

FUZZY ESTIMATION OF ONE VAGUE MISSING VALUE IN TWO-FACTOR EXPERIMENTS

TENG-SAN SHIH¹, JIN-SHIEH SU¹ AND HUEY-MING LEE²

¹Department of Applied Mathematics

²Department of Information Management

Chinese Culture University

55, Hwa-kang Road, Yang-Ming-Shan, Taipei, Taiwan

{ sts; sjs; hmlee }@faculty.pccu.edu.tw

Received July 2008; revised December 2008

ABSTRACT. *In this paper we present the two-factor experimental design, which is suited for one vague missing data value. Its suitability owes to the fact that this vague datum has fuzzy probabilistic measurement properties. Therefore, to estimate the missing value in the fuzzy sense, we used the concept of a probabilistic fuzzy set.*

Keywords: Factor experiment, Vague, Missing data, Fuzzy sets

1. **Introduction.** In statistical analyses, controlled experiments are carefully designed to facilitate the use of straightforward computations. In particular, these experiments correspond to a standard classical experiment design, that is, they employ a standard least squares analysis. In turn, this analysis yields the following: estimates of parameters and the analysis of variance (ANOVA) table. Due to balance in the design, the estimates and ANOVA table that correspond to most designed experiments are easily computed. For example, with two factors, the analysis is particularly simple when the same numbers of observations are taken at each combination of factor levels. However, it occurs occasionally in a randomized block design that an observation is lost in some block. There are various reasons for these phenomena such as, a vial may have broken, mechanical breakdowns may result in an incomplete result from an industrial experiment, or the respondents in a household survey may refuse to report their income. As a consequence, the data set lacks one or more observations. The missing value causes a new problem in the conduct of the analysis because the treatment and the block are not orthogonal. For a single factor, the completely randomized design with a missing value presents no problem because the analysis of variance can be run with an unequal number of observations. However, this is not the case for a two-way analysis as this incompleteness results in a loss of orthogonality.

In the last decade, many researchers who have conducted analyses of two-factor experiments [7] have examined the missing value using probability and statistical methodology [4]. In statistics textbooks [1], the usual procedure involves replacing the missing value with one that minimizes the sum of the squares of the error. Thereafter, these data are treated as a true value. Subsequently, in analyzing variance, an approximate ANOVA is presented. The better statistical computing packages such as SPSS or SAS contain a general linear models (GLM) procedure that will calculate the sum of squares and prepare an appropriate ANOVA summary for a data table with missing values. Let x_{pq} be the missing data values. In effect, model Equation (1), the missing data x_{pq} cannot be expressed as $\mu + \alpha_p + \beta_q + \varepsilon_{pq}$. For missing data x_{pq} , the parameters μ , α_p , and β_q cannot be denoted as a value. We know from other studies and from previous experience that the