for each season only if (1) the computed P value of the χ^2_{season} test *is less than* α', and (2) the computed P value of the χ^2_{station} *exceeds* α', where α' is an a priori specified significance level, say $\alpha' = 0.01$, 0.05, or 0.10, chosen by the investigator. Similarly, the M station statistics (Eq. 17.8) are computed only if the computed P value of χ^2_{station} is less than α' and that for χ^2_{season} is greater than α'. The user of the code can specify the desired value of α'. A default value of $\alpha' = 0.05$ is used if no value is specified.

EXAMPLE 17.2

Table 17.5 gives a set of data collected monthly at 2 stations for 4 years (plotted in Fig. 17.2). These data were simulated on a computer using the lognormal, autoregressive, seasonal cycle model given in Hirsch, Slack, and Smith (1982, p. 112). The data at station 1 have no long-term trend (i.e., they have a slope of zero), whereas station 2 has an upward trend of 0.4 units per year for each season. Hence, seasonal trend directions are homogeneous, but the station trend directions are not.

The chi-square tests are given in Table 17.6. We obtain that $\chi^2_{\text{station}} = 8.16$ has a P value of 0.004. That is, the probability is only 0.004 of obtaining a χ^2_{station} value this large when trends over time at the 2 stations are in the same direction. Hence, the data suggest trend directions are different at the 2 stations, which is the true situation. Both χ^2_{season} and $\chi^2_{\text{station-season}}$ statistics (8.48 and 2.63) are small enough to be nonsignificant. This result is also expected, since trend direction does not change with season.

We chose $\alpha' = 0.05$. Since χ^2 [illegible]

Table 17.5 Simulated Water Quality Using a Lognormal Autoregressive, Seasonal Cycle Model Given by Hirsch, Slack, and Smith (1982, Eq. 14f)

NUMBER OF YEARS = 4
NUMBER OF SEASONS = 12
NUMBER OF STATIONS = 2

STATION 1 YEAR	SEASON	NUMBER OF DATA POINTS n = 48 STATION 1	STATION 2 YEAR	SEASON	NUMBER OF DATA POINTS n = 48 STATION 2
1	1	6.32	1	1	6.29
1	2	6.08	1	2	6.11
1	3	5.16	1	3	5.66
1	4	4.47	1	4	5.16
1	5	4.13	1	5	4.75
1	6	3.65	1	6	6.79
1	7	3.48	1	7	4.51
1	8	3.78	1	8	4.37
1	9	3.94	1	9	4.95
1	10	4.40	1	10	5.22
1	11	4.94	1	11	5.73
1	12	5.32	1	12	6.72
2	1	5.82	2	1	7.42
2	2	5.76	2	2	7.56
2	3	4.88	2	3	6.13
2	4	4.84	2	4	6.24
2	5	4.87	2	5	5.07
2	6	4.13	2	6	4.95
2	7	3.51	2	7	4.59
2	8	4.32	2	8	5.22
2	9	4.06	2	9	5.13
2	10	4.47	2	10	5.69
2	11	5.05	2	11	6.41
2	12	5.20	2	12	7.53
3	1	5.83	3	1	7.02
3	2	5.65	3	2	6.93
3	3	5.32	3	3	6.55
3	4	5.33	3	4	6.66
3	5	4.20	3	5	6.69
3	6	3.85	3	6	5.23
3	7	4.45	3	7	5.14
3	8	3.56	3	8	5.06
3	9	3.85	3	9	5.71
3	10	4.72	3	10	6.17
3	11	5.38	3	11	6.78
3	12	5.33	3	12	7.64
4	1	6.59	4	1	7.46
4	2	5.93	4	2	7.56
4	3	4.98	4	3	7.30
4	4	4.61	4	4	7.22
4	5	4.18	4	5	6.07
4	6	3.79	4	6	5.53

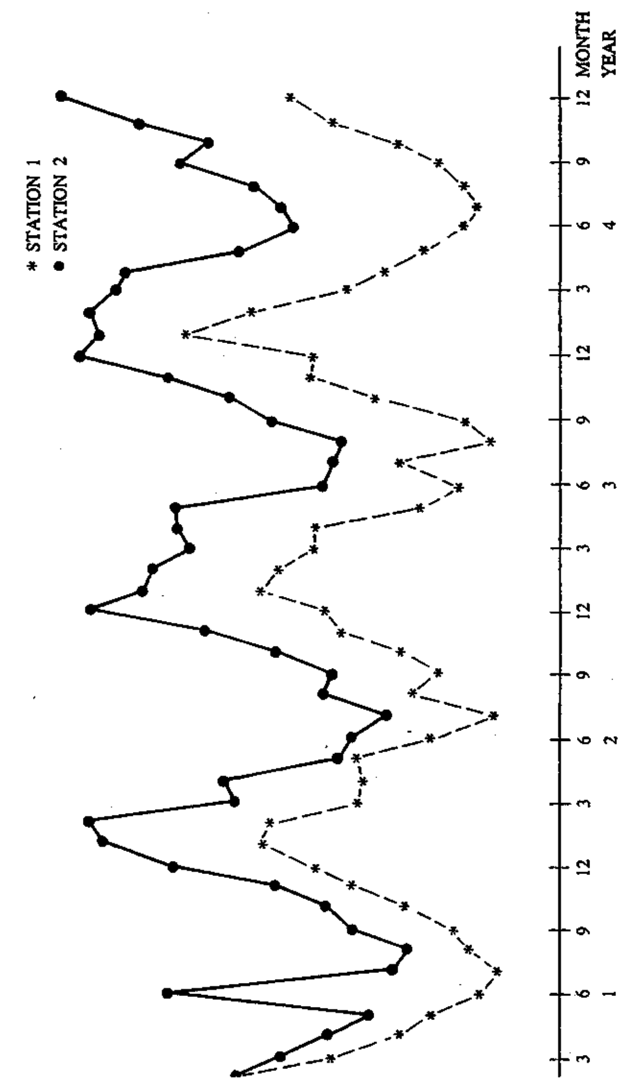


Figure 17.2 Data at two stations each month for four years. Data were simulated using the lognormal autoregressive seasonal model given by Hirsch, Slack, and Smith (1982, Eq. 14f). Simulated data were obtained by D. W. Engel.

Table 17.6 Chi-Square Tests for Homogeneity of Trends over Time between Seasons and between Stations

HOMOGENEITY TEST RESULTS

CHI-SQUARE STATISTICS		df	PROB. OF A LARGER VALUE	
TOTAL	45.02007	24	0.006	
HOMOGENEITY	19.26657	23	0.686	
SEASON	8.48201	11	0.670	
STATION	8.15667	1	0.004	← Trends not equal at the 2 stations
STATION-SEASON	2.62789	11	0.995	
TREND	25.75349	1	0.000	← Not meaningful

INDIVIDUAL STATION TREND

STATION	CHI-SQUARE	df	PROB. OF A LARGER VALUE
1	2.46154	1	0.117
2	31.44863	1	0.000

evidence of a trend at station 1 (P level = 0.117) and a definite trend at station 2 (P level = 0.000).

Table 17.7 gives the seasonal Kendall and Sen aligned rank tests at both stations. These results agree with the true situation. The seasonal Kendall slope estimates are 0.042 and 0.440, which are slightly larger than the actual values of 0.0 and 0.4, respectively. The lower and upper confidence limits on the true slope are also given in Table 17.7. Finally, Table 17.8 gives the individual Mann-Kendall tests for trend over time for each season-station combination. Since n is only 4 for each test, the P values are approximate because they were obtained from the normal distribution (Table A1). The exact P values obtained from Table A18 are also shown in the table. The approximate levels are quite close to the exact. None of the tests for station 1 are significant, and the 12 slope estimates vary from −0.08 to 0.208 (the true value is zero). Seven of the 12 tests for station 2 are significant at the α = 0.10 2-tailed level. If n were greater than 4, more of the tests for station 2 would have been significant. The 12 slope estimates range from −0.070 to 0.623 with a mean of 0.414. Since n is so small, these estimates are quite variable, but their mean is close to the true 0.40. Confidence intervals for the true slope for 4 station-season combinations are shown in Table 17.9. The computer code computes these for all KM combinations.

Table 17.7 Seasonal Kendall and Sen Aligned Ranks Tests for Trend over Time

STATION	SEASONAL KENDALL	n	PROB. OF EXCEEDING THE ABSOLUTE VALUE OF THE KENDALL STATISTIC (TWO-TAILED TEST)
1	1.47087	48	0.141
2	5.51784	48	0.000

STATION	SEN T	n	PROB. OF EXCEEDING THE ABSOLUTE VALUE OF THE SEN T STATISTIC (TWO-TAILED TEST)
1	1.02473	48	0.306
2	4.57814	48	0.000

SEASONAL-KENDALL SLOPE CONFIDENCE INTERVALS

STATION	ALPHA	LOWER LIMIT	SLOPE	UPPER LIMIT
1	0.010	-0.060	0.042	0.111
	0.050	-0.020	0.042	0.085
	0.100	-0.004	0.042	0.081
	0.200	0.007	0.042	0.070
2	0.010	0.345	0.440	0.525
	0.050	0.365	0.440	0.499
	0.100	0.377	0.440	0.486
	0.200	0.380	0.440	0.478

by outliers and gross errors, and missing data or ND values are allowed. However, the tests still require the data to be independent. If they are not, the tests tend to indicate that trends are present more than the allowed $100\alpha\%$ of the time.

EXERCISES

17.1 Use the following data to test for no trend versus a rising trend, using the seasonal Kendall test. Use α = 0.01.

	Season					
Year	*1*	*2*	*3*	*4*	*5*	*6*

Table 17.8 Mann-Kendall Tests for Trend over Time for Each Season at Each Station

STATION	SEASON	MANN-KENDALL S STATISTIC	Z[a] STATISTIC	n	PROB. OF EXCEEDING THE ABSOLUTE VALUE OF THE Z STATISTIC (TWO-TAILED TEST) IF n > 10		SEN SLOPE
1	1	2	0.33968	4	0.734	(0.750)[b]	0.050
	2	-2	-0.33968	4	0.734	(0.750)	-0.080
	3	0	0.00000	4	1.000	(1.000)	-0.005
	4	2	0.33968	4	0.734	(0.750)	0.208
	5	0	0.00000	4	1.000	(1.000)	-0.002
	6	0	0.00000	4	1.000	(1.000)	-0.007
	7	4	1.01905	4	0.308	(0.334)	0.059
	8	-2	-0.33968	4	0.734	(0.750)	-0.057
	9	0	0.00000	4	1.000	(1.000)	0.016
	10	4	1.01905	4	0.308	(0.334)	0.052
	11	4	1.01905	4	0.308	(0.334)	0.090
	12	4	1.01905	4	0.308	(0.334)	0.107
2	1	4	1.01905	4	0.308	(0.334)	0.378
	2	3	0.72232	4	0.470	( )[c]	0.447
	3	6	1.69842	4	0.089	(0.084)	0.508
	4	6	1.69842	4	0.089	(0.084)	0.623
	5	4	1.01905	4	0.308	(0.334)	0.470
	6	0	0.00000	4	1.000	(1.000)	-0.070
	7	6	1.69842	4	0.089	(0.084)	0.445
	8	4	1.01905	4	0.308	(0.334)	0.442
	9	6	1.69842	4	0.089	(0.084)	0.578
	10	6	1.69842	4	0.089	(0.084)	0.435
	11	6	1.69842	4	0.089	(0.084)	0.413
	12	6	1.69842	4	0.089	(0.084)	0.300

[a] ±1 correction factor used to compute the Z statistic.
[b] Exact two-tailed significance levels for the S statistic using Table A18.
[c] Cannot be determined from Table A18 since $S = 3$ resulted because of two tied data in the season.

17.3 Use the results in Exercises 17.1 and 17.2 to compute an 80% confidence interval about the true slope.

17.4 Test for equal trend directions in different seasons, using the data in Exercise 17.1. Use $\alpha = 0.01$. If the trends in the 6 seasons are homogeneous, use chi-square to test for a statistically significant trend at the $\alpha = 0.05$ level.

17.5 Suppose the data in Exercise 17.1 were collected at station 1 and the following data were collected at station 2.

Table 17.9 Sen Slope Estimates and Confidence Intervals for Each Station-Season Combination

STATION	SEASON	ALPHA	SEN SLOPE CONFIDENCE INTERVALS LOWER LIMIT	SLOPE	UPPER LIMIT
1	1	0.010	n too small[a]	0.050	n too small[a]
		0.050	n too small	0.050	0.087
		0.100	n too small	0.050	0.440
		0.200	-0.471	0.050	0.718
	2	0.010	n too small	-0.080	n too small
		0.050	n too small	-0.080	0.032
		0.100	n too small	-0.080	0.162
		0.200	-0.308	-0.080	0.258
2	1	0.010	n too small	0.378	n too small
		0.050	n too small	0.378	-0.171
		0.100	n too small	0.378	0.511
		0.200	-0.353	0.378	1.052
	2	0.010	n too small	0.447	n too small
		0.050	n too small	0.447	0.251
		0.100	n too small	0.447	0.984
		0.200	-0.488	0.447	1.265

[a] The lower and upper limits cannot be computed if n is too small.

Test for homogeneity of trend direction between seasons and between stations, using the chi-square tests in Table 17.4 with $\alpha = 0.01$. Test for a significant common trend at the 2 stations, if appropriate.

ANSWERS

17.1 $\mathrm{Var}(S_i) = 3(2)(11)/18 = 3.667$ for each season. $S' = \Sigma_{i=1}^{6} S_i = 18$, $\mathrm{Var}(S') = 6(3.667) = 22$. From Eq. 17.5, $Z = 17/\sqrt{22} = 3.62$. Since $\alpha = 0.01$ (one-tailed test), $Z_{0.99} = 2.326$. Since $3.62 > 2.326$, we accept the hypothesis of a rising trend.

17.2 The median of the 18 slope estimates is 1.09 units per year.

17.3 $Z_{1-\alpha/2} = Z_{0.90} = 1.282$, $\mathrm{Var}(S') = 22$ from Exercise 17.1. Therefore, $C_\alpha = 1.282\sqrt{22} = 6.0131$, $M_1 = 6$, $M_2 + 1 = 13$. Lower limit = 0.81; upper limit = 1.4.

17.4 From Exercise 17.1 we have $Z_1 = 1.567 = Z_2 = Z_3 = Z_4 = Z_6$. Therefore $\bar{Z} = 1.567$; then $\chi^2_{total} = 14.7$, $\chi^2_{trend} = 14.7$, $\chi^2_{homog} = 0$. Since $\chi^2_{homog} < 15.09$ (from Table A19), we cannot reject the null hypothesis of homogeneous trend direction in all seasons. Hence, test for trend, using $\chi^2_{trend} = 14.7$. Since $14.7 > 3.84$ (from Table A19), we