

JAPANESE
INDUSTRIAL
STANDARD

Translated and Published by
Japanese Standards Association

JIS Z 8101-3 : 1999

(ISO/FDIS 3534-3 : 1999)

Statistics—

Vocabulary and symbols—

Part 3 : Design of experiments

ICS 01.040.03; 03.120.30

Descriptors : vocabulary, marks, symbols, experimental design

Reference number : JIS Z 8101-3 : 1999 (E)

Z 8101-3 : 1999 (ISO/FDIS 3534-3 : 1999)

Foreword

This translation has been made based on the original Japanese Industrial Standard established by the Minister of International Trade and Industry through deliberations at the Japanese Industrial Standards Committee in accordance with the Industrial Standardization Law. By this establishment **JIS Z 8101 : 1981** was withdrawn and replaced with this Standard.

This Standard has been prepared based on **ISO/FDIS 3534-3** issued in 1999.

JIS Z 8101 : 1999 consists of the following parts, under the general title *Statistics—Vocabulary and symbols*:

- Part 1 : *Probability and general statistical terms*
- Part 2 : *Statistical quality control terms*
- Part 3 : *Design of experiments*

Date of Establishment: 1999-05-20

Date of Public Notice in Official Gazette: 1999-05-20

Investigated by: Japanese Industrial Standards Committee

Divisional Council on Basic Items

JIS Z 8101-3:1999, First English edition published in 2000-08

Translated and published by: Japanese Standards Association
4-1-24, Akasaka, Minato-ku, Tokyo, 107-8440 JAPAN

In the event of any doubts arising as to the contents,
the original JIS is to be the final authority.

© JSA 2000

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

Printed in Japan

Contents

	Page
Introduction	1
Scope	2
1 General terms	2
2 Arrangements of experiments	14
3 Methods of analysis	33
Bibliography	41

Statistics—Vocabulary and symbols— Part 3 : Design of experiments

Introduction This Japanese Industrial Standard has been prepared based on ISO/FDIS 3534-3 *Statistics—Vocabulary and symbols—Part 3 : Design of experiments* issued in 1999 as the FDIS edition without changing the technical contents.

The portions underlined with dots in this Standard are not stated in the original International Standard.

Design of experiments is essentially a strategy of planning experiments so that valid and relevant conclusions may be reached efficiently and economically. The selection of the specific experimental plan should depend on the type of question to be addressed, the degree of generality to be attached to the conclusions, and the resources available (experimental material, personnel, time). A properly designed and executed experiment will frequently lead to relatively simple statistical analysis and interpretation of the results.

In recent years, the application of experimental design has flourished, notably due to the recognition that designed experiments are essential for improving the quality of goods and services. Although statistical quality control, management resolve, inspection, and other quality tools also serve this function, experimental design represents the methodology of choice in complex, variable and interactive settings. Historically, design of experiments has evolved and thrived in the agricultural area. Medicine has also enjoyed a long standing history of careful experimental design. Currently, industrial settings are witnessing the considerable benefits of the methodology—due to ease of initiating efforts (user-friendly software packages), improved training, influential advocates, and accumulating successes with experimental design.

Factorial experiments (see 2.1) provide a methodology for studying the interrelationships among multiple factors of interest to the experimenter. These types of experiments can be far more efficient and effective than intuitive one-factor-at-a-time experiments. Factorial experiments are particularly well-suited for determining that a factor behaves differently (as reflected in the experimental response) at different levels of other factors. Frequently, the “breakthrough” in quality comes from the synergism revealed in a study of “interactions” (see 1.17). If the number of factors under consideration is large, then factorial experiments could exceed resources. However, fractional factorial designs (see 2.1.1) offer a possible compromise. Actually, if the initial goal is to identify factors warranting further investigation, then screening designs (see 2.2) can be useful.

In planning an experiment, it is necessary to limit biases introduced by the experimental conditions or assignment of treatments to experimental units. Topics such as “randomization” (see 1.29) and “blocking” (see 1.28) deal with minimizing the effects of nuisance or extraneous elements. Specific blocking strategies include randomized block designs (see 2.3.1), Latin-square designs (see 2.3.2) and variants, and balanced incomplete block designs (see 2.3.4.1).

Viewing design of experiments as an evolutionary process with continuous improvement as a goal, response surface designs (see 2.4) play a pivotal role. By considering multiple levels of key factors, response surface methods neatly accommodate curvilinear effects in the vicinity of optimum points.

Mixture designs (see 2.5) handle situations in which factors constitute proportions of a total, such as ingredients in an alloy. Nested designs (see 2.6) are particularly useful in interlaboratory testing.

Methods of analysis of the collected data are straightforward, if the experiment is carried out according to the plan. Graphical methods (see 3.1) can be particularly effective in revealing overall conclusions. Estimation of parameters from a model (see 1.1 and following) is commonly handled using regression analysis (see 3.3). Regression analysis methods can also handle difficulties with missing data, identification of outliers, and other problems.

Good experimental design should:

- a) incorporate prior knowledge and experience in selection of factors, their levels, and in describing assumptions;
- b) furnish relevant information with minimum effort;
- c) ensure, before starting the experiment, that the design is capable of achieving the objective of the experiment with the desired precision;
- d) reflect the sequential nature of most investigations;
- e) specify both arrangement and sequence of experimental treatments to avoid misunderstandings when the experiment is in progress.

Scope

This part of JIS Z 8101 defines the terms used in the field of design of experiments and may be used in the drafting of other Japanese Industrial Standards.

1 General terms

1.1 model

Description relating the response variable to predictor variable(s) and including attendant assumptions.

Note 1 The model consists of three parts. The first part is the **response** (1.2) that is being modelled. The second part is the deterministic or the systematic part of the model that includes **predictor variable(s)** (1.3). Finally, the third part is the random, error or stochastic part of the model, which can be quite elaborate. For example, the error term can incorporate a **dispersion effect** (1.14) that allows for increasing variability in the response with larger actual values of the response. See also (1.2) and (1.3).

Example 1 The lifetime of a component is related to the environmental conditions that it experiences.

Example 2 A formal model is:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

where,

y_{ij} is the response at level i of factor A and level j of factor B;

μ is the overall mean response;

α_i is the incremental effect of factor A at level i ;

β_j is the incremental effect of factor B at level j ;

ε_{ij} is the error term.

The response part of the model consists simply of y_{ij} . The predictive part of this model is $\mu + \alpha_i + \beta_j$ consisting of an overall mean response and two terms related to the effects of factors. The random or error part of this model consists of ε_{ij} that includes inherent variability in the process which produces the response.

Example 3 A commonly used model is:

$$y_{ijk} = \alpha_i + \beta_j + \tau_{ij} + \varepsilon_{ijk}$$

where,

y_{ijk} is the response of the k th replicate;

α_i is the adjustment due to factor 1;

β_j is the adjustment due to factor 2;

τ_{ij} is the adjustment due to interaction of the factors;

ε_{ijk} is the error term.

The terminology "adjustment" is used instead of "incremental effect" as in example 2, since here the formal mathematical model does not include an overall mean term. Furthermore y_{ijk} (ε_{ijk}) is used in this example rather than y_{ij} (ε_{ij}) to acknowledge the potential existence of replicates.

Example 4 Another formal model is:

$$y_i = e\beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i$$

where,

y_i is the response corresponding to x_i ;

$e\beta_0 + \beta_1 x_i + \beta_2 x_i^2$ represents the mean response corresponding to x_i ;

ε_i is the error term.

Note 2 The above description of a model not only applies to the classical linear models with additive error but also to generalized linear models, where the error can be described by a variety of distributions including the binomial, Poisson, exponential, gamma and normal distributions.

1.2 response variable

Variable representing the outcome of an experiment.

Notes 1 A common synonym is "output variable".

2 The term "dependent variable" is not recommended as a synonym due to potential confusion with independence (JIS Z 8101-1 : 1999, 1.7).

3 It may be that the response variable is vector-valued because several responses are recorded from each experimental unit.

1.3 predictor variable

Variable that can contribute to the explanation of the outcome of an experiment.

Notes 1 Common synonyms include “input variable”.

2 The extent to which a given predictor variable can be controlled dictates its potential role in a designed experiment. Predictor variables can be controllable (fixed), modifiable (controllable only for short duration or at considerable expense) or uncontrollable (random).

3 A predictor variable can include a random element in it or it can, for example, be from a set of qualitative classes which can be observed or assigned without random error.

4 “Independent variable” is not recommended as a synonym due to potential confusion with independence (see JIS Z 8101-1 : 1999, 1.7).

1.4 design region, design space

Set of allowable values for the predictor variables.

Informative reference : It is also called experiment space.

1.5 factor

Predictor variable that is varied with the intent of assessing its effect on the response variable.

Notes 1 A factor may provide an assignable cause for the outcome of an experiment.

2 The use of factor here is more specific than its generic use as a synonym for **predictor variable** (1.3).

3 A factor may be associated with the creation of **blocks** (1.11).

Informative reference : The factor taken for the purpose of choosing the optimal level is called **control factor**. The factor not intending to choose the optimal level but intending to examine the interaction with other control factors is called **indicative factor**.

1.6 level

Potential setting, value or assignment of a factor.

Notes 1 A synonym is the value of a predictor variable.

2 The term “level” is normally associated with a quantitative characteristic. However, it also serves as the term describing the version or setting of qualitative characteristics.

Example : The ordinal-scale levels of a catalyst may be presence and absence. Four levels of a heat treatment may be 100 °C, 120 °C, 140 °C and 160 °C. The nominal-scale variable for a laboratory can have levels A, B and C, corresponding to three facilities.

Note 3 Responses observed at the various levels of a factor provide information for determining the effect of the factor within the range of levels of the experiment. Extrapolation beyond the range of these levels is usually inappropriate without a firm basis for assuming model relationships. Interpolation within the range may depend on the

number of levels and the spacing of these levels. It is usually reasonable to interpolate, although it is possible to have discontinuous or multi-modal relationships that cause abrupt changes within the range of the experiment. The levels may be limited to certain selected fixed values (whether these values are or are not known) or they may represent purely random selection over the range to be studied.

1.7 experimental error

Variation in the response variable beyond that accounted for by the factors, blocks or other attributable sources in the conduct of the experiment.

- Notes 1 It is a common characteristic of experiments that, when repeated, results vary from trial to trial, though the experimental materials, environmental conditions and the experimental operations have been carefully controlled. Thus, experimental error is a common occurrence. This variation introduces a degree of uncertainty into conclusions drawn from these results, and consequently, should be considered when reaching conclusions.
- 2 Specific refinements to this broad conceptual error framework for the individual response variables are provided by the terms **residual** (1.21), **residual error** (1.22) and **pure error** (1.23).
- 3 Of related interest to experimental error are the terms repeatability standard deviation (JIS Z 8101-2 : 1999, 2.12) and reproducibility standard deviation (JIS Z 8101-2 : 1999, 2.15) which apply in the experimental design context directly if the actual design of the experiment is in accordance with repeatability conditions (JIS Z 8101-2 : 1999, 2.13) or reproducibility conditions (JIS Z 8101-2 : 1999, 2.16), respectively.

1.8 variance component

Variance of a random variable describing a factor effect or experimental error.

- Notes 1 In the model, $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, where τ_i is a level chosen at random from an infinite set of values and the distributions of τ_i and ε_{ij} are independent, both τ_i and ε_{ij} are random variables. Once the random selection from the infinite set of possible levels is made, then analysis proceeds on the basis of the realizations of τ_i . In view of the probabilistic structure, it is reasonable to consider an equation involving the variances: $\text{Var}(y_{ij}) = \text{Var}(\tau_i) + \text{Var}(\varepsilon_{ij})$, the right hand side denoted $\sigma_\tau^2 + \sigma_\varepsilon^2$. Symbolically, σ_τ^2 and σ_ε^2 are the variance components of y_{ij} .
- 2 Other models can be envisaged that include nested or crossed factors.

1.9 experimental unit

Entity receiving a particular treatment, subsequently yielding a value of the response variable.

Informative reference : As a result of a treatment given to the experiment unit, the value of response variable is produced.

1.10 treatment

Specific setting of every factor.

1.11 block

Collection of **experimental units** (1.9) more homogeneous than the full set of experimental units (see also 1.28).

Notes 1 The term "block" originated in agricultural experiments in which a field was subdivided into sections having common conditions, such as exposure to the wind, proximity to underground water or thickness of the arable layer. In other situations, blocks are based on batches of raw material, operators, the number of units studied in a day, etc.

2 Generally, recognition of the existence of blocks may affect how the treatments are assigned to experimental units.

1.12 one-factor experiment

Experiment in which a single factor is investigated as to its effect (if any) on the response variable.

Example : Consider the model:

$$y = \mu_i + \varepsilon$$

where,

y is the response variable;

μ_i is the mean response at the i th level of the factor;

ε is a random variable capturing all other effects and sources of variation.

This model relates the response variable y to the effect μ_i (depending on the corresponding level of the factor) and an error term ε . Differences in the μ_i reflect the influence of the factor on the response variable (in this case the mean response value as a function of the level of the factor).

An alternate representation of this model is

$$y = \mu + \alpha_i + \varepsilon$$

where,

y is the response variable;

μ is the overall mean response;

α_i is the incremental effect due to the i th level of the factor;

ε is a random variable capturing all other effects and sources of variation.

1.13 main effect

Influence of a single factor on the mean of the response variable.

Note : For a factor with two levels, the main effect relates to the change in the response in going from one level to the other. If the levels are designated

-1 (for low) and +1 (for high), then the main effect of the factor is estimated as the average response when the factor level is +1 minus the average response when the factor level is -1. Consider the model:

$$y = \mu + \beta X + \varepsilon$$

where y , μ , and ε are as in 1.12, X is +1 or -1 as just described, and β represents the adjustment for the factor X . Note that an estimate of β is equal to one half the main effect for the factor X . If β were equal to zero, then X does not affect the mean of the response variable (it is the same regardless of the level of X being +1 or -1) so that the main effect of X is zero.

1.14 dispersion effect

Influence of a single factor on the variance of the response variable.

Note : It is important to recognize that a factor that does not have much influence on the mean response may have dramatic effects on the variability of the response. In such situations, a particular level of the factor can be much more desirable in terms of achieving low variability or consistency in the responses. It is also possible that a factor can influence both the mean and the variance of the response variable.

1.15 two-factor experiment

Experiment in which two distinct factors are simultaneously investigated for possible effects on the response variable.

Note : If the two factors operate without interfering with each other, the term "main effect" necessarily still applies. Namely, for each factor the main effect is its contribution to the mean of the response variable.

1.16 k -factor experiment

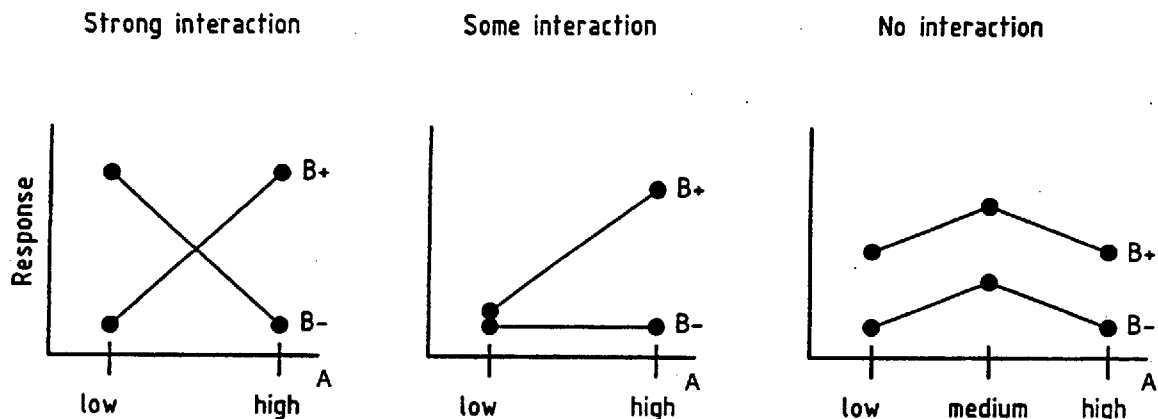
Experiment in which k distinct factors ($k \geq 2$) are simultaneously investigated for possible effects on the response variable.

Note : A synonym is "multi-factor experiment".

1.17 interaction

Effect for which the apparent influence of one factor on the response variable depends upon one or more other factors.

Note 1 Interaction indicates an inconsistency of the main effect of a factor on the response depending on the level of another factor. Differential effect is used as a synonym. The following figure indicates these phenomena.



Notes 2 Most commonly, interactions are considered involving only two factors and are more precisely referred to as either two-way interactions or first order interactions. Of course, it is possible that three factors, for example A, B, and C, interact in the sense that the first order interaction of AB depends on the level of factor C. In this case, there is a second order interaction. Similarly, third, fourth, and higher order interactions can be conceived.

3 Example 3 in 1.1 provides a formal model representation of an experiment with two factors and the two-way or first order interaction t_{ij} between them.

Informative reference : Main effect and interaction are called generically factorial effect.

1.18 confounding

Combining deliberately two or more effects (main and/or interaction) so as to be indistinguishable.

Note : Confounding is an important technique which permits, for example, the effective use of specified blocks in some experimental designs. This is accomplished by deliberately pre-selecting certain effects (main or interactions) as being of little interest, and arranging the design so that it confounds them with block effects, while keeping the other more important effects free from such complications. Confounding may be deliberately used to diminish the number of trials of the **experimental plan** (1.30). Sometimes, however, confounding results from inadvertent changes to a design during the running of an experiment or from incomplete planning of the design, and it serves to diminish, or even to invalidate, the effectiveness of an experiment.

1.19 alias

⟨Statistics⟩ Effect (main or interaction) that is completely confounded with another main effect or interaction due to the nature of the experiment.

1.20 curvature

Departure from a straight line relationship between the response variable and a predictor variable.

- Notes 1 Curvature has meaning with quantitative predictor variables, but not with categorical (nominal) or qualitative (ordinal) predictor variables. Detection of curvature requires more than two levels of the factors. In some instances, replicated centre points (the factor set midway between the high and low settings of the factors) can allow the detection and assessment of curvature. Alternatively, an expanded range of the levels of the factor can be necessary to observe curvature.
- 2 Returning to the model given in the example of 1.12, curvature can be readily modelled via a form such as:

$$Y = \mu + \beta X + \gamma X^2 + \varepsilon$$

If γ deviates from zero, there is evidence of curvature relative to the simple linear relation.

1.21 residual

Difference between an observed value of the response variable and the corresponding predicted value of the response variable.

Note : The predicted value of the response variable is based upon an assumed model, the parameters of which are estimated from the data.

Example 1 $y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$ is the residual corresponding to the experimental unit with factor A set at level i and with factor B set at level j using the model in example 2 from 1.1.

Example 2 $y_{ijk} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\tau}_{ij}$ is a residual corresponding to the model in example 3 from 1.1.

Example 3 $y_i - e^{\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2}$ is a residual corresponding to the model in example 4 from 1.1.

1.22 residual error

Random variable representing the difference between outcomes of the response variable and the corresponding predicted response values based on an assumed model.

Note 1 For the purpose of this definition, the term "predicted response value" is understood to be the estimated response for that treatment determined from the empirical model derived from the data of the experiment using the assumed model.

Example : If $\hat{\mu}$ and $\hat{\beta}$ were the estimators of μ and β , respectively, in the note of 1.13, then $y - \hat{\mu} - \hat{\beta}x$ is the residual error given the observed value of y at the predictor variable value of x .

Notes 2 Residual error includes experimental error and assignable sources of variation not taken into account by the model.

3 The variance of the "residual error" is usually estimated in an experiment by subtracting the pooled sum of squares for terms included in the assumed model from the total sum of squares and dividing by the corresponding difference in "degrees of freedom". (See example 1 in 3.3 and the example in 3.4.)

Informative reference : Making a square sum by adding plural square sums is called **pooling**. The degree of freedom of the square sum after pooling is the sum of the degree of freedom of the corresponding square sum.

1.23 pure error

Random variable reflecting variability associated with replicated observations at a fixed treatment combination.

Note 1 If only the centre point in a design were replicated, then the sample variance of responses at the centre point provides an estimate of the variance of the pure error. If replicates took place at multiple treatment combinations, then an overall estimate of the variance of the pure error can be achieved by pooling the estimates at these treatment combinations.

Example : Returning to example 3 in 1.1, an estimate of the variance of pure error for fixed (i, j) is $\frac{1}{n_{ij}-1} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij})^2$, where $\bar{y}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk}$. If replicates occurred at each (i, j) combination, a pooled estimate of the variance of pure error would be of the form $\frac{1}{N-IJ} \sum_{i,j,k} (y_{ijk} - \bar{y}_{ij})^2$, where $i = 1, \dots, I; j = 1, \dots, J; k = 1, \dots, n_{ij}$.

Note 2 The term pure error is used in practice in two different ways. Sometimes it refers to a population variance (σ^2) in association with a mathematical model. On other occasions, pure error refers to the "sample" or "empirical" pure error which in conjunction with the estimated residual error provides the basis for a lack of fit test of the model. Of the examples illustrating models in 1.1, only example 3 having replicates would facilitate direct estimation of the pure error. From the mathematical perspective, the pure error can be construed as $\text{Var}(\varepsilon_{ij})$ in example 2, $\text{Var}(\varepsilon_{ijk})$ in example 3, and $\text{Var}(\varepsilon_i)$ in example 4.

1.24 contrast

⟨Statistics⟩ Linear function of the response values for which the sum of the coefficients is zero with not all coefficients equal to zero.

Informative reference : It is sufficient if there is any one other than zero in a part of the coefficients.

Note : With observations $y_1, y_2, \dots, y_n$, the linear function $a_1y_1 + a_2y_2 + \dots + a_ny_n$ is a contrast if and only if $a_1 + a_2 + \dots + a_n = 0$ with not all a_i 's equal to zero.

Example 1 A factor is applied at three levels and the results are represented by y_1, y_2 and y_3 . Among the many possible questions to be asked, consider the difference in responses at the first and third level of the experiment (ignoring temporarily the intermediate level). The appropriate contrast for assessing this query is given below (see question 1) and requires the values of y_1 and y_3 . If the levels are equally spaced, a second question (question 2) may be asked as to whether there is evidence that the response pattern shows (a quadratic) curvature rather than a simple linear trend. Here the average of y_1 and y_3 can be compared to y_2 . (If there is no curvature, y_2 should fall close to the line connecting y_1 and y_3 or, in other words, be approximately equal to their average.)

Response	y_1	y_2	y_3
Contrast coefficients for question 1	-1	0	+1
Contrast 1	$-y_1$		$+y_3$
Contrast coefficients for question 2	-1/2	+1	-1/2
Contrast 2	$-y_1/2$	y_2	$-y_3/2$

This example illustrates a regression type study of continuous variables. It is frequently more convenient to use integers rather than fractions for contrast coefficients. In such a case, the coefficients for Contrast 2 would appear as (-1, +2, -1).

Example 2 Another example dealing with discrete levels of a factor may lead to a different pair of questions. Let us suppose there are three sources of supply, one of which, A_1 , uses a new manufacturing technique while the other two, A_2 and A_3 , use the customary one. First, does vendor A_1 with the new technique seem to differ from A_2 and A_3 using the old one? Contrast y_1 with the average of y_2 and y_3 . Second, do the two suppliers using the customary technique differ? Contrast y_2 and y_3 . The pattern of contrast coefficients is similar to that for the previous problem, though the interpretation of the results will differ.

Response	y_1	y_2	y_3
Contrast coefficients for question 1	-2	+1	+1
Contrast 1	$-2y_1$	$+y_2$	$+y_3$
Contrast coefficients for question 2	0	-1	+1
Contrast 2		$-y_2$	$+y_3$

1.25 orthogonal contrast

Set of contrasts whose coefficients satisfy the condition that, if multiplied in corresponding pairs, the sum of the products equals zero.

Example 1

		y_1	y_2	y_3
a_{i1}	Contrast 1	-1	0	+1
a_{i2}	Contrast 2	0	-1	+1
$a_{i1}a_{i2}$		0	0	+1

$\Sigma a_{i1}a_{i2} = 1$ therefore not orthogonal

Example 2

		y_1	y_2	y_3
a_{i1}	Contrast 1	-1	0	+1
a_{i2}	Contrast 2	-1	+2	-1
$a_{i1}a_{i2}$		+1	0	-1

$\Sigma a_{i1}a_{i2} = 0$ therefore orthogonal