Application of biostatistical methods in neuropsychiatric examinations

Ph.D. Thesis

Krisztina Boda

Szeged, 1999

Department of Medical Informatics
Albert Szent-Györgyi Medical University
Szeged, Hungary

Application of biostatistical methods in neuropsychiatric examinations

Ph.D. Thesis

Krisztina Boda

Szeged, 1999

CONTENTS

"Multivariate analysis: a means of finding the answer when you don't know the question."

Stephen Senn: Statistical Issues in Drug Development,
John Wiley & Sons, 1997, p. 146.

"Computers have revolutionised the field of multivariate analysis because they can handle hundreds of variables quickly and in real time. With multiple measurements taken on a system, we have a better basis for understanding and modelling the system, and we can hereby minimise the errors in decision making."

C.R. Rao, Director of the Center for Multivariate Analysis,
http://www.stat.psu.edu/department/grad_handbook/centers/cma.html

PUBLICATIONS OF THE AUTHOR

THE THESIS IS MAINLY BASED ON THE FOLLOWING PUBLICATIONS OF THE AUTHOR

I.

Boda, K. and Pap, Á. 'Diagnostics of pancreatic insufficiency using multivariate statistical and pattern recognition methods', *Computers in Biology and Medicine,* **14**, 91-97 (1984).

II.

Pap, Á. and Boda, K. 'Complex evaluation of secretin pancreosymin test data by multivariate statistical pattern recognition methods', *International Journal of Pancreatology,* **1**, 205-212 (1986).

III.

Lajkó, K. és Boda, K. 'Pánikbetegek környezeti stressz-hatásainak vizsgálata', *Ideggyógyászati Szemle*, **42**, 409-420, (1989).

IV.

Molnár, J., Mucsi, I., Mándi, Y., Béládi, I., Földeák, S. and Boda, K. 'The effect of rutin-n-mustard on the survival of nk/ly ascites tumor bearing mice', *International Journal of Cancer,* **30**, 767-771 (1982).

V.

Gyöngyösi, M., Takács, T., Czakó, L., Jambrik, Z., Boda, K., Farkas, A., Forster, T. and Csanády, M. 'Non-invasive monitoring of haemodynamic changes in acute pancreatitis in rabbits', *Digestive Diseases and Sciences,* **42**, 955-961 (1997).

VI.

Boda, K., Tóth, K. és Boncz, I. 'Többváltozós statisztikai módszerek alkalmazása klinikai adatok kiértékelésére', *NJSZT Kollokvium Kiadványa* (Szerk.:Kozmann, Gy.), NJSZT, Budapest, 53-56 (1996).

INTRODUCTION

Statistics may be defined as a body of methods for learning from experience – usually in the form of numbers from many separate measurements displaying individual variations. Due to the fact that many non-numeric concepts, such as male or female, improved or worse, etc., can be described as counts, rates or proportions, the scope of statistical reasoning and methods is surprisingly broad. Nearly all scientific investigators find that their work sometimes presents statistical problems that demand solutions; similarly, nearly all readers of research reports find that the understanding of the reported results of a study requires an understanding of statistical issues and of the way in which the investigators have addressed those issues.

One characteristic of medical and biological research is that the examinations result in data generally reflected by numbers. Biostatistics provides methods that permit a description and summarisation of such so that consequences may be drawn from them. Biostatistics is an application of mathematical statistics to the evaluation of biological and medical experimental data. It is based on probability theory and mathematical statistics.

Biostatistical methods are widely used in medical research. A scientific paper without such an evaluation is currently almost inconceivable. Moreover, the number of medical papers is increasing very rapidly year by year, while the evaluation of the experiments reported requires increasingly more sensitive methods. Meanwhile, the spreading of up-to-date knowledge is rendered more difficult by the specialisation at present going on throughout the medical profession. It is obvious, therefore, that the quality of statistical methods used in the medical literature has been criticised by a number of authors (Hand, 1985; Rosenthal and Rubin, 1985; Williamson *et al,* 1986; Campbell and Machin, 1993; Seldrup, 1997). All these statements are particularly valid as concerns neuropsychiatry, where the use of statistical methods is often necessary (Ekstrom *et al*, 1990).

The aim of the present work is to interlink two apparently well-separated fields of research: biostatistics and neuropsychiatry. In accordance with this purpose, the two main parts of the work are as follows:

I. The first part presents a short description of the basic biostatistical concepts and methods, and especially the methods used in our experiments. This part provides an overview, beginning with the basic principles of hypothesis tests, followed by several models of analysis of variance (ANOVA) and multivariate statistical methods. The problem of "multiple comparisons" and the increase in Type I error will be demonstrated: the repeated use of a two-sample statistical test for the comparison of several samples may result in incorrect decisions concerning significant differences. To avoid this mistake, corrections (e.g. the *Bonferroni* correction) or special statistical methods must be used. ANOVA is the common name of methods whereby the means of several normally distributed populations are compared. ANOVA has several models. There are models for the comparisons of independent groups (one- and two-way ANOVA), and for related groups, called repeated measurements. Finally, multivariate statistical methods will be described: linear discriminant analysis, which can be regarded as a generalisation of the two-sample t-test or one-way ANOVA, but it can also be used for classification and prediction. In this sense, it is also a statistical pattern recognition method, together with logistic regression. In the presentation of biostatistical methods, neuropsychological relations and examples will be given when appropriate.

II. The second part describes new results as an application with data from two neuropsychiatric experiments. Several such studies that we have published will be cited briefly (Molnár *et al,* 1982; Boda, 1984; Boda and Pap, 1984, Lajkó and Boda, 1989; Kálmán *et al,* 1991; Boda, 1997, Pető *et al*, 1997; Pikó *et al,* 1997). Apart from the demonstration of the application of biostatistical methods in neuropsychiatry, an example will be presented of how to acquire evidence-based results.

The first neuropsychiatric experiment related to the examination of panic disorder on the basis of self-reported questionnaire data. Summarised scores of eleven questionnaires were compared in four independent groups of subjects: control, panic disorder, alcoholic and insane subjects. The advantage of the use of linear discriminant analysis will be illustrated as compared with the univariate two-way ANOVA methods.

The second neuropsychiatric problem was the examination of data furnished by the Jung's word association testing of control and attempted suicide subjects in an alert state and under hypnosis. During the test, 102 words are read by the experimenter to the patient, separately, one after another, and the patient has to be respond to each word with a (another) word, as soon as possible. The reaction time is measured, and the associations are qualified. The performance of this test yields 402 measurements on each subject. The large number of variables leads to difficulty in statistical evaluation. After univariate ANOVA, linear discriminant analysis and logistic regression analysis were used, or univariate methods were applied to summary measures (mean, maximum, median, etc.) computed from the original data. The aim of this analysis was more complex: the effects of hypnosis and the differences between groups were examined. The psychiatric assumption that a pre-suicidal syndrome is similar to a modified state of consciousness characteristic of hypnosis could be demonstrated via the results of different statistical tests.

I. BASIC BIOSTATISTICAL CONCEPTS, UNIVARIATE AND MULTIVARIATE BIOSTATISTICAL METHODS

1.1. Basic concepts and hypothesis tests

Biostatistics is an application of mathematical statistics to the evaluation of biological and medical experimental data. It is based on probability theory and mathematical statistics. Several books provide introduction to biostatistics (e.g. Hajtman, 1971, Altman, 1995); here, only the most important basic concepts will be described.

Probability

In everyday life probability is often expressed as a percentage. In mathematical statistics, probability is a more general concept, defined by axioms. For practical reasons, the probability of some specific outcome of a random event is defined as the proportion of the number of times that this outcome will occur in a very long series of repetitions. Consequently, the value of the probability always lies between 0 and 1.

Population, sample

The purpose of experiments is generally to acquire knowledge about some phenomenon and to establish relationships or differences. If the purpose of an experiment is to learn some characteristic of human beings, we could attain a true picture of this characteristic if we were able to examine the characteristic in every human being of that examination. This is generally impossible because of time limits and for financial reasons. Accordingly, we can observe only a relatively few, arbitrarily chosen individuals, and utilise the data on them to draw conclusions as to the general characteristic. In mathematical statistics, the data on the observed units are called the statistical sample, while the total (perhaps infinite) number of experimental units is called the population. There are several methods of selecting elements randomly from a population in order to obtain a sample (sampling with or without replacement, sequential sampling, etc.). In a random sample, the elements are chosen from the population in such a way that each member of the population has an equal chance of being chosen.

Variable

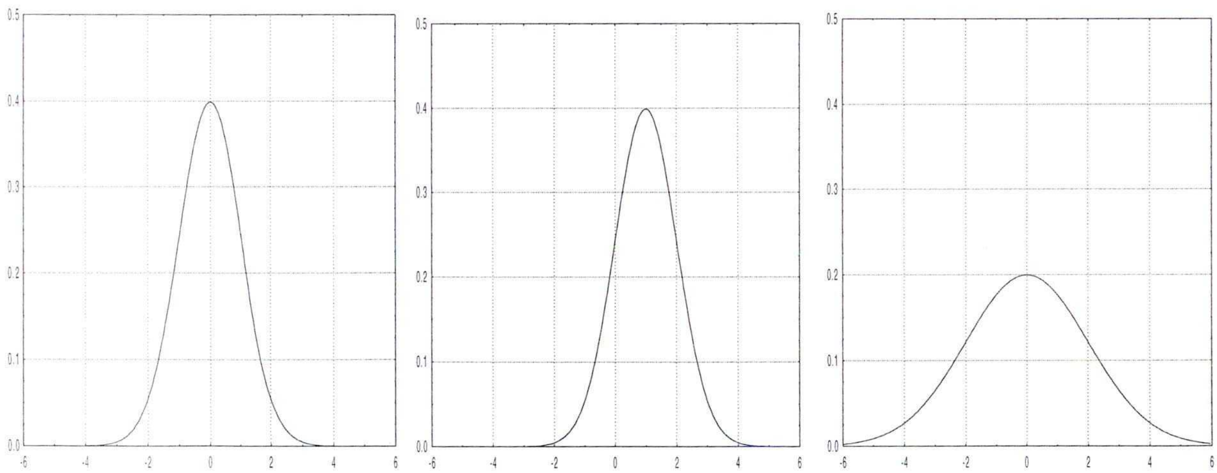
Experimental units (in our case, human beings) are represented by their properties. The examined characteristics of individuals are called variables. A variable can assume different values for different individuals. Variables can be categorical (discrete) or continuous. A categorical variable can assume only finite possible values; a continuous variable can assume an infinite number of values.

Distribution

The distribution of a variable indicates what values it assumes and how often it assumes these values. The distribution of a variable (like the distribution of a population) is generally not known, because we are unable to measure every element of the population. We can merely approximate to the distribution of the population on the basis of the sample. The larger the sample size, the more precise the approximation.

Parameter, normal distribution

There are special theoretical distributions that can be given by formulae. In such cases, constants (the parameters) in the formula characterise the form of the distribution. One such special distribution is the normal or Gaussian distribution, denoted by $N(\mu, \sigma)$ which has two parameters: the mean (μ) and the standard deviation (σ). Figure 1.1 a)-c) depicts the curves of the normal distribution for different parameters. It can be seen that the mean and the standard deviation describe the centre and the spread of the distribution, respectively. In practice, we do not know the exact values of these parameters; they must be approximated to from the sample.



a) b) c)

Figure 1.1. The curve of the normal distribution for different parameters
a): μ=0, σ=1; b) μ=1, σ=1; c) μ=0, σ=2

Mean

Let us denote the sample size by n, and the sample by $y_1, y_2, \ldots, y_n$. The average of these numbers, computed in the ordinary way, approximates to the population mean.

Standard deviation, variance

The standard deviation expresses the dispersion of the data around the mean. Its square is called the variance. The formula of the variance contains the squared difference of the sample elements from the mean in the numerator and n-1 in the denominator. The numerator is also called the sum of squares and the denominator is also called the number of degrees of freedom.

Estimation

The approximated value of a parameter, computed from the sample, is called the estimate of that parameter. For example, the arithmetic mean is the estimate of the population mean. The goodness of estimation can be given by several ways (bias, effectiveness, etc.); these properties will not be discussed here. The larger the sample size, the better the estimation.

Confidence interval

In most statistical applications it is not sufficient to say that the correct value of a parameter is more or less well estimated. When we give a range of values that we think includes the true value of some population parameter, this range of values is called an interval estimate. Such an interval is usually assigned a probability, and it is then called a confidence interval. The higher the probability assigned, the more confident we are that the interval does, in fact, include the true value. The high probability is called the confidence level; its usual values are set at 0.90, 0.95 or 0.99. The higher the sample size and the smaller the confidence level, the narrower the confidence interval.

α

Instead of a "high" probability of confidence we can use a "small" probability of the error, i.e. uncertainty. This "small" probability is generally assigned to the value 1-0.95=0.05 or 1-0.99=0.01 and denoted by α.

Confidence interval for a population mean and the standard error

The standard deviation divided by the square root of the sample size is called the standard error or standard error of the mean. The standard error is used in the computation of the confidence interval for a mean of normal distribution. The interval that contains the true population mean with 95% probability can be computed in such a way that we add to the mean and subtract from the mean twice the standard error. (More precisely, we do not use exactly the twice the standard error but a multiple close to two, called a *t*-value, which can be found in tables of the *t*-distribution. It depends on the number of degrees of freedom and α).

Statistical inference: hypothesis testing

Statistical analysis is concerned not only with summarising data, but also with investigating relationships and drawing conclusions. The majority of statistical analyses involve comparisons, generally between treatments or procedures or between groups of subjects. Conclusions are expressed in statements. A statistical hypothesis is a statement about a population. During the hypothesis tests, we examine whether the statement of interest is true or not. These examinations are called hypothesis tests. In consequence of the great variety of questions of interests and of variables, there can be many hypotheses, which can be examined by means of many different statistical tests. In spite of this variety, the steps of every hypothesis tests are the same.

Null hypothesis, alternative hypothesis

Statisticians usually test the hypothesis, which tells them what to expect by giving a specific value to work with. They refer to this hypothesis as the null hypothesis and symbolise it as H_0. The null hypothesis is often the one that assumes fairness, honesty or equality. The opposite hypothesis is called the alternative hypothesis and is symbolised by H_a. This hypothesis, however, is often the one that is of interest. Some statisticians refer to H_a as the motivated hypothesis.

Steps in a hypothesis test, p-value, significance

The hypothesis test begins with the formulation of two opposing hypotheses and ends with a decision in favour of one of them. The population or its parameters are not known, but it is desirable to make a decision with the smallest possible probability of error. This error is generally fixed as being equal to or less than 0.05 or 0.01, and it is denoted by α (being the probability already mentioned in connection with the confidence intervals). The test starts from the fact that the null hypothesis is true. According to the null hypothesis, a test-statistics calculation is performed on the sample data. On the basis of the test statistics, it is possible to calculate the probability that at least as high test statistics as that observed would occur if H_0 were true. This probability is called the *p*-value or the observed significance level. If this probability is "small" (smaller than α), then we decide to reject the null hypothesis and we say that the difference is significant at the $p<\alpha$ level. The probability of error in this case is α.

Non-significant result

If the *p*-value is large (larger than α), then we do not have a sufficient basis to reject the null hypothesis; thus, we decide to accept it, and say that the difference is not significant, $p>\alpha$; we fail to reject the null hypothesis. In this case, we decide in favour of equality, but in reality this does not always mean equality. It is possible that there really is a difference, but that we do not have sufficient information to reveal it.

Statistical errors

One basic fact which is inseparable from hypothesis testing is that there can never be absolute proof as to which of the two hypotheses is the true one. When we are testing a null hypothesis, we are trying to decide whether it is true or false. However, since statistical hypothesis testing is based on sample information and we cannot be absolutely sure that our decision is correct, we are actually faced with four possible situations.

1. H_0 is true, and the sample data lead to the correct decision that it is true.
2. H_0 is true, but by bad luck the sample data lead to the mistaken decision that it is false.
3. H_0 is false, and the sample data lead to the correct decision that it is false.
4. H_0 is false, but sample data lead to the mistaken decision that it could be true.

In the first and third situations above, the data lead to a correct decision. In the second situation, the true null hypothesis is mistakenly rejected. We refer to this as a **Type I error**. Its probability is just α. In the fourth situation, we fail to reject a false null hypothesis. Statisticians call this a **Type II error**. Its probability is denoted by β. A specific value of β cannot be computed until we have decided upon a specific value from the alternative hypothesis. Statisticians refer to the value $1-\beta$ as the **power** of the test. The power of a test is a measure of how good the test is at rejecting the false null hypothesis. The more "powerful" a test is (the closer is the value of $1-\beta$ to 1), the more likely the test is to reject a false null hypothesis.

Assumptions of statistical tests, robustness

Statistical tests generally involve assumptions. If the result of a statistical test does not change when these assumptions are slightly violated, then the test is said to be robust. For example, the two-sample *t*-test, which compares the means of two independent populations, involve the assumption that samples are drawn from normally distributed populations with equal variances. However, a small departure from normality does not affect the result of the *t*-test: it is robust.

Correlation coefficient (r), the significance of the correlation

The correlation coefficient (r) measures the linear relationship between two variables. Its value always lies between –1 and 1. If we put the values of the two variables on a co-ordinate system, we obtain one dot for each pair of values. Let us draw a graph called a scattergram to investigate this relationship. When the points are arranged in an approximately straight line, we say that there is a good linear correlation between the two variables. When there is no tendency for the points to lie in a straight line, we say that there is no correlation ($r=0$) or only a low correlation (r is near 0). If r is near +1 or –1, we say that we have a high correlation. If $r=1$, we say that there is perfect positive correlation. If $r=-1$, then we say that there is a perfect negative correlation.

We can test whether the observed correlation could have arisen by chance or not. This test gives a t-value and a p-value of the correlation. The assumption underlying the test of significance is that both variables are random samples and at least one has a normal distribution.

The problem of multiple comparisons

We often need to compare several population means or compute several correlation coefficients in the same study. Comparisons might be performed by multiple use of two-sample t-tests, or by the computation of several coefficients of correlation, resulting in several p-values. The expected number of p-values smaller than 0.05 is 1 in 20 tests of the true null hypotheses; the probability that at least one p-value will be smaller than 0.05 therefore increases with the number of tests, even when the null hypothesis is correct for each test. This increase is known as the "multiple-comparisons" problem. When several tests are performed simultaneously, both the calculation and the interpretation of the p-values must be re-examined. Figure 1.2 shows the increase in the probability of error for independent tests, based on *Bonferroni* inequalities (Srivastava and Carter, 1983).

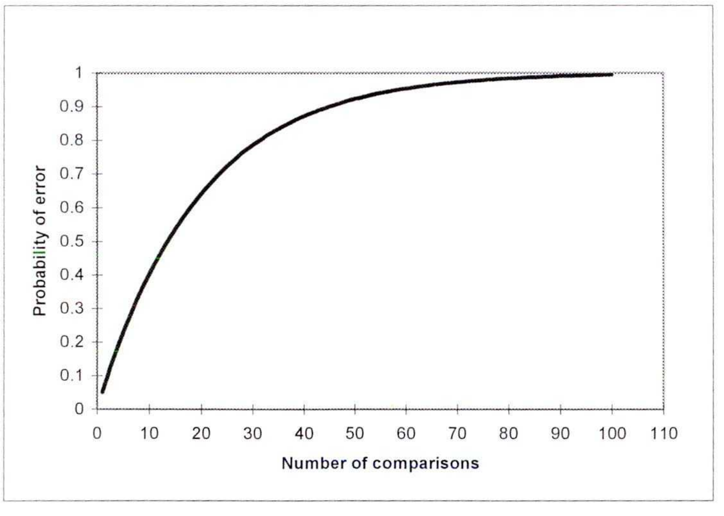


Figure 1.2. Upper limit of error probability with increasing number of comparisons

1.2. Comparison of population means of several normal distributions: ANOVA

The analysis of variance (ANOVA) is a method for the comparison of the means of several, normally distributed random variables. The name comes from the basic principle of the method: the variability (variance) of the overall data can be computed in several ways, and the numerator of the total variance (sum of squares) can be broken down into independent components.

ANOVA is a method for multiple group comparison. It can be used in several experimental designs (Fleiss, 1986). Such an experiment is carried out to study the effects of one or more independent variables on a dependent variable. Independent variables are called factors. The simplest cases of ANOVA methods are the paired and the independent samples *t*-test. The simplest generalisation of the independent sample *t*-test is one-way ANOVA.

1.2.1. One-way ANOVA and multiple comparisons

Let us suppose that we have *t* independent samples (*t* "treatment" groups) drawn from normal populations with equal variances. In other words, let us suppose that we wish to study the effect of one qualitative factor with *t* levels on a dependent variable.

The null and alternative hypotheses are as follows:

H_0: $\mu_1 = \mu_2 = ... = \mu_t$, i.e. the population means are equal, and the samples are drawn from the same population.

H_a: $\mu_i \neq \mu_j$, $1 \leq i, j \leq t, i \neq j$, i.e. the population means are not all equal.

Method: If the null hypothesis is true, then the populations are the same: they are normal, and they have the same mean and the same variance. We will estimate the numerical value of this common variance in two distinct ways: we will compute the "between-groups variance" and the "within-groups variance". If the null hypothesis is true, then these two distinct estimates of the variance should be equal, their equality can be tested by an *F* ratio test. The fact that we can compute variances in two ways stems from the break-down of the total sum of squares into a "between-groups sum of squares" and a "within-groups sum of squares". The total number of degrees of freedom *N*-1, where *N* is the sum of all sample sizes, is also broken down into the appropriate "between-groups" and "within-groups" degrees of freedom: *t*-1 and *N*-*t*, respectively. The results of computations of ANOVA methods are usually tabulated. The rows of such tables give the source of the variance; the columns contain the sum of squares, the number of degrees of freedom, the variances, the *F*-value (variance ratio), and the *p*-value. If the result of the ANOVA is not significant at the specified level, the analysis is complete. We expect all samples be drawn from the same population; the differences between sample means are due to random effects.

Multiple comparisons

If the result of the ANOVA is significant, then we have to accept the alternative hypothesis: there is at least one group different from one of the others. To find these groups, we have to compare each group with each of the others. As the two-sample *t*-test is inappropriate to do this, there are special tests for multiple comparisons that keep the probability of Type I error as α. The most often used multiple comparisons are the modified *t*-tests. Another often used method is the so-called *Bonferroni* method: to achieve a level of not more than α for a set of a number of *c* tests, we need to choose a level α/c for the individual tests. For example for three comparisons a *p*-value less than 0.05/3=0.017 has to be considered significant instead of $p=0.05$. This method is conservative. We know only that the probability does not exceed α for the set. This method can be used in cases involving small numbers of comparisons.

Another multiple comparison method is the *Dunnett* test: a test comparing a given group (control) with the others. The *Scheffé* test performs simultaneous joint pairwise comparisons for all possible pairwise combinations of means. This test can be used to examine special linear combinations of group means, not simply pairwise comparisons.

1.3. Structure of multivariate data, general requirements

Multivariate statistical methods provide a possibility for the analysis of complex sets of data. Multivariate methods can be applied to data sets in which one individual is measured as concerns several characteristics or several times. The data set is not one sequence of numbers, but a table of numbers, in which the rows contain data on individuals, while the columns contain characteristics on different individuals. The variable is called a vector variable. Table 1.1 shows the arrangement of multivariate data. This is a general way to input data into several computer systems.

We generally follow the convention of denoting random variables by upper-case italic letters and observed values by the corresponding lower-case letters; for example, x_{ij} denotes the observed measurement of the *i*-th individual on the *j*-th variable X_j. Variables can be independent (explanatory, predictor) or dependent (outcome) variables, dependent variables are generally denoted by *Y*. Statistical methods depend on the number and measurement scale of the independent and dependent variables. Univariate methods examine variables alone, whereas multivariate methods analyse the total data set simultaneously.

Table 1.1. Arrangement of multivariate data

	X_1	X_2	...	X_j	...	X_k
Individual 1						
individual 2						
...						
individual *i*				x_{ij}		
...						
individual *N*						

There are some requirements or desirable properties as concerns the data set: several multivariate methods require multivariate normal (see section 1.5), or at least symmetric distribution; it is desirable that no data should be missing; variables that are constant multiples of each other should be avoided. As a rule of thumb, there should be approximately 10 times more cases than variables for good results; nevertheless, at a bare minimum, 3 times more cases than independent variables may be used.

If the number of variables is too large, there are methods for the reduction of dimension (factor or cluster analysis, Sváb, 1979; Srivastava, 1983; Dobson, 1991; Altman, 1995).

1.4. Multiple linear regression

Most multivariate methods are based on multiple regression. Regression analyses are a set of statistical techniques, which allow an assessment of the relationship between one dependent variable (Y) and several independent variables (X_1, X_2, ..., X_k). Let y_1, y_2,..., y_N denote the N observed values of Y. The multiple linear regression model can be written as

$$y_i = \mu_i + \varepsilon_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i, \qquad i=1,\ldots,N$$

Here

- $\mu_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}$ is the mean value of the dependent variable when the values of the independent variables are $x_{i1}, x_{i2}, \ldots, x_{ik}$.
- ε_i is an error term that describes the effects on y_i of all factors other than the values of $x_{i1}, x_{i2}, \ldots, x_{ik}$.
- β_0 is the mean value of the dependent variable when the values of the independent variables are zero.
- β_j are the regression coefficients; they indicate the change in the mean value of the dependent variable that is associated with a one-unit increase in the j-th independent variable when the other independent variables do not change.

Here, the Greek letters are used to denote the parameters of the model. The goal is then to derive estimations of the β_j values, given a series of observations. We label these estimates B_j. If we had these estimates, for any subject i, we could predict for each x_{i1}, x_{i2},…,x_{ik}, the value y_i by

$$\hat{y}_i = B_0 + B_1 x_{i1} + B_2 x_{i2} + \ldots + B_k x_{ik}, \qquad i=1,\ldots,N$$

Clearly, we would like to choose B_0, B_1,…,B_k so that y_i and $\hat{y}_i$ are close and hence make our prediction error small. This can be done by minimising the sum $\sum (y_i - \hat{y}_i)^2$. This leads to calling B_0, B_1,..., B_k the least square estimation of the population parameters β_0., β_1.,..., β_k. The differences between predicted and obtained scores, $y_i - \hat{y}_i$, are called residuals.

The assumptions of the analysis are that the residuals are distributed normally around the predicted dependent variable scores, that the residuals are in a linear relationship with the predicted dependent variable scores, and that the variance of the residual about the predicted score is the same for all predicted scores.

Regression may be assessed in a variety of ways, such as:

1. Partial regression and correlation (This isolates the specific effect of a particular independent variable controlling the effects of other independent variables.).
2. Multiple regression and correlation (This is the combined effect of all the variables acting on the dependent variable; for a net, combined effect. The resulting R^2 value provides a reflectance of the goodness of fit of the model).

The regression coefficients reflect the importance of the variable. However, these values are contingent on the other independent variables in the equation. They are also affected by the correlation of the independent variables.

Although regression analyses reveal relationships between variables, this does not imply that the relationships are causal. Demonstration of causality is not a statistical problem, but an experimental and logical issue.

The "optimum" number of independent variables

There are often many independent variables, and it would be desirable to find variables which yield a good estimate of the dependent variable. Although R^2 increases with each independent variable, this does not mean that a model that contains many independent variables provides a better fit to the population. As with a higher number of independent variables than of cases, a regression solution can be found which perfectly predicts the dependent variable for each case. The null hypothesis that the true population value for the change in R^2 is 0 can be tested by using an F statistics also referred to as a partial F test), and it can be formulated in terms of the parameters β.

The observed increase in R^2 does not necessarily reflect a better fit of the model in the population. The inclusion of a large number of independent variables in a regression model is never a good strategy. A model with many variables is difficult to interpret. On the other hand, the inclusion of fewer independent variables than in the true model in the population may result in a poor prediction of the dependent variable. There are several methods to find the optimum number of independent variables. Although there are procedures for computing all possible regression equations, several other methods, the stepwise methods do not require as much computation and are used more frequently.

Stepwise methods

These procedures are forward selection, backward elimination and stepwise selection.

Forward selection

The method can be broken down into a few simple steps:

a) Find the single variable that has the strongest association with the dependent variable according to an entry criterion, and enter it into the model.
b) Find the variable among those not in the model that, when added to the model so far obtained, explains the largest amount of the remaining variability.
c) Repeat step b) until the addition of an extra variable is not statistically significant at some chosen level (entry criterion).

Backward elimination

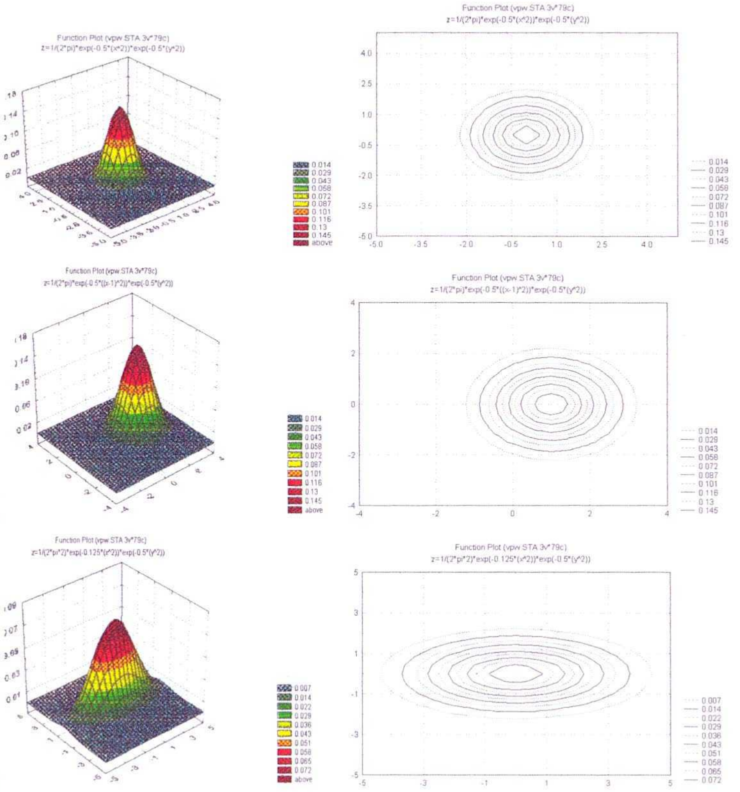
This method starts with all variables in the equation and sequentially removes them. Instead of entry criteria, removal criteria are used.

Stepwise selection of independent variables is really a combination of the backward and forward methods and is probably the most commonly used method. The first variable is selected in the same manner as in forward selection. After the first variable is entered, it is examined to see whether it should be removed according to the removal criterion as in backward elimination. In the next step, variables not in the equation are examined for entry. After each step, the variables in the equation are examined for removal. Variable selection terminates when no more variables meet the entry and removal criteria.

1.5. Multivariate normal distribution

Multivariate normal distribution plays an important role in the multivariate statistics, similarly to the univariate case. Experimental data often are normally distributed, and many statistical methods require normal distribution.

Measurements made on the variable *Y* can be regarded as co-ordinates of an *n*-dimensional vector. If the variables $Y_1, Y_2, \ldots, Y_k$ are normally distributed ($Y_i \sim N(\mu_i, \sigma_i)$) and independent, then the joint distribution of Y_i and Y_j is a multivariate normal distribution. For two variables, the formula of this distribution is $f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}} e^{-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}}$ and its grapf can be seen in Figure 1.3.



$\sigma_1=1$, $\sigma_2=1$, $\mu_1=0$, $\mu_2=0$

$\sigma_1=1$, $\sigma_2=1$, $\mu_1=1$, $\mu_2=0$

$\sigma_1=2$, $\sigma_2=1$, $\mu_1=0$, $\mu_2=0$

Figure 1.3. The bivariate normal distribution for different values of σ_1, σ_2, μ_1 and μ_2

1.6. The general linear model

The general form of the linear model is $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \varepsilon$, or with matrix notations, $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$, where

$\boldsymbol{Y}$ is an Nx1 response vector,

$\boldsymbol{X}$ is an Nxk matrix of constants ("design" matrix), which are mainly values of 0 or 1, and values of independent variables,

$\boldsymbol{\beta}$ is a kx1 vector of parameters, and

$\boldsymbol{\varepsilon}$ is an Nx1 random vector whose elements are independent and all have normal distribution $N(0, \sigma)$.

Estimations of the regression coefficients β_j can differ when several models are fitted to data. Moreover, the test of the hypotheses $\beta_j=0$ depends on which terms were included in the model. Estimates, confidence intervals and hypothesis tests usually depend on which variables are included in the model. There is an exception when matrix X is orthogonal. In that case hypotheses H_1: $\beta_1=0$,...,H_t: $\beta_t=0$ can be tested independently.

Orthogonality is perfect non-association between variables. Independence of variables is desired so that each addition of an independent variable adds to the prediction of the independent variable. If the relationship between independent variables is orthogonal, the overall effect of an independent variable may be partitioned into effects on the dependent variable in an additive fashion.

Unfortunately, orthogonality can be utilised only if we can design the matrix X to have this property. This can be achieved by designing special comparisons of groups or by the use of orthogonal polynomials.

1.7. Models of ANOVA

ANOVA can be modelled by the general linear model.

1.7.1. Model of one-way ANOVA and its relation to multiple linear regression

Let us consider the data in Table 1.1. The model of one-way ANOVA can be written in the following form:

$$y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i = 1,\dots,t, \quad j = 1,\dots,n_t$$

where y_{ij} denotes the i-th element of the j-th sample, μ denotes the "overall population mean", α_i denotes the effect of the i-th treatment, and ε_{ij} denotes the random error, which is assumed to have $N(0, \sigma)$ distribution. The null hypothesis H_0: $\mu_1 = \mu_2 = \dots = \mu_t$ that all population means are equal now corresponds to the null hypothesis that H_0: $\alpha_1 = \alpha_2 = \dots = \alpha_t$.

This is a linear model and it can be rewritten in a form of a linear regression:

$y_{ij} = \mu_i + \varepsilon_{ij} = \beta_0 + \beta_1 D_{i2} + \beta_2 D_{i3} + \dots + \beta_{t-1} D_t + \varepsilon_{ij}$, $i=1,2,..,t$, $j=1,2,\dots, n_t$,

where the D_i-s are "dummy" variables formed from the independent variables, for example, in the following way:

Let the first group be a "reference" group. Then, let

$D_{i2}=1$ if an observation belongs in group 2; otherwise let $D_{i2}=0$.

$D_{i3}=1$ if an observation belongs in group 3; otherwise let $D_{i3}=0$.

...

$D_{it}=1$ if an observation belongs in group t; otherwise let $D_{it}=0$.

Then, if an observation belongs in group 1, $\mu_1 = \beta_0 + \beta_1(0) + \beta_2(0) + \dots + \beta_{t-1}(0)$, i.e. $\mu_1 = \beta_0$.

If an observation belongs in group 2, then $\mu_2 = \beta_0 + \beta_1(1) + \beta_2(0) + \dots + \beta_{t-1}(0) = \beta_0 + \beta_1$; hence $\mu_2 = \mu_1 + \beta_1$, and $\beta_1 = \mu_2 - \mu_1$.

Similarly, the other coefficients are $\beta_1 = \mu_2 - \mu_1$, $\beta_{t-1} = \mu_t - \mu_1$; i.e. regression coefficients are estimates of the differences between group means.

The test of the null hypothesis H_0: $\mu_1 = \mu_2 = ... = \mu_t$ is equivalent to the test of the hypothesis $\beta_1 = \beta_2 = ... = \beta_{t-1}$=0.

1.7.2. Two-way analysis of variance

In two-way analysis, we wish to assess the effects of two qualitative factors (independent variables) on a dependent variable. We call the groups of a factor the **levels** of that factor. The goal of two-factor analysis is to estimate and compare the effects of the different factors to the dependent variable. Depending on the particular situation, we may wish to learn whether there are statistically significant differences

a) between the effects of the different levels of factor 1,

b) between the effects of the different levels of factor 2, or

c) between the effects of the different combinations of a level of factor 1 and a level of factor 2. Factors 1 and 2 interact if the relationship between the mean response and the different levels of one factor depends upon the level of the other factor.

Let us denote the numbers of levels of factors 1 and 2 by t and l, respectively, and by N the total number of observations. The two-way ANOVA model is:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \Theta_{ij} + \varepsilon_{ijk}, \quad i = 1,...,t \quad j = 1,...,l, \quad k = 1,...,n_{ij}$$

where we use the following notations:

y_{ijk} = the k-th observed value of the dependent variable when we are using level i of factor 1 and level j of factor 2,

μ = an overall mean, (unknown constant),

α_i = the effect due to level i of factor 1 (an unknown constant),

β_j = the effect due to level j of factor 2, (an unknown constant),

Θij = the effect due to the interaction of level i of factor 1 and level j of factor 2 (an unknown constant),

ε_{ijk} = the k-th error term when we are using level i of factor 1 and level j of factor 2 (assumed to be distributed as $N(0, \sigma)$).

According to the above questions, the following null hypothcses can be tested:

a) H_1 : $\alpha_1 = \alpha_2 = ... = \alpha_t$

b) $H_2: \quad \beta_1 = \beta_2 = \ldots = \beta_l$

c) $H_3: \quad \Theta_{ij} = 0$

In two-way ANOVA, the total sum of squares is decomposed into four terms, according to the effects in the model. The results are generally written into an ANOVA table which contains rows for the effects of factors 1 and 2, the interaction and the error with $(t\text{-}1)$, $(l\text{-}1)$, $(t\text{-}1)(l\text{-}1)$ and $(N\text{-}tl)$ degrees of freedom, respectively.

The rows of this tables give the components for the effects of factor 1, factor 2, the interaction and the error term, while the columns contain the sum of squares, the number of degrees of freedom ($(t\text{-}1)$, $(l\text{-}1)$, $(t\text{-}1)(l\text{-}1)$ and $(N\text{-}tl)$), the variances, the F-values (variance ratio), and the p-value of F.

There are three F-values in this table according to the three hypotheses.

Question c), i.e. the interaction, is tested first. If it is not significant, the significance of each of factors 1 and 2 can be tested separately. If H_1 is rejected, we can say that at least two of the factor 1 means differ. If t, the number of levels of factor 1, is more than two, we again have to use multiple comparisons to find pairwise differences.

If the interaction is significant, then the relationship between the means of factor 1 depends on the level of factor 2. Multiple comparisons can be performed for each combination of one factor with a given level of the other factor. There are special methods against the increase of Type I error, and the use of t-tests independently is an incorrect solution.

1.8. Assumptions of ANOVA, transformations

The data are presumed to be random samples from a normal population with the same variances. There are tests to check normality (the Kolmogorov-Smirnov test, or the Shapiro-Wilk test for samples with small numbers of observations). ANOVA is robust to departures from normality, although the data should be symmetric. However, if the assumptions are highly violated, transformations or non-parametric methods can be used.

The distribution of the data is often skewed to the right; this causes departure from normality and different variances in the different groups. In such cases, a logarithmic transformation can normalise the distribution and can homogenise the group variances. Other often used transformations are the square root, the reciprocal and the arcsine transformations.

1.9. ANOVA with repeated measurements

The response to a drug treatment, for example, is often measured several times during or after administration of the drug, the intention being to compare treatments with respect to the trends in their effects over time and with respect to their mean levels of response. A widely used and general term is repeated measures data, which refers to data measured repeatedly on subjects either under different conditions, or at different times, or both. In ANOVA with repeated measurements, the repetition is expressed as a factor in the analysis, called the within-subject factor. Multivariate data refer to the case where the same subject is measured on more than one outcome variable. ANOVA with repeated measurements can be modelled by using a univariate or multivariate approach. The results of the two approaches are not necessarily the same.

Univariate approach (split-plot, mixed model)

The univariate approach regards the dependent variable as response to levels of within-subject factors. Let us suppose that t treatment groups are compared by means of a repeated measurement study, in which each subject is measured b times during or after treatment. The classical ANOVA table will contain the following rows (source of variation): groups, times, interaction, subjects and residual with $(t-1)$, $(b-1)$, $(t-1)(b-1)$, $(N-t)$ and $(N-1)(b-1)$ degrees of freedom, respectively. The test statistics F for testing whether the differences between the mean levels of the t groups are statistically significant is equal to the ratio of the mean squares for the groups and the mean squares for the subjects. The denominator shows that a subject is now characterised by the mean of b repeated measurements and not simply by a single measurement. The sum of squares for time measures whether the mean level of the response varies in any systematic way over time. The sum of squares for the group by time interaction measures the extents to which the patterns of variation over time for the individual groups differ from the overall mean pattern for all groups and thus from one another. In other words, the sum of squares for time measures the degree to which a trend exists in the mean level of response, and the sum of squares for interaction measures the degree to which the t individual trends are not parallel.

Huynh and Feldt (1970) found a general condition under which this model can be used: multivariate normality, and the equality of all $b(b-1)$ correlations between pairs of responses. This condition will rarely be satisfied in a repeated measurements study. There are several corrections in this case, adjustments to the numerator and denumerator degrees of freedom, called epsilon (e.g. Greenhouse-Geisser, 1958).

Multivariate approach

The multivariate approach (multivariate ANOVA, called MANOVA) considers the measurement on a subject to be a sample from a multivariate normal distribution, and the variance-covariance matrices are the same across the cells formed by the between-subject effects (groups). Hypotheses concerning the time factor (whether there is any overall mean trend over time and whether the trends for the t groups are parallel) can be tested by defining matrices. Here, we have matrices in the MANOVA table, and obtain a multivariate F-value (Wilks' lambda) based on a comparison of the effect covariance matrix and the error covariance matrix.

1.10. Discriminant analysis

Discriminant analysis, first introduced by Fisher (1936), is a statistical method for using several variables to help distinguish groups. The usual situation is that we wish to be able to find some combination of variables that classifies a large proportion of subjects into the correct group, so that we can have a good chance of allocating new subjects correctly. Simultaneously, we usually wish to choose for the discrimination a subset of useful variables from a larger set of candidates. Assumptions: each group must be a sample from a multivariate normal population, and the population covariance matrices must all be equal.

The concept of discriminant analysis with two groups

Let us suppose that we have two groups with n_i observations in each group, and let us suppose that we observe k characteristics of each observation, that is, we have k independent variables: X_1, X_2,..., X_k. In discriminant analysis, a linear combination of the independent variables is formed and serves as the basis for assigning cases to groups. Thus, information contained in multiple independent variables is summarised in a single index. The linear discriminant equation has the form

$$Z = w_0 + w_1 X_1 + w_2 X_2 + \ldots + w_k X_k$$

The coefficients w have to be estimated from the data in such a way that the values of the discriminant function differ as much as possible between the groups, i.e. the ratio

$$\frac{between\text{-}groups\ \ sum\ of\ \ squares}{within\text{-}groups\ \ sum\ of\ \ squares}$$

is maximum. The coefficients are solutions of a system of equations, and are called unstandardised discriminant function coefficients. On the basis of these coefficients, it is possible to calculate the discriminant score for each case.

The concept of discriminant analysis with more than two groups

If there are several (t) groups, t-1 discriminant functions can be calculated. The idea is as follows. The first function has the largest ratio of between-groups to within-groups sums of squares. The second function is uncorrelated with the first and has the next largest ratio, etc. These are also called canonical discriminant functions.

Classification of cases

In possession of the discriminant functions, we have several possibilities to classify cases:

1) In the two-group case, we may prepare a histogram from the Z scores, and us find the point at which the overlap is minimum.
2) We may use the Bayes rule for the classification. According to this rule, the probability that a case with a discriminant score of Z belongs in group i is estimated by means of

$$P(G_i \mid Z) = \frac{P(Z \mid G_i)P(G_i)}{\sum_{i=1}^{k} P(Z \mid G_i)P(G_i)}$$

where

$P(G_i)$ is the prior probability of group i, i.e. the likelihood that a case belongs in a particular group when no information about it is available. The prior probability can be estimated in several ways, from earlier knowledge, or on the basis of sample sizes.

$P(Z|G_i)$ is the conditional probability of Z given the group. To calculate this probability, the case is assumed to belong in a particular group, and the probability of the observed score is estimated.

$P(G_i|Z)$ is the posterior probability of group i, given Z. This is an estimate of how likely membership in group i is, given the available information. It can be computed by using the Bayes rule. A case is classified in the group for which the posterior probability is the largest.

Classification results

A table can be constructed to demonstrate the relationship between the original groups and the groups of classifications.

Estimating misclassification rates

The probability of incorrect classification can be estimated as the proportion of correctly classified cases to the total number of cases. This is an inflated estimate of the true performance in the population, since the same cases are used both for estimating and testing the function.

If the sample size is large enough to be randomly split into two parts, then we can use one part to determine the discriminant function (learning set) and the other part to test it (test set).

Another technique for obtaining an improved estimate of the misclassification rate is the jack-knife method. This involves leaving out each of the cases in turn, calculating the function based on the remaining cases, and then classifying the omitted case.

Other discriminant function statistics

1) A "good" discriminant function is one that has much between-groups variability as compared with within-groups variability. Thus, after an ANOVA of the discriminant scores, we can get an "eigenvalue" to each discriminant function

$$eigenvalue = \frac{between\text{ - }groups\ sum\ of\ squares}{within\text{ - }groups\ sum\ of\ squares}$$

Large eigenvalues are associated with good functions.

2) According to the eigenvalues, the proportion of total between-groups variability attributable to each function can be calculated. The first function always has the largest between-groups variability. The remaining functions have successively less between-groups variability.

3) Wilks' lambda is the ratio of the within-groups sum of squares to the total sum of squares. Small values of lambda are associated with functions that display much variability between groups and little variability within groups. A lambda value of 1 occurs when the mean of the discriminant scores is the same in all groups.

4) A test of the null hypothesis that in the population from which the samples are drawn there is no difference between the group means can be based on Wilks' lambda. Lambda is transformed to a variable that has approximately χ^2 distribution.

Relationship to multiple linear regression

Two-group linear discriminant analysis is closely related to multiple linear regression analysis. If the binary grouping variable is considered the dependent variable of a multiple regression (Y=0 for the first and Y=1 for the second group), the multiple regression coefficients are constant multiples of the discriminant coefficients. This is true only for two-group discrimination.

Relationship to two-sample t-test and one-way ANOVA

Linear discriminant analysis is a generalisation of the two-sample *t*-test and one-way ANOVA, if it is regarded as a test of the null hypothesis that the group means are equal. If there is only one independent variable in discriminant analysis, then the *p*-value of Wilks' lambda is approximately equal to the *p*-value of the *t*-test (two groups) or of one-way ANOVA (several groups).

Checking assumptions of discriminant analysis

The assumptions of discriminant analysis can be checked. It is generally agreed that some departure from multivariate normal distribution is acceptable. If assumptions are highly violated, e.g. there are categorical or binary variables between the independent variables, then the use of logistic regression may be better.

1.11. Logistic regression

Logistic regression is useful for situations in which we wish to be able to predict the presence or absence of some outcome (such as being suicidal or not) on the basis of the values of a set of predictor variables. It is similar to a linear regression model, but is suited to models where the dependent variable is categorical and most usually dichotomous. This model requires far fewer assumptions than in discriminant analysis.

Several multivariate statistical techniques can be used to predict a dependent variable from a set of independent variables, e.g. multiple regression or discriminant analysis. However, these techniques pose difficulties when the dependent variable can have only two values: an event occurring or not occurring. When the dependent variable can have only two values, the assumptions necessary for hypothesis testing in regression analysis are necessarily violated. For example, it is unreasonable to assume that the distribution of errors is normal. Another difficulty with multiple regression analysis is that predicted values cannot be interpreted as probabilities. They are not constrained to fall in the interval between 0 and 1. Linear discriminant analysis does allow the direct prediction of group membership, but the assumption of the normality of the independent variables, and of equal variance-covariance matrices in the two groups, is required. Logistic regression is applicable to a broader range of research situations than discriminant analysis.

Example. What lifestyle characteristics are risk factors for suicide? Given a sample of subjects measured with regard to family pattern, reaction time, associations, etc., a model could be constructed by using these variables to predict the likelihood of suicide in a sample of subjects. The model could then be used to derive, for example, the degree to which people with abnormal associations are more likely to be suicidal than people with normal associations.

Odds, odds ratio

If p denotes the probability of an outcome (e.g. the probability of being suicidal), the odds of the same outcome are $p/(1-p)$. The logarithm of the odds is called log-odds. The odds ratio is the ratio of the odds of an outcome when a particular predictor is, say 1, compared with the odds of the outcome when that predictor equals, say, 0. The odds ratio, always predicted by a particular independent variable, is the increase (or decrease) in the odds of the outcome when that independent variable increases by 1.

The model of the logistic regression is written in terms of the logarithm of the odds ratio of an outcome:

$$L = \log(\frac{p}{1-p}) = B_0 + B_1 X_1 + B_2 X_2 + ... + B_k X_k$$

This can be rewritten as

$$Prob(event) = p = \frac{e^{B_0+B_1X_1+B_2X_2+...+B_kX_k}}{1+e^{B_0+B_1X_1+B_2X_2+...+B_kX_k}} = \frac{1}{1+e^{-(B_0+B_1X_1+B_2X_2+...+B_kX_k)}}$$

The first form exhibits similarity with multiple regression, but the method for the estimation of parameters is different. In linear regression and in discriminant analysis, parameters are estimated by using the least square method. In logistic regression, the parameters of the model are estimated by using the maximum likelihood method. That is, the coefficients that make the observed results most "likely" are selected. Since the logistic regression model is non-linear, an iterative algorithm is necessary for parameter estimation.

The interpretation of the resulting coefficients is similar to that in multiple regression with a continuous independent variable. With a binary independent variable, the coefficients B_i indicate the expected log-odds increase when that particular independent variable increases by 1 (other variables remaining constant). Thus, the odds ratio is the increase in odds, and the B values are the increase in log-odds, for a given unit increase in a particular independent variable.

1.12. Multivariate statistical and pattern recognition methods

"Multivariate analysis plays an important role in diagnostic medicine because considering all relevant variables before coming to a conclusion is what making a diagnosis is about. A physician examining a patient can test and evaluate several individual factors such as blood pressure and cholesterol count. But by employing multivariate analysis techniques, the physician can take many variables into account (i.e., family history and entire body chemistry) to help arrive at an accurate diagnosis. Medical diagnosis through the use of computer programs that employ multivariate analysis will continue to grow in importance.")[1]

---

[1] (http://www.stat.psu.edu/department/grad_handbook/centers/cma.html)

The aim of pattern recognition is to assign objects (experimental units or patterns) to one of a given number of classes. Objects are given by (measured) characteristics or features. Mathematically, pattern recognition is a classification problem. Let us consider the recognition of characters. We wish to design a system such that a hand-written symbol will be recognised as an "A", a "B", etc. In other words, the “machine” that we design must classify the observed hand-written character into one of 26 classes. The hand-written characters are often ambiguous, and there will be misclassified characters. The major goal in designing a pattern recognition “machine” is to have a low probability of misclassification.

There are many problems that can be formulated as pattern recognition problems, e.g. in weather prediction, in signal analysis or in image analysis. Pattern recognition methods can likewise be applied to medical diagnostics problems; here, the patterns are the patients with their symptoms and the classes are the possible diagnoses. On the basis of the symptoms, we have to classify a patient into the correct class, i.e. we have to find the correct diagnosis.

There are several pattern recognition techniques. Statistical pattern recognition techniques discuss the problem in the framework of hypothesis testing.

In any pattern recognition problem, one very important task is to find characteristics (independent variables) that provide a good description of the variability of the data, and to reduce the number of these characteristics. This phase is called feature extraction. The most commonly used feature extraction methods are the factor and cluster analysis of variables and the stepwise techniques. The classification generally involves two steps. In the first step, called the learning step, we have to find a rule that assigns a category to any pattern with the smallest possible misclassification. The second step can be the test of that rule.

Statistical pattern recognition methods are also multivariate methods. Factor and cluster analyses are used in future extraction. Discriminant analysis and logistic regression are methods that can be used in the learning phase and in the classification step. I have presented the application of such methods to various medical diagnostic problems in several publications (Boda and Pap, 1984; Pap and Boda, 1986; Nádasdy *et al*, 1991, Boda and Tóth, 1996, Gyöngyösi *et al*, 1997) and in my thesis at Attila József University, Szeged (Boda, JATE, 1984).

II. USE AND RESULTS OF BIOSTATISTICAL METHODS IN TWO NEUROPSYCHIATRIC PROBLEMS

Several of my publications have illustrated the use of biostatistical methods in neuropsychiatry (Molnár *et al,* 1982; Boda, 1984; Boda and Pap, 1984; Kálmán *et al,* 1991; Pető *et al,* 1997, Pikó *et al,* 1997). Two examinations will be demonstrated in detail here, where univariate methods alone gave results that were insufficient to interpret the complicated relationships between the data. In both cases, the computations were performed by using the SPSS software (SPSS Inc., 1977).

2.1. Examination of characteristics of panic disorder on the basis of psychological questionnaires

It is obvious that panic disorder is becoming increasingly more frequent among the distress states diagnosed earlier as neurosis. Because of its frequency and serious consequences, panic disorder imposes a great burden on psychiatric attendance and on somatic remedy (Weissman, 1989).

Our investigations relating to panic disorders have been reported in several papers (Lajkó and Boda, 1989, 1990; Boda and Lajkó, 1991). In the present study, self-report questionnaires were selected on the basis of our previous clinical experience; these questionnaires were considered to reflect the psychological difficulties manifested by panic patients. Statistical methods were used to identify those questionnaires which were characteristic of panic patients in general, and which were, additionally most suitable to create new dimensions within this group.

2.1.1. Subjects and methods

The subjects in the examinations were assigned to four groups: the first group consisted of control persons: manual workers, members of the working staff, and of professional staff of two industrial companies, who did not suffer from any disorder or psychosomatic syndrome. This group comprised 66 subjects (12 males and 54 females). Patients with panic disorders were assigned to the second group. They were patients at the Albert Szent-Györgyi Medical University and the local Psychiatric Outpatient Centre. Diagnoses were made by two psychiatrists independently of each other. This group consisted of 68 patients (13 males and 55 females). Among them, 10 patients were disability-pensioners, and the other were clerks or manual workers. Members of the third group were chosen from among the alcoholic patients undergoing treatment at the Alcohol-Withdrawal Centre at Nagyfa. 32 males and 11 females were involved in this group. The fourth group comprised psychiatric patients receiving treatment at the Department of Psychiatry of Albert Szent-Györgyi Medical University; there were 14 males and 16 females in this group; all of them were disability-pensioners.

Questionnaires in the examination

Eleven self-report questionnaires were selected from the book by Corcolan (1987). They were presumed to be the most adequate to characterise the symptoms of panic disorder. These questionnaires were as follows:

A: Frequency of Self-Reinforcement Questionnaire,
B: Provision of Social Relations,
C: Self-Efficacy Scale,
D: Brief Fear of Negative Evaluation,
E: Index of Self-Esteem,
F: Verbal Aggressiveness Scale,
G: Revised Martin-Larsen Approval Motivation,
H: Liking People Scale,
I: Interpersonal Dependency Inventory,
J: Assertion Anxiety Inventory
K: Assertion Inventory

Statistical methods

The scores for the items in the questionnaires were summed; these sums formed 11 variables. They were analysed by two-way ANOVA to find differences between groups and sexes. To make a complex comparison of the four groups on the basis of each questionnaire, linear discriminant analysis was used.

2.1.2. Results

Table 2.1.1 contains the means and standard deviations of the scores of the questionnaires in each group.

Table 2.1.1. Summary statistics of sums of scores of questionnaires in the four examined groups

	Control		Panic		Alcoholic		Insane	
	Mean	S.D.	Mean	S. D.	Mean	S.D.	Mean	S. D.
A	47.34	8.84	42.8	8.73	51.65	8.53	45	10.15
B	35.95	9.19	40.23	9.66	41.81	9.68	43.6	9.7
C	78.34	11.14	68	14.6	79.5	12.09	65.08	17.89
D	34.86	7.16	39.7	9.16	30.23	7.03	37.28	10.49
E	66.29	14.56	74.85	17.19	61.46	17.68	79.84	17.3
F	46.81	11.17	46.61	10.65	49.38	7.51	51.92	11.41
G	51.86	10.06	56.49	10.16	51.77	9.17	58.28	10.6
H	55.88	7.35	51.59	9.74	52.85	10.75	50.28	10.05
I	45.41	9.08	51	9.91	39.46	8.96	45.96	11.06
J	101.76	21.63	108.89	24.13	90.77	21.39	102.72	21.68
K	107.75	20.3	110.1	19.35	105.31	17.62	110.44	19.78

The two-way ANOVA did not result in a significant interaction between sex and group; i.e. differences between the groups do not depend on the sex. Moreover, there were no significant differences between males and females.

The linear discriminant analysis resulted in three discriminant functions, the first two of which were significant by a $\chi 2$ test, as shown in Table 2.1.2. The first two eigenvalues account for 90.6% of the total between-groups variability. The co-ordinate system of discriminant functions is now two-dimensional: the first and second discriminant functions form the abscissa and the ordinate, respectively. Figure 2.1.1 shows the group means in this new co-ordinate system. Each group has its own place on one of the axes, i.e. the first discriminant function distinguishes groups 2 and 3, while the second function distinguishes groups 1 and 4.

Table 2.1.2. Eigenvalues and Wilks' lambda values

a)

Function	Eigen value	% of variance	Cumulative %	Canonical correlation
1	0.356	61.8	61.8	0.512
2	0.166	28.8	90.6	0.377
3	0.054	9.4	100.0	0.227

b)

Test of function(s)	Wilks' lambda	Chi-square	df	Sig.
1 through 3	0.600	82.989	33	0.000
2 through 3	0.814	33.522	20	0.030
3	0.949	8.591	9	0.476

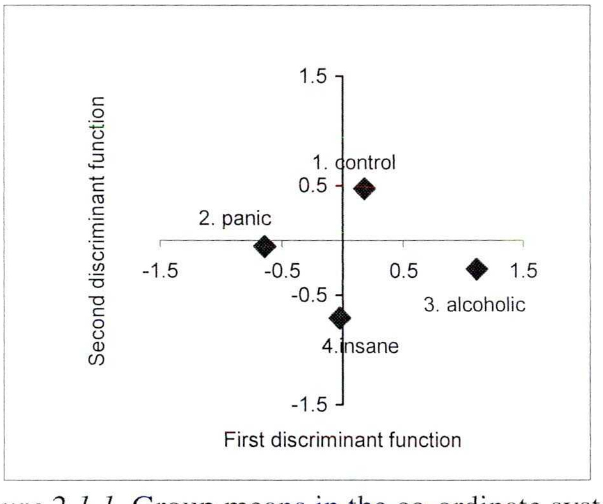


Figure 2.1.1. Group means in the co-ordinate system of the first two discriminant functions

Table 2.1.3 presents standardised discriminant coefficients. The coefficients of the first function show how the questionnaires feature in the separation of groups 2 and 3. As the mean for group 2 is almost on the axis, on the negative side, the negative coefficients with the higher absolute values discriminate better between groups 2 and 3. Consequently, questionnaire I distinguishes mostly the panic group from the other three best, and questionnaire B characterises group 3, having the highest positive coefficient. Similarly, questionnaires C and G discriminate best the first and fourth groups, respectively.

This study was regarded as a starting point for further examinations. It is important that information yielded by self-report questionnaires can not be trustworthy as regards the prediction of the behaviour of an individual. This study may help by reflecting mainly the cognitive aspects of the psychic functions. Results of statistical methods must be tested in practice. For example, questionnaire I best characterised the panic disorder. This is in accordance with clinical practice, as patients with panic disorder have more demand on dependence and are hard hit by losses.

Table 2.1.3. Standardised discriminant function coefficients

	Function 1.	Function 2.
A	0.330	-0.073
B	0.490	-0.399
C	0.488	0.591
D	-0.209	0.221
E	-0.018	0.155
F	0.249	-0.153
G	0.155	-0.538
H	0.271	0.175
I	-0.561	0.121
J	-0.015	0.344
K	0.087	0.026

2.2. Statistical evaluation of data furnished by Jung's word association testing of attempted suicide and control subjects in hypnotic and alert states

2.2.1. Basic assumptions and goals

It is a well-known and regrettable fact that Hungary and Csongrád county provide one of the highest figures in the world's statistics as concerns the suicide rate. The present study was designed to find characteristic features of Jung's word association test (Jung, 1906) applied to 30 attempted suicide subjects and 30 control persons in an alert state and under hypnosis.

The medical aim was to define parameters revealing the effect of hypnosis and the difference between groups. The objective of this work was to demonstrate the use of univariate and multivariate methods for the comparison of the two groups on the basis of several parameters.

The suicide act was long regarded by researchers as a "black box", i.e. it was thought that only the suicide attempt and the previous communication processes could be examined directly, and inferences could be drawn on the "intrapsychic processes" within the personality.

According to Ringel (1953), every suicidal act or serious suicidal attempt is preceded by the same psychological constellation. This is the pre-suicidal syndrome. Because of its three characteristics, it is often referred to as the suicidal triad, consisting of the following features:

I. Narrowing, which has two forms:

a) Situative narrowing, when external personal opportunities are narrowed.

b) Dynamic narrowing, when the perception, behaviour, emotion, thinking and mentality of a person are strictly unidirectional; the other directions are bottled up. In this state, the individual behaviour becomes colourless, uniform and independent of the situation.

II. Aggression is blocked and directed towards the self.

III. Suicide imaginations, with escape into the world of fantasy.

According to Ringel's theory, dynamic narrowing means the rigid decline of apperception and of interest in the outside world. This corresponds to an altered state of consciousness. Since hypnosis too is considered to be a peculiar altered state of consciousness, we made the psychological assumption that the pre-suicidal syndrome is very similar to or even identical to the altered state of consciousness typical in hypnosis.

The aims of the present study were to acquire a better understanding of he intrapsychic processes of persons who have attempted suicide, and to examine whether the assumption that Ringel's pre-suicidal syndrome is similar to a modified state of consciousness characteristic of hypnosis can be proven for attempted suicide and healthy control subjects in an alert state and under hypnosis.

In this respect, the susceptibility of persons to attempt or repeat attempts at suicide was examined, and the Jung's word association test was employed in an alert state and under hypnosis. This method is a basic psychological device to characterise the subconscious processes of persons. The association test is carried out as follows: the experimenter pronounces words, separately, one after the other, and it is the task of the subject to respond to each word as soon as possible with the first word that comes into his/her mind. The reaction time (RT) is measured and the association mistakes of the reply are categorised for each stimulus word. Statistical analysis was performed on these data. The questions considered in the statistical analysis were as follows:

1. Does the mean RT depend on the alert or the hypnotic state, on belonging in the suicide or the control group or on the stimulus word?
2. Does the category of the association mistake depend on the alert or the hypnotic state, on belonging in the suicide or control group or on the stimulus word?
3. Are the RTs and the categories of the answers associated with environmental factors?

2.2.2. Subjects and methods of examination

Jung's word association testing was performed on 60 persons. Of these, 30 individuals were patients at the Department of Neuropsychiatry, Albert Szent-Györgyi Medical University, who had previously attempted suicide on one or more occasions; apart from this, however, they did not suffer from any psychiatric disorder or any psychosomatic illness. As a method of suicide, 26 subjects chose a drug overdose, 3 persons cut their wrists with a knife or blade, and there was one attempt by hanging. The other 30 subjects were healthy controls, matched to the attempted suicide group according to age, sex, education, residence and marital status.

The battery of the original Jung's association test is a list of 100 words frequently used in the German language. In the Hungarian version, the Hungarian translations of these words were used, together with two words of great importance from the aspect of suicide: hopeless and the tool of suicide, i.e. medicament, knife or rope. The list of the 102 stimulus words is to be seen in Appendix A.

During testing, these words were read by the experimenter to the patient, separately, one after the other, and the patient had to respond to each word as soon as possible with the first word that came into his/her mind. The measured RT was the time interval between the last syllable of the stimulus word and the first syllable of the answer. These RT values provided the numerical data for the statistical evaluation.

The answer words were recorded and graded later according to special association mistakes. In normal associative processes, association is based on similarity, contradiction, space and time relations. If the association arises by some other law or relationship, we speak about pathological associations or association-mistakes. Association mistakes were characterised and classified by Rapaport (1946). The same categories were used, supplemented by the category "association with death and suicide". In the statistical analysis, these categories formed the categorical variables.

2.2.3. Statistical methods and results

The five sociological characteristics in the two groups were compared by means of univariate tests. The summary statistics are shown in Table 2.2.1. The difference was not significant for the first four characteristics ($\chi 2$ test, $p>0.05$), but only in the case of age, where the mean difference of 7.5 years was statistically significant. However, this is not important from the aspect of our examination.

Mean RTs were compared by two-way ANOVA with repeated measurements and the (Rapaport kind) categories of word association mistakes were compared by means of the $\chi 2$

test or Fisher's exact test. Univariate analyses applied to each word separately increase the probability of experimentwise Type I error, but they can serve as the starting-points of the multivariate methods. A stepwise linear discriminant analysis was applied to the RT data. With the use of logistic regression, the other variables, such as association mistakes and sociological characteristics, were also taken into account (with the exception of the RT data).

Table 2.2.1. Sociological characteristics of the examined groups

		Attempted suicide	Control
SEX	male	10	11
	female	20	19
MARITAL STATUS	married	16	12
	single	14	18
RESIDENCE	village	6	5
	town	24	23
EDUCATION	elementary	14	10
	secondary	14	17
	high	2	3
AGE	mean ± std. dev.	36.73 ± 14.76	29.26 ± 10.44

2.2.3.1. Examination of reaction times, outliers, transformations

One typical property of the RT data is a skewness to the upper tail. A common strategy in modelling RT data is to use outlier exclusion, but the restricted mean value introduces a bias in estimating the mean of a skewed population (Miller, 1991). Another strategy is to transform all of the data (Ulrich and Miller, 1993). To avoid the omission of data in RTs through outlier elimination, a logarithmic transformation was used. In this way, all the data were available for further analysis and the skewness of the RT distribution was eliminated.

2.2.3.2. Application and results of two-way ANOVA with repeated measurements

a) Comparison of mean RTs word by word

A repeated measurement ANOVA was performed on the logarithm of each RT. One between-subject effect (group), one within-subject effect (hypnosis) and their interaction were included in the model. With this model, the following questions could be answered:

- Is there a significant difference between the means for the attempted suicide group and the control group?
- Is there a significant difference between the means measured in an alert state and under hypnosis? and
- Does the difference between the means measured in the alert state and under hypnosis depend on the measured group, i.e. is there a group-by-hypnosis interaction?

ANOVA was performed on the natural logarithms of the original RT data. Detailed results are to be found in Appendix B.

Examination of the mean RT data revealed that the mean RTs are generally longer in an alert state than under hypnosis, and they are longer in the attempted suicide group than in the control group. The differences between the means for the individual words varied: for some words, the group effect was significant, for some words the hypnosis effect was significant, and for some words both effects or neither effect was significant. Surprisingly, depending on the significance of the effects, the words could be arranged into thematic groups as regards their semantic connections.

1. Both effects were significant for words that related to questions of life and death (e.g. death, long, old, ill, door, to sleep, to die). For these words, all the means were "far" from each other, as can be seen, for example, in the case of the word to die in Figure 2.2.1. The significant difference in RT suggests that life as an indicator of existence can be related to word stimuli indicative of the transitoriness of life.

2. The effect of hypnosis was significant only for words where the common feature was the regression back to childhood (ink, pencil, exercise book, book, to ask). For these words, the difference between the means for the control and attempted suicide groups is very small, as for instance, in the case of the word pencil, as shown in Figure 2.2.2.

3. The group effect was significant only for words relating to the question of choice and existence (head, water, ship, village, salt, bird, flower, part, to choose). For these words, the difference in means between the alert and hypnotic states was very small, as may be seen for the word to choose in Figure 2.2.3.

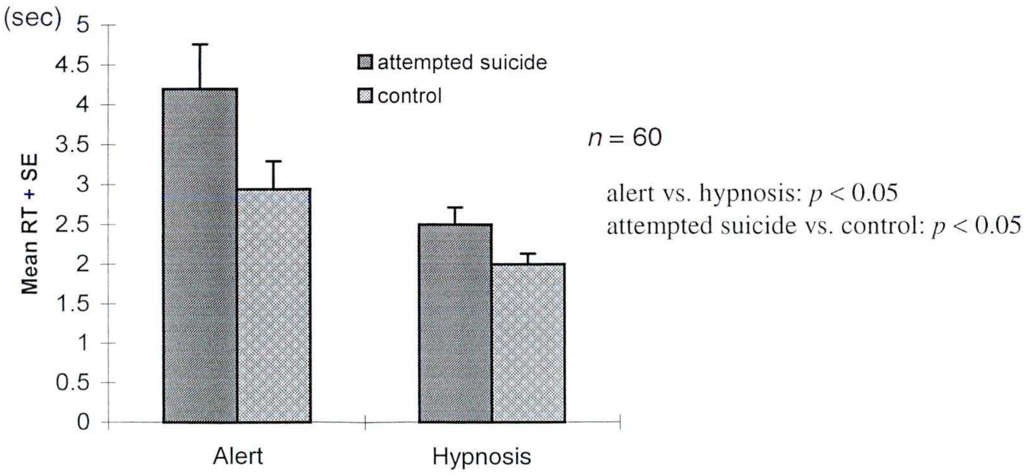


Figure 2.2.1: Mean reaction times to the word to die

4. A significant interaction was found for words relating to richness (money, rich). For these words, the differences between the alert and hypnotic states were much higher in the attempted suicide group than in the control group; the "direction" is opposite. For an example, see Figure 2.2.4.

Words with no significant effects were not identified; they did not show any common feature. It is important, that words relating to the family (child, brother, wife, family, bride), resulted in a significant difference in the control vs. attempted suicide comparison.

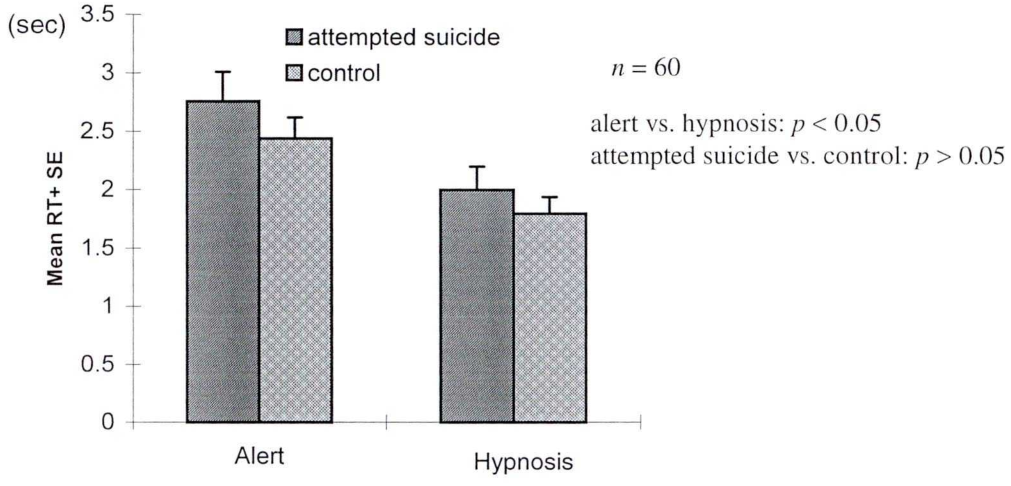


Figure 2.2.2. Mean reaction times to the word pencil

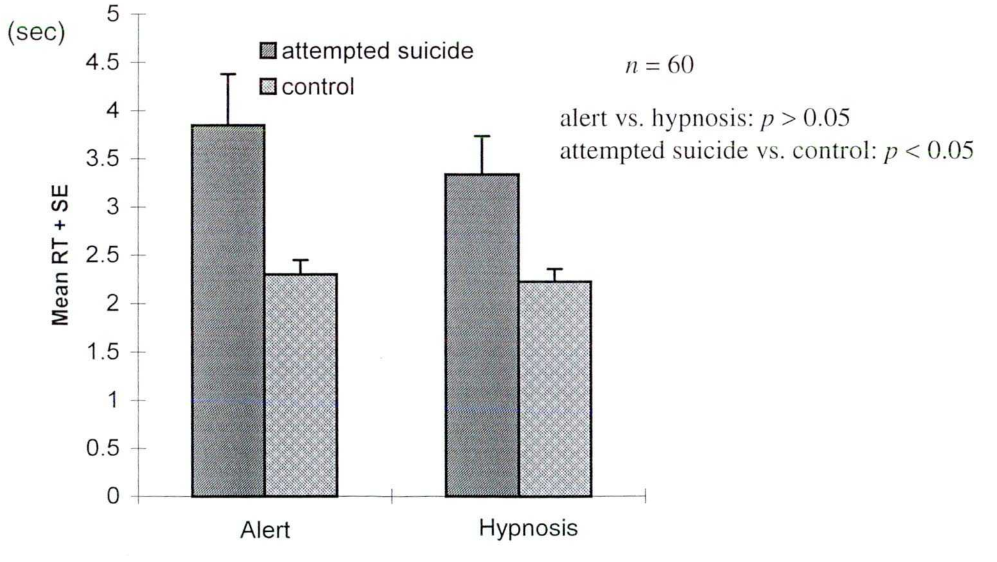


Figure 2.2.3. Mean reaction times to the word to choose

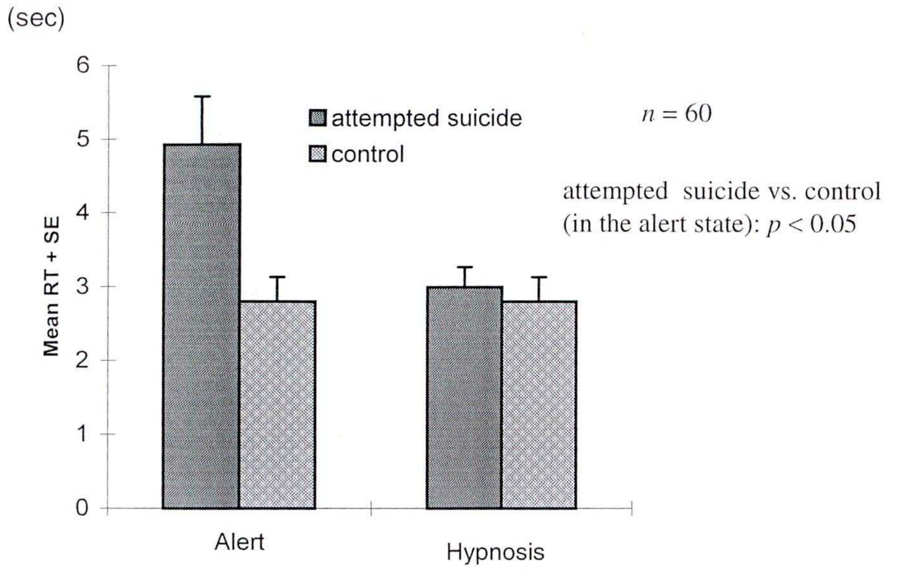


Figure 2.2.4. Mean reaction times to the word money

b) Comparison of means of all RTs

The application of ANOVA to each word separately does not take into account the correlation between variables. As the same persons were examined, the variables must somehow be correlated. An attempt was made to assign to each person only one value, which can represent the person's RT. The mean of all 102 RTs was therefore calculated for each patient, and the two-way ANOVA was applied to these "summarised" values.

When the mean of all RTs was taken and the same ANOVA model was used, significant mean effects were found. The differences were significant between the attempted suicide group and the control group ($F_{1,58}$=27.26, p<0.001), and between the alert and hypnotic states ($F_{1,58}$=16.316, p<0.001). That is, both the alert state and a suicidal tendency had an "increasing" effect on the mean of the RTs. These results seem to contradict our original psychological assumption that suicide is an altered state of consciousness similar to a hypnotic state. In this case, we were expecting nearly identical mean RTs in the attempted suicide and control groups under hypnosis. This apparent contradiction can be solved by considering the small p-value of the interaction (p=0.08) and analysing Figure 2.2.5. The data suggest that these effects are not simply added, but are greater in the alert state in the attempted suicide group. This is not quite significant at the generally used 5% level. The power of the test (the probability of demonstrating at least such a large difference) is 0.408. The sample size is not sufficient to reveal the existence of an interaction. If we presume the existence of this interaction, the pairwise comparisons yield significant differences between the suicide and control groups in the alert and hypnotic states, but the alert and hypnotic means are not significantly different in the control group (Table 2.2.1).

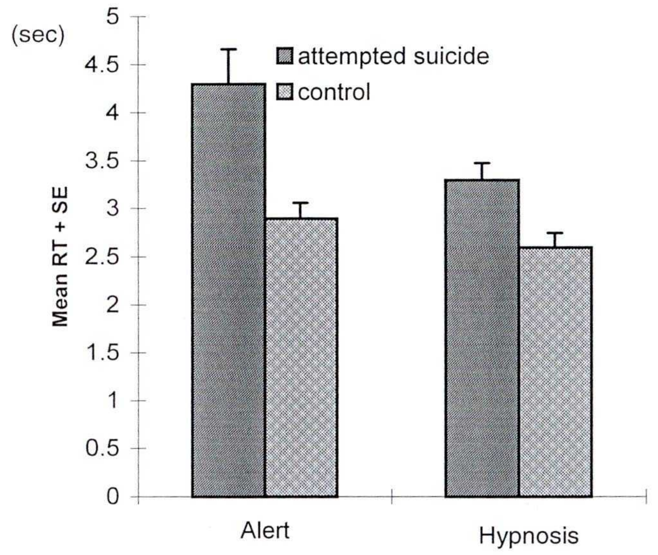


Figure 2.2.5. Means of all reaction times

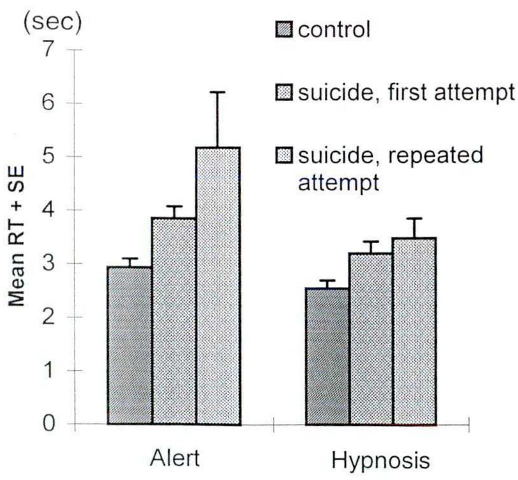
Table 2.2.1. Results of pairwise comparisons on the assumption of a significant interaction: a) comparison of attempted suicide and control groups in both states; b) comparison of alert and hypnotic states in both groups

a)

	Mean Diff. (I-J)	SE	Sig.	95% Confidence Interval for Difference	
Alert	1.321	0.396	0.001	0.529	2.113
Hypnosis	0.736	0.243	0.004	0.251	1.222

b)

	Mean Diff. (I-J)	SE	Sig.	95% Confidence Interval for Difference	
Attempted suicide	0.970	0.235	0.000	0.499	1.441
Control	0.386	0.235	0.107	-0.085	0.857



comparison	alert	hypnosis
control vs. suic. 1.	$p < 0.05$	$p < 0.05$
control vs. suic. 2.	$p < 0.05$	$p < 0.05$
suic.1. vs. suic. 2.	$p < 0.05$	n.s.

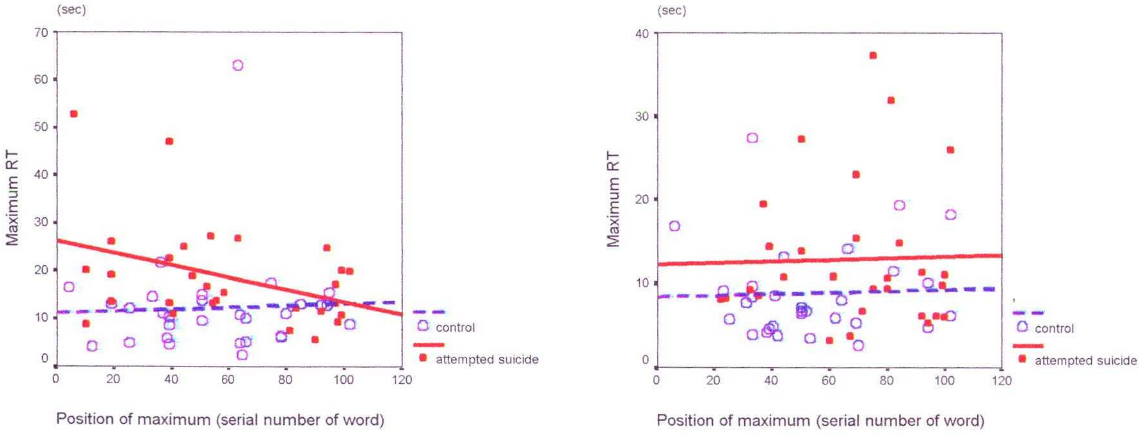
Alert-hypnosis difference: $p < 0.05$

Figure 2.2.6. Mean reaction times in the control and attempted suicide groups

2.2.3.3. Serial measurements analysis: use of summary measures

We can consider the RT data on a subject to be serial measurements, because the sequence of the words was the same in each case. In this way, we lose information about the words by taking into account only the sequence of the words instead of their meaning. One possible method of analysis is the use of summary measures (Matthews *et al*, 1990). This method reduces the repeated measures data to a univariate summary measure. A natural summary measure is the sum or mean of all RT data; this was used in the previous section.

Now, two summary measures were calculated: the maximum RT and the serial number of the word where the maximum RT was found. The relation between these two values in the alert state and under hypnosis was examined in a scatterplot. Figure 2.2.7 gives the scatterplot with the regression lines by groups. Under hypnosis (Figure b), the maximum RTs were given after the 20th word in both groups and they were distributed uniformly. In the alert state (Figure a), the maximum RTs were given earlier, and in the attempted suicide group, there were higher maximum RTs at the beginning than by the end of the experiment. This can be expressed by a significant correlation (r=-0.39, p=0.039). In the other cases, the correlations are not significant (alert state, control group: r=0.028, p=0.884, hypnosis, attempted suicide group: r=0.049, p=0.797, control group: r=0.038, p=0.843)



a) b)

Figure 2.2.7. Scatterplot of maximum RT by the serial number of the word a) in alert state and b) under hypnosis.

2.2.3.4. Examination of the classification of answers and their relation to family factors

a) Examination of the effect of hypnosis in the two groups word by word

Rapaport's classification of word association mistakes results in categorical variables. The frequencies of these categories were computed for each word. Because of the high number of these categories, statistical comparison was performed only as concerns normal and abnormal associations. The frequencies of normal and abnormal associations in the control and attempted suicide groups were compared by means of Fisher's exact test in the alert state and under hypnosis. The results are shown in Appendix A. The results for word stimuli connected with the family are given in Table 2.2.2. It can be seen that an abnormal association occurred more frequently in the attempted suicide group. This relation was significant for the word family both in alert the state and under hypnosis.

Table 2.2.2: The numbers of normal and abnormal word associations in the two groups in response to word stimuli connected with the family a) in the alert state and b) under hypnosis

a)

	Attempted suicide				*Control*		
	Normal		Abnormal		Normal		Abnormal
Child	12	⇒	18		17	⇐	13
Family	10	⇒	20	*	19	⇐	11
Brother	12	⇒	18		13	⇒	17
Bride	14	⇒	16		22	⇐	8
Wife	9	⇒	21	*	19	⇐	11

b)

	Attempted suicide				*Control*		
	Normal		Abnormal		Normal		abnormal
Child	12	⇒	18	*	21	⇐	9
Family	10	⇒	20	**	23	⇐	7
Brother	14	⇒	16		15		15
Bride	15		15	*	20	⇐	10
Wife	11	⇒	19		17	⇐	13

*: Fisher's test, $p < 0.05$
**: Fisher's test, $p < 0.001$

b) Comparison of categories of association mistakes given to all words

To reduce the number of variables containing categories of association mistakes, we had to combine these variables so that each person could be characterised by one variable containing information about the answers given to the 102 words. This was done in several ways.

a) The number of words that responded to with an abnormal answer (ranges between 0 and 102) or a special abnormal answer (e.g. a mental block, an association with death, or self-involvement) was computed.

b) Only the occurrence or not of an abnormal or special abnormal answer was considered; this resulted in a binary variable (1= the person gave at least one abnormal answer, 0=the person did not give an abnormal answer to any of the 102 words).

The mean numbers of abnormal associations were compared by two-way repeated measures ANOVA. The frequencies of binary variables in the two groups were compared by means of

Fisher's exact test and the changes in answers given in the alert and hypnotic states were compared via the McNemar test.

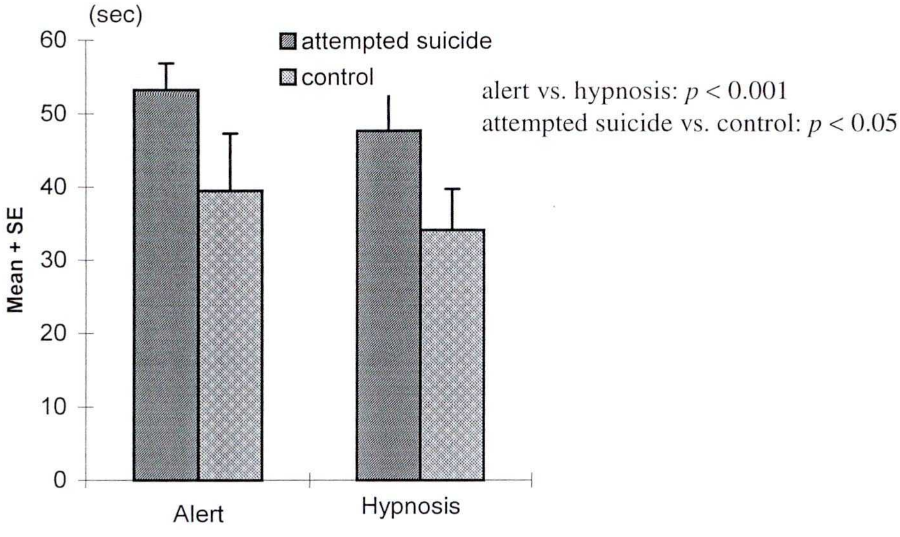


Figure 2.2.8. Mean numbers of abnormal associations

The mean number of abnormal associations was higher in the attempted suicide group both in the alert state, and under hypnosis (Figure 2.2.8); the group effect of the ANOVA was significant (p=0.015). The effect of hypnosis was also significant (p=0.001): the number of answers with abnormal associations was higher in the alert state that under hypnosis. The group by hypnosis interaction was not significant (p= 0.932), i.e. differences between groups do not depend on the alert or hypnotic state, the value of the mean difference being almost the same.

The frequencies of the occurrence of special abnormal answers in the two groups were compared by Fisher's exact test; the changes of answers given in the alert and hypnotic states were compared by the McNemar test. Association with death was a special abnormal association; it occurred when the answers referred to death (for example, the answer to the word window was „to tumble out"). Self-involvement was another special abnormal association; it occurred when answers referred to the subject himself or herself (for example, the answer to the word to love was „me"). The difference concerning the changes in the answers connected with self-involvement was significant between the alert and hypnotic states, i.e. fewer patients gave answers with self-involvement under hypnosis (McNemar test, p=0.008). The association with death was significantly higher in the attempted suicide group, both in the alert state and under hypnosis. (Fisher's exact test, $p < 0.001$). Although there were more such associations under hypnosis, the difference was not significant relative to the alert state (McNemar test, $p > 0.05$, Table 2.2.3).

Table 2.2.3. The number of patients with abnormal associations relating to self-involvement and death a) in the alert state and b) under hypnosis

a)

Abnormal association	*Attempted suicide* No		*Attempted suicide* Yes		*Control* No		*Control* Yes
Self-involvement	9	⇒	21		10	⇒	20
Death, suicide	14	⇒	16	**	27	⇐	3

b)

Abnormal association	*Attempted suicide* No		*Attempted suicide* Yes		*Control* No		*Control* Yes
Self-involvement	17	⇒	13		12	⇒	18
Death, suicide	13	⇒	17	**	28	⇐	2

*: Fisher's test, $p < 0.05$
**: Fisher's test, $p < 0.001$

2.2.3.5. Application and results of linear discriminant analysis and logistic regression

Univariate methods are useful for the acquisition of knowledge relating to data, but the significance furnished by univariate methods can be misleading when the overall experiment is considered. Univariate methods do not take into account the correlation between the variables, and their repeated use increases the probability of Type I error, i.e. results may be regarded as significant when in reality there is no difference. The use of multivariate methods can help in the avoidance of this problem, although new problems arise in finding the best-fitting model.

Table 2.2.3. Results of discriminant analysis

Classification Results

		Predicted group membership: Control	Predicted group membership: Attempted suicide	Total
Original	Control	26	4	30
	Attempted suicide	6	24	30
Cross-validated	Control	23	7	30
	Attempted suicide	6	24	30

a Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
b 83.3% of original grouped cases correctly classified.
c 78.3% of cross-validated grouped cases correctly classified.

Stepwise discriminant analysis was performed on the differences of the logarithms of the RTs. This method resulted in a significant difference between the two groups; the mean discriminant function was –0.993 and 1.009 in the attempted suicide group and in the control group, respectively. The words remaining in the analysis were mainly words that gave rise to a significant interaction effect in the univariate ANOVA: month (-0.693), money (-1.007), new (0.521), house (0.511), to choose (0.441), stork (0.414). The proportions of correctly classified cases were 83.3% and 78.3% on the original data and with the jack-knife (cross-validated) method, respectively (Table 2.2.3).

The result of discriminant analysis is not surprising, in the knowledge of the univariate results. The interpretation is similar (e.g. the negative weight of the word money shows that subjects in the attempted suicide group exhibited higher differences between the alert and hypnotic states than did those in the control group). However, it is difficult to interpret the presence of the word stork or new which were not significantly different in univariate ANOVA. For such a large number of independent variables, stepwise methods can "run", but the automatic procedure can produce surprising results. Further, because of the risk that the model may be over-optimistic, it was desirable to assess the predictive capability of the model on a new, independent set of data, but this was not yet possible. Other disadvantages of the use of discriminant analysis are that its assumptions are slightly violated (the Box M test was significant p=0.032), and, that categorical variables such as the categories of association mistakes cannot be taken into account. For these data, in spite of the good discrimination, the results of this analysis do not seem to be applicable in practice. To attain a better model, logistic regression was performed.

In logistic regression, the maximum likelihood iterative algorithm was used for the estimation of coefficients (instead of the least square method). Here, it is impossible to put all 408 independent variables in the stepwise model, where there are only 60 individual cases. Here, an important starting point is to find independent variables. In the first step, variables were chosen by means of several heuristic methods (based on the results of univariate tests (t-test, $\chi 2$ test), in accord with the special interest of the physician. As a starting point, we included the following variables in the model:

1. The suicide pattern in the family.
2. The RT to the word money in the alert state.
3. The RT to the word money under hypnosis, because there was a significant interaction in the two-way ANOVA.
4. The interaction of these RTs.
5. The category of the answer to the word money (normal or abnormal association) in the alert state.
6. The category of the answer to the word money (normal or abnormal association) under hypnosis.
7. The interaction of these categorical variables.
8. The RT to the word pencil in the alert state.
9. The RT to the word pencil under hypnosis, because there was a significant test effect in the two-way ANOVA.
10. The interaction of these RTs.
11. Normal or abnormal associations with death in alert state.
12. Normal or abnormal associations with death in hypnosis.

13. The interaction of these associations.

The forward selection method of logistic regression with the likelihood ratio test criterion was used, and resulted in a model with only three variables:

$$Prob(suicide) = \frac{1}{1+e^{-(1.7037*\text{FAMILY}+1.917*\text{RTMONEY}+2.2343*\text{DEATH}-3.9)}}$$

where

FAMILY indicates the suicide pattern in the family (1=yes, 0=no),

RTMONEY is the RT to the word money in the alert state, and

DEATH indicates an abnormal association with death (1=yes, 0=no).

According to this equation, the probability of being suicidal can be calculated, as shown in Table 2.2.4. It can be seen that this probability is higher when there is a suicide pattern in family, an abnormal association with death and a high RT in the alert state to the word money. For example, when there has already been a suicide attempt in the family (FAMILY=1), with a mean RT to money (RTMONEY=1.2), and an abnormal association with death (DEATH=1), this probability is p=0.912, while without the suicide pattern (FAMILY=0) and with a normal association (DEATH=0) the probability is only p=0.168.

It can be seen that the interactions of the variables between the alert and hypnotic states were not retained in the model. Similarly, variables under hypnosis were not included in the model either, showing that the difference between the control and the attempted suicide group disappears under hypnosis.

Table 2.2.4. Probability of being suicidal for different values of independent variables, according to the logistic regression model

Suicide attempt in the family	Reaction time given to the word money	Abnormal association with death	Probability of suicide
Yes (=1)	1.2	Yes	0.912008
Yes	1.6	Yes	0.957107
Yes	0.8	Yes	0.828012
No	1.6	No	0.303053
No	0.8	No	0.085771
Yes	1.6	No	0.704933
Yes	0.8	No	0.340133

On the basis of the predicted probabilities, a classification of subjects is possible. For example, if a predicted probability for a subject is greater than 0.5, this subject is predicted to the suicide group; otherwise, he or she is predicted as belonging in the control group. Table 2.2.5 presents the classification table comparing the original and predicted group memberships. The success of separation (the proportion of correctly classified cases) was 78.33%.

Table 2.2.5. Classification table of the logistic regression

		Predicted group membership		
		Control	Suicide	Total
Original	Control	24	6	30
	Suicide	7	23	30

The odds ratio for the suicide pattern in the family is 5.49 (=$e^{-1.7037}$), which means that the risk of a suicide attempt is more than 5 times higher when there is a suicide pattern in the family. The 95% confidence interval (1.3375, 22.5699) is quite wide (because of the relatively small sample size), but it does not contain the value of 1. Thus it can be stated that the odds of being suicidal is higher if there has already been a suicide attempt in the family. The odds ratio for the abnormal association with death is 9.3403. The coefficient of the RT to money is 1.917. This means that the higher the RT, the higher the predicted probability.

The goodness of fit of the model can be expressed via what is called likelihood: the probability of the observed results, given the parameter estimates. It is customary to use –2 times the log of the likelihood. Another measure of how well the model fits is the likelihood-ratio $\chi 2$ statistic. The advantage of the likelihood-ratio $\chi 2$ is that it can be subdivided into interpretable parts that add up to the total. For general purposes, the significance value is more important than the actual value of the statistic. In our case, the model $\chi 2$ statistic was 27.451 with 2 degrees of freedom, which is highly significant.

On the basis of this model, the success of separation was 79%. However, this is not too high and it also has the disadvantage that it was tested on the same sample as that for which the "rule" was found. Another frequently used measure of the ability of a model to discriminate between groups is the *c* statistic. This can be interpreted as the proportion of pairs of cases with different observed outcomes, in which the model results in a higher probability for the attempted suicide cases than for the control cases. (Hanley and McNeil, 1982). Our *c* statistic was 0.8.

Naturally, this model seems too simple relative to the very complicated question involved. However, it demonstrates that, to distinguish control and attempted suicide cases, the answers given in the alert state are sufficient (perhaps verifying the psychological assumption that hypnosis and the pre-suicidal syndrome are both special, altered state of consciousness), and it reveals the fact that suicide attempts are highly influenced by earlier such attempts in the family and by a poor socio-economic situation.

III. DISCUSSION

Biostatistics is the application of mathematical statistical methods to biological data. The use of biostatistics is becoming increasingly more frequent in medicine and particularly so in neuropsychiatry.

This work demonstrates the possibilities inherent in the application of biostatistical methods. After a short description of the basic statistical concepts and univariate methods, an overview is given of multivariate methods applicable to neuropsychiatry. Finally, methods, results and consequences stemming from studies of two neuropsychological problems are shown.

The main statements connected with statistical methods

- Univariate methods are used in cases of one response or dependent variable that depends on another (independent) variable. If both variables are continuous, then their relations can be examined by means of regression and correlation analysis. If the independent variable is categorical, then we can examine the dependent variable (generally the mean of the dependent variable) in groups according to the categories (*t*-test, ANOVA). If both variables are categorical, then the frequency distribution of one variable in groups of the other variable can be examined via the $\chi 2$ test.

  The repeated application of two-sample univariate methods to several groups can lead to an increased probability of Type I error, i.e. significant differences are erroneously believed to exist between groups that do not differ in reality. This mistake can be avoided through the use of special correction or more complicated models.

- Multivariate methods examine several response variables or one response variable that depends on several independent variables.

  If the data set contains several variables, the separate application of univariate methods to each variable can also increase the probability of error, the correlation between the variables not being taken into account. This mistake can be avoided by the use of multivariate methods.

- ANOVA is a generalisation of the *t*-test for several samples. It is generally used if we have one, normally distributed variable that depends on one or more categorical variables (one and multiway ANOVA), or if the same variables were measured repeatedly on the same subject (repeated measurement ANOVA). The general assumption in the use of ANOVA, besides normality, is the equality of variances in groups. Slight violation of these assumptions does not affect the results of ANOVA because of the robustness of this method. Nevertheless, in the event of higher departure from the assumptions, ANOVA can be performed on appropriately transformed data, or the use of non-parametric testing is recommended. ANOVA is a special multivariate method and it is also special multiple regression. These models are called general linear models, because they are linear in the parameters to be estimated.

- Another generalisation of the two-sample *t*-test is the multivariate or Hotelling *t*-test. This is used if several (normally distributed) continuous variables have to be compared in groups of one categorical variable. The problem is also called linear discriminant analysis; it can additionally be applied to several groups for the comparison of means in multidimensional space, and can be used for the classification of cases. Discriminant analysis is a special pattern recognition method and is well applicable in medical diagnostics problems.
- Logistic regression is a further classification method and is similar to discriminant analysis, but the variables in the model can be continuous, normally distributed, or even categorical. This is no longer a linear model; it is a so-called generalised linear model.
- Multivariate methods need a large sample size. There are special methods to predict the necessary sample size for each method.

Main statements, neuropsychiatric results, practical consequences of the use of statistical methods in neuropsychiatric problems

From among several neropsychiatric examinations, the methods and results relating to only two problems are demonstrated: the examination of panic disorder on the basis of self-reported questionnaire data, and the examination of data resulting from the study of attempted suicide and control subjects in hypnotic and alert states by Jung's word association test.

- In both examinations, descriptive statistics and univariate methods were utilised first. The results of these methods were taken as the starting-point for the use of multivariate methods.
- In the study of panic disorder, linear discriminant analysis led to results that could be interpreted and employed in practice in spite of the weak discrimination. It was helpful to find questionnaires that best characterised each group. This result was verified by later observations.
- In the evaluation of Jung's word association test data, univariate methods were again used first. Linear discriminant analysis was only partially helpful; logistic regression was more appropriate, because categorical independent variables could also be taken into account.

- Statistical methods seem to prove the psychological assumption that the pre-suicidal syndrome is an altered state of consciousness.
- The mean RT was found to be longer in the attempted suicide group in both the alert and hypnotic states. In the attempted suicide group the stimulus words affect the "complex". For special stimulus words, the association chain breaks and unusual associations appear. Some words attain the emotionally affected category, and the emotional stress elongates the association processes.
- The mean RT is shorter under hypnosis. This can be explained by the selective attention in the altered state of consciousness and by the uncluttered concentration possibilities. The RT is longer also in depressive patients and it is well known that a depressive change of mood is one characteristic of suicide attempters.
- Examination of the categories of the association mistakes revealed that the frequencies of abnormal associations were higher in the suicide group in both the alert and hypnotic states. A special change in the outer stimuli is presumable in the attempted suicide group.
- The results of the word association test verify the fact that the family environment plays an important role in the psychodynamics of suicide. A harmonic family and relatives can be a very important factor in maintaining life. Moreover, fellowship or a religious community can provide important help in a crisis.
- In clinical practice, there is a need for a method that, besides allowing the exploration of patients in crisis, can predict the danger of suicide with high probability. These data were utilised to develop a logistic regression model. Testing of the model is currently under way.
- It was found that the patient's intrapsychic struggle in the pre-suicidal syndrome for and against life is measurable in time. In most cases, however, only the suicidal formula that is rooted deeply in the person's psyche can be employed.

Application of statistical methods can help to reveal undiscovered and hidden relationships of neuropsychological processes and may influence the practical course of healing. These results may be used in prevention; here, the responsibility of the mass media is very important. They may also affect the therapy through the use of suggestive and cognitive techniques. These results may possibly be used in the early warning of a suicidal danger by employing prevention tests. On the basis of our results, we hope that increasing knowledge of the intrapsychic processes of suicide (together with effective prevention, professional therapy and aftercare) may improve the hitherto unfavourable situation in Hungary.

Acknowledgement

I am deeply indebted to Professor Zoltán Janka, head of the Department of Psychiatry, my supervisor, for his continuous support.

I am grateful to Professor Gábor Jancsó, Department of Physiology, for his tremendous support and for his extremely helpful suggestions relating to this thesis.

I express my sincere gratitude to Professor Zoltán Hantos, head of the Department of Medical Informatics, for providing me with the facilities necessary to accomplish this thesis.

I am grateful to Dr. Károly Lajkó and to Dr. István Boncz, Department of Psychiatry, for providing me with the possibility to collaborate with them.

My thanks are due to all my colleagues at the Department of Medical Informatics for their kind support.

I am especially grateful to my father, Domokos Boda, Professor of Pediatrics, Professor Emeritus and also to my mother, for interesting discussions, and, for their enormous encouragement.

My husband and my son must likewise be praised for their never-ending patience and help.

REFERENCES

1. Altman, D.G. *Practical Statistics for Medical Research,* Chapman & Hall, London, 1995.
2. Bowerman, B.L. and O'Connell, R.T. *Linear Statistical Models. An Applied Approach,* PWS-KENT Publishing Company, Boston, 1990.
3. Campbell, M.J. and Machin, D. *Medical Statistics,* John Wiley & Sons, Chichester, 1993 [Chapter 10: 126-136].
4. Corcolan, K. and Fischer, J. *Measures for Clinical Practice. A Sourcebook,* The Free Press, New York, 1987.
5. Dobson, A.J. *An Introduction to Generalized Linear Models.* Chapman & Hall, London, 1991.
6. Ekstrom, D., Quade, D. and Golden, R.N. 'Statistical analysis of repeated measures in psychiatric research', *Archives of General Psychiatry,* **47**, 770-772 (1990).
7. Fisher, R.A. 'The use of multiple measurements in taxonomic problems', *Annals of Eugenics,* **7**, 179-188 (1936).
8. Fleiss, J.L. *The Design and Analysis of Clinical Experiments.* John Wiley & Sons, New York, 1986.
9. Greenhouse, S.W. and Geisser, S. 'Extension of Box 's results on the use of the F distribution in multivariate analysis', *Annals of Mathematical Statistics,* **29**, 95-112 (1958).
10. Hajtman, B. *Bevezetés a matematikai statisztikába,* Akadémiai Kiadó, Budapest, 1971.
11. Hand, D.J. 'The role of statistics in psychiatry', *Psychological Medicine,* **15**, 471-476 (1985).
12. Hanley, J.A. and McNeil, B.J. 'The meaning and use of the area under a receiver operating characteristic (ROC) curve', *Radiology,* **143**, 29-36 (1982).
13. Huynh, H. and Feldt, L.S. 'Conditions under which mean square ratios in repeated measures designs have exact F distributions', *Journal of the American Statistical Association,* **65**, 1582-1589 (1970).
14. Jung, C. G. *Diagnostische Assoziationsstudien.* Bd.L.Borth, Leipzig, 1906.
15. Matthews, J. N. S., Altman, D.G., Campbell, M.J. and Royston, P. 'Analysis of serial measurements in medical research', *British Medical Journal* , **300**, 230-235 (1990).
16. Miller, J. 'Reaction time analysis with outlier exclusion: Bias varies with sample size', *The Quarterly Journal of Experimental Psychology,* **43**A, 907-912 (1991).
17. Rapaport, D. 'The word association test'. In: Rapaport, D., Gill, M. and Schafer, R. *Diagnostic Psychological Testing,* Vol. II, 13-84, Year Book Publ., Chicago, 1946.
18. Ringel, E. *Der Selstmord Abschluss einer Krankhaften psychischen Entwicklung.* W. Maudrich, Vienna,1953.
19. Rosenthal, R. and Rubin, D.B. 'Statistical analysis summarizing evidence versus establishing facts', Part I, *Psychiatry Research,* **23**, 295-299 (1985).
20. Seldrup, J. 'Whatever happened to the t-test?', *Drug Information Journal,* **31**, 745-750 (1997).
21. SPSS Inc. *SPSS Advanced Statistics™* 7.5 Copyright © 1977 by SPSS Inc., ISBN 0-13-6569i27-7.
22. Srivastava, M.S. and Carter. E.M. *An Introduction to Applied Multivariate Statistics,* North-Holland, New York, 1983.
23. Sváb, J. *Többváltozós módszerek a biometriában,* Mezőgazdasági Kiadó, Budapest, 1979.
24. Ulrich, R. and Miller, J. 'Information processing models generating lognormally distributed reaction times', *Journal of Mathematical Psychology,* **37**, 513-525 (1993).
25. Weissman, M.M., Klerman, G.L., Markowitz, J.S. and Ouellette, R. 'Suicidal ideation and suicide attempts in panic disorder and attacks', *The New England Journal of Medicine,* **321**, 1209-1214 (1989).

26. Williamson, J.W., Goldschmidt, P.G. and Colton, T. 'The quality of medical literature: An analysis of validation assessments'. In: *Medical Uses of Statistics*, Ed. by Bailar, J.C. and Mosteller, F., NEJM Books, Waltham, Massachusetts, 1986.

Papers in connection with the thesis

1. Molnár, J., Béládi, I., Domonkos, K., Földeák, S., Boda, K. and Veckenstedt, A. 'Antitumor activity of flavonoids on NK/LY ascites tumor cells', *Neoplasma (Bratisl.)*, **28**, 11-18 (1981).
2. Molnár, J., Mucsi, I., Mándi, Y., Béládi, I., Földeák, S., Boda, K. 'The effect of rutin-n-mustard on the survival of nk/ly ascites tumor bearing mice', *International Journal of Cancer*, **30**, 767-771 (1982).
3. Molnár, J., Mándi, Y., Boda, K. and Földeák, S. 'The effect of ring-substituted and quaternary chlorpromazine derivatives on the survival of the nk/ly ascites tumor bearing mice', *Neoplasma*, **29**, 215-222 (1982).
4. Boda, K. 'Statisztikus alak-felismerési módszerek alkalmazása orvos-diagnosztikai problémákra', *Egyetemi doktori disszertáció*, JATE (1984).
5. Boda, K. and Pap, Á. 'Diagnostics of pancreatic insufficiency using multivariate statistical and pattern recognition methods', *Computers in Biology and Medicine*, **14**, 91-97 (1984).
6. Pap, Á. and Boda, K. 'Complex evaluation of secretin pancreosymin test data by multivariate statistical pattern recognition methods', *International Journal of Pancreatology*, **1**, 205-212 (1986).
7. Lajkó, K. és Boda, K. 'Pánikbetegek környezeti stressz-hatásainak vizsgálata' *Ideggyógyászati Szemle*, **42**, 409-420 (1989).
8. Lajkó, K. és Boda, K. 'Pánikbetegség vizsgálata statisztikai módszerekkel pszichológiai kérdõívek alapján', *Számítástechnikai és Kibernetikai Módszerek Alkalmazása az Orvostudományban és a Biológiában, NJSZT Medicomp Kollokvium, Szeged,* 204-205 (1990).
9. Bandl, E., Boda, K. and Sonkodi, S. 'Hypertension screening in a dental surgery: a Hungarian study', *Journal of Human Hypertension*, **4**, 253-257 (1990).
10. Kálmán, J., Rimanóczy, Á., Szekeres, Gy., Boda, K. és Janka, Z. 'Biológiai markerek vizsgálata krónikus alkoholizmusban: I.Trombocita monoamino oxidáz-B aktivitás és sziálsav tartalom', *Ideggyógyászati Szemle*, **44**, 303-309 (1991).
11. Nádasdy, T., Krenács, T., Kalmár-Nagy, K., Csajbók, T., Boda, K. and Ormos, J. 'Importance of plasma cells in the infiltrate of renal allografta. An immunhistochemical study', *Pathology, Research And Practice*, **187**, 178-183 (1991).
12. Petõ, Z. és Boda, K. 'Testkép és tápláltság egészségügyi szakközépiskolás lányoknál. *A Pszichoterápiás Relaxáció Elsõ Európai Kongresszusa*, Budapest, 127-129 (1991).
13. Petõ, Z., Boda, K., Szabó, P. és Tury, F. 'Deviancia-jelenségek középiskolás korban', *Szociális Munka*, **IV**, 91-111 (1992).
14. Szabó, S. és Boda, K. 'A szorongással összefüggõ tényezõk serdülõ fiatalok körében Makón II', *Egészségnevelés*, **34**, 156-160 (1993).
15. Szabó, S. és Boda, K. 'A serdülõ fiatalok mentálhigiénés vizsgálata Csongrád megye négy városában, a schizoid viselkedéssel összefüggõ tényezõ', *Egészségnevelés*, **35**, 4-7 (1993).
16. Boda, K., Gyõri, I., Mizsei, I., Pavelka, Z. és Szofrán, T. 'Számítástechnika, biometria és informatika oktatás a SZOTE-n', *Informatika a felsõoktatásban. Országos Konferencia. Debreceni Universitas, Debrecen,* 369-373 (1993).

17. Boda, K., Tóth, K. és Boncz, I. 'Többváltozós statisztikai módszerek alkalmazása klinikai adatok kiértékelésére', *NJSZT Kollokvium Kiadványa*, (Szerk. Kozmann, Gy.), NJSZT, Budapest, 53-56 (1996).
18. Gyöngyösi, M., Takács, T., Czakó, L., Jambrik, Z., Boda, K., Farkas, A., Forster, T. and Csanády, M. 'Non-invasive monitoring of haemodynamic changes in acute pancreatitis in rabbits', *Digestive Diseases and Sciences* **42**, 955-961 (1997).
19. Boda, K. 'Kérdõívek kiértékelésére alkalmas statisztikai módszerek és számítógépes programrendszerek' In: *Devianciajelenségek Serdülõkorban.* (Szerk. Petõ, Z.) Edutech Kiadó Sopron, 5-28, (1997).
20. Petõ, Z., Boda, K., Szabó, P. és Túry, F. 'Devianciajelenségek a középiskolás korban' (Könyvfejezet)In: *Devianciajelenségek Serdülõkorban.*(Szerk. Petõ, Z.) Edutech Kiadó Sopron, 68-77, (1997).
21. Pikó, B., Barabás, K. and Boda, K. 'Frequency of common psychosomatic symptoms and its influence on self-perceived health in a Hungarian student population', *European Journal of Public Health*, **7**, 243-247 (1997).

Presentations in connection with the thesis

1. Lajkó, K. és Boda, K. 'Pánikbetegség vizsgálata statisztikai módszerekkel pszichológiai kérdõívek alapján', *Számítástechnikai és Kibernetikai Módszerek Alkalmazása az Orvostudományban és a Biológiában, NJSZT Medicomp Kollokvium, Szeged,* 1990.
2. Boda, K., Lajkó, K. 'Characteristics of panic patients on self-report questionnaires revealed by mathematical statistical methods', *MIE'91. Satellite Conference on Computer Modelling and National Meeting on Biomedical Informatics, Budapest,* 1991.
3. Boda, K. 'Az SPSS, STATGRAPHICS és CSS:STATISTICA használatával szerzett tapasztalataink' (abstract).*MEDICOMP'92, Szeged. NJSZT XVI. Kollokviumának kiadványa,* 45, 1992.
4. Ungi, I., Thury, A., Szánthó, H., Boda, K., Forster, T. és Csanádi, M. 'Holter monitorozás adatainak spontán vaariabilitása és összefügése echocardiographiás paraméterekkel' (abstract). *Cardiologia Hungarica. A Magyar Kardiológusok Társasága Tudományos Kongresszusa, Budapest,* 1993.
5. Boda, K. 'Computer Assisted Teaching of Biostatistics at the Medical University of Szeged' (abstract). *Proceedings of the Fourth International Conference on Teaching Statistics*, Marrakech, Morocco, 450, 1994.
6. Boda, K. 'Számítógéppel segített biostatisztika oktatás a Szent-Györgyi Albert Orvostudományi Egyetemen' (abstract). *MEDICOMP'94, Szeged. NJSZT XVII. Kollokviumának Kiadványa,* 159, 1994.
7. Boda, K. and Tóth, K. 'Multivariate Analysis of Pedobarographic Data of Rheumatoid Arthritis Patients'(abstract). *16th Meeting of the International Society for Clinical Biostatistics, Barcelona,* 1995.
8. Várkonyi, T., Lengyel, Cs., Madácsi, L., Velõsi, B., Kempler, P., Fazekas, T., Lonovics, J., Csernay, L., Such, Gy., Hidvégi, Zs. and Boda, K. 'Impairement of Caerulein-induced Gall Bladder Emptying in Diabetic Patients with Autonomic and Sensory Neuropathy' (abstract). *Neurodiab. V. Diabetic Neuropathy Study Group of the EASD. Stockholm,* 1995.
9. Fazekas, T., Lengyel, Cs., Várkonyi, T. and Boda, K.' Increased Spatial Dispersion of Ventricular Repolarization in Diabetes Mellitus', *13th Congress of the European Association of Internal Medicine, Athens,* 1995.

10. Boda, K. 'Statisztikai szoftverek alkalmazási tapasztalatai orvosegyetemi környezetben', *Kemometia'95. Magyar Kémikusok Egyesülete és MTA konferencia, Veszprém,* 1995.
11. Boda, K.,and Tóth, K. 'Multivariate analysis with large number of variables: evaluation of orthopaedic data' (abstract). *17th Meeting of the International Society for Clinical Biostatistics, Budapest,* 1996.
12. Boda, K., Tóth,. K. és Kellerman, P. 'A SAS rendszer alkalmazása ortopédiai adatok kiértékelésére többváltozós statisztikai módszerekkel', *SAS Klub.* 1996.
13. Boncz, I., Vidács, J., Boda, K., Fabricz, K., Kovács, Z. and Janka, Z.: 'Word Association Trials with Attempted Suicide Patients in Hypnotic and Awake States. Characteristic Features of Dream Content Analysis' (abstract). *7th European Congress of Hypnosis, Budapest,* 18, 1996.
14. Várkonyi, T., Lengyel, Cs., Rosztóczy, A., Fehér, A., Boda, K., Kempler, P., Fazekas, T., Wittman, T. and Lonovics, J. 'Esophageal, Gastric and Ano-Rectal Motility Disorders in Diabetic Patients with Autonomic and Sensory Neuropathy', *Neurodiab VI., Diabetic Neuropathy Study Group of the EASD, Baden,* 1996.
15. Lengyel, Cs., Thury, A., Várkonyi, T., Ungi, I., Boda, K., Fazekas, T. és Csanády, M. 'A szívfrekvencia-variabilitás és a cardiovascularis reflextesztek összehasonlító vizsgálata autonom neuropathiával szövõdõ inzulindependens diabetes mellitusban' (abstract). *Magyar Belorvosi Archivum Supplementum 1996/2,* 174, 1996.
16. Jambrik, Z., Gyöngyösi, M., Takács, T., Czakó, L., Boda, K., Forster, T. és Csanády, M. 'Kísérletes akut pancreatitis okozta hemodinamikai változások noninvazív monitorizálása', *A Magyar Kardiológusok Társasága Tudományos Kongresszusa,Cardiol. Hung.,* 1996.
17. Jambrik, Z., Gyöngyösi, M., Takács, T., Czakó, L., Boda, K., Forster, T. and Csanády, M. 'Noninvasive Monitoring of Hemodinamic Chamges in Experimental Acute Pancreatitis', *International Society for Heart Research, XVII. European Congress. Bologna, 1996. Journal of Molecular and Cellular Cardiology,* **28**, A 393, 100, 1996.
18. Boncz, I., Boda, K., Vidács, J., Maglóczky, E., Kovács, Z. és Janka, Z. 'A Jung-féle szóasszociációs próba alkalmazása öngyilkosságot megkíséreltteknél éberen és hipnózisban', *VIII. Magyar Hipnózis Találkozó, Lillafüred,* 1997.
19. Boncz, I., Boda, K., Szendi, I., Vidács, J., Maglóczky, E. és Janka, Z. 'Szóasszociációs próbával végzett vizsgálatok öngyilkosságot megkíséreltteknél éberen és hipnózisban', *Magyar Pszichiátriai Társaság IV. Nemzeti Kongresszus, Budapest,* 22, 1998.
20. Boda, K., Boncz, I 'Multivariate Analysis of Reaction Times using the Jung's Word Association Test with Attempted Suicide Patients under Hypnotic and Alert States' (abstract). *19th Meeting of the International Society for Clinical Biostatistics, Dundee, Scotland, UK,* 1998.

Appendix A. Frequencies of normal and abnormal answers

	Alert state					Hypnosis				
	Attempted suicide		Control		p	Attempted suicide		Control		p
Word stimuli	Normal	Abnormal	Normal	Abnormal		Normal	Abnormal	Normal	Abnormal	
1. head	21	9	25	5	0.360	18	12	23	7	0.267
2. green	22	8	27	3	0.181	25	5	28	2	0.424
3. water	12	18	22	8	0.018	22	8	26	4	0.333
4. to sing	19	11	19	11	1.000	22	8	23	7	1.000
5. death	11	19	18	12	0.120	13	17	20	10	0.119
6. long	15	15	16	14	1.000	13	17	18	12	0.301
7. ship	21	9	22	8	1.000	21	9	25	5	0.360
8. to pay	17	13	23	7	0.170	26	4	24	6	0.731
9. window	12	18	22	8	0.018	19	11	24	6	0.252
10. friendly	16	14	18	12	0.795	17	13	20	10	0.596
11. table	18	12	21	9	0.589	19	11	23	7	0.399
12. to ask	17	13	23	7	0.170	22	8	20	10	0.779
13. village	12	18	20	10	0.069	13	17	21	9	0.067
14. cold	17	13	24	6	0.095	20	10	23	7	0.567
15. stem	20	10	25	5	0.233	21	9	28	2	0.042
16. dance	19	11	21	9	0.785	22	8	23	7	1.000
17. lake	17	13	20	10	0.596	21	9	22	8	1.000
18. sick	12	18	19	11	0.120	10	20	19	11	0.038
19. pride	10	20	12	18	0.789	9	21	14	16	0.288
20. to cook	10	20	16	14	0.192	13	17	16	14	0.606
21. ink	19	11	27	3	0.030	22	8	26	4	0.333
22. angry	17	13	19	11	0.792	20	10	20	10	1.000
23. needle	20	10	25	5	0.233	26	4	26	4	1.000
24. to swim	18	12	17	13	1.000	17	13	17	13	1.000
25. journey	15	15	16	14	1.000	15	15	16	14	1.000
26. blue	13	17	23	7	0.017	19	11	29	1	0.002
27. lamp	25	5	27	3	0.706	24	6	28	2	0.254
28. to sin	14	16	15	15	1.000	16	14	17	13	1.000
29. bread	17	13	16	14	1.000	16	14	18	12	0.795
30. rich	11	19	17	13	0.195	13	17	17	13	0.439
31. tree	12	18	19	11	0.120	19	11	20	10	1.000
32. to prick	18	12	20	10	0.789	19	11	25	5	0.143
33. pity	14	16	21	9	0.115	15	15	22	8	0.110
34. yellow	7	23	18	12	0.008	11	19	23	7	0.004
35. mountain	18	12	19	11	1.000	21	9	22	8	1.000
36. to die	6	24	12	18	0.158	11	19	17	13	0.195
37. salt	9	21	16	14	0.115	6	24	18	12	0.003
38. new	12	18	11	19	1.000	7	23	6	24	1.000
39. custom	9	21	12	18	0.589	9	21	10	20	1.000
40. to pray	14	16	17	13	0.606	19	11	17	13	0.792
41. money	8	22	16	14	0.064	9	21	15	15	0.187
42. stupid	19	11	24	6	0.252	23	7	22	8	1.000
43. exercise-book	21	9	23	7	0.771	23	7	26	4	0.506
44. to despise	13	17	9	21	0.422	13	17	15	15	0.796
45. finger	8	22	18	12	0.018	15	15	19	11	0.435
46. expensive	9	21	12	18	0.589	11	19	17	13	0.195
47. bird	15	15	23	7	0.060	19	11	28	2	0.010
48. to fall	16	14	16	14	1.000	19	11	19	11	1.000
49. book	21	9	23	7	0.771	19	11	25	5	0.143
50. unjust	9	21	12	18	0.589	13	17	11	19	0.792
51. frog	10	20	18	12	0.069	15	15	23	7	0.060

Appendix A (cont.)

	Alert state					Hypnosis				
	Attempted suicide		Control		p	Attempted suicide		Control		p
Word stimuli	Normal	Abnormal	Normal	Abnormal		Normal	Abnormal	Normal	Abnormal	
52. to part	3	27	11	19	0.030	5	25	13	17	0.047
53. hunger	9	21	11	19	0.785	9	21	11	19	0.785
54. white	15	15	22	8	0.110	16	14	22	8	0.180
55. child	12	18	17	13	0.301	12	18	21	9	0.037
56. to pay attention	14	16	20	10	0.192	14	16	20	10	0.192
57. pencil	18	12	25	5	0.084	23	7	27	3	0.299
58. sad	18	12	18	12	1.000	21	9	18	12	0.589
59. plum	14	16	20	10	0.192	15	15	24	6	0.029
60. to marry	12	18	14	16	0.795	12	18	19	11	0.120
61. house	15	15	19	11	0.435	15	15	23	7	0.060
62. dear (be loved)	17	13	16	14	1.000	18	12	18	12	1.000
63. glass	19	11	28	2	0.010	20	10	27	3	0.057
64. to quarrel	18	12	13	17	0.301	16	14	15	15	1.000
65. fur	20	10	22	8	0.779	20	10	27	3	0.057
66. big	18	12	18	12	1.000	19	11	22	8	0.580
67. carrot	12	18	20	10	0.069	10	20	22	8	0.004
68. to paint	15	15	16	13	0.796	16	14	16	13	1.000
69. part	6	24	16	14	0.015	10	20	19	11	0.038
70. old	18	12	19	11	1.000	16	14	21	9	0.288
71. flower	8	22	20	10	0.004	11	19	22	8	0.009
72. to beat	13	17	16	14	0.606	16	14	21	9	0.288
73. box	11	19	21	9	0.019	13	17	22	8	0.035
74. wild	20	10	22	8	0.779	22	8	24	6	0.761
75. family	10	20	19	11	0.038	10	20	23	7	0.002
76. to wash	13	17	20	10	0.119	16	14	22	8	0.180
77. cow	21	9	25	5	0.360	24	6	25	5	1.000
78. strange	10	20	8	22	0.779	15	15	9	21	0.187
79. happiness	9	21	17	13	0.067	11	19	12	18	1.000
80. to lie (tell untruth)	4	26	7	23	0.506	4	26	10	20	0.125
81. deportment	8	22	12	18	0.412	11	19	14	16	0.601
82. narrow	16	14	16	14	1.000	17	13	19	11	0.792
83. brother	12	18	13	17	1.000	14	16	15	15	1.000
84. to fear	16	14	14	16	0.797	11	19	14	16	0.601
85. stork	16	14	27	3	0.003	14	16	27	3	0.001
86. false	20	10	17	13	0.596	16	14	22	8	0.180
87. anxiety	18	12	18	12	1.000	19	11	19	11	1.000
88. to kiss	19	11	14	16	0.299	16	14	17	13	1.000
89. bride	14	16	22	8	0.064	15	15	20	10	0.295
90. pure	14	16	18	12	0.438	15	15	23	7	0.060
91. door	15	15	23	7	0.060	17	13	23	7	0.170
92. to choose	11	19	13	17	0.792	10	20	11	19	1.000
93. hay	15	15	24	6	0.029	15	15	25	5	0.013
94. contented	8	22	18	12	0.018	9	21	16	14	0.115
95. ridicule	6	24	12	18	0.158	8	22	11	19	0.580
96. to sleep	16	14	18	12	0.795	19	11	22	8	0.580
97. month	17	13	19	11	0.792	22	8	22	8	1.000
98. pretty	25	5	19	11	0.143	21	9	24	6	0.552
99. wife	9	21	19	11	0.019	11	19	17	13	0.195
100. to abuse	16	14	15	15	1.000	20	10	17	13	0.596
101. medicament	11	19	19	11	0.070	14	16	18	12	0.438
102. hopeless, knife, rope	11	19	9	21	0.785	16	14	9	21	0.115

Appendix B. Results of two-way ANOVA ordered by the significance of the factors

Word stimuli	HIPN	TEST	HIPN * TEST	Thematic group
MONTH	0.000	0.000	0.000	4.
HAY	0.000	0.000	0.002	
COW	0.006	0.041	0.005	
MONEY	0.128	0.012	0.008	
LAMP	0.003	0.014	0.017	
TO FALL	0.000	0.005	0.019	
CONTENTED	0.055	0.036	0.021	
RICH	0.058	0.039	0.055	
TO DIE	0.000	0.001	0.177	1.
BLUE	0.000	0.010	0.056	
WIFE	0.000	0.011	0.716	
DEATH	0.000	0.012	0.475	
STEM	0.000	0.016	0.208	
TO SING	0.000	0.026	0.205	
TO BEAT	0.000	0.070	0.202	
TO DIVORCE	0.001	0.001	0.496	
CUSTOM	0.001	0.005	0.895	
TO DESPISE	0.001	0.009	0.125	
FROG	0.002	0.000	0.226	
SEA	0.002	0.001	0.061	
DOOR	0.002	0.002	0.574	
CHILD	0.002	0.010	0.741	
FRIENDLY	0.007	0.006	0.512	
TO WASH	0.007	0.016	0.323	
MOUNTAIN	0.007	0.037	0.538	
TO SIN	0.009	0.008	0.260	
TABLE	0.009	0.025	0.501	
TO LIE (TELL UNTRUTH)	0.010	0.062	0.622	
SICK	0.012	0.005	0.211	
TO SLEEP	0.012	0.045	0.698	
BREAD	0.015	0.001	0.929	
LONG	0.016	0.001	0.188	
TO PAY	0.017	0.048	0.312	
STRANGE	0.018	0.027	0.391	
WINDOW	0.023	0.027	0.427	
PRIDE	0.024	0.001	0.717	
BROTHER	0.027	0.028	0.620	
FINGER	0.033	0.000	0.324	
NEEDLE	0.035	0.015	0.564	
TO PRAY	0.039	0.037	0.744	
PURE	0.041	0.015	0.479	
TO PAY ATTENTION	0.043	0.004	0.953	
OLD	0.043	0.054	0.104	
TO KISS	0.002	0.060	0.828	2.
RIDICULE	0.000	0.127	0.713	
JOURNEY	0.002	0.341	0.324	
INK	0.002	0.396	0.217	
TO ASK	0.004	0.137	0.746	
TO PRICK	0.005	0.123	0.555	
BOX	0.005	0.510	0.406	
SAD	0.008	0.144	0.503	
NARROW	0.013	0.161	0.369	
PENCIL	0.020	0.379	0.666	

Appendix B (cont.)

Word stimuli	HIPN	TEST	HIPN * TEST	Thematic group
GLASS	0.022	0.349	0.622	
BOOK	0.029	0.112	0.472	
EXERCISE-BOOK	0.034	0.460	0.485	
PRETTY	0.047	0.636	0.429	
PART	0.201	0.000	0.879	3.
VILLAGE	0.458	0.000	0.672	
FLOWER	0.659	0.000	0.478	
MEDICAMENT	0.172	0.001	0.461	
TREE	0.276	0.001	0.442	
TO DANCE	0.106	0.002	0.204	
SHIP	0.106	0.002	0.354	
HEAD	0.139	0.002	0.522	
DEPORTMENT	0.166	0.002	0.701	
GREEN	0.291	0.002	0.170	
HUNGER	0.430	0.002	0.437	
UNJUST	0.079	0.004	0.553	
HAPPINESS	0.400	0.004	0.296	
WHITE	0.070	0.007	0.900	
SALT	0.457	0.007	0.911	
BIRD	0.131	0.009	0.275	
TO FEAR	0.109	0.010	0.888	
HOUSE	0.733	0.010	0.066	
TO COOK	0.473	0.012	0.371	
WATER	0.253	0.015	0.303	
FAMILY	0.271	0.015	0.480	
WILD	0.055	0.016	0.549	
BRIDE	0.061	0.021	0.812	
PLUM	0.051	0.025	0.985	
TO MARRY	0.196	0.035	0.967	
TO CHOOSE	0.152	0.039	0.187	
CARROT	0.944	0.054	0.279	
TO PAINT	0.909	0.056	0.847	
COLD	0.572	0.066	0.369	
STUPID	0.665	0.072	0.269	
HOPELESS	0.780	0.087	0.826	
EXPENSIVE	0.845	0.089	0.738	
TO ABUSE	0.096	0.096	0.172	
TO SWIM	0.116	0.120	0.296	
PITY	0.268	0.154	0.202	
NEW	0.301	0.186	0.475	
FUR	0.168	0.192	0.297	
ANXIETY	0.054	0.199	0.631	
ANGRY	0.225	0.208	0.677	
YELLOW	0.107	0.278	0.215	
FALSE	0.315	0.418	0.516	
BIG	0.161	0.442	0.255	
STORK	0.135	0.467	0.356	
DEAR (BE LOVED)	0.077	0.861	0.752	
TO QUARREL	0.089	0.907	0.471	