# Chapter 4
# STATISTICS USED IN THE EXCERPTS

---



---

Text by Wim Kalmijn

Findings on happiness can be quantified in various ways. There are many statistical techniques, which apply in different situations and which have their strong and weak points in that various situations. The finding-excerpts specify the statistical techniques that have been used. Current summary statistics are coded and appear in the field 'statistics'. Further measures and methods are mentioned in the field 'remarks'.

This chapter starts with a short 'refresher' on some basic notions of descriptive statistics and statistical analysis (section 4/1). Next the three kind of statistical measures used in the excerpts are discussed. The statistics that are used for describing how happy people are in a particular population are dealt with in section 4/2. The statistics used for characterising/quantifying the relationship with other variables ('correlates') are enumerated in section 4/3. The statistics for testing hypotheses about associations (mostly the null hypothesis of no association at all in the population from which the sample has been drawn) are depicted in section 4/4. Finally a survey of the various statistics and statistical techniques that have been used in empirical studies on happiness is provided in section 4/5. This overview is alphabetically ordered and consists of shorthand descriptions of the statistics. The statistics used in the excerpts link to these descriptions.

# 4/1 SOME BASICS OF STATISTICAL ANALYSIS



## 4/1.1 Statistics

In the excerpts in the World Database of Happiness (further abbreviated as WDH) various statistics from the study concerned may be incorporated.

*What is a 'statistic'?*
Anything, which can be calculated from observed data in a sample, is called a ***statistic***. All studies included in the WDH are considered to be a description and analysis of a sample from a larger public and/or in time.

*Kinds of statistics.*
In this survey the various statistics are grouped in three categories:
- Statistics of distribution, Section 4/2
- Statistics of association, Section 4/3
- Statistics for testing a hypothesis, Section 4/4.

*Overview of statistics.*
The various statistics and statistical procedures are summarized in a standardized way in Section 4/5. These summaries can be reached from the abstracts of research findings using a hyperlink. From these summaries one can return to the general explanation in this text.

## 4/1.2 Level of measurements and types of interrelationship

This Catalogue of Correlational Findings contains information on the interrelationship between happiness and other characteristics of the same subjects, which characteristics (variables) are here labelled ***correlates***.
Correlates can be measured at one of three ***levels of measurement*:**

- The ***metric*** level, including both the interval and the ratio level, e.g. income
- The ***ordinal*** level, e.g. educational level
- The ***nominal*** level, e.g. nationality

Although happiness is measured at the ordinal level, the results are often treated as metric. If however, the correlate is measured at the ordinal level, the usual statistical approach is also to handle the happiness results as ordinal.

The existence of a statistical relationship between two variables is called ***correlation***. An other word for the same is ***association***. Both words are also used in a more specific sense and denote then a particular kind of statistical relationship. The wide and the narrow meanings of these words are placed in the scheme below.

**Types of statistical interrelation**
(Called correlation or association in the wide sense)

| *Measurement level of the correlate* | *Interrelationship type* |
|---|---|
| Metric | Correlation (in the narrow sense) |
| Ordinal | Rank correlation |
| Nominal | Association (in the narrower sense) |

Various measures have been proposed for the quantification of the degree of association. Generally speaking each combination of the above three levels of measurement has its own association statistics. The more frequently used ones are mentioned below.

If the correlate can adopt two values only, it is called "***dichotomous"***. Sometimes these two possible outcomes are coded with the values 0 and 1; therefore such a correlate is also referred to as a "binary variable" or as an "indicator variable". If e.g. the correlate is the sex of a subject (WDH-label: MALE SEX), the obvious standard coding in WDH is female = 0 and male =1. In the case of "Yes" or "No" responses, the usual code is "No" = 0 and "Yes" = 1.

For a dichotomous correlate the distinction nominal/ordinal/metric virtually vanishes. Hence for this subclass a wide variety of interrelationship measures are reported in the various studies. Often the interrelationship question is reduced to a straightforward two-sample difference problem with happiness as the dependent variable, measured at the metric level. For this case, various standard statistical solutions are available.
A special case is that of the ***double dichotomy*** (or 2x2-frequency table), where both happiness and the correlate have been measured at the dichotomous level (happiness as unhappy/happy).

NOTE 1: *The level of measurement* described above should be well distinguished from the *level of variation* of a correlate. If six different values of a (nominal or ordinal) correlate are involved in a study, it is said that the correlate is varied at six levels; the symbol c will be used for this number of levels; in this example c=6. In case of a dichotomy c=2.

NOTE 2: In this context an *observation* is defined as the set of all variables measured on a single subject. If the sample consists of 400 subjects, and the happiness and value of two correlates have been measured for each subject, there are 1200 data (measurements), but only 400 observations.

## 4/1.3 Sample size

The ***total sample size*** (symbol: $N_s$) is defined as the number of subjects that is requested to answer the questions raised by the investigator. However, not everybody will respond. If the ***number of respondents*** is denoted as N, the ***nonresponse percentage*** NR equals

$$100(N_s - N)/N_s \ \%.$$

Among the N respondents some will report, "don't know " or will give no answer at all. This category of responses is denoted ***DK/NA***; those respondents are ignored in the statistical analysis. This analysis is performed on the basis of the ***effective sample size***, which is denoted as $N_e$ and defined as N minus the number of "DK/NA" respondents. The ***DK/NA percentage***, as reported in the excerpts, is defined as $100(N - N_e)/N$ %.

## 4/1.4 Assumptions

All statistics are based on assumptions about the nature of data. These assumptions are often implicit. Most statistics used in this Catalogue of Correlational findings are bases on the assumption of *metric measurement* and/or the *normal distribution*.

### *Metric measurement*

The distinction between the various levels of measurement does not only apply to the correlates, it also applies but to that of happiness measurement.
Although theoretically happiness is always measured at the ordinal rating level of measurement, most studies use the rating scores as metric, i.e. at least the interval level of measurement: consecutive scale points are declared to be equidistant, which makes sense only for k>2.

### *Normal distribution*

Individual happiness ratings have a discrete and generally skewed frequency distribution. Nevertheless, the mean of larger numbers of ratings can be considered to have a ***normal probability distribution***, at least approximately. This approximation becomes better as the

number of individual ratings increases, which justifies the application of a number of statistical procedures based on the normality assumption, such as the construction of a confidence interval for the true but unknown mean (see 4/1.6.3 ).

## 4/1.5 Descriptive vs. inferential statistics

In many cases statistics of association are used only as a number, which is reported as a characteristic of the interrelationship. Such statistics may allow the reader to compare the strength of that interrelationship to that of others reported in similar studies. If statistics are used in this way only, they are referred to as ***descriptive statistics.***

There are, however, also situations in which statistics of association act as 'inferential statistics' or 'test statistics'; these terms are synonymous. A ***test statistic*** is usually defined as a statistic, that is not only descriptive, but on which also the decision whether or not to reject a given ***hypothesis*** can be based.

If a statistic is used as a test statistic, there is always an underlying ***statistical model*** assumed. This model concerns the distribution of the variables and/or their interrelationship. Such a model is characterized by both its structure and by one or more 'model parameters'. In its descriptive role, a test statistic acts as an ***estimator*** of one of the model parameters. The hypotheses mentioned above may concern either the structure or one of the model parameters.

Only very exceptionally are such "given" hypotheses formulated explicitly in the study to be excerpted. It can be argued, however, that statistical tests in such studies are often applied to test the null hypothesis (see 4/1.6 ), that some parameter of the underlying statistical model has a zero value. Three examples are: (a) correlation coefficients, (b) differences between means and (c) association test statistics in which the null hypothesis is tested that there is no association at all. WDH assumes that in such cases there are in fact given hypotheses, albeit implicitly, and that there are test statistics applied or at least applicable in these situations. There is, however, one fundamental problem. All null-hypotheses as described above concern statements, which, from a scientific, psychological/sociological, point of view, are already most unlikely, and their rejection does not really enlarge our knowledge. Such null hypotheses should always be rejected, but whether or not this really will happen is mainly a matter of sample size. Small samples will hardly ever, and sufficiently large samples will always result into the rejection of the null hypothesis. Hence the result of such a study is eventually controlled by the study budget and this may give rise to some criticism.

The use of a statistic as a test statistic implies that certain assumptions have to be made, which are not necessary as long as the statistic is considered as to be only descriptive. As an example: implicit significance testing of a correlation coefficient between two variables is nothing other than formulating a statement on the algebraic sign (+ or -) of the model parameter rho ($\rho$). Such a statement or conclusion is based on the value of the measured correlation coefficient r, in that case considered as an estimator of $\rho$, together with the (effective) sample size. The statistical test assumes that the joint distribution between the two variables is a bivariate normal one, at least approximately. If, however, one of the variables, e.g. the correlate, is varied at two levels only (say 0 and 1), there is no bivariate normal distribution, not even approximately, so in that case the usual significance tests for $\rho$

are highly debatable. In this situation, the measured correlation coefficient r can only be used as a descriptive statistic.
This example illustrates that the same statistic, in our case the measured correlation coefficient r, may be merely descriptive in one situation, but inferential in other situations. In the specific situation of a dichotomous (0/1) correlate, the (descriptive) measured coefficient of correlation is referred to as the "point biserial coefficient of correlation" (WDH code: rpb) rather than r; this is done to avoid confusion.

## 4/1.6 Significance testing

If two or more random samples are drawn from the same population, each of them will have its own composition. Generally the observations in the different samples are not identical, nor will be the statistics derived from them. In other words: statistics have *a* ***statistical (or probability) distribution*** (in this context also called "sampling distribution")**.** The parameters of this distribution depend on the (unknown) parameters of the distribution of the population. Therefore, sample statistics can be used as a basis of conclusions about the population, to be more precise, about the parameters of the model, which describes the distribution in the population. Significance testing is part of this process.

The procedure of significance testing can be summarized briefly as follows.
A ***'null hypothesis'*** can be formulated, e.g. that there is no interrelationship between happiness and the correlate; in other words, that the correlation in the population from which the sample was drawn is zero. "Under this hypothesis", i.e. if that hypothesis were true, the probability distribution of the relevant test statistic is known.

### 4/1.6.1 Probability distribution of the test statistic; p-value.

Some numerical values of this test statistic are then considered to be extreme, i.e. the probability that such a value or even a more extreme one is obtained, that is their 'tail probability' or ***p-value *,*** is small. If this probability is sufficiently small, say 5 % or less, this may be a reason to reject the null hypothesis in favour of the alternative, that there is some interrelationship with the correlate; such an extreme value of the test statistic is called ***significant***. The value can be significant at the 5 % level as above, but also at lower levels, e.g. 1 % or even 0.1 % if the value is very extreme. Instead of a 5 % significance level of the test statistic, sometimes the conclusion that there is some nonzero relationship is reported "with a confidence of at least 95 %". The ***confidence level*** equals 100 % minus the significance level.

If the test is performed as a one-sided test, the p-value is reported as a p1-value; a p2-value refers to a two-sided testing.

* In the excerpts p-values can be notated in two ways.

#### *Full p-value*

The first one is to report the ***p-value itself***, i.e. the probability of finding a value equal to the test statistic value as obtained or more extreme, provided the null hypothesis holds. This value can be reported in the standard notation and/or in the scientific or E-format, e.g. as $0.000\ 24 = 2.4 * 10^{-4}$, which is written as 2.4E-04. The latter notation is recommended for p-values < 0.001.
It is WDH's policy to report the p-values whenever possible, i.e. either reported by the author(s) as such or computable from the data in the publication.

#### *p-value classes*

The second notation makes a distinction into a few ***p-value classes***.

If one of the conclusions of the columns 1, 2 or 3 is reported by the author(s), it is included in the WDH according to the corresponding notation in the fourth column.

| | | | |
|---|---|---|---|
| not significant | $p > 0.05$ | NS | Ns |
| Significant | $0.01 < p \leq 0.05$ | * | 05 |
| Strongly significant | $0.001 < p \leq 0.01$ | ** | 01 |
| Very significant | $p \leq 0.001$ | *** | 00 |

## 4/1.6.2 Test methods

Some test statistics are specific ones: they are applied only in a specific test situation (e.g. Wilcoxon's two sample test statistic). Other test statistics have a much broader scope of application and are called "standard test statistics". The latter class will be dealt with in section 4/4.

### *Specific tests*

Sometimes the test statistic has - under the null hypothesis - a specific probability distribution, which is tabulated exactly, albeit only if the effective sample size $N_e$ is relatively small.. In such cases the WDH sometimes may report the test statistic as U, i.e. unclassified, and reports the test method as such, e.g. Fisher's exact test for 2x2 tables.

Tables are available for small values of $N_e$, with critical values of some test statistics, e.g. for

- Kendall's tau (a) test ( $N_e \leq 10$)
- Spearman's rank correlation test ($N_e \leq 30$)
- Wilcoxon (Mann-Whitney) two sample test ($N_e \leq 50$)

Moreover statistics also exist, which have critical points of their distribution tabulated for various parameters. An example is Duncan's (New) Multiple Range Test.

### *Standard test statistics*

Much more frequently the test statistic can be transformed into some other test statistic, which under the null hypothesis has a distribution that is exactly or at least approximately equal to the distribution of one of the '***standard test statistics***'. In the case of 'approximately equal', the agreement usually gets better as the sample size increases. These standard test statistics are:

- The standard normal distribution
- The chi-square distribution
- Student's t-distribution
- Snedecor's F-distribution
- The Studentized Range

They will be dealt with in Section 4/4 and the conversion of a test statistic into its appropriate standard test statistic in Section 4/5.

The difference between association statistics and standard test statistics can be illustrated for the case of a correlation coefficient. Establishing a correlation coefficient's significance is nothing other than concluding whether in the population the correlation coefficient (rho) is "significantly positive" if $r > 0$ (or "significantly negative" if $r < 0$). This test can be performed e.g. as a t-test; in this case t is the standard test statistic. This test uses the fact that the test statistic r can be transformed by

$$r \rightarrow \sqrt{((N_e - 2)/(1-r^2))},$$

which has approximately a t-distribution with Ne-2 degrees of freedom under the null hypothesis that there is a zero correlation.

### 4/1.6.3 Confidence intervals.

The use of '*confidence intervals*' for the true but unknown population parameter is much more informative than significance testing.

Consider e.g. the measurement of the mean happiness in some country at a given moment, i.e. within a short period. Let the sample mean have a value of say 7.24 on a scale [0 -10]. What is the value of this result?
In such situations, one can compute a lower and an upper 95 % confidence limit. Let the outcomes be 7.06 and 7.42 respectively. In this case the 95 % confidence interval for the true but unknown population mean value is reported as:

CI95 = [7.06; 7.42]

This means that any mean happiness value between 7.06 and 7.42 in the population is sufficiently well in agreement with the observed sample mean 7.24, and any value outside that interval is not. Among all 95 % confidence intervals reported in scientific literature, about 95% cover the true mean. Nobody knows which ones, only the long-run frequency at which it occurs is known approximately. In other words: the "95 % label" concerns always the ***construction procedure*** and never the ***individual result***.

The ***relationship with hypothesis testing*** is as follows. For any value v* within the confidence interval, the null hypothesis that the true but unknown population mean value is equal to v* ***cannot*** be rejected on the basis of the observations. If however v* is not covered by the interval, this null hypothesis can be rejected and this conclusion has a confidence of at least 95 %.

## 4/1.7 Statistical Procedures

This Catalogue of Correlational Findings on happiness sometimes reports statistical indications, which are neither a descriptive nor a test statistic, but a ***statistical analysis procedure***, such as:

- Analysis of Variance (ANOVA)
- Multivariate Analysis of Variance (MANOVA)
- Analysis of Covariance (ANCOVA)
- Loglinear Models (LLM)
- Multiple Comparisons procedure (MCP)
- Multiple Regression Analysis (MRA).

Such procedures will be mentioned in the 'remarks' field of the excerpt whenever applied, but they will. not be mentioned in the 'statistics' field.

# 4/2 STATISTICS OF DISTRIBUTION



Statistics of distribution concern one variable only and summarize the pattern of occurrence in a sample. In the excerpts these statistics are used for condensing responses to queries about happiness in a sample and for summarizing how the correlated variable is distributed in the sample.

## 4/2.1 Happiness measurement scales

A primary description of observed happiness in a sample includes at least:

- The ***scale of measurement***, especially the level of variation, expressed number of points (k) of the rating scale
- The ***observed frequencies*****,** either absolute or as a percentage, of each scale value of the scale applied in the study.

## 4/2.2 Measures (statistics) to characterize the distribution.

The results of the primary description may be condensed to one or more of the following secondary (derived) statistics:

### 4/2.2.1 Statistics for central tendency

Statistics for the central tendency include:

- The (arithmetic) ***mean happiness value*** (M), being the weighted average of the k transformed scale values and using the frequencies as weights. This statistic is defined only if the happiness scale of measurement is assumed to be a (pseudo-) metric one
- A ***95% confidence interval*** (CI95) for the true but unknown (transformed) mean happiness value in the public that is represented by the study sample. The calculation of these confidence limits requires knowledge of the mean value, the standard deviation (see below) and the effective sample size $N_e$. Their calculation is based on an approximation. If the distribution can be assumed to be normal and $N_e > 30$, for t the

value 2 can be adopted; in that case the error in the confidence interval width remains < 2 %.
- The median value (Me)
- The mode Mo, i.e. the rating with the highest occurring frequency.

### 4/2.2.2 Statistics for dispersion

Statistics for the dispersion of a distribution include
- The ***standard deviation*** (SD) as a measure of the dispersion of the individual happiness ratings about the mean value as defined above
- The lower (Q1) and the upper quartile (Q3) and the interquartile range (Q3-Q1),

The standard deviation is based on the assumption that the scale of measurement can be handled as metric.

### 4/2.2.3 Transformation of happiness scores to a standard 0-10 scale

To enable a comparison of the results from different studies using different scales (e.g. different k-values), all 'happiness' scores are subjected to a ***(linear) transformation onto a* [0;10] *scale***, where a higher degree of happiness corresponds with a higher value on the "transformed" scale. Such transformation is admissible only if the scale of measurement is (pseudo-) metric.

The general formula for this linear transformation is

$$y \rightarrow 10 * (y-U)/((H-U),$$

where H and U are the two end points of the original scale, i.e. the scale on which y has been measured, H corresponding to extreme happiness and U to extreme unhappiness respectively.
If on the original scale U = 1 and H = k, the transformed value of y equals 10*(y-1)/(k-1).

*Example*: the scale values of a five-point scale, e.g. {1, 2, 3, 4, 5}, are transformed into the values 0.0, 2.5, 5.0, 7.5 and 10.0 respectively.

*Thurstone* has introduced a second transformation method. In this approach each possible answer to a happiness question is presented to a panel of experts. Each panel member is asked to assign a score to each answer option on the [0;10] interval. If the individual expert scores do not differ too much, their average value is adopted as the transformed value of the original rating.
Scores obtained in this way are dependent on the exact formulation of the answers, including the language in which they are formulated. When available, such expert scores are entered into the database, together with the happiness query concerned. Thurstone transformations are used standard in the Catalogue of Happiness in Nations and are incidentally reported in this Catalogue of Correlational Findings.

## 4/2.3 Notation policy about statistics of distribution

It is the policy to include at least the following distribution statistics in the excerpts, provided they are available, either reported by the author or computable from the observational data as reported:

- The total sample size $N_s$
- The non-response percentage (NR)
- The number of scale points k, which is included implicitly already in the specification of the question(s) as raised in the study
- The distribution itself, expressed as the relative frequency (as a percentage) for each point of the original happiness scale, ignoring "don't know"/"no answer" responses

If the scale is used for the measurement of happiness, then also

- The average "happiness" value Mt (i.e. after transformation onto a [0; 10] scale);
- The standard deviation SDt (ditto)
- The 95 % confidence limits for the true, but unknown mean happiness value in the public represented by the sample of the study (CI95; ditto)

## 4/2.4 Notation Example

---

| | | |
|---|---|---|
| ***Study*** | FICTI 1998 | *Page in report*: 13 |
| *Population* | 18+ aged general public, Fantasyland, 1995 | |
| *Sample* | Probability multistage cluster sample, NR 45%, N = 1000 | |

| | |
|---|---|
| ***Query on happiness*** | Code: O-DT/u/sq/v/7/aa |
| *Type* | Self-report on single question |
| *Text* | "How do you feel about your life as a whole.....?" |
| | 7 delighted |
| | 6 pleased |
| | 5 mostly satisfied |
| | 4 mixed |
| | 3 mostly dissatisfied |
| | 2 unhappy |
| | 1 terrible |
| | DK/NA |

***Observed responses***

*Frequencies (valid)* **1**: 4 %, **2**: 8%, **3**: 12%, **4**: 15%, **5**:20 %, **6**: 31%, **7**: 9% (total 100%)

No answer, don't know 3.2%

| | *on original 1-7 scale* | *transformed to range 0-10* | |
|---|---|---|---|
| *Mean* | M = 4.70 | Mt = 6.16 | CI95 [6.04; 6.27] |
| *Standard Deviation* | SD = 1.61 | SDt = 2.7 | |

***Error estimates***

Repeat-correlation after one week: r = +.62 (follow up of 50 Ss)

***Remarks***

This question was preceded by several questions on satisfaction with life domains

# 4/3 STATISTICS OF ASSOCIATION



## 4/3.1 Association in bivariate situations.

***Bivariate situations*** are those where the interrelationship is investigated between a single correlate and a single happiness rating.
The statistical methods that are commonly used for such situations are depicted in section 4/3.1. This section opens with a schematic overview that provides links to subsections and to the shorthand description of the methods in section 4/5.

All other situations are referred to as '***multivariate'*** in this context and are dealt with in section. 4/3.2. This section also opens with a schematic overview with hyperlinks to the details

# Overview of statistics and procedures for bi-variate situations

| ***Level of measurement of the correlate*** | ***Level of measurement of the happiness response*** | | |
|---|---|---|---|
| | *Ordinal* | *Metric* | *Dichotomous* |
| *Metric* (§ 4/3.1.3.) | | Correlation coeff. (r)<br>Correlation ratio ($R^2$)<br>Regression coeff. (b)<br>Beta coefficient (β) | |
| *Ordinal* (§ 4/3.1.2.) | Spearman's rank corr. coeff.(rs)<br>Kendall's tau-a (ta)<br>Kendall's tau-b (tb)<br>Kendall's tau-c (tc)<br>Goodman/Kruskal's tau<br>Gamma (G)<br>Somers' D (Dyx) | | |
| *Nominal* (§ 4/3.1.1) | Chi-square ($\chi^2$)<br>Pearson's C<br>Cramér's V<br>Tschuprow's T | One-Way Analysis of Variance with multiple comparisons (AoV) (BMCT, DMRT, SNK)<br>Correlation ratio ($E^2$) | |
| *Dichotomous* (§4/3.1.1) | Difference in % (D%)<br>Difference modus (DMo) | Difference means (DM)<br>- transformed (DMt)<br>- standardized (DMs)<br>Critical ratio (CR; CR=DMs)<br>Cohen's d<br>Hedges's gH<br>Point biserial correlation (rpb).<br>Correlation ratio ($E^2$) | Fisher's 2x2 test<br>Odds ratio (OR)<br>Yule's Q<br>Yule's Y<br>Logit coefficient (lgt)<br>Gamma (G) |

### 4/3.1.1 Association with 'nominal correlates'

Bivariate findings are generally available as rxc ***cross-classifications***, c and r being the number of columns and rows respectively. In WDH it is standard to use the columns for the various levels of the correlate (the lowest level of the correlate in the left-hand column); if a correlate can adopt c distinct values, the correlate is said to be varied at c levels. The rows are used for the various happiness scale values with the lowest value on the bottom row, so for a k-point scale r = k.
If the correlate is measured at the ordinal level, the various levels of the correlate have a 'natural' order, which is also applied in the cross-tabulation from low (left) to high (right). Correlates measured at the nominal level, however, do not have such a natural order; it is recommended to order the various columns according to increasing mean happiness from left to right.

#### Statistics for association

Various association statistics are in use for such tabulations. The most frequently reported ones and their range of possible values are:

- The chi-square ($\chi^2$) statistic with (c-1)*(r-1) degrees of freedom
  $0 \leq \chi^2 \leq N_e(s-1)$, where $s = \min(c,r)$
- The phi coefficient ($\varphi$)
- Phi-square ($\varphi^2$) $0 \leq \varphi^2 \leq (s-1)$
- Pearson's Contingency Coefficient (C) $0 \leq C^2 \leq C \leq \sqrt{(1-1/s)} < 1$
- Cramér's V statistic
- Cramér's $V^2$ statistic $0 \leq V^2 \leq V \leq 1$
- Tschuprow's T- or $T^2$ -statistic $0 \leq T^2 \leq \sqrt{[[\min(c,r)-1]/[\max(c,r)-1]]} \leq 1$

All these statistics ***are equivalent***, i.e. each of them can be calculated from any other one. For comparison reasons it is attractive to select one of them as a primary association statistic to characterize the strength of association and to convert the other (secondary) ones to the primary elected. **Cramér's V-statistic** has the advantage that its range is always [0;1] and is a good candidate for it.
Since the transformation of chi-square to Cramér's V is monotonous, the p1-value of the chi-square test statistic also applies to the corresponding value of V.

It should be borne in mind that all above statistics do **not** assume that happiness is measured at the (pseudo)-metric level, in contrast to the next method.

The chi-square test may reject the hypothesis of complete independence, and demonstrate some dependency, but in this case, it does not specify what dependency exactly is present. A second consequence is even more serious. As has been pointed out in Section 4/1.4, the significance of these test statistics depends strongly on the sample size. If the sample is very small, one will never find a significant result, and if the sample is sufficiently large, a significant result will always be obtained. Therefore, significance tests of this class do not really contribute to scientific understanding.

#### Analysis of variance and multiple comparisons

A much more informative approach, which overcomes the problems raised in the preceding section, is to evaluate the data in a ***one-way analysis of variance*** with the correlate as a

variable at $c \geq 3$ levels.

In any Analysis of Variance (ANOVA), the total variability in the response variable y (in our case the happiness rating) is characterized by the total sum of squares (SST), which is equal to $\sum(y-\bar{y})^2$. The summation is made over all $N_e$ measurement results, $\bar{y}$ being the 'overall average value', i.e. the average of all $N_e$ y-values.
In the simplest form of an ANOVA (including the one-way ANOVA), SST can be written as the sum of two components. One of them (SSX) is that part of the total variability for which the correlate(s) are held accountable, where as the residual variance (SSE = SST - SSX) is assigned to all other influences on Y. In the Analysis of Variance, SST is partitioned in such a way that SSX is maximized and consequently SSE is minimized.
All variances are estimates and the precision of each of them is characterized by its number of degrees of freedom (df); those df are also additive.
Basically such an ANOVA table has the structure below:

| Source of variability | SS | df | MS := SS/df | F |
|---|---|---|---|---|
| Correlate(s) | SSX | c-1 | MSX = SSX/(c-1) | F = MSX/MSE |
| Residual ("error") | SSE | $N_e$-c | MSE = SSE/( $N_e$-c) | |
| Total | SST | $N_e$-1 | | |

The mean squares MS are obtained by dividing the sum of squares SS by the corresponding number of df. In this case, the value of c is the number of correlate classes if the correlate is measured at the nonmetric level; it is the number of regressors if a regression analysis is performed. The variance ratio F = MSX/MSE in this case is a statistic with c-1 and $N_e$-c df respectively.
A significantly large F-value indicates that there are one or more significant differences in happiness between the various levels of the correlate. In this case, the largest of the c average happiness values is systematically larger than the smallest one, but nothing is said about the ½c(c-1)-1 other differences.

Multiple inference, i.e. inferences about *all* differences between the various correlate levels, can be made by applying a ***multiple comparison procedure*** to the set of the mean happiness values of the c levels of the correlate as varied in this study. The conclusions are a set of statements, and a confidence of 95 % refers to the total 'family' of statements and not to each of the individual members.
Various multiple comparison methods are available, each having its own theoretical and practical pro's and cons. Three of them are:

- Bonferroni's method of the t-statistics
- Duncan's (New) Multiple Range method
- The Student-Newman-Keuls, also referred to as Newman-Keuls method;

Some alternatives are:

- The Dunnett-method
- Fisher's LSD-method (LSD = Least Significant Difference)
- The Scheffé-method
- The Tukey-method

It should be borne in mind, however, that this approach is admissible only if the assumption is justified that the happiness rating scale can be treated as (pseudo-) metric. As has already been pointed out, the chi-square related association statistics do not require this assumption.

## The correlation ratio.

One of the measures to express the strength of the statistical association between happiness and one or more correlates is the correlation ratio. Its application requires that the happiness as the 'dependent variable' is measured at the metric level of measurement.

Generally speaking, this strength is reflected in the extent to which knowledge of the value of the correlate reduces one's uncertainty about the happiness rating (Y). This uncertainty is characterized by $\sigma^2$ (the variance in Y). If the correlate value is fully unknown or ignored, this variance is denoted $\sigma^2(Y)$; if the value X of the correlate is known, the variance of Y is $\sigma^2(Y|X)$; read: Y, given X. If there is some association $\sigma^2(Y|X) < \sigma^2(Y)$. The *relative reduction* of the uncertainty is defined as $(\sigma^2(Y) - \sigma^2(Y|X))/(\sigma^2(Y))$ and is usually denoted as $\omega^2$ (omega-square).

For the estimation of this characteristic, at least three statistics are in use. Confusion arises from the fact that not all authors use the same symbols, where different authors use the same symbol for different statistics.
All estimators of $\omega^2$ are related to the F-value in the ANOVA situation as described above. The most frequently occurring one – which is also called the variance ratio – will be denoted here as $E^2$ and is defined as the ratio SSX/SST in the above ANOVA-model
This statistic, however, is a biased estimator of $\omega^2$ ; it is systematically too large. This is due to the fact that SSX is a measure of the variability between the mean happiness values at the c different levels of the correlate. Hence it does not only include the contribution of the correlate(s), but also some 'error variance' that is incorporated in the mean values.
An unbiased estimator of $\omega^2$ is given by the statistic $\varepsilon^2$ (epsilon-square, but incidentally written as $\eta^2$, eta-square), which is defined as $\varepsilon^2 = 1 - MSE/MST$, where
$MST = SST/(N_e-1)$ and all other symbols defined as in the above ANOVA-model.
In the WDH, this statistic will be written as EPS.
A third statistic is also referred to as *omega-square* and denoted here as $w^2$ . It is defined through $\varepsilon^2$ by the relationship $\varepsilon^2 = w^2 (1 + MSE/SST)$. Clearly $w^2$ is also a biased estimator of $\omega^2$, since $w^2 < \varepsilon^2$ and $\varepsilon^2$ is unbiased.
By converting them into the appropriate F-value , all these three statistics can be used for testing the null hypothesis of no association at all.

## Notation policy

It is the policy of the WDH to include at least the following information, provided it is either available or computable from the data reported in the excerpted publication or provided additionally by the author(s):

- The **complete cross-classification** with absolute, when necessary approximate, frequencies, with the various happiness levels as rows (starting with the highest level) and the various levels (/values) of the correlate as columns
- The (transformed) **means, their standard errors and 95 % confidence intervals** for the true but unknown mean happiness value **in each sub-population**
- A **multiple comparison test**, applied to the means of the various correlate levels and performed at a 95 % overall confidence level
- The **correlation ratio** $E^2$.
- (Only if the means and their standard errors are *not* available) the value of the **chi-square test** statistic for independence together with **its number of degrees of freedom**(df) and its **p1-value.** The chi-square statistic should be replaced with Cramér's

V by the appropriate transformation:

$$\chi^2 \rightarrow V = \sqrt{(\chi^2 / (N_e(s-1)))},$$

s being the lesser of the number of columns and the number of rows of the cross tabulation

Other statistics will be stored only if they are provided by the author(s), irrespective whether they are applicable or not. This may include one of the association statistics mentioned above and related to the chi-square test for independence.

## Notation example

***Study*** FICTI 1998 *Page* 2
*Population* 18+ aged, general public, Fantasyland, 1995
*Sample* Probability multistage cluster sample, NR = 45%, $N_s$ = 1000

***Measured correlate***
*Correlate class:* Source of current income *Correlate code*: I 1.4.2

| | |
|---|---|
| *Measurement:*<br>Single question:<br>'What is your main source of income?'<br>A work<br>B capital rents<br>C pension, social security<br>D support by spouse or parents | *Observed responses:*<br>A: 26%, B: 18 %, C: 18%, D: 38% |
| | *Error estimates:*<br>Apparent inconsistency with response to question on occupation: 5% |
| | *Remarks:* |

***Observed association with happiness***

| *Happiness query* | *Statistics* | *Elaboration/remarks* |
|---|---|---|
| O-DT/c/sq/v/7/aa | F = 4,71<br>P1 = 0.003<br>$E^2$ = 0.025 | Statistics computed from cross table<br>M Mt CI95<br>A: 5.1. 6.8 [6.46; 7.22]<br>B: 5.1 6.8 [6.48; 7.20]<br>C: 4.6 6.1 [5.62; 6.49]<br>D: 4.7 6.1 [5.81; 6.47]<br>Multiple comparison test: (C,D)<(B,A) |

Cross table

| | A | B | C | D |
|---|---|---|---|---|
| 7 | 19 | 3 | 10 | 8 |
| 6 | 50 | 26 | 63 | 23 |
| 5 | 31 | 31 | 55 | 35 |
| 4 | 26 | 24 | 36 | 18 |
| 3 | 12 | 9 | 23 | 4 |
| 2 | 5 | 5 | 14 | 1 |
| 1 | 2 | 3 | 6 | 1 |
| SUM | 145 | 101 | 207 | 97 |

### 4/3.1.2 Association statistics for 'ordinal correlates'

If the correlate is measured at the ordinal level (e.g. the educational level), it is common practice to handle happiness also as ordinal (rather than metric). Of course, this approach is fully correct.

#### *Rank correlation statistics*

Several methods are available for the analysis of a situation with two ordinal variables. The most widely reported ***rank correlation statistics*** are:

- Spearman's rank correlation coefficient rs — WDH-notation: rs
- Kendall's rank correlation coefficient $\tau_a$ (tau-a) — WDH-notation: ta
- Kendall's rank correlation coefficient $\tau_b$ (tau-b) — WDH-notation: tb
- Kendall's rank correlation coefficient $\tau_c$ (tau-c) — WDH-notation: tc
- Goodman and Kruskal's $\tau$ (tau) — WDH notation: tau
- Goodman and Kruskal's $\gamma$ (Gamma) — WDH-notation: G
- Somers' D-statistics. — WDH-notation: Dxy

All these statistics can adopt values in the interval [-1; +1], but often only a part of that interval is really available.

Spearman's rank correlation coefficient is applicable when a small set of N objects can be ranked as 1 to N according to two different characteristics. In this way each object gets two rankings, one for each characteristic. Spearman's $r_s$ is the correlation coefficient between both rankings.
Unambiguous ranking is not always feasible. If e.g. it is not possible to decide for one characteristic which object has ranking 3 and which has 4, the two different objects have to be ranked ex aequo and both objects are given the average ranking (3.5) for the specific characteristic. The two objects are called "tied" for that characteristic. The occurrence of too many such "***ties***" can be a problem for the application of Spearman's test.

If the number of objects increases, inevitably a limited number of 'classes' will be created and a class will be assigned to each object, where all objects within the same class are tied by definition.

The other methods given above have a common underlying approach.
If there are N objects (or subjects), there are ½N (N-1) ***pairs*** of objects. Such pairs are called "***concordant"***, if the object with the higher ranking for one characteristic also has the higher ranking for the other characteristic. If the object with the higher ranking for one characteristic has the lower ranking for the other characteristic, the pair is called "***discordant"***.
Pairs can be concordant, discordant or tied for one of the characteristics, but also for both at the same time.

Kendall's tau-a method is applicable to the same situations where Spearman's method is used.
The ranks, however, are processed in a different way. ***Tau-a*** is defined as the difference between the number (C) of concordant pairs and the number (D) of discordant pairs, divided by the total number of pairs ½N (N-1), so including also all pairs with ties.
Kendall's ***tau-b*** makes a correction in the denominator for the numbers of ties, but it can attain unity only if the numbers of possible different rankings for both characteristics are equal (in other words: c = r) and all observations lie on the main diagonal.
In ***tau-c***, the difference C-D is divided by its maximum attainable value, which equals ½N²(1-1/s), where s = the lesser of the number (r) of rows and the number (c) of columns.
Hence $\tau_c = 2s(C-D)/(N^2(s-1))$.
Tau-c is mainly of use for 'rectangular' contingency tables i.e. when $c \neq r$.

Goodman & Kruskal's gamma is defined almost identically to Kendall's tau-a, but in the denominator all pairs with ties are ignored, so $\gamma = (C-D)/(C+D)$. Gamma is the only rank correlation statistic for which the complete interval [-1; +1] is really available under all conditions.

The relevant Somers' D-statistic for the study of correlation with happiness is $D_{yx} = (C-D)/(C+D+T_C)$, where $T_C$ = the number of pairs with equal 'happiness' ratings and unequal correlate rankings at the same time. The WDH-notation is Dyx.

The underlying approach of Goodman & Kruskal's tau is different. It belongs to the ***proportional reduction in error*** (PRE) statistics class and it is an indicator of the degree at which knowledge of the value of the correlate improves the prediction quality of the happiness rating in terms of the expected percentage of correct predictions.

The various association statistics described in this section are ***not equi*valent**. They are indicators for different approaches of the association and cannot be converted into other ones in a unique way.

## Computation policy

For cross-classification tables - if available in the original publication or if the author supplies them afterwards - it is useful to select and compute a common test statistic for comparison reasons. In that case, tau-**c** is attractive, since (a) it has a finite range (-1; +1)

where both boundary values are attainable, and (b) it does not disregard any information on tied ranks.

## Notation policy

It is policy to include at least the following information in the excerpts, provided it is either available or computable from the data reported in the excerpted publication, or provided additionally by the author(s).

- The ***complete cross classification*** with absolute, when necessary approximate, frequencies, with the various happiness levels as rows (starting with the highest level) and the various levels (/values) of the correlate as columns (with the left hand column corresponding with lowest-ranked correlate level);
- The value of ***tau-c*** and its ***p2-value*;**
- The (transformed) ***means*, *their standard errors* *and* *95 % confidence intervals*** for the true but unknown mean happiness value ***in each sub-population*,** if there are different ones.

Other statistics will be stored only if they are provided by the author(s), irrespective of whether they are applicable or not. This may include one of the rank correlation statistics mentioned above and their statistical significance, expressed as the relevant p-value.

## Notation example

---

***Study*** FICTI 1998 *Page* 2
*Population* 18+ aged, general public, Fantasyland, 1995
*Sample* Probability multistage cluster sample, NR = 30 %, N = 1000

***Measured correlate***
*Correlate class:* Relative income *correlate code*: I 1.6.3

| | |
|---|---|
| *Measurement:*<br>Single question:<br>'How do you rate your household income relative to compatriots?<br>4 high<br>3 upper middle<br>2 lower middle<br>1 low | *Observed responses:*<br>1: 26 %, 2: 29 %, 3: 26 %, 4: 19%. |
| | *Error estimates:*<br>Consistency with similar question: r =+.80 |
| | *Remarks:*<br>This item was preceded by questions on sources of income |

***Observed association with happiness***

| *Happiness query* | *Statistics* | *Elaboration/remarks* |
|---|---|---|
| O-DT/c/sq/v/7/aa | DMt = +<br><br><br>Tauc = + .072<br>P2 = .014 | Statistics computed from cross table<br>M Mt CI95<br>4: 5.31 7.19 [6.86;7.51]<br>3: 5.48 7.47 [7.15;7.78]<br>2: 4.97 6.61 [6.26;6.97]<br>1: 5.14 6.89 [6.57;7.22] |

Cross table

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 7 | 17 | 20 | 26 | 12 |
| 6 | 69 | 75 | 92 | 58 |
| 5 | 54 | 49 | 35 | 39 |
| 4 | 22 | 23 | 11 | 14 |
| 3 | 10 | 15 | 8 | 8 |
| 2 | 8 | 11 | 6 | 2 |
| 1 | 3 | 9 | 3 | 1 |
| SUM | 183 | 202 | 181 | 134 |

## 4/3.1.3 Association statistics and statistical procedures for 'metric correlates'.

### *Linear regression and correlation analysis*

If the correlate is measured at the metric level and the happiness is considered as also metric, the most widely applied statistical method is (linear) ***regression analysis***.
In this approach, variation in the correlate is held responsible for, at least a part of, the variability in the happiness measurements. Therefore, the correlate is often referred to as the ***explanatory variable*** or as the ***independent variable***, whereas 'happiness' is called then the ***dependent variable***. This common terminology falsely suggests a causal relationship, while this method addresses only statistical interrelation. In a more neutral way the correlate is sometimes referred to as the ***x-variable*** or as the ***regressor***, and the happiness as the ***y-variable*** or as the ***regressand***.

On the basis of the regression analysis the following statistics can be computed:

- The ***regression coefficient*** (b)
- The ***standardized regression coefficient*** (beta) also referred to as the ***beta-coefficient*** or the ***beta weight***; this statistic is the regression coefficient, which is obtained when the regression technique is applied to the 'standardized variables'.
  Variables are ***standardized*** by subtracting their mean value, followed by division by their standard deviation.
  Standardization of variables is useful since in this way variables are obtained which are no longer dependent on the scale at which they have been measured. If e.g. the annual income has been measured in NLG, this also applies to the mean and the standard deviation. By standardization, a variable is obtained which is not an amount of money, but 'only' a number, which is independent of the currency in which the measurement has been made at that time. The same holds for the beta coefficients. This procedure allows one to compare directly the results in countries with different currencies (or in the same country before and after January 1st 2002). In more general terms: standardization of variables makes them insensitive to linear transformations.
- Note that the *variables* are standardized and *not* the regression coefficient, as is suggested by its confusing name.
- The ***correlation coefficient*** (r, its full name is "Pearson's product moment correlation coefficient") and its square. In the case of one correlate only, r is equal to the beta coefficient .

### *Notation policy*

In the case of a single metric correlate, it is policy to incorporate the following results of the regression analysis in the excerpts:

- The ***regression equation*** $Y = a + bx$, where
  Y = the 'fitted' value of the happiness rating after transformation onto a [0; 10] scale,
  x = value of the correlate, and
  b = the regression coefficient;
- A 95 % ***confidence interval (CI95) for the true but unknown value of the regression coefficient.*** These confidence limits can be computed as $b \pm t.\ SE\ (b)$, where t is the t-value with N-2 df and a p2-value of 0.05, while SE (b) is the standard error of b; if $N \geq 30$, for t the value 2 can be adopted as an acceptable approximation
- The ***correlation ratio*** $R^2$ (which in the case of one correlate equals $r^2$), being the fraction

of the variation in happiness for which the variation of that correlate can be made accountable. Note that this $R^2$ is the same statistic as $E^2$ used in the case of nonmetric correlates;

- (Only if the joint distribution of the correlate and the happiness can be considered more or less as bivariate normal) the ***95% confidence interval (CI95) for the true but unknown value of rho*** of that distribution**.** These confidence limits can be calculated e.g. by using Fisher's z-transformation or by transformation of r into a t-value and vice versa.
- Otherwise the value of the ***correlation coefficient*** r can be reported as a descriptive statistic.

## *Notation example*

---

**Study** FICTI 1993 Page 4
*Population* 18+ aged, general public, Fantasyland, 1993
*Sample* Probability multistage cluster sample, NR = 45%, N = 1000

***Measured correlate***
Correlate class: Household income Correlate code: I 1.1.2

| | |
|---|---|
| *Measurement:*<br>Single question:<br>'What is the monthly income of your family? (in FAD) | *Observed responses:*<br>Mean: FAD 550, SD: FAD 50 |
| | *Error estimates:*<br>Consistency with similar question: r =+.80 |
| | *Remarks:*<br>This item was preceded by questions on sources of income |

***Observed association with happiness***

| *Happiness query* | *Statistics* | *Elaboration/remarks* |
|---|---|---|
| O-DT/c/sq/v/7/aa | r = +.45<br>$r^2$ = .18 | 0-10 happiness = 4.3 +.0045* income<br>CI95 for b [0.0041;0.0050]<br>CI95 for rho [0.38;0.51] |

## 4/3.1.4 Association with dichotomous correlates.

If the correlate is varied at two levels only; the correlate is said to be measured at ***the dichotomous level***.

### *Association statistics and methods*

Association statistics for a dichotomous correlate include:

- The difference in means (DM), i.e. the difference between the average happiness ratings at the two correlate levels
- The difference in transformed means (DMt), the same as DM, but after transformation onto a [0;10] scale
- The standardized difference in means (DMs), i.e. the difference in means, divided by its standard error. This statistic is insensitive to any linear transformation (e.g. onto a [0;10] scale) of the individual ratings and of their means. Incidentally, the statistic is also referred to as the "critical ratio" (WDH notation: CR)
- Cohen's d-statistic, which is equal to Hedges's g-statistic, but for a factor $\sqrt{(N_e /( N_e -2))}$, where $N_e$ is the effective number of subjects in the sample
- The point biserial coefficient of correlation (rpb), which is obtained as the Pearson product moment coefficient of correlation between happiness and correlate, where the correlate is given the values 0 and 1 respectively.

All the above statistics are based on the assumption that happiness is measured at a metric level. There are also some statistics which do not require this assumption, and for which it would be sufficient that the happiness is assessed at the ordinal level. Some of those statistics are:

- The difference in percentages (D%). If the % of the subjects with the same happiness rating on the same scale (e.g. the rating 5 on the [0;10] scale) are 38 and 26 respectively at the two correlate levels, then *for that specific rating* D% = 38 - 26 = 12%
- The difference in Modus (DMo), the difference between the most frequently occurring rating at the two correlate levels.

Within this class of studies, one subset deserves special attention, the 2x2 contingency tables (cross-classifications). Here happiness is also reported as a dichotomy e.g. happy/unhappy. A special test (Fisher's exact 2x2 test) and a set of special association statistics have been devised for this situation on the basis of the 'odds ratio' (OR), which is defined as $n_{11}$ x$n_{22}$ /$n_{12}$x$n_{21}$, where the n's are the four cell frequencies. The value of OR will adopt one of its most extreme values (0 and infinite) whenever (at least) one of the cell frequencies has a zero value. In this case at (at least) one correlate level a perfect prediction of the (binary) happiness rating is possible.
The odds ratio can be converted into Yule's Q or into the logit coefficient (lgt). Yule's Q as a measure of association is defined as (OR-1)/(OR+1); instead also his 'measure of colligation' Y may be reported, which is defined as $(\sqrt{OR} - 1)/(\sqrt{OR} + 1)$. The logit coefficient is defined as the natural logarithm of OR.

In a 2x2 cross table, both happiness and the correlate can be considered as ordinal variables, albeit sometimes slightly artificially. The association statistics for this situation (see section 4/3.1.3) can also be applied. One of them, Goodman & Kruskal's GAMMA, is then identical to Yule's Q.

In a 2x2 contingency table c = r = s = 2. As a consequence $0 \leq \varphi^2 = T^2 = V^2 = \chi^2/N_e \leq 1$. These relationships may be useful for converting other association statistics into Cramér's V.

## *Notation policy*

In this situation, it is policy to include at least the following information in the excerpts;, . at least when this information is available, either reported as such or computable from the data in the study::

- The (transformed) ***means, their standard errors and 95% confidence intervals*** for the true but unknown mean happiness value in both sub-publics
- ***Hedges's gH* statistic** as a so-called 'Effect Size Indicator' for the 'happiness' difference between both correlate levels
- A 95 % ***confidence interval*** for the true, but unknown difference of the two (transformed) means
- (Only if the happiness is measured at the pseudo-metric level) the ***correlation ratio*** $E^2$, i.e. the fraction (or proportion) of the happiness variability for which the difference in the correlate value is accountable.

Other statistics will be included only if they are reported by the author(s), irrespective if they are applicable or not.

In case of a ***double dichotomy***, either the p1-value from the (exact) Fisher distribution will be included (if $N_e < 30$), or the value of the chi-square statistic with 1 df, after application of Yates' correction for continuity. Other statistics reported by the author(s) will also be included.

## Notation example

---

***Study*** FICTI 1993 Page 5
*Population* 18+ aged, general public, Fantasyland, 1995
*Sample* Probability multistage cluster sample, NR = 45%, N = 1000

***Measured correlate***
Correlate class: own income or not Correlate code**: I 1.4.1**

| | |
|---|---|
| *Measurement:*<br>Single question:<br>'Do you have an income of your own?'<br>1 No<br>2 Yes | *Observed responses:*<br>1: 20%; 2: 80% |
| | *Error estimates:*<br>Apparent inconsistency with question on occupation: 2% |
| | *Remarks:*<br>This item was preceded by questions on sources of income |

***Observed association with happiness***

| *Happiness query* | *Statistics* | *Elaboration/remarks* |
|---|---|---|
| O-DT/c/sq/v/7/aa | D% = +<br><br>$Chi^2$ = 13.3<br>P1 = 2.7E-04 | % happy (computed from cross table)<br>no 13<br>yes 71 |

Cross table

| Own income ?<br>Happiness | No | Yes | Sum |
|---|---|---|---|
| Happy | 103 | 343 | 446 |
| Unhappy | 7 | 97 | 104 |
| Sum | 110 | 440 | 550 |

## 4/3.2 Association in multivariate situations

In this context, multivariate situations are those where the simultaneous interrelationship is investigated between:

- Two or more correlates and a single happiness rating as response
- A single correlate and two or more happiness ratings as response

- Two or more correlates and two or more happiness ratings as response.

In a 'simultaneous' method of analysis, the two or more happiness measures are not analyzed separately, but they are treated as *a set* of correlated response variables. Their correlation, due to the fact that the various happiness measures have been measured at the same subject, is also taken into account in the statistical analysis. The same holds for two or more correlates.

This category includes a wide class of situations, since not only the numbers of correlates (denoted here as m; m>1). may be different, but also the level of measurement of each of them. Hence, the overview and description in the next sections is in not exhaustive. However, in practice the vast majority of all the multivariate cases are covered by a limited number of cases, and these will be discussed briefly in this section albeit in general terms.

NOTE 1: The number of correlates (m) should not be confused with the number of levels of each of them $\{c_1, c_2, ....., c_m\}$.

NOTE 2: In the statistical literature the term "*multiple*" is generally used if more than one x-variable is involved, whereas the term "*multivariate*" – in the narrower sense - is reserved for the situation of more than one response (y-) variable. In the catalogue of correlates, "multiple" would refer to more correlates and "multivariate" to more than one happiness measure

### 4/3.2.1 Overview of statistics and procedures for multivariate situations

.

| | *Statistical procedure* | *Statistics* |
|---|---|---|
| *Two or more correlates, all-metric*<br><br>*Single happiness response* | Multiple Regression Analysis | Multiple correlation coefficient (R)<br>Coefficient of determination ($R^2$)<br>Adjusted coefficient of determination ($R_a^2$)<br>Partial regression coefficient (b)<br>Standardized partial regression coeff. (β)<br>Partial correlation coefficients (rpc) |
| *At least one correlate metric and at least one nominal*<br><br>*Single happiness response* | Analysis of Covariance (ANCOVA) | Difference in adjusted means (DMa)<br>Coefficient of determination ($R^2$)<br>Regression coefficient (b)<br>Standardized regression coefficient (β) |
| *Two or more nominal correlates* | One-Way Analysis of Variance (ANOVA) | F-test and multiple comparisons. |

*to be continued*

| | | |
|---|---|---|
| *Single happiness response* | Loglinear models | |
| | Contingency Tables | Chi-square |
| *Two or more happiness responses*. | Multivariate Analysis of Variance (MANOVA) | Hotelling's $T^2$ |

### 4/3.2.2 Statistical procedures in multivariate situations

A relatively simple subclass is that where all correlates are measured at the ***nominal level***. If m correlates are involved in the study, the obvious approach is to apply *an* ***m-way Analysis of Variance***. This procedure enables one not only to estimate the main effects influence of each correlate separately, but also their interactions, if they exist.
*Example*: if one correlate is the sex of a subject (male/female) and the other one is the country where the individual lives, one may raise the question whether the difference of happiness between males and females depends on the country in which it has been investigated. If so, an interaction (in the statistical sense of the word) is said to be present.

Application of a chi-square test to a cross tabulation may only reveal *that* there is some association without specifying which one. If the sample size is sufficiently large, it will always reveal the existence of some association. Therefore, it is hardly meaningful to report this uninformative test, certainly not in situations where an alternative is available.
If the number of happiness scale points k in a contingency table is small (say $k \leq 4$), the application of a ***Loglinear Model*** may be considered as an alternative, which does not require the assumption of a (pseudo-) metric level of happiness measurement.

If all correlates are measured at the ***metric level***, ***multiple (linear) regression*** **analysis** is the standard approach. Each of the correlates acts as a regressor and has its own partial ***regression coefficient.*** These can be either ***non-standardized*** or ***standardized***; the standardized ones are also referred to as "***beta coefficients"***.
Usually the various correlates are correlated among themselves. In this case, the calculation of ***partial correlation coefficients*** will be appropriate. A partial correlation between happiness and one of the correlates is the correlation, which remains after accounting for the variation of the influence of the other ones (or some of them). Sometimes this is expressed in shorthand as "the correlation between A and B, controlling for C and D". However, the term "controlled" is confusing in that sense that in observational studies the variables are usually not controllable at all. Similarly, it should be borne in mind, that the interpretation of partial correlation coefficients as correlation coefficients is limited to situations in which the values of other correlates can be kept constant, which is normally not the case. All partial correlation coefficients can be calculated from all 'ordinary' correlation coefficients, including those between the correlates.
Incidentally reference is made to a technique called **"*stepwise regression*".** In this procedure, one identifies the minimum size subset of regressors, which is sufficient to describe the happiness variation, whereas the other regressors can be considered to be superfluous, since they do not make a real additional contribution to 'explain' the variation in

the happiness score. This is achieved by successively entering and/or removing the various regressors into/from the regression equation in turn.
Moreover, it is meaningful to compute thec ***coefficient of multiple correlation*** of the happiness (R). The square of this quantity is called the ***coefficient of multiple determination*** $R^2$ and it is even more informative, as it has the same meaning as the correlation ratio $E^2$ in the case of correlates measured at the nonmetric level of measurement.

If the number of observations (N) is relatively small and that of the independent variables (the number of correlates m) is large, sometimes the ***adjusted coefficient of multiple determination*** ($R_a^2$) is reported, which is equal to $1-(1-R^2)(N_e-1)/(N_e-m)$. Generally speaking, this is only meaningful for the studies incorporated in the WDH concerning country means rather than individual subjects. A suitable rule of thumb is that this adjustment makes sense only if $N_e < 100m$.

In several investigations a number of correlates is involved. Sometimes these correlates are strongly interrelated and can be combined to a subset, e.g. socio-demographic variables. Such studies enable to discriminate between relative strong and weak classes of influences on happiness. In this Catalogue, the findings are coded S15 for "Summed effects" (combination of correlates).

Occasionally, the correlates have been subjected to a factor analysis, followed by a multiple regression analysis. In this case, the factors act as regressors and they are statistically independent by definition. Consequently the squared value of each correlation coefficient is the proportion of the total happiness variability for which this factor is accountable.
If the regression analysis is applied to the correlates as such, the recommended practice is to compute all first order partial correlation coefficients, i.e. the correlation between happiness and the relevant correlate *after accounting for the influence of all other correlates*. Their squared values are ranked from large to small, and enable the researcher to judge their relative importance for explaining happiness variability.
The joint effect of all correlates in the study is characterized by the coefficient of determination $R^2$ or its adjusted value $R_a^2$. Its statistical significance is characterized by an F-test, but it is much more relevant as an indicator of *scientific significance,* which should be judged on the basis of its numerical value, being the proportion of the happiness variability that is 'explained' by all correlates jointly.

If some correlates have been measured at the **nominal** and others at the ***metric*** level, an ***Analysis of Covariance*** (ANCOVA) may be appropriate. The most frequently occurring case is that with m = 2, where one correlate is measured at the nominal level and the other one at the metric level; this second correlate is referred to as a ***covariable***, The means at the various levels of the first correlate are compared after having taken into account the influence of the covariable; these means are called ***adjusted means*** (DMa). Adjusted means have no meaning in themselves, only their mutual differences are relevant results of the study.

If happiness is measured by multiple indicators, e.g. queries on mood and contentment the obvious statistical approach is the ***Multivariate Analysis of Variance (MANOVA)***. Testing for statistical significance can be done by using Hotelling's $T^2$-statistic, which is the multivariate equivalent of the F-test in the univariate (ANOVA) situation.

## 4/3.2.3 Notation policy

The most frequently occurring case is the situation in which all correlates are metric. Let m = the number of regressors, generally one for each correlate, so m>1. In this case, it is to incorporate at least the following results of the multiple regression analysis in the excerpts:

- The ***regression equation*** $Y = a + b_1x_1 + b_2x_2 + b_3x_3 + \ldots\ldots\ldots$, where
  Y = the 'fitted' happiness rating after transformation onto an [0; 10] scale
  $x_1$ = value of the first correlate
  $x_2$ = value of the second correlate,
  $x_3$ = value of the third correlate, etc.
  $b_1$, $b_2$, $b_3$, …= the (partial) regression coefficients;
- The ***coefficient of multiple determination*** $R^2$, being the fraction of the variation in happiness for which the variation of the correlate can be made accountable (the correlation ratio).

Optional statistics are:

- The various standardized (partial) regression coefficients, also called 'beta coefficients'
- Partial correlation coefficients which are judged to be informative
- (In case of moderate values of N) the adjusted coefficient of determination $R^2_a$.

Moreover all other statistics reported by the author(s) will be included.

## 4/3.2.4 Notation examples

---

***Study*** FICTI 1993 *Page* 4
*Population* 18+ aged, general public, Fantasyland, 1993
*Sample* Probability multistage cluster sample, NR = 45%, N = 1000

***Measured correlate***
*Correlate class:* Household income *Correlate code*: I 1.1.2

| | |
|---|---|
| *Measurement:*<br>Single question:<br>'What is the monthly income of your family? (in FAD)" | *Observed responses:*<br>Mean: FAD 550, SD: FAD 50 |
| | *Error estimates:*<br>Consistency with similar question: r =+.80 |
| | *Remarks:*<br>This item was preceded by questions on sources of income |

***Observed Association with Happiness***

| Happiness query | Statistics | Elaboration remarks |
|---|---|---|
| O-DT/c/sq/v/7/aa | r = +.45<br><br>beta = +.25<br><br>$R^2$ = . 05 | 0-10 happiness = 4.3 + 0.0045 income in FAD<br>CI95 for b [0.0041;0.0050]<br><br>Beta after accounting for:<br>a) Age<br>b) Marital status<br>c) Occupational level |

---

***Study*** FICTI 1996 *Page* 5
*Population* 18+ aged, general public, Fantasyland, 1995
*Sample* Probability multistage cluster sample, NR = 45%, N = 1000

***Measured correlate***
*Correlate class:* Joint effects *Correlate code:* S 15.2

| | |
|---|---|
| *Measurement:*<br>Single questions:<br>'What is the monthly income of your family? (in FAD)<br><br>Are you married now? ("no" = 0; "yes" =1)<br><br>"What is your age ?" | *Observed responses:*<br>Mean: FAD 550, SD: FAD 50 |
| | *Error estimates:*<br>Consistency with similar question: r |
| | *Remarks:*<br>This item was preceded by questions on sources of income |

***Observed association with happiness***

| *Happiness query* | *Statistics* | *Elaboration/remarks* |
|---|---|---|
| O-DT/c/sq/v/7/a | r = +.15<br><br>beta = +.25<br><br>$R^2$ = . 05 | 0-10 happiness = 4.3 + 0.0045 income in FAD<br>CI95 for b [0.0041;0.0050]<br><br>Beta after accounting for:<br>a) Age<br>b) Marital status |

# 4/4 STANDARD TEST STATISTICS



In the sections 4/1.5 and 4/3 association statistics have been introduced , which could be either descriptive or inferential.
There is a second class of statistics, which is referred to here as *'standard test statistics'*. These are inferential statistics as well (allow to reject or not a statistical hypothesis), but they have a wider area of application than inference about association only. Many association statistics can be transformed to one of these 'standard test statistics.' Since the probability distribution of the latter has been tabulated extensively in almost every textbook on Inferential Statistics, they enable to test association hypotheses in an indirect way. This very important property justifies their description in this context.

Basically two types of test statistics can be distinguished*:* ***symmetric*** (4/4.2) and ***asymmetric*** (4/4.3) test statistics.
All but one standard test statistics have at least one parameter, which is called "***the number of degrees of freedom"*** (further abbreviated as df).

## 4/4.1 Degrees of freedom.

For the chi-square test statistic, as applied to a contingency table, i.e. a cross frequency table with c columns and r rows, the number of degrees of freedom has to do with the number of cells. Df is equal to the number of cells that can be given a free choice for the cell frequency, given the marginal total frequencies of all rows and columns, as long as the cell frequency exceeds neither the column nor the row sum of that cell.
Oviously in that case df = (c-1)(r-1).

In the ***t-test***, the denominator is a standard deviation, or, to be more precise, an estimator of it. The number of df is an ***indicator of the quality of that estimation***: the more df, the more precise the estimator is. If the standard deviation is estimated from a sample with size n, the number of df = n-1. This will be clear if one realizes that the standard deviation is computed from n differences, each of them between one of the observed values and the mean of all n values together. There are n differences, but if n-1 differences are given, the remainder is also known, since the n differences sum up to zero. Actually the n observations are a set of n pieces of information, which is transformed into a new set of n elements: one element for the average value, and the other n-1 for the characterization of the dispersion in the set of

observed data. If the standard deviation is exactly known, the number of df is infinite.
In the ***F-test***, the statistic is the ratio of two variance estimators, each with the corresponding number of degrees of freedom, so we have two df-values: one for the numerator (variance) and one for the denominator (variance).

## 4/4.2 Symmetric test statistics.

Symmetric standard test statistics have a probability distribution, which is symmetric about zero. This subclass includes:

- The ***standard normal test statistic*** (in this context "standard" means: zero mean and unit variance), which is sometimes referred to as the z-statistic and its distribution as N (0,1);
- ***Student's t-statistic*** with $\nu$ degrees of freedom.

As a rule, a ***p2-value*** is appropriate, since implicitly testing in such cases should always be two-sided. Only if the hypotheses are described explicitly in the study excerpted and if those hypotheses have a basis in the relevant theory, might one-sided testing be an admissible option and is a p1-vaqlue relevant.

## 4/4.3 Asymmetric test statistics.

The following test statistics are asymmetric, since their probability distribution covers the range [0; $\infty$).This subclass includes:

- The ***chi-square test statistic*** with $\nu$ degrees of freedom,
- ***Snedecor's F-statistic*** with $\nu_1$ and $\nu_2$ degrees of freedom,
- The ***Studentized Range statistic*** with parameters n and $\nu$.

The latter test statistic plays a role in some multiple comparison procedures and is defined as the ratio W/s, where

- W = the range (the difference of the largest and the smallest value) of a sample of size n from a normally distributed population;
- s = an estimator (with $\nu$ df) of the population standard deviation $\sigma$, and
- W and s are statistically independent.

These asymmetric test statistics can adopt nonnegative values only and their application in significance testing is normally right-hand one-sided, hence the use of a p1-value is obvious.

## 4/4.4 Relationships between standard test statistics

**Four relationships** exist between the various standard test statistics:

- For $\nu = \infty$: $t(\nu) = z$
- For $\nu = 1$: $\chi^2(\nu) = z^2$
- For $\nu_1 = 1$: $F(\nu_1, \nu_2) = t^2(\nu_2)$
- For $\nu_2 = \infty$: $F(\nu_1, \nu_2) = \chi^2(\nu_1)/\nu_1$

## 4/4.5 Transformation of test statistics into standard test statistics.

Many statistics can be transformed into some other one, which for larger N (read $N_e$ ) is distributed exactly or approximately as one of the above test statistics. As an example: the coefficient of determination $R^2$ can be transformed into an other statistic, which has an F-distribution.
In section 4/5, this will be described in for the various association test statistics.
An overview of the standard statistics that are applied for testing significance is given below.

*Z-scores*

- Goodman & Kruskal's gamma
- Kendall's rank correlation coefficient Tau-A
- Kendall's rank correlation coefficient Tau-B
- Kendall's rank correlation coefficient Tau-C
- Partial correlation coefficient
- Product-moment correlation coefficient
- Somers' asymmetric test statistic
- Spearman's rank correlation coefficient
- Standardized regression coefficient
- Wilcoxon (Mann-Whitney) two sample test

*Chi-square*

- Chi-square statistic for goodness of fit
- Cramér's V
- Pearson's contingency coefficient C
- Tschuprow's T

*Student's t*

- Standardized difference of means
- Regression coefficient (non-standardized)

*Snedecor's F*

- Coefficient of determination
- Correlation ratio
- Epsilon square

# 4/5 DESCRIPTION OF STATISTICS USED IN EXCERPTS

## 4/5.1 Most used statistical techniques

| | |
|---|---|
| $R_a$ | ADJUSTED COEFFICIENT OF MULTIPLE CORRELATION |
| AoC | ANALYSIS of COVARIANCE (ANCOVA) |
| AoV | ANALYSIS of VARIANCE (ANOVA) |
| Mt | AVERAGE HAPPINESS VALUE AFTER TRANSFORMATION |
| SDt | AVERAGE STANDARD DEVIATION IN HAPPINESS AFTER TRANSFORMATION |
| BMCT | BONFERRONI's MULTIPLE COMPARISON TEST |
| $\chi^2$ | CHI-SQUARE STATISTIC (I) |
| $Chi^2$ | CHI-SQUARE STATISTIC (II) |
| $R^2$ | COEFFICIENT of DETERMINATION |
| R | COEFFICIENT OF MULTIPLE CORRELATION |
| d | Cohen's COEFFICIENT FOR GROUP DIFFERENCES |
| C | CONTINGENCY COEFFICIENT<br>See: Pearson's CONTINGENCY COEFFICIENT (C) |
| r | CORRELATION COEFFICIENT<br>See: PRODUCT-MOMENT CORRELATION COEFFICIENT |
| $E^2$ | CORRELATION RATIO |
| V | Cramér's V |
| CR | CRITICAL RATIO<br>See: STANDARDIZED DIFFERENCE of MEANS |
| OR | CROSS-PRODUCT RATIO<br>See: ODDS RATIO |
| DMr | DIFFERENCE in MEAN RIDITS |
| D% | DIFFERENCE in PERCENTAGES |

DM DIFFERENCE of MEANS

DMt DIFFERENCE of MEANS AFTER TRANSFORMATION

DMRT DUNCAN's MULTIPLE RANGE TEST

EPS EPSILON SQUARE

F F STATISTIC

Fisher's EXACT 2x2 TEST

G GOODMAN & Kruskal's GAMMA.

Tau GOODMAN & Kruskal's TAU

$h^2$ see: $E^2$ (CORRELATION RATIO)

gH Hedges's COEFFICIENT FOR GROUP DIFFERENCES

ta KENDALL'S RANK CORRELATION COEFFICIENT TAU-A

tb KENDALL'S RANK CORRELATION COEFFICIENT TAU-B

tc KENDALL'S TAU-C

lgt LOGIT COEFFICIENT

MANN-WHITNEY TWO-SAMPLE TEST
See: WILCOXON (MANN-WHITNEY) TWO-SAMPLE TEST

MAoV MULTIVARIATE ANALYSIS of VARIANCE (MANOVA).

Mt MEAN HAPPINESS VALUE AFTER TRANSFORMATION
See: AVERAGE HAPPINESS VALUE AFTER TRANSFORMATION

MULTIPLE CORRELATION COEFFICIENT
See: COEFFICIENT OF MULTIPLE CORRELATION

NEWMAN-KEULS
See: STUDENT-NEWMAN-KEULS

OR ODDS RATIO

$w^2$ OMEGA-SQUARE

rpc PARTIAL CORRELATION COEFFICIENT

| | |
|---|---|
| B | PARTIAL REGRESSION COEFFICIENT (non-standardized).<br>See: REGRESSION COEFFICIENT (non-standardized) |
| C | Pearson's CONTINGENCY COEFFICIENT C |
| r | Pearson's (PRODUCT MOMENT) CORRELATION COEFFICIENT<br>See: PRODUCT-MOMENT CORRELATION COEFFICIENT |
| rpb | POINT BISERIAL COEFFICIENT of CORRELATION |
| r | PRODUCT-MOMENT CORRELATION COEFFICIENT |
| b | REGRESSION COEFFICIENT (non-standardized) |
| F | SNEDECOR'S F STATISTIC.<br>See: F STATISTIC |
| Dyx | SOMERS' ASYMMETRIC TEST STATISTIC |
| rs | SPEARMAN'S RANK CORRELATION COEFFICIENT |
| DMs | STANDARDIZED DIFFERENCE of MEANS |
| Beta | STANDARDIZED REGRESSION COEFFICIENTS |
| | STEPWISE MULTIPLE REGRESSION ANALYSIS |
| tc | STUART'S TAU-C<br>See: KENDALL'S TAU-C |
| SNK | STUDENT-NEWMAN-KEULS |
| t | STUDENT'S t-STATISTIC<br>See: t-STATISTIC |
| SDtt | THURSTONE TRANSFORMATED HAPPINESS STANDARD DEVIATION |
| Mtt | THURSTONE TRANSFORMATED MEAN HAPPINESS |
| TT | THURSTONE TRANSFORMATION. |
| t | t STATISTIC (often referred to as Student's t-statistic) |
| T | TSCHUPROW'S T |
| MW | WILCOXON (MANN-WHITNEY) TWO-SAMPLE TEST |
| Y | Yule's COEFFICIENT of COLLIGATION (Y-STATISTIC) |
| Q | Yule's Q-STATISTIC |

# 4/5.1 Current statistical techniques in alphabetic order

Ranges of a statistic are usually specified by their boundaries; in that case, they are denoted with [ or ] if the relevant boundary is included in the range (a so-called "closed interval") and with ( or ) if it is not, i.e. in case of an "open interval ".
In both 'Correlate level' and 'Happiness level', the term 'level' refers to 'level of measurement'.

**$R_a$** **ADJUSTED COEFFICIENT OF MULTIPLE CORRELATION** scheme 4/3.2

Type: descriptive statistic only
WDH Symbol: $R_a$
Correlates level: all metric
Happiness level: metric

Range: $0 \leq R^2_a \leq R_a < 1$.
Meaning:
$R_a = 0 \leftrightarrow$ not any association
$R_a \rightarrow 1 \leftrightarrow$ strongest possible association
Recommended conversion: $R_a \rightarrow R^2_a$, the adjusted coefficient of determination

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 229.

**AoC** **ANALYSIS of COVARIANCE (ANCOVA)** scheme 4/3.1

Type: statistical procedure
WDH-symbol: AoC
Correlates level: at least one nominal and at least one metric.
Happiness level: metric.

Just as in an ANOVA, in an ANCOVA the total happiness variability, expressed as the sum of squares, is split into several parts, each of which is assigned to a source of variability. At least two of those sources are the variability of the correlates, in case there is one for each correlate, and always one other is the residual variability, which includes all unspecified influences on the happiness variable. Each sum of squares has its own number of degrees of freedom (df), which sum up to $N_e - 1$ for the total variability. If a

sum of squares (SS) is divided by its own number of df, a **mean square** (MS) is obtained. The ratio of two correctly selected mean squares has an F-distribution under the hypothesis that the corresponding association has a zero-value.

In an Analysis of Covariance, the treatment means for all levels of the nominal correlate are 'adjusted' for differences in the mean values of the metric correlate.

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 685-718.

**AoV** **ANALYSIS of VARIANCE (ANOVA)** scheme 4/3.1

Type: statistical procedure
WDH-symbol: AoV
Correlate(s) level: nominal
Happiness level: metric.

In an ANOVA, the total happiness variability, expressed as the sum of squares, is split into two or more parts, each of which is assigned to a source of variability. At least one of those sources is the variability of the correlate, in case there is only one, and always one other is the residual variability, which includes all unspecified influences on the happiness variable. Each sum of squares has its own number of degrees of freedom (df), which sum up to $N_e$ -1 for the total variability. If a sum of squares (SS) is divided by the corresponding number of df, a **mean square** (MS) is obtained. The ratio of two correctly selected mean squares has an F-distribution under the hypothesis that the corresponding association has a zero-value.

NOTE: A significantly high F-value only indicates that, in case of a single correlate, the largest of the c mean values is systematically larger than the smallest one. Conclusions about the other pairs of means require the application of a Multiple Comparisons Procedure (see e.g. BONFERRONI's MULTIPLE COMPARISON TEST, DUNCAN's MULTIPLE RANGE TEST or STUDENT-NEWMAN-KEULS)

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 436 sq.

**Mt** **AVERAGE HAPPINESS VALUE AFTER TRANSFORMATION**

Type: distribution statistic
WDH-symbol: Mt
Happiness level: metric.
Theoretical range: [0; 10]

This statistic is the average (or mean) value of all happiness scores in a sample after linear transformation onto a [0; 10] scale.

Meaning:
Mt = 0 ↔ extremely unhappy/unsatisfied etc..
Mt = 10 ↔ extremely happy/satisfied etc.

Recommendation:
From Mt and the corresponding standard deviation SDt in the same sample, one can

calculate the 95 % confidence limits (CI95) for the true but unknown mean happiness in the population, which is represented by that sample.
For $N_e > 30$ the lower confidence limit is (approximately) $Mt - 2*SDt/\sqrt{N_e}$ and the upper one is $Mt + 2*SDt/\sqrt{N_e}$. For smaller $N_e$, the value 2 has to be replaced with the appropriate value of Student's t with $N_e$-1 degrees of freedom , whereas $N_e$ has to be replaced with $N_e$-1.

**BETA** See: STANDARDIZED REGRESSION COEFFICIENTS

**BMCT** **BONFERRONI's MULTIPLE COMPARISON TEST** scheme 4/3.1

Type: statistical procedure
WDH Symbol: BMCT
Correlate level: nominal
Happiness level: metric

Meaning: if the correlate is measured at c levels, the c mean happiness values can be ranked from low to high. A multiple comparison procedure judges for each of the ½c(c-1) pairs whether or not they differ significantly. A convenient way to represent the results is by ranking the c means and by underlining them in such a way that means which have a common underlining do **NOT** differ significantly.

Test: Bonferroni's method uses a Student's t-test for each pair at an adjusted confidence level. In this procedure, the confidence of the total package of statements is at least 95 %.

Ref: Miller, R.G.; *Simultaneous Statistical Inference,* McGraw-Hill Book Company, (1966), New York, USA, 67-70.

## $\chi^2$ CHI-SQUARE STATISTIC (I)

Type: asymmetric standard test statistic
WDH Symbol $\chi^2$
One parameter: $\nu$ (= number of degrees of freedom); range $\nu$: [1; + $\infty$ )

Range: $0 \leq\leq \chi^2 < \infty$
Distribution: the statistic has a (skew) probability distribution with a mean value = $\nu$ and a variance = $2\nu$.
The critical values are tabulated extensively in almost any textbook on Statistics, and in

Ref: Pearson, E.S.; Hartley, H.O. ; *Biometrika Tables for Statisticians vol I*; Cambridge University Press; 1966[8]; Cambridge, UK; 9, 128.

NOTE: The standard statistic is applied widely in judging whether or not a frequency distribution is in agreement with a given statistical model, a 'goodness-of-fit-test'; see also: CHI-SQUARE STATISTIC (II).

## $Chi^2$ CHI-SQUARE STATISTIC (II)

scheme 4/3.1

Type: test statistic
WDH Symbol: $Chi^2$
Correlate level: nominal
Happiness level: ordinal

Range: $0 \leq \chi^2 \leq N_e$ *[min(c,r) – 1], where c and r are the numbers of columns and rows respectively in a cross tabulation.
Meaning:
$Chi^2$ = (c-1)*(r-1) $\leftrightarrow$ no association
$Chi^2$ >> (c-1)*(r-1) $\leftrightarrow$ strong association
Recommended conversion: transformation to Cramér's V = $\sqrt{(Chi^2/(N_e(s-1)))}$, where s = the smaller of c and r, the number of columns and rows respectively, and $N_e$ = effective total number of observations.

The test statistic is computed as the sum over all cells of a contingency table of the value of $(o-e)^2/e$, where o = observed frequency and e= expected frequency in the same cell under the assumption of complete independence of happiness and correlate. For a 2x2-contingency table, Yates, correction is recommended, which means that (o-e) is replaced with the value of ABS(o-e)-0.5.
Test: Under the null hypothesis (no association at all), the test statistic approximately has a chi-square distribution with (c-1)(r-1) df, where c and r are the numbers of columns and rows respectively of the contingency table.

Ref: Dixon, W.J.; Massey, F.J. ; *Introduction to Statistical Analysis*, McGraw Hill Inc., 1969[3], New York, US; 242.

## $R^2$ COEFFICIENT of DETERMINATION

scheme 4/3.2

Type: test statistic
WDH Symbol: $R^2$

Correlates level: all metric
Happiness level: metric

Range: [0; 1]
Meaning:
$R^2 = 0 \leftrightarrow$ no influence of any correlate in this study has been established.
$R^2 = 1 \leftrightarrow$ the correlates determine the happiness completely.
Recommended conversion: none.

Test: $R^2$ can be transformed into an F standard test statistic with m-1 and $N_e$ -m df, where $N_e$ = effective number of observations and m = number of regressors, usually the number of correlates. Transformation $R^2 \rightarrow F = (N_e\text{-}m)R^2 / ((m\text{-}1)(1\text{-}R^2))$

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 228, 408

## R COEFFICIENT OF MULTIPLE CORRELATION

scheme 4/3.2

Type: descriptive statistic
WDH Symbol: R
Correlates level: all metric
Happiness level: metric

Range: [0; 1]
Meaning:
$R = 0 \leftrightarrow$ no influence of any correlate in this study has been established.
$R = 1 \leftrightarrow$ the correlates determine the happiness completely.
Recommended conversion: $R \rightarrow R^2$, the coefficient of determination, which is a test statistic,

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 228, 408.

## C CONTINGENCY COEFFICIENT

See: Pearson's CONTINGENCY COEFFICIENT (C)

## d Cohen's COEFFICIENT FOR GROUP DIFFERENCES

scheme 4/3.1

Type: descriptive statistic only
WDH Symbol: d
Correlate level: dichotomous
Happiness level: metric

Theoretical range: $(-\infty ; +\infty)$
Meaning: effect size indicator for difference of means.
Recommended conversion: Cohen's d $\rightarrow$ Hedges'g by multiplying d with $\sqrt{((N-2)/N)}$.

Ref: Rosenthal, R; *Meta-analytic Procedures for Social Research*, Sage Publications, Inc., 1984, Beverly Hills Ca, USA, ISBN 0-8039-2033-4 and 0-8039-2034-2; 39.

**r** **CORRELATION COEFFICIENT**
See: PRODUCT-MOMENT CORRELATION COEFFICIENT

**$E^2$** **CORRELATION RATIO** scheme 4/3.1

Type: test statistic
WDH-symbols: $E^2$, sometimes $h^2$
Correlate level: nominal or ordinal.
Happiness level: metric

NOTE: if the correlate is measured at the metric level, the correlation ratio $E^2$ is replaced with the coefficient of multiple determination $R^2$, which has the same numerical value.

Range: [0; 1]
Meaning: correlate is accountable for $E^2$ x 100 % of the variation in happiness.
$E^2 = 0 \leftrightarrow$ knowledge of the correlate value does not improve the prediction quality of the happiness rating.
$E^2 = 1 \leftrightarrow$ knowledge of the correlate value enables an exact prediction of the happiness rating
Recommended transformation: none

Test: $E^2$ can be converted into an F standard test statistic.
Under the null hypothesis that there is not any association, the statistic

$$((N_e-c)E^2)/((c-1)(1-E^2))$$

has an F distribution with c-1 and $N_e$ -c degrees of freedom, where $N_e$ = effective number of observations and c = number of levels at which the correlate is varied.

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 373.

**V** **Cramér's V** scheme 4/3.1

Type: test statistic
WDH-symbol: V
Correlate level: nominal
Happiness level: ordinal

Range: [0; 1]
Meaning:
$V = 0 \leftrightarrow$ no association
$V = 1 \leftrightarrow$ strongest possible association
Recommended transformation: none

Test: Under the null hypothesis of no association at all, $N_e(s-1)V^2$ is distributed as a chi-square standard test statistic with df = (c-1)(r-1). In this formula, s the lesser of c and r, the number of columns and rows respectively.

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 305

NOTE: sometimes the value of $V^2$ is reported instead.
Recommended conversion in that case: $V^2 \rightarrow V$, the value of which has been incorporated in the WDH.

**CR** **CRITICAL RATIO**
See: STANDARDIZED DIFFERENCE of MEANS

**OR** **CROSS-PRODUCT RATIO**
See: ODDS RATIO

**DM** **DIFFERENCE of MEANS** scheme 4/3.1

Type: Descriptive statistic only.
WDH Symbol: DM
Correlate level: dichotomous
Happiness level: metric

Range: depending on the happiness rating scale of the author; range symmetric about zero.
Meaning: the difference of the mean happiness, as measured on the author's rating scale, between the two correlate levels.
Recommended conversion: to DMt by transformation onto a [0; 10] scale:
DM $\rightarrow$ DMt = 10*(DM-U)/(H-U), where H and U are the two end points of the original rating scale, H corresponding with extreme happiness and U with extreme unhappiness respectively.
Often the original scale is a k-point scale with U = 1 and H = k; in that case the transformation is DM $\rightarrow$ DMt = 10 *(DM-1)/(k-1).

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 224

**DMt** **DIFFERENCE of MEANS AFTER TRANSFORMATION** scheme 4/3.1

Type: Descriptive statistic only.
WDH Symbol: DMt
Correlate level: dichotomous
Happiness level: metric

Theoretical range: [-10; +10]
Meaning: the difference of the mean happiness (happiness measured at a 0-10 rating scale) between the two correlate levels.

Test: significance testing is possible after conversion to DMs (STANDARDIZED DIFFERENCE of MEANS)
DMt $\rightarrow$ DMs = (DMt/SDt).$\sqrt{(n_1 n_2 / N_e)}$, where SDt is the pooled transformed standard deviation within a correlate level and $n_1$ and $n_2$ are the numbers of observations of both correlate levels ($n_1 + n_2 = N_e$).

Recommendation: Reporting of a CI95 for the true, but unknown (transformed) happiness difference between the two correlate levels is strongly recommended.
The (approximate) confidence limits are DMt - $2SDt\sqrt{(N_e / (n_1 n_2))}$ and

DMt + 2SDt$\sqrt{(N_e /(n_1 n_2))}$ respectively.

Ref: Dixon, W.J.; Massey, F.J.; *Introduction to Statistical Analysis*, McGraw Hill Inc., 1969[3], New York, US; 116.

**D%** **DIFFERENCE in PERCENTAGES** scheme 4/3.1

Type: Descriptive statistic only.
WDH Symbol: D%
Correlate level: dichotomous, but nominal or ordinal theoretically possible as well
Happiness level: dichotomous

Range: [-100; +100]
Meaning: the difference of the percentages happy people at two correlate levels.

NOTE: if the correlate level of measurement is dichotomous, the symbol ε (epsilon) is usual.

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979[3]), 232

**DMr** **DIFFERENCE IN MEAN RIDITS**
Type: test statistic
Measurement level: Happiness ordinal
Range: 0 to +1

Meaning of Mr (Mean Ridit):
Mr < .50: happiness in this subgroup lower than in the larger population
Mr = .50: happiness in this subgroup the same as in the larger population
Mr < .50: happiness in this subgroup higher than in the larger population

'Ridit analysis' compares the distribution of happiness in subgroups with its distribution in the entire sample (Relative to an Identified Distribution)

Test for significance: Bross Confidence Interval (BCI). Tests probability of Mr ≠ 50 in a subgroup, while there is no true difference with the larger population.

References: Bross, I.D. *How to use ridit analysis, Biometrics*, (1958), vol. 14, pp. 18-38
Blalock H.M., *Social Statistics*, McGraw Hill (1979[3]), 232

**DMRT** **DUNCAN's MULTIPLE RANGE TEST** scheme 4/3.1

Type: statistical procedure
WDH Symbol: DMRT
Correlate level: nominal
Happiness level: metric

Meaning: if a single correlate is measured at c levels, the c mean happiness values can be ranked from low to high. A multiple comparison procedure judges for each of the ½c(c-1) pairs whether or not they differ significantly. A convenient way to represent the result is by ranking the c means and by underlining them in such a way that means which have a

common underline do **NOT** differ significantly.

Test: The Duncan procedure uses the studentized range as a standard test statistic, taking into account the number of other average values between both elements of each pair. In this procedure, the confidence of the total package of statements is 95 %.

Ref: Miller, R.G.; *Simultaneous Statistical Inference,* McGraw-Hill Book Company, (1966), New York, USA, 81-90 and 243-246 (Tables).

NOTE 1: Duncan's procedure is similar to that of Student-Newman-Keuls, but it is slightly more discriminating between the various average happiness values. For the consequences, see the above reference.
NOTE 2: The tables for this test are quite complicated and have been improved. The newer tables are referred to as Tables for "Duncan's *New* Multiple Range Test".

## EPS EPSILON SQUARE

Type: Test statistic.
WDB Symbol: EPS
Correlate level: nonmetric
Happiness level: metric

Range: [0; 1]
Meaning: correlate is accountable for EPS x100 % of the variation in happiness.
EPS = 0 $\leftrightarrow$ knowledge of the correlate value does not improve the prediction quality of the happiness rating.
EPS = 1 $\leftrightarrow$ knowledge of the correlate value enables an error-free prediction of the happiness rating.
EPS is an unbiased estimator of the relative reduction of the uncertainty about happiness by knowledge of the correlate value
Recommended conversion: none

Test: EPS can be transformed into an F standard test statistic. Under the null hypothesis of no association at all, the statistic $((N_e\text{-}c)*EPS + c - 1)/((c\text{-}1)(1\text{-}EPS))$ has an F distribution with c-1 and $N_e$-c degrees of freedom, where $N_e$ = effective number of observations and c = number of levels at which the correlate is varied.

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 373.
NOTE: In this edition the statistic is given the symbol $\eta^2$ (eta square), whereas the 1972 edition uses the symbol $\varepsilon^2$ !

## F-STATISTIC

Type: asymmetric standard test statistic.
WDH Symbol: none.
Two parameters: $\nu_1$ (= number of degrees of freedom for the numerator; range df: [1; + $\infty$) and $\nu_2$ (= number of degrees of freedom for the denominator; range df: [1;+ $\infty$)

Range: [0; + $\infty$)
Meaning : the test statistic is also called the "Variance Ratio" and is the ratio of two

independent estimators of the same variance with $\nu_1$ and $\nu_2$ degrees of freedom respectively.
The critical values of its probability distribution are tabulated extensively in almost any textbook on Statistics, and in

Ref: Pearson, E.S.; Hartley, H.O.; *Biometrika Tables for Statisticians vol I*; Cambridge University Press; 19668; Cambridge, UK; 36, 169.

## Fisher's EXACT 2x2 TEST

Type: statistical procedure.
WDH Symbol: FET.
Correlate level: dichotomous.
Happiness level: dichotomous.

Test: The test is performed on the basis of the exact distribution under the null hypothesis that there is not any association. This exact distribution is tabulated for $N < 30$.
For larger N-values a chi-square test is the usual test for independence of the correlate and happiness. In this case, the application of the Yates' correction for continuity is recommended.

Ref: Pearson, E.S.; Hartley, H.O.; *Biometrika Tables for Statisticians vol I*; Cambridge University Press; 19703; Cambridge, UK; 71, 212.
Dixon, W.J.; Massey, F.J.; *Introduction to Statistical Analysis*, McGraw Hill Inc., 1969[3], New York, US; 242.

## G GOODMAN & Kruskal's GAMMA. Scheme 4/3.1

Type: test statistic
WDH Symbol: G
Correlate level: ordinal
Happiness level: ordinal

Range: [-1; +1]
Meaning:
GAMMA = 0 ↔ no rank correlation
GAMMA = +1 ↔ strongest possible rank correlation, where high correlate values correspond with high happiness ratings.
GAMMA = -1 ↔ strongest possible rank correlation, where high correlate values correspond with low happiness ratings.
Recommended conversion: none
Test: GAMMA can be transformed into a z-statistic.

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 442.

## Tau GOODMAN & Kruskal's TAU scheme 4/3.1

Type: descriptive statistic only.
WDH-symbol: tau

Correlate level: nominal
Happiness level: ordinal

Range: [0; +1]
Meaning:
tau = 0 ↔ knowledge of the correlate value does not improve the prediction quality of the happiness rating.
tau = 1 ↔ knowledge of the correlate value enables a perfect (error-free) happiness rating.
Recommended conversion: none

Ref: Goodman, L.A, Kruskal, W.H., *Measures of Association for Cross-Classification,* J. Am. Statistical Association. 32 (1954), 732 -

NOTE: Avoid confusion with the τ-values introduced by Kendall and Stuart, which can adopt negative values as well.
This confusion is compounded by authors who use the symbols $\tau_a$ or $\tau_b$ for Goodman & Kruskal's tau !

**gH** **Hedges's COEFFICIENT FOR GROUP DIFFERENCES** scheme 4/3.1

Type: descriptive statistic only.
WDH Symbol: gH
Correlate level : dichotomous
Happiness level: metric

Theoretical range: (- ∞; + ∞).
Meaning: effect size indicator for difference of means.
Recommended conversion: none

Ref: Rosenthal, R; *Meta-analytic Procedures for Social Research*, Sage Publications, Inc., 1984, Beverly Hills Ca, US, ISBN 0-8039-2033-4 and 0-8039-2034-2; 39.

**h²** see: **E² (CORRELATION RATIO)**

**ta** **KENDALL'S RANK CORRELATION COEFFICIENT TAU-A** scheme 4/3.1
Type: test statistic
WDH Symbol: ta
Correlate level: ordinal
Happiness level: ordinal

Range: [-1; +1], but if ties are present -1< ta< +1.
Meaning:
ta = 0 ↔ no rank correlation
ta = 1 ↔ perfect rank correlation, high happiness at high correlate value levels.

ta =-1 ↔ perfect rank correlation, high happiness at low correlate value levels.

Recommended conversion: none, however, as the number of tied ranking increases, there is a stronger reason to apply Kendall's tau-b instead.
Test: for $N \leq 10$, the exact probability distribution of C-D has been tabulated for the null-hypothesis that concordant and discordant pairs are equally likely to occur, where C and D are the numbers of concordant and discordant pairs respectively. For large N-values, the distribution of $ta*\sqrt{((9N(N-1))/(4N+10))}$ under the null hypothesis of no rank correlation, can be approximated by a standard normal (z) distribution. For moderate N-values a much more complicated approximation is required. Obviously for N the value of $N_e$ has to be used.

Ref: Kendall, M.G., Rank *Correlation Methods,* Ch. Griffin & Company Ltd., London (1962³), 3, 49.

**tb** **KENDALL'S RANK CORRELATION COEFFICIENT TAU-B** scheme 4/3.1

Type: test statistic
WDH Symbol: tb
Correlate level: ordinal
Happiness level: ordinal

Range: [-1; +1]
Meaning:
tb = 0 ↔ no rank correlation
tb = 1 ↔ perfect rank correlation, where high values of the correlate correspond with high happiness ratings.
tb = -1 ↔ perfect rank correlation, where high values of the correlate correspond with low happiness ratings.
Recommended conversion: none

Test: for large N-values, the distribution of $tb*\sqrt{((9N(N-1))/(4N+10))}$ under the null hypothesis of no rank correlation, can be approximated by a standard normal (z) distribution. For moderate N-values a much more complicated approximation is required. Obviously for N the value of $N_e$ has to be used.

NOTE: the application of tau-b is recommended only if the number of rows and columns are equal. Otherwise, tau-c is to be preferred.

Reference: Kendall, M.G., Rank *Correlation Methods,* Ch. Griffin & Company Ltd., London (1962³), 36.

**tc** **KENDALL'S TAU-C** scheme 4/3.1

Type: test statistic
WDH Symbol: tc
Correlate level: ordinal
Happiness level: ordinal

Range: [-1; +1]
Meaning:

tc = 0 ↔ no rank correlation
tc = 1 ↔ perfect rank correlation, where high values of the correlate correspond with high happiness ratings.
tc = -1 ↔ perfect rank correlation, where high values of the correlate correspond with low happiness ratings.
Recommended conversion: none

Test: for large N-values, the distribution of $tc*\sqrt{((9N(N-1))/(4N+10))}$ under the null hypothesis of a zero rank correlation at all, can be approximated by a standard normal (z) distribution. A much more complicated approximation is required for moderate N-values. Obviously for N the value of $N_e$ has to be used.

Ref: Kendall, M.G., Rank *Correlation Methods,* Ch. Griffin & Company Ltd., London (1962³), 47.

NOTE: This statistic is also referred to as Stuart's tau-c.

**lgt** **LOGIT COEFFICIENT** Scheme 4/3.1

Type: descriptive statistic only.
WDH-symbol: lgt
Correlate level: dichotomous.
Happiness level: dichotomous

Range: $(-\infty; +\infty)$
Meaning: lgt = 0 ↔ no association at all;
lgt = -/+ ∞ ↔ at least one level of the correlate allows a perfect prediction of the happiness.
Recommended transformation: none.

Ref: Bishop, Y.M.M., Fienberg, S.E., Holland, P.W., *Discrete Multivariate Analysis: Theory and Practice,* MIT Press, Cambridge Mass. (US), 1975, ISBN 0-262-02113-7; 22, 30.

**MANN-WHITNEY TWO-SAMPLE TEST**
See: WILCOXON (MANN-WHITNEY) TWO-SAMPLE TEST

**MAoV** **MULTIVARIATE ANALYSIS of VARIANCE (MANOVA).**

Multivariate Analysis of Variance is very similar to the ordinary Analysis of Variance. The 'multivariate' difference is that more than one response variable, i.c. happiness measure, is involved. This technique is meaningful only if the response variables are correlated, which in happiness studies generally will be the case. Mathematically, the input data is not a column vector but a matrix. The equivalent of the F-test for significant differences between the correlate levels is e.g. Hotelling's $T^2$-test.

Ref: Morrison, D.F., *Multivariate Statistical Methods*, McGrawHill, New York (1976²).

**Mt** **MEAN HAPPINESS VALUE AFTER TRANSFORMATION**

See: average happiness value after transformation

**R** **MULTIPLE CORRELATION COEFFICIENT**
See: COEFFICIENT OF MULTIPLE CORRELATION

**SNK** **NEWMAN-KEULS**
See: STUDENT-NEWMAN-KEULS

**OR** **ODDS RATIO** scheme 4/3.1

Type: descriptive statistic only.
WDH-symbol: OR
Correlate level: dichotomous.
Happiness level: dichotomous

Range: (0; ∞)
Meaning: OR = 1 ↔ no association at all;
OR = 0 or ∞ ↔ at least one level of the correlate allows an error-free prediction of the happiness.
Recommended transformation: none

Ref: Bishop, Y.M.M., Fienberg, S.E. , Holland, P.W., *Discrete Multivariate Analysis: Theory and Practice,* The MIT Press, Cambridge Mass. (US), 1975, ISBN 0-262-02113-7; 13.

NOTE: sometimes this statistic is referred to as "cross-product ratio".

**$w^2$** **OMEGA-SQUARE**

Type: descriptive statistic
WDH-symbol: $w^2$
Correlate level: nominal or ordinal
Happiness level: metric

Range: [0; 1]
Meaning: $w^2$ is an estimator of the relative reduction in uncertainty about the happiness by knowledge of the correlate value.
$w^2 = 0$ ↔ knowledge of the correlate value does not improve the prediction quality of the happiness rating.
$w^2 = 1$ ↔ knowledge of the correlate value enables an exact prediction of the happiness rating
Recommended transformation: none

If there is no association between happiness and the correlate(s), the statistic
$$(N_e w^2 + (1 - w^2)(c - 1))/((1 - w^2)(c - 1))$$
has an F-distribution with c-1 and $N_e$-c df respectively.
In this formula $N_e$ = effective number of observations and c = number of levels at which the correlate is varied.

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 373.

NOTE: This omega square is a biased estimator: the corresponding population parameter is systematically underestimated by this omega square. An unbiased estimator of the population parameter is found in epsilon square (see: EPSILON SQUARE).

**rpc** **PARTIAL CORRELATION COEFFICIENT** scheme 4/3.2

Type: test statistic
WDH Symbol: rpc
Correlate level: metric
Happiness level: metric

Range: [-1; +1]
Meaning: a partial correlation between happiness and one of the correlates is that correlation, which remains after
accounting for the variation of the influence of the other ones, or some of them. Under that conditions
rpc > 0 $\leftrightarrow$ a higher correlate level corresponds with a higher happiness rating, and
rpc < 0 $\leftrightarrow$ a higher correlate level corresponds with a lower happiness rating,
both after accounting for the influence of one or more other correlates.
Recommended conversion: none.

Test: a significance test implies a conclusion about the sign (- or +) of the partial correlation coefficient. It is performed by the so-called Fisher transformation:

$$rpc \rightarrow z = (\tfrac{1}{2}\sqrt{(N_e - q - 3)}).^e\log((1+rpc)/(1-rpc)).$$

Under the null hypothesis of no partial correlation at all, this z has a standard normal distribution, where $N_e$ = number of observations and q = the order of the partial correlation coefficient, i.e. the number of correlates that have been accounted for; to say it simply: the number of indices "behind the dot".

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 267, 409.

**B** **PARTIAL REGRESSION COEFFICIENT** (non-standardized).
See: REGRESSION COEFFICIENT (non-standardized)

**C** **Pearson's CONTINGENCY COEFFICIENT C** scheme 4/3.1

Type: test statistic
WDH Symbol: C
Correlate level: nominal
Happiness level: ordinal

Range: $0 \leq C^2 \leq C < \sqrt{(1-1/s)} < 1$, where s = the lesser of c and r, the number of columns and rows respectively.
Meaning: C = 0 $\leftrightarrow$ no association.
C $\rightarrow$ 1 $\leftrightarrow$ strongest possible association

Recommended conversion: to Cramér's V according to C $\rightarrow$ $V = \sqrt{(C^2/((1-C^2)(s-1)))}$,
Test: Under the null hypothesis of no association at all, $N_eC^2/(1-C^2)$ is distributed as a chi-square standard test statistic with df = (c-1)(r-1).

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 305.

**r** **Pearson's (PRODUCT MOMENT) CORRELATION COEFFICIENT**
See: PRODUCT-MOMENT CORRELATION COEFFICIENT

**rpb** **POINT BISERIAL COEFFICIENT of CORRELATION** scheme 4/3.1

Type: Descriptive statistic only
WDH Symbol: rpb
Correlate level: dichotomous
Happiness level: metric

Range: [-1; +1]
Meaning: The point bi-serial coefficient of correlation is the Pearson's coefficient correlation which is obtained if the correlate is an "indicator variable", i.e. a variable which only adopts the values 0 and 1.
rpb = 0 $\leftrightarrow$ no correlation established
rpb = 1 $\leftrightarrow$ at correlate level 1 only higher happiness ratings and at correlate level 0 only lower happiness ratings
rpb = -1 $\leftrightarrow$ at correlate level 1 only lower happiness ratings and at correlate level 0 only higher happiness ratings.

Recommended conversion: none

Ref: Tate, R.F.; The *theory of correlation between two continuous variables when one is dichotomized,* Biometrika, 1955, Vol 42, 205-216.

**r** **PRODUCT-MOMENT CORRELATION COEFFICIENT** scheme 4/3.1

Also referred to as "correlation coefficient" simply or as "Pearson's correlation coefficient".
Type: Test statistic.
WDH Symbol: r
Correlate level: metric
Happiness level: metric

Range: [-1; +1]
Meaning:
r = 0 $\leftrightarrow$ no correlation ,
r = 1 $\leftrightarrow$ perfect correlation, where high correlate values correspond with high happiness values, and
r = -1 $\leftrightarrow$ perfect correlation, where high correlate values correspond with low happiness values.
Recommended conversion: none.

Test: a significance test implies a conclusion about the sign (- or +) of the correlation

coefficient rho. It is performed by using the 'Fisher transformation':

$$r \rightarrow z = (½\sqrt{(N_e - 3)}).^{e}\log((1+r)/(1-r)).$$

Under the null hypothesis of no correlation at all (rho = 0), this z has a standard normal distribution, where $N_e$ = effective number of observations.

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 90, 396, 404.
Pearson, E.S.; Hartley, H.O.; *Biometrika Tables for Statisticians vol I*; Cambridge University Press; 19703; Cambridge, UK; 28, 146.

Recommendation: the construction of 95 % confidence interval (CI95) for the true but unknown value of rho is recommended whenever appropriate, i.e. whenever the joint distribution of the correlate and the happiness can be considered more or less as bi-variate normal. The procedure is based on the Fisher transformation as given above.

The lower confidence limit for rho is equal to

$$(\exp((z-1.96)/(½\sqrt{(N_e - 3)}))-1)/(\exp((z-1.96)/(½\sqrt{(N_e - 3)}))+1)$$

and the upper limit equals

$$(\exp((z+1.96)/(½\sqrt{(N_e - 3)}))-1)/(\exp((z+1.96)/(½\sqrt{(N_e - 3)}))+1),$$

where z = Fisher-transformed value of r.

## b REGRESSION COEFFICIENT (non-standardized) scheme 4/3.2

Type: test statistic
WDH Symbol: b
Correlate level: metric
Happiness level: metric

Theoretical range: $(-\infty ; +\infty)$
Meaning:
$b > 0 \leftrightarrow$ a higher correlate level corresponds with, on an average, higher happiness rating.
$b < 0 \leftrightarrow$ a higher correlate level corresponds with, on an average, lower happiness rating.
$b = 0 \leftrightarrow$ not any correlation with the relevant correlate.
Recommended conversion: none.

Test: a significance test implies a conclusion about the sign (- or +) of the regression coefficient and is performed as a Student's t-test. Under the null hypothesis of not any correlation, the ratio of b and its standard error, which equals $(sr)/((s(x)).\sqrt{(Ne - 1)})$, has a Student's t-distribution with $N_e - 2$ df.
In the above formula, $N_e$ = effective number of observations, s(x) = standard deviation of the observed correlate values and sr = estimated residual standard deviation

NOTE: if there is more than one regressor, the regression coefficients are also referred to as "non-standardized *partial* regression coefficients".

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 38, 264.

**F** **SNEDECOR'S F-STATISTIC.**
See: F-STATISTIC

**Dyx** **SOMERS' ASYMMETRIC TEST STATISTIC** scheme 4/3.1

Type: test statistic
WDH Symbol: Dyx
Correlate level: ordinal
Happiness level: ordinal

Range: [-1; +1]
Meaning:
Dyx = 0 ↔ no rank correlation
Dyx = +1 ↔ strongest possible rank correlation, where high correlate values correspond with high happiness ratings.
Dyx = -1 ↔ strongest possible rank correlation, where high correlate values correspond with low happiness ratings.
Recommended conversion: none
Test: Dyx can be transformed into a standard test statistic

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 443.

**rs** **SPEARMAN'S RANK CORRELATION COEFFICIENT** scheme 4/3.1

Type: test statistic
WDH Symbol: rs
Correlate level: ordinal
Happiness level: ordinal.

Range: [-1; +1]
Meaning:
rs = 0 ↔ no rank correlation
rs = 1 ↔ perfect rank correlation, where high correlate values are associated with high happiness ratings
rs =-1 ↔ perfect rank correlation, where high correlate values are associated with low happiness ratings
Recommended conversion: none.

Test: For sufficiently large samples and under the null hypothesis if no rank correlation, the statistic $rs/(\sqrt{(N_e - 1)})$ has a probability distribution, which is approximately standard normal (z). For smaller $N_e$-values ($n_1 = n_2 < 30$) exact critical values of the test statistic are tabulated, e.g. in

Ref: Beyer, W.H. (ed); *Handbook of Tables for Probability and Statistics*, The Chemical Rubber Co, 1968², Cleveland Oh, US, ISBN 0-8493-0692-2; 445-448.

**SDt** **STANDARD DEVIATION IN HAPPINESS AFTER TRANSFORMATION**

Type: distribution statistic
WDH-symbol: SDt
Happiness level: metric.
Theoretical range: [0; 5*√($N_e$ /($N_e$-1))], in practice [0;5]

This statistic is the standard deviation of all happiness scores in a sample after linear transformation onto a [0; 10] scale.

Meaning:
SDt = 0 ↔ all respondents in the sample are equally happy/satisfied etc..
SDt = 5 ↔ completely split happiness, where half the sample is extremely happy/satisfied, etc. and the other half is extremely unhappy/unsatisfied etc.
Recommended conversion: none.

**DMs** **STANDARDIZED DIFFERENCE of MEANS** scheme 4/3.1

Type: test statistic.
WDH Symbol: DMs, previously DMsd
Correlate level: dichotomous
Happiness level: metric
.
Theoretical range: (- ∞ ; + ∞).
Meaning: DMs is the ratio of the difference between the - either untransformed or transformed - means and the standard error of that difference.
Recommended conversion: none
Test: Under the null hypothesis that there is no influence of the correlate, DMs has a Student's t-distribution with $N_e$ -2 degrees of freedom.

Recommendation: Reporting of a CI95 for the true, but unknown (transformed) happiness difference between the two correlate levels is strongly recommended.
The (approximate) confidence limits are

$$DMt - 2SDt\sqrt{(N_e/(n_1n_2))} \text{ and}$$
$$DMt + 2SDt\sqrt{(N_e/(n_1n_2))} \text{ respectively,}$$

where SDt is the pooled transformed standard deviation within a correlate level and $n_1$ and $n_2$ are the numbers of observations of both correlate levels ($n_1 + n_2 = N_e$).

Ref: Dixon, W.J.; Massey, F.J. ; *Introduction to Statistical Analysis*, McGraw Hill Inc., 1969[3] , New York, US; 116.

NOTE: This statistic is reported incidentally as the 'Critical Ratio'.

**Beta** **STANDARDIZED REGRESSION COEFFICIENTS** scheme 4/3.2

Type: test statistic
WDH Symbol: beta
Correlates level: all metric
Happiness level: metric.

Range: [-1 ; +1]
Meaning:

beta > 0 ↔ a higher correlate level corresponds with, on an average, higher happiness rating.
beta < 0 ↔ a higher correlate level corresponds with, on an average, lower happiness rating.
beta = 0 ↔ no correlation.
beta = + 1 or -1 ↔ perfect correlation.

Recommended conversion: none

Test: In the case of one regressor only, beta is equal to Pearson's PRODUCT-MOMENT CORRELATION COEFFICIENT and its significance test has been described under that statistic.

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 268.

NOTE 1: The term is somewhat confusing, since it is not the regression coefficient that is standardized, but the variables involved in the regression analysis.

NOTE 2: Standardized regression coefficients are also referred to as Beta coefficients or as Beta weights.

**STEPWISE MULTIPLE REGRESSION ANALYSIS**

Type: statistical procedure
WDH Symbol: none
Correlates level: all metric
Happiness level: metric

In stepwise regression analysis, one identifies the minimum size subset of regressors that is sufficient to describe the happiness variation, whereas the remaining regressors can be considered as superfluous, since they do not make a real additional contribution. This is achieved by successively entering and/or removing the various regressors into/from the regression equation in turn.

Ref: Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1, 371-386.

**tc** **STUART'S TAU-C**
See: KENDALL'S TAU-C scheme 4/3.1

**t** **STUDENT'S t-STATISTIC**
See: t-STATISTIC

**SNK** **STUDENT-NEWMAN-KEULS**

Type: Statistical procedure
WDH Symbol: SNK
Correlate level: nominal

Happiness level : metric

Meaning: if the correlate is measured at c levels, the c mean happiness values can be ranked from low to high. A multiple comparison procedure judges for each of the ½c(c-1) pairs whether or not they differ significantly. A convenient way to represent the results is by ranking the c means and by underlining them in such a way that means which have a common underlining do **NOT** differ significantly.

Test: the Student-Newman-Keuls procedure uses the Studentized Range as a standard statistic, taking into account the number of other average values between both elements of each pair.
In this procedure, the confidence of the *total package* of statements is at least 95 %.

Ref: Miller, R.G.; *Simultaneous Statistical Inference,* McGraw-Hill Book Company, (1966), New York, USA, 81 - 90.

**TT** **THURSTONE TRANSFORMATION**.

Type: transformation procedure for happiness scores.
WDH Symbol: TT
Correlate level: not applicable
Happiness level: from ordinal to pseudo-metric.

Meaning: *Thurstone* has introduced a transformation method, which is applied in WDH to happiness scores. In this approach each possible answer to a happiness question is presented to a panel of people who are considered to be experts. Each panel member is asked to assign a score on the [0; 10] interval to each answer on the original (primary) scale. If the individual scores for the same answer do not differ too much, their average value is adopted as the transformed value of the original rating/label.
Since the scores obtained in this way are dependent on the exact formulation of the answers, including the language in which they are formulated, application of this transformation requires that the average expert scores for each rating be entered in to the excerpt.

Ref: Torgerson, W.S.; *Theory and Methods of Scaling,* J. Wiley USA, (1958), New York, USA.
.

**Mtt** **THURSTONE TRANSFORMATED MEAN HAPPINESS**

Type: descriptive distribution statistic
WDH Symbol: Mtt
Correlate level: not applicable
Happiness level: ordinal transformed to pseudo-metric

Theoretical range: [0; 10]
Meaning: see under THURSTONE TRANSFORMATION. This statistic is obtained as

the mean of all scores after transformation of the primary scores according to the Thurstone transformation.

**SDtt** **THURSTONE TRANSFORMATED HAPPINESS STANDARD DEVIATION**

Type: descriptive distribution statistic
WDH Symbol: SDtt
Correlate level: not applicable
Happiness level: ordinal transformed to pseudo-metric

Range: [0; < 10)
Meaning: see under THURSTONE TRANSFORMATION. This statistic is obtained as the standard deviation of all scores after transformation of the primary scores according to the Thurstone transformation.

**t** **t-STATISTIC** (often referred to as Student's t-statistic)

Type: symmetric standard test statistic.
WDH Symbol: t.
One parameter: $\nu$ (= number of degrees of freedom; range df: [1; + $\infty$)

Range: (- $\infty$ ; + $\infty$)

Meaning : the test statistic is the ratio of a difference between a statistic and its expected value under the null hypothesis and its (estimated) standard error with $\nu$ degrees of freedom.
The critical values of its probability distribution are tabulated extensively in almost any textbook on Statistics, and in

Ref: Pearson, E.S.; Hartley, H.O.; *Biometrika Tables for Statisticians vol I*; Cambridge University Press; 19668; Cambridge, UK; 19, 146.

**T** **TSCHUPROW'S T** scheme 4/3.1

Type: test statistic.
WDH Symbol: T
Correlate level: nominal
Happiness level: ordinal

Range: $0 \leq T^2 \leq \sqrt{[[\min(r,c)-1]/[\max(r,c)-1]]} \leq 1$.
Meaning: $T = 0 \leftrightarrow$ no association
$T \rightarrow 1 \leftrightarrow$ strongest possible association.
Recommended conversion: T $\rightarrow$ Cramér's V = $T\sqrt{((\sqrt{((c-1)(r-1)))/(s-1)})}$.
In this formula, s the lesser of c and r, the number of columns and rows respectively.
Test: Under the null hypothesis of no association at all, $T^2 \times N_e\sqrt{((c-1)(r-1))}$ is distributed as a chi-square standard test statistic with df = (c-1)(r-1).

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 304

NOTE 1: Instead of T, incidentally the author reports $T^2$.
In that case the conversion $T^2 \rightarrow T$ is recommended and the value of T has been incorporated in the WDH.
NOTE 2: Avoid confusion between this reference (T) and Hotelling's $T^2$ as used in Multivariate Analysis of Variance (MANOVA).

**MW** **WILCOXON (MANN-WHITNEY) TWO-SAMPLE TEST**

Type: statistical procedure.
WDH Symbol procedure: MW
Correlate level: dichotomous
Happiness level: ordinal

Symbol test statistic: W (Wilcoxon) or U (Mann-Whitney).
Relationship: $W = U + \frac{1}{2}n_1(n_1 + 1)$,
where $n_1$ and $n_2 = N_e - n_1$, while $n_1 \leq n_2$, are the sizes of the two samples.

Range U [ 0; $n_1n_2$ ] and W [ $\frac{1}{2}n_1(n_1 + 1)$; $\frac{1}{2}n_1(n_1 + 2n_2 + 1]$.
Meaning:
At the limit values of U and W there is the strongest possible association.
$U = \frac{1}{2}n_1n_2 \leftrightarrow W = \frac{1}{2}n_1(N + 1) \leftrightarrow$ no association.
Recommended conversion: none

Test: For $N_e < 50$ an exact test is available;
For $N_e > 7$, both W and U can be transformed into a test statistic, which approximately has a standard normal distribution.
For U this transformation is $U \rightarrow (U - \frac{1}{2}n_1n_2)/(\sqrt{(n_1n_2(n_1+n_2+1)/12)})$.

Ref: Pearson, E.S.; Hartley, H.O.; *Biometrika Tables for Statisticians vol II*; Cambridge University Press; 1972; Cambridge, UK; ISBN 0 521 06937 8, 46, 227.

**Q** **Yule's Q-STATISTIC** scheme 4/3.1

Type: descriptive statistic only.
WDH-symbol: Q.
Correlate level: dichotomous
Happiness level: dichotomous

Range: [-1; +1]
Meaning: $Q = 0 \leftrightarrow$ no association
$Q = -1$ or $+1 \leftrightarrow$ at least one level of the correlate allows a perfect prediction of the happiness.
Recommended transformation: none

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 373.

**Y** **Yule's COEFFICIENT of COLLIGATION** (Y-STATISTIC) scheme 4/3.1

Type: descriptive statistic only.

WDH-symbol: Y
Correlate level: dichotomous
Happiness level: dichotomous

Range: [-1;1]
Meaning: Y = 0 ↔ no association
Y= -1 or + 1 ↔ at least one level of the correlate allows a perfect prediction of the happiness.
Recommended transformation: none

Ref: Blalock H.M., *Social Statistics*, McGraw Hill (1979³), 373.

## 4/5.2 Incidentally used statistical techniques

Some uncommon statistical techniques have been used in some empirical studies of happiness. These exceptions are not included in the above list, since this would blur the overview. When a statistical technique is applied in only a few cases, we suffice with a short remark and leave it to the user to dig up the details of the method in the research report. An overview of these miscellaneous statistics and the studies in which they have been applied is given below.

| *Symbol* | *Name* | *Applied in study* |
|---|---|---|
| a | Factor loading | KAMMA 1983/1 |
| βL | LISREL Path coefficient | BEALS 1985 |
| BCI | Bross' confidence interval for average ridits | BRADB 1969<br>HEADE 1989<br>ORMEL 1980 |
| d | Discriminant Coefficient | BRAEN 1991 |
| DMa | Difference in Adjusted Means | THOMA 1986 |
| DMo | Difference in modus | BAHR 1980 |
| DMr | Difference in mean correlation | LOUNS 1979 |
| DMsd | Difference from a reference mean in sd | GEHMA 1989B |
| DMR | Difference in mean ridits | BERKM 1971<br>BRADB 1969<br>LOUNS 1979 |
| Gs | Partial Gamma | BRADB 1969<br>BRENN 1979 |
| mc | Guttman's monotonicity coefficient | LEVY 1975, 1975/1 and 1975/2 |
| rp | Polychoric correlation | BATIS 1996 |
| W | Wilcoxon's signed rank test | CHARN 2000/1<br>FUGL 1991 |

## APPENDIX A: FOUR LEVELS of MEASUREMENT.

*Measurement at the ratio level.*

In physical and related sciences, measurement is a key activity. Generally speaking measurement is a **process of comparison**. Measuring the length of a stick is just comparing that length with some other length, which has been adopted as length unit. Between 1799 and 1960, this unit (the so-called "standard meter"), was defined as the distance between two parallel scratches in a particular bar made from an alloy of platinum and iridium and stored at a temperature of 0 °C in a protected site (The "Bureau International des Poids et Mesures" at Sèvres in France). After 1960, this standard meter has been replaced with a new standard meter on the basis of a more accurate measurement of a more stable standard material.
The expression that "the length of the stick is 0.83 meter" means nothing else than that the ratio of the length of the stick and that of the corresponding unit (the standard meter) equals the number 0.83. Hence this type of measurement is called **"measurement at the ratio level".**
The set of all ordered possible outcomes of length measurements together is called a **measurement scale,** in this case for length measurement. Such scales have **two standard points.** One of them is the length of the standard meter as described above. The other one is the zero. This corresponds to a zero length, being the non-existent, but theoretically smallest conceivable length.

*Measurement at the interval level.*

The vast majority of all measurements in physical and related sciences are performed at the ratio level, but not all of them. A different level is found when the temperature of some substance is measured using the centigrade (°C) as unit of temperature measurement. At this level there is also a zero value, but, in contrast to the ratio level, this zero is neither 'absolute' nor natural, but chosen fully arbitrarily (in this case the temperature of melting pure $H_2O$ at a pressure of 1,0132 bar). As a consequence of this particular choice, also negative temperature values may be measured. The second standard temperature is chosen as the temperature of boiling pure $H_2O$ at the same a pressure of 1,0132 bar. The interval between the two standard temperatures is divided by 100, giving the centigrade as unity - also called '**scale unit' -** of temperature (difference). In this case, the comparison process is essentially to establish the value of the ratio of two intervals; hence this level of measurement is referred to as "**the interval level of measurement**". If a substance has a temperature of 31 °C, the temperature differences (substance - melting ice) and (boiling water - melting ice) have a ratio 31/100.
Measurement at the interval level occurs relatively infrequently; another example is time in the sense of a moment, e.g. "the actual time is December 25th 2001 16:04:27".

*Measurement at the ordinal level.*

A third level occurs almost equally infrequently (in physical and related sciences). The classical example is the measurement of hardness of minerals according to Mohs.
F. Mohs (1773 - 1839) proposed a hardness scale with 10 standard points: talc (soapstone) = 1; gypsum = 2; calcite = 3; fluorite = 4; apatite = 5; orthoclase = 6; quartz = 7; topaz = 8; corundum = 9 and diamond = 10. Any mineral can scratch a mineral with a lower hardness

number and can be scratched only by minerals with higher hardnesses.

The above ten minerals are ranked according to increasing hardness (in the sense of scratching behaviour), but nothing is said about the mutual differences. Although numerically 4 - 3 = 8 - 7 (=1), this does not at all imply that the hardness difference between fluorite and calcite is equal to that between topaz and quartz. If hardness is reported using decimals, e.g. 7.5, this means that the mineral is harder than quartz, but not as hard as topaz. Scales like that of Mohs (another example is the wind speed scale of Beaufort) are called "ranking scales" and the corresponding level of measurement is referred to as the "**ordinal level of measurement**".

Ordinal scales have more than two standard values. 'Scale unit' as *the* difference between two consecutive scale points is an undefined concept at an ordinal scale. The standard values can be expressed in numbers, but these numbers are to be considered as code numbers, e.g. 8 for topaz. Standard values can be expressed also in verbal terms. A scale with five standard values{very poor, poor, adequate, good, very good} is a five-point ordinal scale. This example illustrates that measurement does not necessarily result in assigning numbers to items, but that other, i.c. verbal labels are possible outcomes as well. Pictures, as applied in "smiley scales" are also an alternative. In other words: measurement is not necessarily a quantitative process, but can also be qualitative.

*Measurement at the nominal level.*

Finally, a fourth level of measurement can be distinguished. At this level measurement is usually qualitative. This may be the reason why most physicists and other scientists who are only familiar with quantitative measurements, do not recognise this as a level of measurement.
At this level, the investigator gives a label to the object of investigation; e.g. a chemist will identify a liquid as methanol. One can establish that the colour of a cat is 'black'. The label can be either verbal (as in the examples) or numerical. If, however, three or more subjects receive different numbers (e.g. social security numbers), those numbers can be ranked, but this ranking does not include any information about any ranking of the subjects in any dimension. This fourth level of measurement is referred to as the **"nominal level"**, since it is a matter of giving a name (nomen) or other label to elements of the study sample.

*Hierarchy of the levels of measurement.*

The four levels of measurement can be ordered from high to low as follows:
- ratio level
- interval level
- ordinal level
- nominal level.

The ratio level and the interval level are also referred to as *"metric levels of measurement",* whereas the other two are the *"nonmetric levels*".
Any operation that is admissible at some level is also admissible at higher levels.

At the *nominal level*, the only possible comparison of two objects is to establish whether they are equal or different. Equal objects can be collected into 'classes'. The only possible operation is counting the number of equal elements in each class.

At the *ordinal level* the various elements of the set of items under investigation can be ranked according to the relevant criterion, either from high to low or the reverse. Transformations are allowed as long as they are monotonous.
At the *interval level* equal differences between measured values have a meaning: if the measured value of objects A is denoted as m (A), then from m (A) - m (B) = m(C) - m (D) it follows that the difference between A and B is equal to that between C and D.
Linear transformation of measured values is admissible at this level.
At the *ratio level* also other transformations, e.g. logarithmic ones are admissible. Moreover, relative and percentual variables are meaningful concepts, which they are not at the interval level.

The above distinctions are highly relevant in behavioural sciences, in which nonmetric measurement is occurring much more frequently. Measurement of happiness is only one of the many examples. The selection of the most suitable - or the least unsuitable - label of such a **rating scale** is always a matter of judgement by the subject. This subjective judging is only possible at the nonmetric level and this has consequences for the way individual scores have to be processed.

## References

Ref: Beyer, W.H. (ed); *Handbook of Tables for Probability and Statistics*, The Chemical Rubber Co, 1968[2], Cleveland Oh, US, ISBN 0-8493-0692-2.

Bishop, Y.M.M., Fienberg, S.E., Holland, P.W., *Discrete Multivariate Analysis: Theory and Practice,* MIT Press, Cambridge Mass. (US), 1975, ISBN 0-262-02113-7;

Blalock H.M., *Social Statistics*, McGraw Hill (1979³)

Bross, I.D. *How to use ridit analysis, Biometrics*, (1958), vol. 14.

Dixon, W.J.; Massey, F.J. ; *Introduction to Statistical Analysis*, McGraw Hill Inc., 1969[3], New York, US;

Goodman, L.A, Kruskal, W.H., *Measures of Association for Cross-Classification,* J. Am. Statistical Association. **32** (1954)

Kendall, M.G., Rank *Correlation Methods,* Ch. Griffin & Company Ltd., London (1962³)

Miller, R.G.; *Simultaneous Statistical Inference,* McGraw-Hill Book Company, (1966), New York, USA.

Morrison, D.F., *Multivariate Statistical Methods*, McGrawHill, New York (1976²).

Neter, J.; Wasserman, W.; *Applied Linear Statistical Models*; R.D. Irwin Inc., 1974, Homewood. Ill., USA, ISBN 0-256-1498-1.

Pearson, E.S.; Hartley, H.O. ; *Biometrika Tables for Statisticians vol I*; Cambridge University Press; 1966[8]; Cambridge, UK.

Rosenthal, R; *Meta-analytic Procedures for Social Research*, Sage Publications, Inc., 1984, Beverly Hills Ca, USA, ISBN 0-8039-2033-4 and 0-8039-2034-2.

Tate, R.F.; The *theory of correlation between two continuous variables when one is dichotomized,* Biometrika, 1955, Vol 42.

Torgerson, W.S.; *Theory and Methods of Scaling,* J. Wiley USA, (1958), New York, USA.