Juergen Schwarz

Correlation Coefficients According to Bravais-Pearson, Spearman, and Kendall

ABSTRACT *This paper discusses computer applications in the use of correlation coefficients. Programs for calculating the rank correlation coefficients, since they are distribution free, can be frequently used for ordinally scaled data.*

Keywords: Statistics, empirical covariance, correlation, correlation coefficients according to Bravais-Pearson, Spearman's rank correlation coefficient, Kendall's rank correlation coefficient, partial correlation. **Hardware:** HP 9845B with structured programming ROM and advanced programming ROM.

1 Introduction

In the analysis of statistical data it is important to know whether interdependences and interrelations exist among various characteristics. Sometimes we can be sure from theoretical consideration that such relationships exist. If these can be quantified then further consideration is unnecessary. If there is no such knowledge then another approach will have to be adopted in order to obtain qualitative and quantitative information.

If studies are limited to two-dimensional numerical data, then the known correlation coefficients according to Bravais-Pearson, Spearman, and Kendall (Reference 1) provide a measure of the linear relationship between the two characteristics. On the use of correlation coefficients, it is essential to ensure that a logical relationship exists between the two characteristics of a series of observations; it is all too easy to calculate "nonsense correlations" if this precaution is not observed (Reference 1). It is also general knowledge that small (non-significant) correlation coefficients do not necessarily mean the characteristics are independent since only linear correlations can be established with the above-mentioned coefficients (Reference 2). Drawing a scatter diagram can provide additional information in such cases (Figure 1e).

Dr.-Ing. Juergen Schwarz, AEG Aktiengesellschaft, Power Electronics and Plant Equipment, Culemeyerstrasse 1, D-1000 Berlin 48, FRG.

2 Covariance and Correlation Coefficients

The correlation coefficient $\rho(X,Y)$ of the random variables X and Y is defined as:

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}} = \frac{\sigma(X, Y)}{\sigma(X) \cdot \sigma(Y)} \quad (1)$$

Since the covariance Cov(X,Y) for independent random variables X and Y vanishes, the correlation coefficient of independent random variables is equal to zero. Such random variables are called uncorrelated. However, being uncorrelated does not mean that the random variables are independent. This follows from the form (Reference 3) of the covariance (Figure 1e):

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) \quad (2)$$

For random variables X and Y having a normal distribution the converse conclusion is admissible, i.e. if X and Y are normally distributed and $\rho(X,Y)$ is zero, then X and Y are independent.

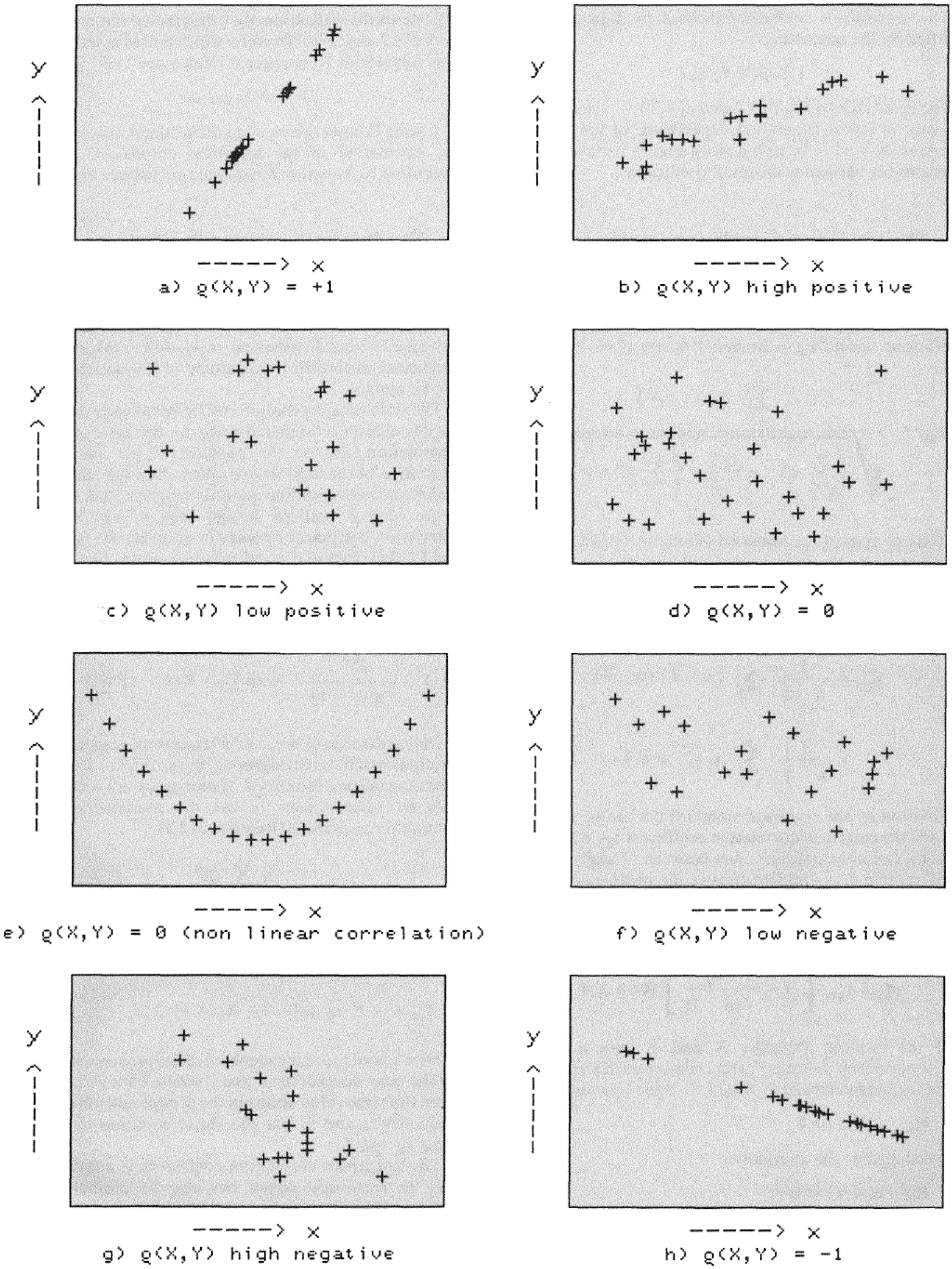


Figure 1 *Correlation coefficients ρ(X, Y) and scatter diagrams.*

The correlation coefficient defined by Equation 1 is satisfied by the inequality:

$$-1 \leq \rho(X,Y) \leq 1 \tag{3}$$

The equal sign ($\rho(X,Y) = 1$ and $\rho(X,Y) = -1$) applies in the case of linear functional dependence of the random numbers ($X = aY + b$, with $a \neq 0$). Figure 1 shows scatter diagrams for various correlation coefficients.

3 The Empirical Correlation Coefficient According to Bravais-Pearson

If the characteristics (X,Y) to be linked are on an interval or ratio (cardinal) scale then the empirical correlation coefficient, according to Bravais-Pearson, given by:

$$r_{XY} = \frac{\sum_{i=1}^{n} x_i y_i - n \bar{x} \bar{y}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - n \bar{x}^2\right) \left(\sum_{i=1}^{n} y_i^2 - n \bar{y}^2\right)}} \tag{4}$$

affords an asymptotic unbiased estimator for $\rho(X,Y)$. Here (x_i,y_i) are the i-th observation from a series of n measurements and $\bar{x}$ and $\bar{y}$ are the arithmetic means of x_i and y_i, respectively. The numerator of Equation 4 is essentially the empirical covariance between X and Y defined as:

$$s_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} x_i y_i - n \bar{x} \bar{y} \tag{5}$$

Division by the empirical standard deviations s_X and s_Y affords the empirical correlation coefficient r_{XY} which, like $\rho(X,Y)$, can only assume values between -1 and +1. For a small sample n, r_{XY} underestimates the parameter $\rho(X,Y)$. An improved estimation of $\rho(X,Y)$ with a smaller variance around the expected value is given (Reference 2) by:

$$r^*_{XY} = r_{XY} \left[1 + \frac{1 - r^2_{XY}}{2(n-3)} \right] \text{ for } n \geq 8 \tag{6}$$

If the random variables X and Y have a normal distribution then the empirical correlation coefficient allows a test for independence of X and Y. If the hypothesis:

H_0: $\rho(X,Y) = 0$

is tested against the alternative:

H_1: $\rho(X,Y) \neq 0$

at a confidence level α, then:

$$t = \frac{r_{XY}\sqrt{n-2}}{\sqrt{1 - r^2_{XY}}} \tag{7}$$

represents the realization of a t-distributed random variable with $n - 2$ degrees of freedom which is used as test criterion. The hypothesis H_0 is rejected (Reference 1) if:

$$|t| > t_{n-2;1-\alpha/2} \tag{8}$$

Listing 1 shows two versions of the functional subprogram for calculation of the empirical correlation coefficient according to Equation 4 requiring no further explanation.

4 The Rank Correlation Coefficient According to Spearman

The Pearson correlation coefficient should not be used to estimate the correlation between random numbers not having a normal distribution; instead the rank correlation coefficient, according to Spearman or Kendall (Section 5), can be applied.

The term rank correlation coefficient already reveals that the correlation is estimated solely on the basis of the rank information, i.e. on the position of the value in the realization of the random quantities. As a rule, the smallest realization of the random quantities is given rank 1, the next larger rank 2, and the largest rank n. The Spearman correlation coefficient r results in inserting the ranks $R(x_i)$ and $R(y_i)$ in place of the realizations x_i and y_i into Equation 4 for the Pearson correlation coefficient. After conversion we obtain:

$$r_S = 1 - \frac{6 \sum_{i=1}^{n} D_i^2}{n(n^2 - 1)} \quad \text{with } D_i = R(x_i) - R(y_i) \tag{9}$$

On occurrence of ties, i.e. when several realizations of a random variable are congruous, the arithmetic mean of the ranks in question is inserted. If numerous ties occur, then it may be advantageous to use the modified Spearman correlation coefficient (References 2,7):

$$r_{S,B} = 1 - \frac{6 \sum_{i=1}^{n} D_i^2}{n(n^2 - 1) - (T_{x'} + T_{y'})} \tag{10}$$

with:

$$T_{x'} = ½ \sum t_{x'} (t^2_{x'} - 1); \quad T_{y'} = ½ \sum t_{y'} (t^2_{y'} - 1)$$

Here $t_{x'}$ and $t_{y'}$ are the numbers of ties in successive groups (of the same rank) of the x' and y' series. Thus we count how often the same value occurs in the groups and introduce the frequency $t_{x'}$ and $t_{y'}$ into the above equation to form the sums $T_{x'}$ and $T_{y'}$.

The Spearman correlation coefficient is applicable not only to cardinally scaled but also to ordinally scaled characteristics. It affords useful information also for small samples and non-normal distributions. Moreover, the effect of outliers, which can have a pronounced influence on the magnitude of the Pearson correlation coefficient, is weakened. However, it only utilizes 91% of the observations (Reference 2).

To test the significance of r_S, i.e. to test the hypothesis:

H_0: the series of measurements are independent.

against the alternative:

H_1: the series of measurements are not independent

we have to take recourse to tabulated values for samples $n < 30$ (References 1,2).

For $n \geq 30$ Equations 7 and 8 can be used approximately also for r_S and $r_{S,B}$ (Reference 2).

Listing 2 shows a functional subprogram for calculation of the Spearman rank correlation coefficients. The measured values are stored in lines 250 and 260 and then sorted in increasing order. Then the rank values of X(*) are inserted into the vector D(*) in lines 320 to 420. Ties are considered by the REPEAT-UNTIL loop in lines 340 to 370. The calculation of the rank values of Y(*) proceeds analogously and insertion of the difference in rank into the vector D(*) in line 510. The result according to Equations 9 and 10 is subsequently calculated in lines 570 to 580.

5 The Rank Correlation Coefficient According to Kendall

Less frequently encountered than the Spearman rank correlation coefficient r_S is the correlation coefficient according to Kendall (Reference 1). It is calculated from ranks assigned in exactly the same way as for Spearman's r_S. Corresponding pairs of rank numbers ($R(x_i)$, $R(y_i)$) are calculated from the n natural pairs of observations (x_i, y_i), ..., (x_n,y_n). The pairs of rank are then ordered according to $R(x_i)$. In this way the order of the rank numbers of the realization y_i, ..., y_n of the random variable Y is determined. From this series the number q_i of the rank numbers $R(y_j)$, which are less than or equal to $R(y_i)$ and come behind $R(y_i)$ in the series, is calculated for every rank number $R(y_i)$. The Kendall τ value is then:

$$\tau = 1 - \frac{4 \sum_{i=1}^{n} q_i}{n(n-1)} \qquad (11)$$

The test criterion for the independence of the two series of measurements is the value:

$$K = ½\ n(n-1)\ \tau = ½\ n(n-1) - 2 \sum_{i=1}^{n} q_i \qquad (12)$$

which is compared with the Kendall K statistic.

The hypothesis:

H_0: the series of measurements are independent

is rejected in favor of the alternative:

H_1: the series of measurements are not independent

at the confidence level α, if:

$$| K | > K_{n;1-\alpha/2} \qquad (13)$$

A table of the critical values $K_{n;1-\alpha/2}$ is found, for example, in Reference 1. For approximate calculations the test quantity:

$$K^* = \frac{K}{\sqrt{n(n-1)(2n+5)/18}} \qquad (14)$$

can be used instead of the table, and the null hypothesis is rejected if:

$$|K^*| > u_{1-\alpha/2} \qquad (15)$$

which is the quantile of the standard normal distribution.

A functional subprogram for calculation of Kendall's τ is shown in Listing 3. After intermediate storage of the vector Y(*) in line 210 it is sorted in the next line. The rank values $R(y_i)$ of the vector Y_prime(*) are determined in the following REPEAT-UNTIL loop and inserted in the vector Y_prime_rank(*). Ties are considered by the LOOP construction in lines 260 to 300. This kind of rank calculation could also have been used in the Spearman functional subprogram (Listing 2) just as the algorithm (Listing 2) is also applicable here.

The X values are now temporarily stored and sorted (lines 360 to 380). The index J for each rank I is first determined in lines 400 to 430. Then the corresponding index K of the Y vector is determined. The actual elements are then filled with an auxiliary value $9.999\,999\,999\,99 \times 10^{99}$ to preclude ties. This auxiliary value may not appear in the starting fields. The intermediate result (rank $R(y_i)$ with index I sorted according to increasing $R(x_i)$) into the vector Y_rank(*) n will be inserted in line 470.

The final result is provided by lines 510 to 570. In accord with the procedure for calculating Kendall's τ the sum of the number of the rank values in the sorted vectors Y_rank(*) which for an increasing r index are lower than or equal to the actual rank, is inserted into Q. The result according to Equation 11 is ultimately transferred to line 570.

6 Partial Correlation

A correlation between the characteristics X and Y frequently occurs only because the two characteristics are both correlated with a third characteristic Z. The correlation calculated between X and Y is then only an illusory correlation. Partial correlation of X and Y with constant Z, according to:

$$\rho_{(X,Y)/Z} = \frac{\rho_{XY} - \rho_{XZ}\ \rho_{YZ}}{\sqrt{(1-\rho_{XZ}^2)(1-\rho_{YZ}^2)}} \qquad (16)$$

provides a means of "working out" such a third influence.

7 The Empirical Partial Correlation Coefficient

For a normal distribution of characteristics X,Y, and Z of the parent population, it is possible to estimate $\rho_{(X,Y)/Z}$ by estimating the correlation coefficients ρ_{XY}, ρ_{XZ}, and ρ_{YZ} with the aid of the empirical correlation coefficients according to Bravais-Pearson. The estimator for correlation between the characteristics X and Y on partialization of the characteristic Z thus results as:

$$r_{(X,Y)/Z} = \frac{r_{XY} - r_{XZ}\ r_{YZ}}{\sqrt{(1-r_{XZ}^2)(1-r_{YZ}^2)}} \qquad (17)$$

A test for partial independence of X and Y with Z at the confidence level α can be performed with the test criterion:

$$t = \frac{r_{(X,Y)/Z}\sqrt{n-3}}{\sqrt{1-r^2_{(X,Y)/Z}}} \tag{18}$$

The hypothesis:

H_0: $\rho_{(X,Y)/Z} = 0$

is rejected by the test in favor of the alternative:

H_1: $\rho_{(X,Y)/Z} \neq 0$

if:

$$|t| > t_{n-3;1-\alpha/2} \tag{19}$$

is the quantile of the t distribution with n − 3 degrees of freedom (Reference 1).

8 The Partial Rank Correlation Coefficient According to Kendall

Kendall's partial rank correlation coefficient is an estimator for the partial correlation $\rho_{(X,Y)/Z}$ which is suitable for characteristics X, Y, and Z, which are at least ordinally scaled, but is otherwise distribution free. It is calculated from the individual Kendall τ's between the pairs of characteristics (X,Y), (X,Z), and (Y,Z), designated τ_{XY}, τ_{XZ}, and τ_{YZ} in a manner similar to Equation 17:

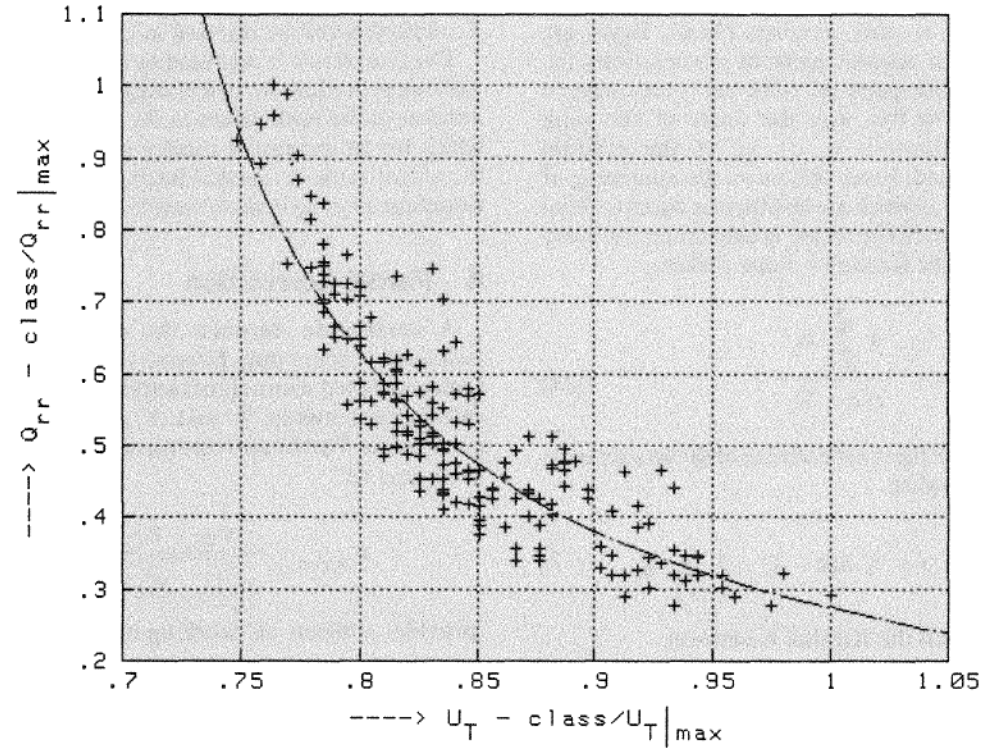
$$\tau_{(X,Y)/Z} = \frac{\tau_{XY} - \tau_{XZ}\,\tau_{YZ}}{\sqrt{(1-\tau^2_{XZ})\,(1-\tau^2_{YZ})}} \tag{20}$$

is thus the estimator for the partial correlation between the characteristics X and Y with Z. No tests of significance are yet known for partial τ (References 1,7).

9 A Detailed Example

The somewhat theoretically oriented subject matter presented so far will now be illustrated by an example from the laboratory.

Modern equipment used for inverters incorporates semiconductor components as control elements. These semiconductor components (generally thyristors, less often diodes) are relatively sensitive to overloading. One of the decisive criteria in dimensioning these inverters is the




```
Spearman's rank correlation coefficient without ties                    r_S,B = -0.891
Spearman's rank correlation coefficient with ties                       r_S,B = -0.892
empirical correlation coefficient according to Bravais-Pearson          r     = -0.842

Remark:   In this figure a empirical regression function of the kind of
          Q_rr = a_regr/(b_regr+U_T) is shown. The coefficients are calculated
          by the method of least squares.
```

Figure 2 U_T-Q_{rr} *scatter diagram of 180 thyristors T 2200 N 3600 ... 4400 (AEG).*

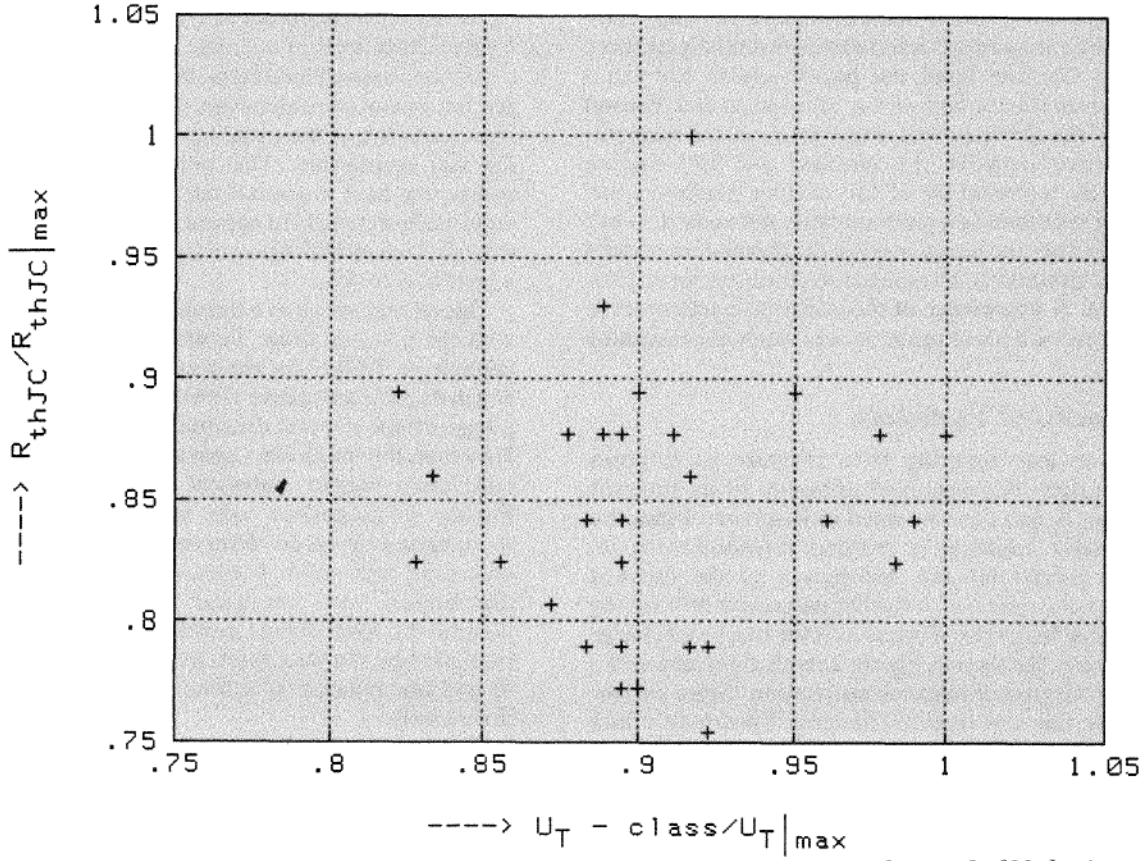


Figure 3 *Relation between electrical and thermal parameters [U_T – R_{thJC} scatter diagram] of 32 thyristors T 2200 N 3600 . . . 4400 (AEG).*

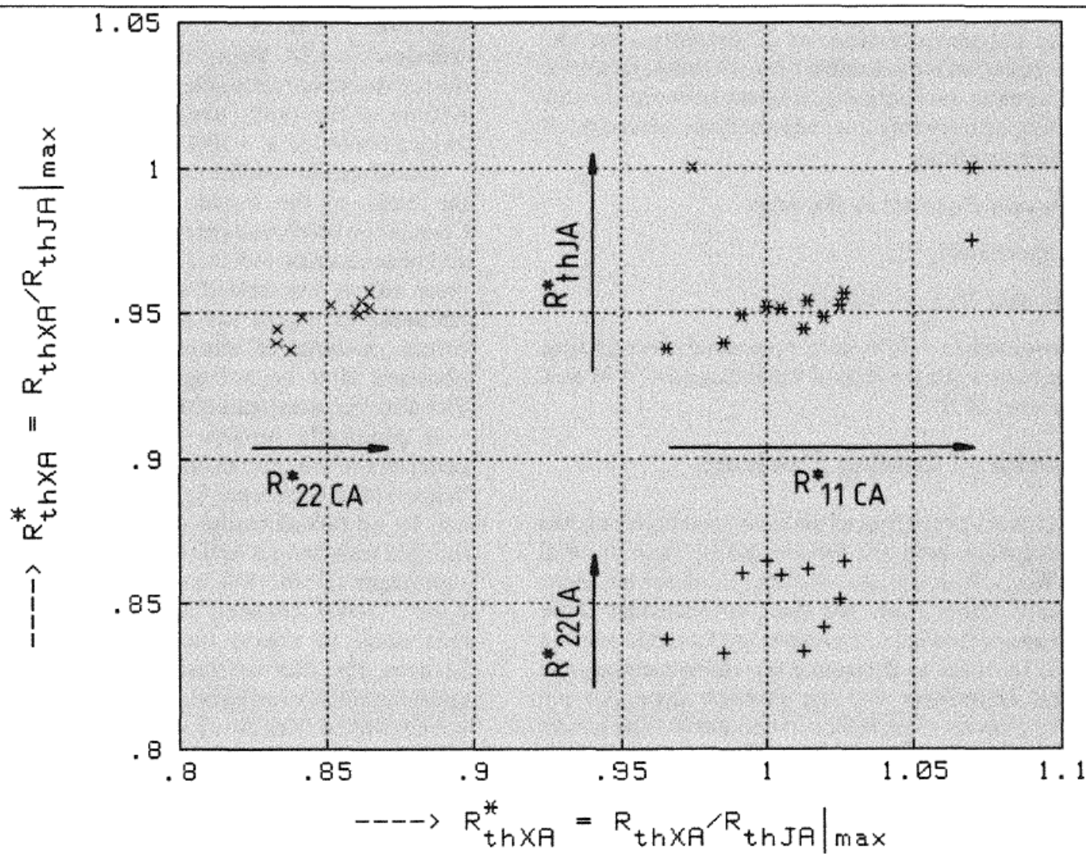


Figure 4 *Scatter diagram of the steady state thermal behavior of 11 AEG heat sinks (diam. 100 mm) in combination with thyristors T 2200 N 3600 ... 4400 (AEG) (diam. 100 mm).*

temperature at the junction of the thyristors. In steady state operation this temperature is influenced essentially by three parameters. On one hand the power loss in the semiconductor component and on the other hand the thermal resistance of the thyristor heat sink determines the temperature difference between the junction and the cooling medium. The temperature of the cooling medium must generally be regarded as a given quantity not subject to our influence. Neither the power loss nor the thermal resistances are constant quantities, but subject to considerable statistical variation. A knowledge of the statistical parameters of these quantities is a prerequisite for optimum dimensioning of inverters.

10 Behavior of Thyristors

The power loss resulting in a thyristor at a given operating current is determined primarily by its forward voltage drop. It was demonstrated in Reference 4 that the three-parameter logarithmic normal distribution is an appropriate model for the distribution of the forward voltage drop of thyristors. A further parameter influencing the operational behavior of inverters is the recovered charge of the thyristor. As seen in Figure 2, both these quantities show a high degree of negative correlation. Since neither characteristic has a normal distribution, Spearman's rank correlation coefficient is used to test for independence. According to Equation 7 the test quantity is t = −26.3 and, as expected from this figure, confirms the interdependence at a significance level of 1%:

$$|t| = 26.3 > t_{178;0.95} = 2.605$$

Figure 3 shows the scatter diagram of forward voltage drop-internal thermal resistance of 32 thyristors. As expected, no correlation is discernible between these quantities (Table 1). The rank correlation coefficient between U_T and R_{thJC} is already remarkably low; however, the criterion for testing of the hypothesis:

H_0 : *Q_{rr} and R_{thJC} are independent*

against the alternative

H_1 : *Q_{rr} and R_{thJC} are correlated*

which turns out to be t = 1.06, gives no grounds for rejecting H_0 in comparison with the critical value $t_{30;0.95}$ = 1.697 at a significance level of 10%.

11 Behavior of Cooling Elements

The steady-state properties of the water-cooled heat sinks under investigation here are determined by four thermal resistances R_{11CA}, R_{12CA}, R_{21CA}, and R_{22CA} and the behavior of the coolant circulation. We shall not consider these points in greater detail here; the reader is referred instead to Reference 5. In order to determine the characteristics, the four thermal resistances and the pressure drop Δp are measured for a given volume flow of coolant $\dot{V}$. The results obtained are shown with the resulting thermal resistance of the overall set-up R_{thJA} in Figure 4. Here an asterisk corresponds to a pair of values R_{thJA} − R_{11CA}, a plus sign to a pair of values R_{11CA} − R_{22CA}, and a multiplication sign to a pair of values R_{thJA} − R_{22CA}. The first point to notice is that three pairs of values which all belong to the same heat sink differ from the rest of the spectrum (the three values at top right). The Grubbs outlier test (Reference 1) confirms this at a significance level of $\alpha < 1\%$.

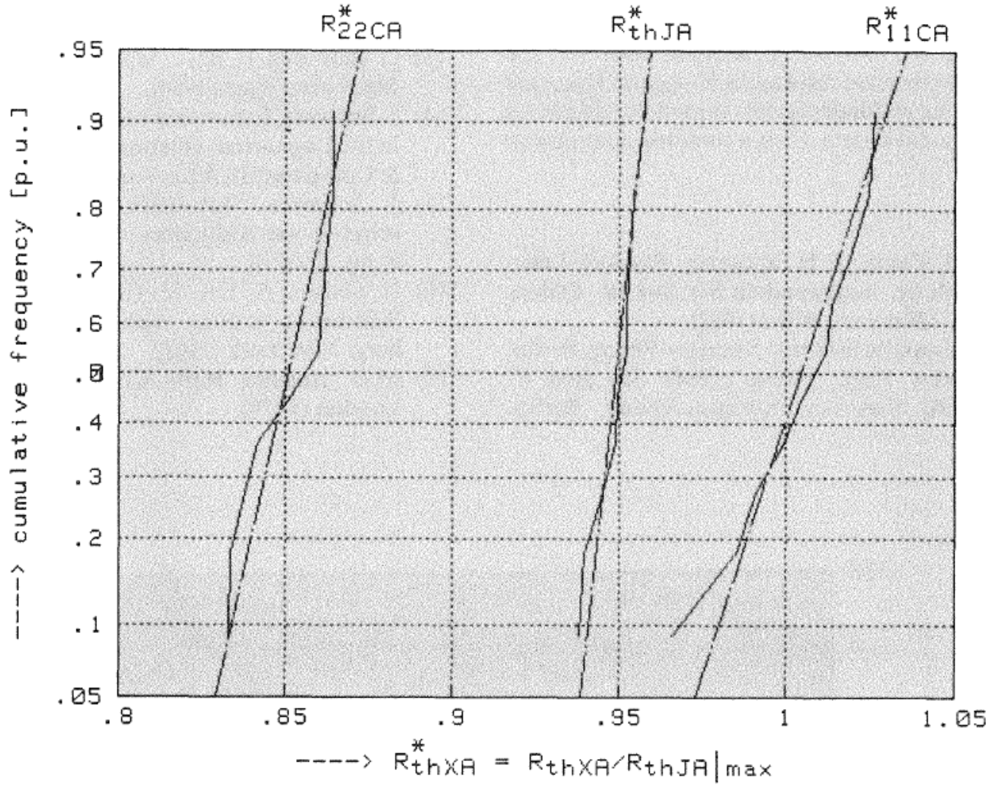
In the following treatment, the values for these heat sinks are left out of consideration. Figure 5 shows a graphical representation of the empirical distribution function of the thermal resistances. The null hypothesis that thermal resistances have a normal distribution with respect to the empirically determined parameters x and s is not rejected by the Lilliefors modification of the Kolmogorov fitness test at a level of $\alpha > 5\%$.

One cannot speak of a normal distribution in connection with the pressure drop. Figure 6 shows the density curve taken from Table 2 for the pressure drop of 363 heat sinks, which displays a negative skewness. It suggests that a test for a logarithmic normal distribution might not be pointless. However, the rough class assignment of Table 2 precludes a reasonable further statistical evaluation of these data. Further investigations were undertaken on the basis of more accuate pressure drop measurements, performed on only nine heat sinks. Figure 7 depicts the pressure drop distribution with empirical parameters of the three-parameters logarithmic normal distribution which were estimated by the maximum-likelihood method (Reference 4) and not rejected at a level of $\alpha > 20\%$ (Kolmogorov fitness test).

The decisive question to be answered here is whether the measured thermal parameters R_{11CA} and R_{22CA} of the heat sinks, which dominate the result, are mutually independent. Table 3 shows the various calculated correlation coefficients. Since there is no evidence against assumption of normal distribution of thermal resistances R_{11CA} and R_{22CA}, the Bravais-Pearson correlation coefficient can be applied. Equation 7 with $r* = 0.375$ and n = 10 affords the test criterion t = 1.14. The critical value is $t_{8;0.95}$ = 1.86. Hence the test does not reject $\rho(R_{11CA}; R_{22CA}) = 0$ at a level of 10%. Neither of the rank correlation coefficients is significant (critical values at $\alpha = 10\%$; $r_S = 0.4426$ and $K_{10;0.95} = 19$).

The situation is different when the heat sinks omitted on the basis of the outlier test are considered after all. Correlation coefficients of $r* = 0.796$, $r_{S,B} = 0.538$, $\tau = 0.382$, and hence a value of K = 21, are then obtained. At $\alpha = 10\%$ these values are indeed significant. Now it is generally advisable to regard the results of statistical tests with a certain amount of caution when different results are obtained after neglecting outliers. These results should therefore be investigated further.

A physically positive correlation is known to exist between the pressure drop and the heat transfer coefficient in heat sinks (Reference 6). If the pressure drop is increased, e.g. by an (accidentally) greater roughness, then a lower thermal resistance is to be expected. This is confirmd by the significant (at the 5% level) rank correlation coefficient $r_{S,B} = -0.647$ between R_{thJA} and Δp. It would appear reasonable to regard the suspected positive correlation between the thermal resistances R_{11CA} and R_{22CA} as a spurious correlation because both of the thermal resistances are correlated with Δp. Table 3 confirms this hypothesis in so far as not only R_{thJA} but also R_{22CA} is significantly correlated with Δp (this correlation is not so pronounced in the case of R_{11CA}) and if the partial rank correlation coefficients calculated between R_{11CA} and R_{22CA} with Δp, according to Kendall, then the significantly decreasing values are seen as an indication of the suspected spurious correlation.



Remark: probability grid of the normal distribution

Figure 5 *Representation of the distributions of the thermal resistances of 10 AEG heat sinks (diam. 100 mm) in combination with thyristors T 2200 N 3600 ... 4400 (AEG) (diam. 100 mm).*

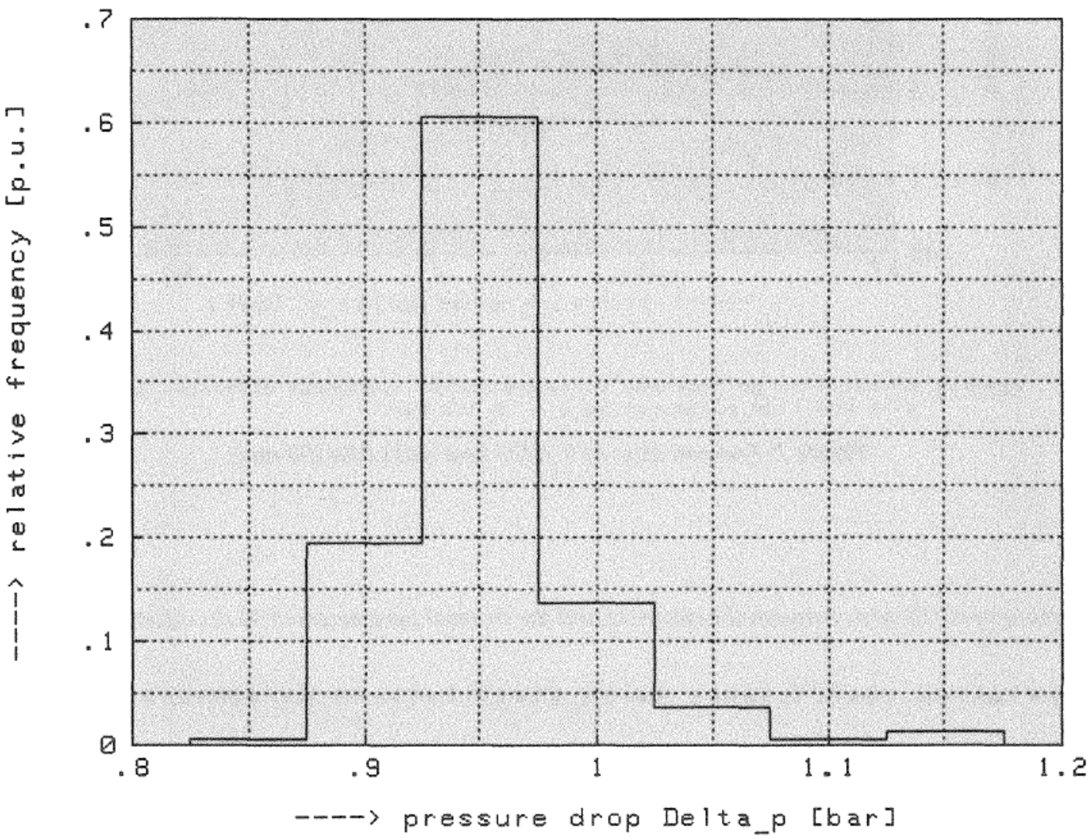


Figure 6 *Empirical frequency function of the pressure drop of 363 AEG heat sinks (diam. 100 mm).*

Unfortunately this sample is not very large, and very precise statements are therefore precluded. However, the hypothesis that the thermal resistances R_{11CA} and R_{22CA} are normally distributed and independent cannot be rejected on the basis of the numerical data. Only a common relationship via Δp exists.

References

(1) J. Hartung, B. Elpelt, K.H. Klösener, Statistik-Lehr- und Handbuch der angewandten Statistik. R. Oldenbourg Verlag, München, Wien (1982).

(2) L. Sachs, Angewandte Statistik. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo (1984). See also: L. Sachs, Applied Statistics. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo (1984).

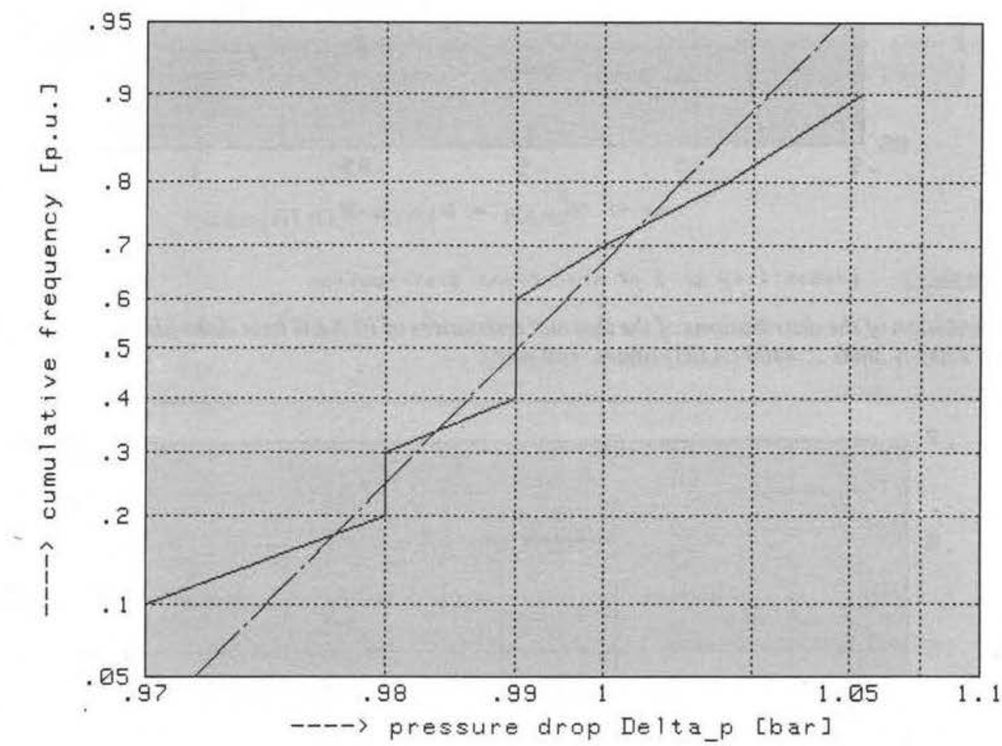
(3) G. Bamberg, F. Baur, Statistik. R. Oldenbourg Verlag, München, Wien (1984).

(4) J. Schwarz, Estimation of Parameters and Fitness Tests in the Lognormal Distribution. Intelligent Instruments & Computers (CAL), Vol. 4 (1986) No. 6, pp. 297-310.

(5) J. Schwarz, Stationäre thermodynamische Eigenschaften von Kühldosen, etz-Archiv, Vol. 8 (1986) No. 6, pp. 207-212.

(6) H. Gröber, S. Erk, U. Grigull, Die Grundgesetze der Wärmeübertragung. Springer-Verlag, Berlin, Heidelberg, New York (1981).

(7) M.G. Kendall, Rank Correlation Methods. Griffin, London (1970).



Remark: probability grid of the three-parameter lognormal distribution with position parameter Δp_0 = 0.962 bar.

Figure 7 *Pressure drop of 9 AEG heat sinks (dia 100 mm).*

Table 1 *Correlation coefficients between the electrical and the thermal parameters of 32 thyristors T 2200 N 3600 ... 4400 (AEG).*

```
Rank correlation coefficients (with ties) between parameters:

size of sample                                      n     =    32

correlation between U_T  and Q_rr              r_S,B = -0.785
correlation between U_T  and R_th_JC           r_S,B = -0.007
correlation between Q_rr and R_th_JC           r_S,B = +0.190
```

Table 2 *Measured values of the pressure drops of 363 AEG heat sinks.*

#	class boundaries				frequency		
	begin [bar]	mark [bar]	end [bar]	length [bar]	absolute [piece]	relative [%]	cumulative [%]
1	0.825	0.850	0.875	0.050	2	0.6	0.6
2	0.875	0.900	0.925	0.050	71	19.6	20.1
3	0.925	0.950	0.975	0.050	220	60.6	80.7
4	0.975	1.000	1.025	0.050	50	13.8	94.5
5	1.025	1.050	1.075	0.050	13	3.6	98.1
6	1.075	1.100	1.125	0.050	2	0.6	98.6
7	1.125				5	1.4	100.0

Table 3 *Correlation relations between the parameters of the AEG heat sinks.*

Relation between R_th_JA and Δ_p:

size of sample	n	= 9
Spearman's rank correlation coefficient with ties	r_S,B	= - 0.647
Kendall's rank correlation coefficient	τ	= - 0.556
test value for the significance of τ	K	= - 20.00

Relation between R_11_CA and R_22_CA:

size of sample	n	= 10
correlation coefficient by Bravais-Pearson	r	= 0.353
improved correlation coefficient by Bravais-Pearson	r^x	= 0.375
Spearman's rank correlation coefficient with ties	r_S,B	= 0.383
Kendall's rank correlation coefficient	τ	= 0.244
test value for the significance of τ	K	= 11.00

Relation between R_11_CA and Δ_p:

size of sample	n	= 9
Spearman's rank correlation coefficient with ties	r_S,B	= - 0.264
Kendall's rank correlation coefficient	τ	= - 0.333
test value for the significance of τ	K	= - 12.00

Relation between R_22_CA and Δ_p:

size of sample	n	= 9
Spearman's rank correlation coefficient with ties	r_S,B	= - 0.562
Kendall's rank correlation coefficient	τ	= - 0.500
test value for the significance of τ	K	= - 18.00

Relation between R_11_CA and R_22_CA under Δ_p:

size of sample	n	= 9
Kendall's rank correlation coefficient	τ	= 0.389
test value for the significance of τ	K	= 14.00
Kendall's partial rank correlation coefficient	τ	= 0.272

Listing 1 *Two versions of functional subprograms for calculation of the correlation coefficients according to Bravais-Pearson.*

```
10     DEF FNPearson(INTEGER N,REAL X(*),Y(*))
20     !
30     ! Functional subprogram for calculation of the
40     ! empirical correlation coefficients r according to Bravais-Pearson
50     !
60     ! (c) 1987 by Jürgen Schwarz        version/date:              1.1 / 07.04.87
70     ! filename:             englRh      memory:                cartridges 77/78
80     ! language:           HP-BASIC      computer: HP 9845 B with SP and AP ROM
90     !
100      INTEGER I
110      REAL X,Y,Xq,Yq,Xx,Yy,Xy
120      IF N<2 THEN RETURN 2                                     ! bad data check
130      REDIM X(1:N),Y(1:N)
140      !
150      Xq=Yq=Xx=Yy=Xy=0
160      FOR I=1 TO N
170        X=X(I)
180        Y=Y(I)
190        Xq=Xq+X
200        Yq=Yq+Y
210        Xx=Xx+X*X
220        Yy=Yy+Y*Y
230        Xy=Xy+X*Y
240      NEXT I                                               ! empirical corre-
250      RETURN (Xy-Xq*Yq/N)/SQR((Xx-Xq*Xq/N)*(Yy-Yq*Yq/N))   ! lation coefficient
260    FNEND
```

```
10     DEF FNBravais(INTEGER N,REAL X(*),Y(*))
20     !
30     ! Functional subprogram for calculation of the
40     ! empirical correlation coefficients r according to Bravais-Pearson
50     !
60     ! (c) 1987 by Jürgen Schwarz        version/date:              1.0 / 23.03.87
70     ! filename:             englRh      memory:                cartridges 77/78
80     ! language:           HP-BASIC      computer: HP 9845 B with SP and AP ROM
90     !
100      REAL Xq,Yq
110      IF N<2 THEN RETURN 2                                     ! bad data check
120      REDIM X(1:N),Y(1:N)
130      !
140      Xq=SUM(X)  ! SUM function returns the sum of all elements in an array
150      Yq=SUM(Y)  ! SUM function returns the sum of all elements in an array
160                 ! DOT function returns the inner (dot) product of two vectors
170      RETURN (DOT(X,Y)-Xq*Yq/N)/SQR((DOT(X,X)-Xq*Xq/N)*(DOT(Y,Y)-Yq*Yq/N))
180    FNEND
```

Listing 2 *Functional subprogram for calculation of the rank correlation coefficients according to Spearman.*

```
10     DEF FNSpearman(INTEGER N,Ties,REAL X(*),Y(*))
20     !
30     ! Functional subprogram for calculation of the
40     ! Spearman's rank correlation coefficient r_S without
50     ! and r_S,B with consideration of occured ties
60     !
70     ! Ties=0  ==> r_S    ==> without consideration of occured ties
80     ! Ties<>0 ==> r_S,B ==> correction of the estimation
90     !
100    ! (c) 1984 by Jürgen Schwarz       version/date:            1.0 / 25.06.84
110    ! filename:          englRh       memory:                cartridges 77/78
120    ! language:         HP-BASIC       computer: HP 9845 B with SP and AP ROM
130    !
140    ! Referenc: Lothar Sachs: "Applied Statistic."
150    !           Springer-Verlag, New York, Heidelberg, Berlin, Tokyo (1984).
160    !
170      INTEGER I,J,K
180      REAL T_x,T_y
190      REAL X_prime(1:N+1),Y_prime(1:N+1),D(1:N)
200      !
210      IF N<2 THEN RETURN 2                                   ! bad data check
220      REDIM X(1:N),Y(1:N),X_prime(1:N),Y_prime(1:N)
230      !
240      T_x=T_y=0
250      MAT X_prime=X
260      MAT Y_prime=Y
270      MAT SORT X_prime                             ! sorting the vector X_prime
280      MAT SORT Y_prime                             ! sorting the vector Y_prime
290      REDIM X_prime(1:N+1),Y_prime(1:N+1)
300      X_prime(N+1)=Y_prime(N+1)=9.99999999999E99
310      !
320      J=1
330      REPEAT
340        I=J
350        REPEAT
360          J=J+1
370        UNTIL X_prime(J)>X_prime(I)
380        FOR K=1 TO N
390          IF X(K)=X_prime(I) THEN D(K)=.5*(I+J-1)                  ! rank of X(K)
400        NEXT K
410        T_x=T_x+(J-I)*((J-I)*(J-I)-1)
420      UNTIL (I=N) OR (J>N)
430      !
440      J=1
450      REPEAT
460        I=J
470        REPEAT
480          J=J+1
490        UNTIL Y_prime(J)>Y_prime(I)
500        FOR K=1 TO N
510          IF Y(K)=Y_prime(I) THEN D(K)=D(K)-.5*(I+J-1)     ! difference in rank
520        NEXT K
530        T_y=T_y+(J-I)*((J-I)*(J-I)-1)
540      UNTIL (I=N) OR (J>N)
550      !
560      MAT D=D.D                              ! calculation from D(i)=D(i)*D(i)
570      IF Ties THEN RETURN 1-6*SUM(D)/(N*(N*N-1)-.5*(T_x+T_y))
580      IF NOT Ties THEN RETURN 1-6*SUM(D)/(N*(N*N-1))
590    SUBEND
```

Listing 3 *Functional subprogram for calculation of the rank correlation coefficients according to Kendall.*

```
10    DEF FNKendalls_tau(INTEGER N,REAL X(*),Y(*))
20    !
30    ! Functional subprogram for calculation of the
40    ! Kendall's correlation coefficient tau
50    !
60    ! (c) 1984 by Jürgen Schwarz        version/date:              1.0 / 19.07.84
70    ! filename:            englRh       memory:                  cartridges 77/78
80    ! language:          HP-BASIC       computer: HP 9845 B with SP and AP ROM
90    !
100   ! Referenc: Maurice G. Kendall:
110   !           "Rank Correlation Methods."
120   !           Charles Griffin & Company Limited, London (1970).
130   !
140     INTEGER I,J,K
150     REAL Q,X_prime(1:N),X_cross(1:N)
160     REAL Y_prime(1:N),Y_prime_rank(1:N),Y_rank(1:N)
170     !
180     IF N<2 THEN RETURN 2                                   ! bad data check
190     REDIM X(1:N),Y(1:N)
200     !
210     MAT Y_prime=Y
220     MAT SORT Y_prime                          ! sorting the vector Y_prime
230     J=0
240     REPEAT
250       I=J=J+1
260       LOOP
270       EXIT IF J=N
280       EXIT IF Y_prime(J+1)>Y_prime(J)
290         J=J+1
300       END LOOP
310       FOR K=I TO J
320         Y_prime_rank(K)=.5*(I+J)
330       NEXT K
340     UNTIL J=N
350     !
360     MAT X_prime=X
370     MAT SORT X_prime                          ! sorting the vector X_prime
380     MAT X_cross=X
390     FOR I=1 TO N
400       J=K=0
410       REPEAT
420         J=J+1
430       UNTIL X_cross(J)=X_prime(I)
440       REPEAT
450         K=K+1
460       UNTIL Y(J)=Y_prime(K)
470       Y_rank(I)=Y_prime_rank(K)
480       X_cross(J)=Y_prime(K)=9.99999999999E99
490     NEXT I
500     !
510     Q=0
520     FOR I=1 TO N-1
530       FOR J=I+1 TO N
540         IF Y_rank(J)<=Y_rank(I) THEN Q=Q+1
550       NEXT J
560     NEXT I
570     RETURN 1-4*Q/(N*(N-1))
580   SUBEND
```

■