Statistical testing procedure for the interaction effects of several controllable factors in two-valued input-output systems

Satoshi Aoki
Department of Mathematics and Computer Science
Kagoshima University
and
Masami Miyakawa
Department of Industrial Engineering and Management
Tokyo Institute of Technology

June, 2007

Abstract

Suppose several two-valued input-output systems are designed by setting the levels of several controllable factors. For this situation, Taguchi method has proposed to assign the controllable factors to the orthogonal array and use ANOVA model for the standardized SN ratio, which is a natural measure for evaluating the performance of each input-output system. Though this procedure is simple and useful in application indeed, the result can be unreliable when the estimated standard errors of the standardized SN ratios are unbalanced. In this paper, we treat the data arising from the full factorial or fractional factorial designs of several controllable factors as the frequencies of high-dimensional contingency tables, and propose a general testing procedure for the main effects or the interaction effects of the controllable factors.

Keywords: Confoundings, Contingency tables, Controllable factors, Covariate matrix, Generalized linear models, Hierarchical models, Fractional factorial designs, Full factorial designs, Standardized SN ratio, Sufficient statistics, Two-valued input-output systems.

1 Introduction

In this paper, we consider evaluating performance of several two-valued input-output systems. Before introducing our motivated problem, first we give a typical example of a *single* two-valued input-output system and review the measure for evaluating its performance. Suppose a vending machine judges an inserted coin as fair coin or false coin. In this system, the input is the *true state* of the inserted coin, {fair, false}, whereas the output is the *judged state* of the coin, {fair, false}. For evaluating performance of this machine, prepare n_1 fair coins and n_2 false coins and insert them to the machine. The result of the judgment by this machine is summarized as a 2×2 contingency table as follows.

	$y = 1$	$y = 2$	total
$M = 1$	n_{11}	n_{12}	n_1
$M = 2$	n_{21}	n_{22}	n_2

In the above table, two-valued signal M is the true state of the coin, and y is the judged state by this machine (1: fair, 2: false). A natural statistical model for this experiment is two independent binomial distributions. We introduce random variables, N_{11}, N_{21}, and parameters $p_{ij},\ i,j = 1,2$, and write $N_{i1} \sim Bin(n_i, p_{i1}),\ i = 1,2$. p_{ij} is the probability that a coin of true state i is judged as a state j. Here $p_{i1} + p_{i2} = 1, i = 1,2$ holds. For this type of the systems, the error of judging a fair coin as false (type I error) and the error of judging a false coin as fair (type II error) are in the trade-off relation. Therefore we have to take into account the two types of errors for evaluating the performance of the system. In Taguchi method, a standardized Signal-Noise (SN) ratio,

$$\hat{\eta} = -\log\left[\frac{1}{(1-2\hat{p}_0)^2} - 1\right], \tag{1}$$

is proposed as a measure to quantitatively evaluate performance of this type of system, where

$$\hat{p}_0 = \frac{1}{1+\sqrt{\hat{\theta}}}$$

is the estimated common error rate, and

$$\hat{\theta} = \frac{n_{11}n_{22}}{n_{12}n_{21}} \tag{2}$$

is the sample odds ratio. See Taguchi (1987) and Taguchi (1991). Note that the standardized SN ratio (1) is a function of the sample odds ratio.

Now we suppose several two-valued input-output systems are designed by setting the levels of several controllable factors, which is a situation that we mainly consider in this paper. Table 1 is an example of such situations shown in Section 7.3.2 of Miyakawa (2006). In this experiment, 40 normal products and 20 failure products are judges whether normal or failure, by a testing inspection machine, which makes use of the leaking helium gas, in the casing process of a compressor. There are three controllable factors which are considered to have influence on performance of the inspection, each of which is settled on one of the two levels. In Table 1, the meaning of n_{ij} is the same to the previous example. Note that the sample odds ratio (2) is modified as

$$\hat{\theta} = \frac{(n_{11}+0.5)(n_{22}+0.5)}{(n_{12}+0.5)(n_{21}+0.5)}, \tag{3}$$

Table 1: The result of an experiment judging 40 good products and 20 bad products as good or bad by a testing inspection machine shown in Miyakawa (2006)

No.	A	B	C	n_{11}	n_{12}	n_{21}	n_{22}	$\hat{p}_0$	$\hat{\eta}$
1	1	1	1	27	13	11	9	0.435	-17.689
2	1	1	2	25	15	3	17	0.259	-5.169
3	1	2	1	17	23	8	12	0.489	-32.873
4	1	2	2	15	25	2	18	0.320	-8.295
5	2	1	1	40	0	12	8	0.119	1.427
6	2	1	2	38	2	10	10	0.203	-2.638
7	2	2	1	40	0	11	9	0.109	1.974
8	2	2	2	37	3	10	10	0.234	-4.037

Table 2: Results of ANOVA for Table 1

Factors	df	Sum sq	Mean sq	F value	p value
A	1	461.35	461.35	18.3578	0.02336$*$
B	1	45.90	45.90	1.8263	0.26944
C	1	91.27	91.27	3.6319	0.15276
A $\times$ C	1	278.17	278.17	11.0689	0.04482$*$
Residuals	3	75.39	25.13		

since some zeros are included in the table. For this type of the data, Taguchi method has proposed to use ANOVA model for the standardized SN ratio. For the data of Table 1, the main effects and the interaction effects of the controllable factors are evaluated as Table 2. Note that we include the interaction effects A $\times$ B and B $\times$ C into the residuals. This result suggests the existence of the interaction effect A $\times$ C in addition to the main effects A and C. Therefore an optimal condition is considered as $\mathrm{A}_2\mathrm{C}_1$, and the controllable factor B does not have a significant influence on performance of the inspection.

The above procedure is simple and useful in applications. However, it seems that the result can be unreliable when the estimated standard errors of the standardized SN ratios are unbalanced. This is caused by the fact that the standardized SN ratio is only a function of the sample odds ratio, and therefore the sample size for each run is not appropriately taken into account in the procedure. To clarify the problem, divide the sample sizes for the runs No. 5 and 6 by 4 and 2, respectively, which yields Table 3. Of course, the result of ANOVA is almost the same to Table 2 for this situation. Note that the difference is only caused by the modification (3). However, it is obvious that the relatively good contributions of the interaction effect A $\times$ C in the run 5 should be underestimate. Consequently, the p value for the interaction effect A $\times$ C can increase.

In this paper, we consider inference for the main and the interaction effects of several controllable factors in two-valued input-output systems. For the designed experiments with counts data, the theory of the generalized linear models (McCullagh and Nelder,

Table 3: Imaginary data set

No.	A	B	C	n_{11}	n_{12}	n_{21}	n_{22}
1	1	1	1	27	13	11	9
2	1	1	2	25	15	3	17
3	1	2	1	17	23	8	12
4	1	2	2	15	25	2	18
5	2	1	1	10	0	3	2
6	2	1	2	19	1	5	5
7	2	2	1	40	0	11	9
8	2	2	2	37	3	10	10

1989) are frequently used. See Hamada and Nelder (1997), Chapter 13 of Wu and Hamada (2000) or Aoki and Takemura (2006) for example. We also rely on the general theory of the generalized linear models and construct a general testing procedure for various hypotheses of the interaction effects in Section 2. Note that our settings are not limited to the full factorial designs. Considering aliasing relations carefully, fractional factorial designs are also treated by our procedure. In particular, we focus on the relation to the models for the high-dimensional contingency tables in Section 3. For example, as we will show, the data of Table 1 can be treated as 2^5 contingency table, and the model A $\times$ C is shown to be equivalent to one of the hierarchical models for the five-dimensional contingency tables. In Section 4, we give some numerical examples and show the effectiveness of our procedure.

2 Conditional tests for the interaction effects of the multiple controllable factors

Suppose there are K two-valued input-output systems, each of which is constructed by setting the levels of several controllable factors. As we have seen in Section 1, the observation for each run is summarized as 2×2 contingency table. We write the observation for the kth run as

$$n_{11k}, n_{12k}, n_{21k}, n_{22k}. \tag{4}$$

It is natural to suppose an independent binomial model for the observations. We write the occurrence parameters as $p_{ijk},\ i,j = 1,2;\ k = 1,\dots,K$, where $p_{i1k} + p_{i2k} = 1$ for $i = 1,2;\ k = 1,\dots,K$. The likelihood function is written as

$$\prod_{k=1}^{K}\prod_{i=1}^{2}\prod_{j=1}^{2}\binom{n_{1\cdot k}}{n_{11k}}\binom{n_{2\cdot k}}{n_{21k}} p_{ijk}^{n_{ijk}},$$

where $n_{i\cdot k} = n_{i1k} + n_{i2k}$. In this paper, following the convention of the analysis of the contingency tables, we use the similar dot notations to express various marginal totals of the observations. For example, we write

$$n_{ij\cdot} = \sum_{k=1}^{K} n_{ijk},\ n_{i\cdot\cdot} = \sum_{j=1}^{2}\sum_{k=1}^{K} n_{ijk}$$

and so on.

To express various models for the parameter p_{ijk}, we use the theory of the generalized linear models. Since we know that the odds ratio

$$\theta_k = \frac{p_{11k}p_{22k}}{p_{12k}p_{21k}}$$

is a good measure for evaluating performance of each system, we assume the structure

$$\log\theta_k = \beta_0 + \beta_1 x_{k1} + \cdots + \beta_{\nu-1}x_{k\nu-1},\ k = 1,\ldots,K$$

where $x_{k1},\ldots,x_{k\nu-1}$ are the covariates. We write the ν-dimensional parameter β and the covariate matrix X as

$$\beta = (\beta_0, \beta_1, \ldots, \beta_{\nu-1})' \tag{5}$$

and

$$X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1\nu-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{K1} & \cdots & x_{K\nu-1} \end{pmatrix} = \begin{pmatrix} \mathbf{1}_K & \mathbf{x}_1 & \cdots & \mathbf{x}_{\nu-1}, \end{pmatrix}$$

where $\mathbf{1}_K = (1,\ldots,1)'$ is the K-dimensional column vector. We also define the K-dimensional frequency vector

$$\mathbf{n} = (n_{111},\ldots,n_{11K})'.$$

Then the likelihood function is written as

$$\begin{aligned}
& \prod_{k=1}^{K}\prod_{i=1}^{2}\prod_{j=1}^{2}\binom{n_{1\cdot k}}{n_{11k}}\binom{n_{2\cdot k}}{n_{21k}}p_{ijk}^{n_{ijk}} \\
= & \prod_{k=1}^{K}\binom{n_{1\cdot k}}{n_{11k}}\binom{n_{2\cdot k}}{n_{21k}}p_{12k}^{n_{1\cdot k}}p_{22k}^{n_{2\cdot k}}\left(\frac{p_{21k}}{p_{22k}}\right)^{n_{\cdot 1k}}\exp(n_{11k}\log\theta_k) \\
= & \left[\prod_{k=1}^{K}\binom{n_{1\cdot k}}{n_{11k}}\binom{n_{2\cdot k}}{n_{21k}}p_{12k}^{n_{1\cdot k}}p_{22k}^{n_{2\cdot k}}\left(\frac{p_{21k}}{p_{22k}}\right)^{n_{\cdot 1k}}\right]\exp\left(\sum_{k=1}^{K}n_{11k}\log\theta_k\right) \\
= & \left[\prod_{k=1}^{K}\binom{n_{1\cdot k}}{n_{11k}}\binom{n_{2\cdot k}}{n_{21k}}p_{12k}^{n_{1\cdot k}}p_{22k}^{n_{2\cdot k}}\left(\frac{p_{21k}}{p_{22k}}\right)^{n_{\cdot 1k}}\right]\exp\left(\beta_0\mathbf{1}_K'\mathbf{n} + \sum_{j=1}^{\nu-1}\beta_j\mathbf{x}_j'\mathbf{n}\right) \\
= & \left[\prod_{k=1}^{K}\binom{n_{1\cdot k}}{n_{11k}}\binom{n_{2\cdot k}}{n_{21k}}p_{12k}^{n_{1\cdot k}}p_{22k}^{n_{2\cdot k}}\left(\frac{p_{21k}}{p_{22k}}\right)^{n_{\cdot 1k}}\right]\exp\left(\beta' X'\mathbf{n}\right),
\end{aligned}$$

which implies that the sufficient statistic for the parameter is $n_{i\cdot k}, n_{\cdot jk}, i,j = 1,2, k = 1,\ldots,K$, which are the marginal totals of the 2×2 table of (4), and $X'\mathbf{n} = (\mathbf{1}_K'\mathbf{n}, \mathbf{x}_1'\mathbf{n}, \ldots, \mathbf{x}_{\nu-1}'\mathbf{n})$.

Now we consider the covariate matrix X. In this paper we consider the typical situation that the run sequence of the experimental units is allocated to each row of an orthogonal design matrix. For example, the run sequence of Table 1 in Section 1 is allocated to the 2^3 full factorial design. We write the orthogonal design matrix where the run sequence of the experimental units is allocated as $K \times p$ matrix D, where each element is $+1$ or -1. For example of Table 1, we have

$$D = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ -1 & -1 & -1 \end{pmatrix}. \tag{6}$$

In this paper, we only consider the situation that each controllable factor has two levels. For later use, we write $D = (d_{ij}) = (\mathbf{d}_1, \ldots, \mathbf{d}_p)$ where $\mathbf{d}_j = (d_{1j}, \ldots, d_{Kj}) \in \{-1, +1\}^K$ is the j-th column vector of D. We also define a simple convention

$$\mathbf{d}_{st} = (d_{1s}d_{1t}, \ldots, d_{Ks}d_{Kt})'$$

and

$$\mathbf{d}_{stu} = (d_{1s}d_{1t}d_{1u}, \ldots, d_{Ks}d_{Kt}d_{Ku})'$$

for $1 \leq s < t < u \leq p$.

The matrix X is constructed from the design matrix D to reflect the main effects of the controllable factors and their interactions which we intend to measure. For example, a simple model which only includes the main effects of each controllable factor is given as $X = (\mathbf{1}_K \ D)$. Of course, we can consider more complicated models containing various interaction effects. In particular, the saturated model includes K parameters. When K is a power of 2, the covariate matrix X for the saturated model is the Hadamard matrix of order K. In the case of 2^3 full factorial design (6), for example, the covariate matrix for the saturated model is given as

$$X = (\mathbf{1}_K \ \mathbf{d}_1 \ \mathbf{d}_2 \ \mathbf{d}_{12} \ \mathbf{d}_3 \ \mathbf{d}_{13} \ \mathbf{d}_{23} \ \mathbf{d}_{123}). \tag{7}$$

Since the saturated model cannot be tested, we consider an appropriate submodel of the saturated model. For the purpose of illustration, we again focus on the example of Table 1. Since the analysis in Section 1 implies the model of the two main effects A, C and the interaction effect A $\times$ C, we treat this model as the *null model* and consider significance tests. Hereafter, we write this model as AC by the manner of the hierarchical models.

This null hypothesis can be described by the parameter β as follows. Permuting the columns of (7), we partition the covariate matrix X of the saturated model as

$$\begin{aligned} &X = (X_0 \ X_1), \\ &X_0 = (\mathbf{1}_K \ \mathbf{d}_1 \ \mathbf{d}_3 \ \mathbf{d}_{13}) = (\mathbf{1}_K \ \mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3), \\ &X_1 = (\mathbf{d}_2 \ \mathbf{d}_{12} \ \mathbf{d}_{23} \ \mathbf{d}_{123}) = (\mathbf{x}_4 \ \mathbf{x}_5 \ \mathbf{x}_6 \ \mathbf{x}_7), \end{aligned}$$

and consider the corresponding parameter $\beta = (\beta_0, \beta_1, \ldots, \beta_7)$. Note that $\nu - 1 = 3$ in (5) in this case. Then the null hypothesis is described as

$$\mathrm{H}_0 : \ \beta_\nu = \cdots = \beta_{K-1} = 0.$$

Under the null hypothesis H_0, the nuisance parameters are $\beta_0, \ldots, \beta_{\nu-1}$ and the sufficient statistic for the nuisance parameters is written as

$$\{n_{i\cdot k}\},\ \{n_{\cdot jk}\},\ i,j = 1,2;\ k = 1,\ldots,K,\ X_0'\mathbf{n}. \tag{8}$$

Then by the theory of the similar tests, we can consider the conditional tests based on the conditional distribution of $\mathbf{n}$ given (8).

Now we consider significance tests of null hypothesis H_0, against various alternative hypothesis H_1. Again for the purpose of illustration, we consider the example of Table 1. In this case, an important alternative is to test the effect of a single additional effect, the main effect of B. This alternative hypothesis is written as

$$\mathrm{H}_1 : \ \beta_\nu \neq 0,\ \beta_{\nu+1} = \cdots = \beta_{K-1} = 0. \tag{9}$$

Or we can also consider the goodness-of-fit of the null hypothesis. In this case, the alternative hypothesis is written as

$$\mathrm{H}_1 : \ (\beta_\nu = \cdots = \beta_{K-1}) \neq (0, \ldots, 0).$$

Depending on the alternative hypothesis, we can use appropriate test statistic $T(\mathbf{n})$. For the alternative hypothesis written as (9), for example, a typical test statistic is a deviance function

$$2 \sum_{i,j,k} n_{ijk} \log \frac{\widetilde{n_{ijk}}}{\widehat{n_{ijk}}},$$

where $\widehat{n_{ijk}}$ and $\widetilde{n_{ijk}}$ are the fitted values under H_0 and H_1, respectively. Note that $\{\widehat{n_{ijk}}\}$ and $\widetilde{n_{ijk}}$ are calculated from the sufficient statistics under the hypothesis, i.e.,

$$\{n_{i\cdot k}\},\ \{n_{\cdot jk}\},\ i,j = 1,2;\ k = 1,\ldots,K,\ X_0'\mathbf{n}$$

and

$$\{n_{i\cdot k}\},\ \{n_{\cdot jk}\},\ i,j = 1,2;\ k = 1,\ldots,K,\ X_0'\mathbf{n},\ \mathbf{x}_\nu'\mathbf{n},$$

respectively. In Section 4, we perform various tests for Table 1.

Finally of this section, we give a brief remark on the case that the design is fractional factorial designs. From the arguments above, it is obvious that our procedure is also applicable for the case of fractional factorial designs. All that we have to pay attention is the consideration on the aliasing relation when we construct a model including interaction effects. To illustrate, we again consider Table 1. Suppose there is another controllable factor D. If we only consider eight-run design ($K = 8$), the main effect of D has to be confounded to some interaction effect of $\mathrm{A}, \mathrm{B}, \mathrm{C}$, i.e., 2^{4-1} fractional factorial design is considered. When we define D = AC, for example, at most only one of the main effect D and the interaction effect A $\times$ C is estimable. Similarly, the interaction effects A $\times$ D and C $\times$ D are also not estimable when the main effects C and A exist, respectively. Then the resolution of such design is III. On the other hand, if we define D = ABC, we can estimate all the two-factor interaction effects, under the constraints that at most only one of $(\mathrm{A} \times \mathrm{B}, \mathrm{C} \times \mathrm{D})$, $(\mathrm{A} \times \mathrm{C}, \mathrm{B} \times \mathrm{D})$ and $(\mathrm{A} \times \mathrm{D}, \mathrm{B} \times \mathrm{C})$ is included in the model. The resolution of the design is IV. In Section 4, we consider such situations for Table 1.

3 Relation to the high-dimensional contingency tables

In Section 2, we give a general procedure to describe models by the covariate matrix X and the parameter β, and to describe the sufficient statistic under the models. As we have seen, statistical tests are based on some discrepancy measure between the fitted values under the null and the alternative hypotheses, which are calculated from the sufficient statistics of the form (8). In this section, we give an interpretation of the sufficient statistics by considering the high-dimensional contingency tables. The main concepts of the arguments in this section are first shown in the previous work by one of the authors, Aoki and Takemura (2006).

Full factorial designs When the experiment is allocated to the full factorial design, there is a direct correspondence to the high-dimensional contingency tables as follows. Suppose the design is 2^p full factorial design, where $2^p = K$. In this case, we can treat the observations $n_{ijk},\ i,j = 1,2;\ k = 1,\ldots,K$ as if they are the frequencies of 2^{p+2} contingency tables. To illustrate this point, we rewrite the cell indices of the frequency as $n_{ija_1a_2\cdots a_p}$, where $i,j,a_1,\ldots,a_p = 1,2$. Consequently, the observation of Table 1, i.e., the case of $p = 3$, is written as follows.

No.	A	B	C	n_{11}	n_{12}	n_{21}	n_{22}
1	1	1	1	n_{11111}	n_{12111}	n_{21111}	n_{22111}
2	1	1	2	n_{11112}	n_{12112}	n_{21112}	n_{22112}
3	1	2	1	n_{11121}	n_{12121}	n_{21121}	n_{22121}
4	1	2	2	n_{11122}	n_{12122}	n_{21122}	n_{22122}
5	2	1	1	n_{11211}	n_{12211}	n_{21211}	n_{22211}
6	2	1	2	n_{11212}	n_{12212}	n_{21212}	n_{22212}
7	2	2	1	n_{11221}	n_{12221}	n_{21221}	n_{22221}
8	2	2	2	n_{11222}	n_{12222}	n_{21222}	n_{22222}

Now we focus on the sufficient statistics under various models in this notation. For notation convenience, we use a, b, c instead of a_1, a_2, a_3 here. We have already seen that the marginal totals

$$\{n_{i\cdot abc}\},\ \{n_{\cdot jabc}\}$$

are included in the sufficient statistic for every model. Under the main effect model A/B/C, the sufficient statistic is given as

$$\{n_{i\cdot abc}\},\ \{n_{\cdot jabc}\},\ \{n_{ija\cdot\cdot}\},\ \{n_{ij\cdot b\cdot}\},\ \{n_{ij\cdot\cdot c}\}.$$

We know that this is the sufficient statistics of the hierarchical model

$$M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{A}/My\mathrm{B}/My\mathrm{C}$$

in the five-way contingency tables. On the other hand, under the model of AC, i.e., two main effects A, C and the interaction effect A $\times$ C, the sufficient statistic is given as

$$\{n_{i\cdot abc}\},\ \{n_{\cdot jabc}\},\ \{n_{ija\cdot c}\},$$

which is the sufficient statistic of the hierarchical model

$$M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{AC}$$

in the five-way contingency tables. In the same way, all the hierarchical models of the effects of the controllable factors $\mathrm{A, B, C}$ can be characterized as the corresponding hierarchical models in the 2^5 contingency tables. We give the relations in Table 4. Similarly,

Table 4: Hierarchical models of the effects of the controllable factors $\mathrm{A, B, C}$ and their corresponding hierarchical models in 2^5 contingency tables (2^3 full factorial design)

Models for $\mathrm{A, B, C}$	Models for $M, y, \mathrm{A, B, C}$
A	$M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{A}$
A/B	$M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{A}/My\mathrm{B}$
AB	$M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{AB}$
A/B/C	$M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{A}/My\mathrm{B}/My\mathrm{C}$
AB/C	$M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{AB}/My\mathrm{C}$
AB/AC	$M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{AB}/My\mathrm{AC}$
AB/AC/BC	$M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{AB}/My\mathrm{AC}/My\mathrm{BC}$

we can consider full factorial designs with 16 runs, 32 runs and so on. All the hierarchical models of full factorial designs with 2^p runs can be treated as the corresponding hierarchical models in 2^{p+2} contingency tables.

Fractional factorial designs Next we consider the case of fractional factorial design. Again we consider the design with 8 runs for illustration. In the case of $p = 4$, it is natural to define the aliasing relation as D = ABC since this gives the design of resolution IV. The observation is written as follows.

No.	A	B	C	D	n_{11}	n_{12}	n_{21}	n_{22}
1	1	1	1	1	n_{11111}	n_{12111}	n_{21111}	n_{22111}
2	1	1	2	2	n_{11112}	n_{12112}	n_{21112}	n_{22112}
3	1	2	1	2	n_{11121}	n_{12121}	n_{21121}	n_{22121}
4	1	2	2	1	n_{11122}	n_{12122}	n_{21122}	n_{22122}
5	2	1	1	2	n_{11211}	n_{12211}	n_{21211}	n_{22211}
6	2	1	2	1	n_{11212}	n_{12212}	n_{21212}	n_{22212}
7	2	2	1	1	n_{11221}	n_{12221}	n_{21221}	n_{22221}
8	2	2	2	2	n_{11222}	n_{12222}	n_{21222}	n_{22222}

In this case, because of the relation $\mathrm{D} = \mathrm{ABC}$, some hierarchical models of the controllable factors $\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}$ does not have a corresponding hierarchical model in 2^5 contingency tables. For example, the sufficient statistic for the main effect model A/B/C/D is written as follows.

$$\begin{aligned}&\{n_{i\cdot abc}\},\ \{n_{\cdot jabc}\},\ \{n_{ija\cdot\cdot}\},\ \{n_{ij\cdot b\cdot}\},\ \{n_{ij\cdot\cdot c}\},\\ &\{n_{ij111} + n_{ij122} + n_{ij212} + n_{ij221}\},\ \{n_{ij112} + n_{ij121} + n_{ij211} + n_{ij222}\},\end{aligned}$$

which does not correspond to the sufficient statistic for any hierarchical model in 2^5 contingency tables. Similarly, the sufficient statistic for the model AB/C/D, i.e., the model of the four main effects and the interaction effect $\mathrm{A} \times \mathrm{B}$, is written as

$$\begin{aligned}&\{n_{i\cdot abc}\},\ \{n_{\cdot jabc}\},\ \{n_{ijab\cdot}\},\ \{n_{ij\cdot\cdot c}\},\\ &\{n_{ij111} + n_{ij122} + n_{ij212} + n_{ij221}\},\ \{n_{ij112} + n_{ij121} + n_{ij211} + n_{ij222}\},\end{aligned}$$

which again does not correspond to the sufficient statistic for any hierarchical model in 2^5 contingency tables. Note here that the set $\{\mathbf{d}_{12}'\mathbf{n}, \mathbf{d}_1'\mathbf{n}, \mathbf{d}_2'\mathbf{n}\}$, i.e.,

$$\{n_{ija\cdot\cdot}\},\ \{n_{ij\cdot b\cdot}\},\ \{n_{ij11\cdot} + n_{ij22\cdot}\},\ \{n_{ij12\cdot} + n_{ij21\cdot}\},$$

is simply written as $\{n_{ijab\cdot}\}$ by the relation

$$n_{ijab\cdot} = \frac{n_{ija\cdot\cdot} + n_{ij\cdot b\cdot} - (n_{ijab^*\cdot} + n_{ija^*b\cdot})}{2} \qquad (10)$$

where $\{a, a^*\}$ and $\{b, b^*\}$ are the distinct indices, respectively. We see that only the saturated model AB/AC/BC/D among the models including the main effect of D has a corresponding hierarchical models of 2^5 contingency tables.

We also consider the case of $p = 5, 6$. If there are 5 controllable factors, it is natural to define the aliasing relation as $\mathrm{D} = \mathrm{AB}, \mathrm{E} = \mathrm{AC}$, which yields the design of resolution III. In this case, the sufficient statistic for the main effect model A/B/C/D/E is written as

$$\{n_{i\cdot abc}\},\ \{n_{\cdot jabc}\},\ \{n_{ijab\cdot}\},\ \{n_{ija\cdot c}\},$$

which is the sufficient statistic for the hierarchical model $M\mathrm{ABC}/y\mathrm{ABC}/My\mathrm{AB}/My\mathrm{AC}$ of 2^5 contingency tables. Similarly, consider the models containing interaction effect. From the aliasing relation, the estimable interaction is $\mathrm{B} \times \mathrm{C}$, $\mathrm{B} \times \mathrm{E}$ or $\mathrm{C} \times \mathrm{D}$, where the later two are also confounded. We see that the sufficient statistic for the model A/BC/D/E is written as

$$\{n_{i\cdot abc}\},\ \{n_{\cdot jabc}\},\ \{n_{ijab\cdot}\},\ \{n_{ija\cdot c}\}\ \{n_{ij\cdot bc}\}, \tag{11}$$

which is the sufficient statistic for the hierarchical model MABC/yABC/MyAB/MyAC/MyBC of 2^5 contingency tables. On the other hand, the sufficient statistic for the model A/BE/C/D is written as

$$\begin{array}{l}\{n_{i\cdot abc}\},\ \{n_{\cdot jabc}\},\ \{n_{ijab\cdot}\},\\ \{n_{ij111}+n_{ij122}+n_{ij212}+n_{ij221}\},\ \{n_{ij112}+n_{ij121}+n_{ij211}+n_{ij222}\},\end{array}$$

which does not correspond to the sufficient statistic for any hierarchical model in 2^5 contingency tables. Finally, in the case of $p = 6$, consider the design defined as $\mathrm{D} = \mathrm{AB}, \mathrm{E} = \mathrm{AC}, \mathrm{F} = \mathrm{BC}$. In this case, the sufficient statistic for the main effect model A/B/C/D/E/F is written as (11), which is also the sufficient statistic for the hierarchical model MABC/yABC/MyAB/MyAC/MyBC of 2^5 contingency tables.

In the same way, we can consider the fractional factorial designs with 16 runs, 32 runs and so on. The arguments and results are almost similar. For example, if some controllable factor F is defined as F = ABC, the sufficient statistic for the model containing the main effect F includes the marginal total of the type that some three-way marginal tables $\{n_{ijk}\}$ is written as

$$n_{111}+n_{112}+n_{212}+n_{221},\ n_{112}+n_{121}+n_{211}+n_{222}.$$

We see that, if all the two-way marginal tables $\{n_{ij\cdot}\}, \{n_{i\cdot k}\}, \{n_{\cdot jk}\}$ are given, then all the three-way marginal totals $\{n_{ijk}\}$ can be derived from the similar relation to (10) such as

$$\begin{aligned} n_{111} = \ & \frac{1}{8}\left[3(n_{11\cdot}+n_{1\cdot 1}+n_{\cdot 11})+(n_{22\cdot}+n_{2\cdot 2}+n_{\cdot 22})\right. \\ & \left. -(n_{111}+n_{122}+n_{212}+n_{221})-3(n_{112}+n_{121}+n_{211}+n_{222})\right].\end{aligned}$$

4 Numerical examples

In this section we investigate the data of Table 1. As we have seen in Section 1, a simple ANOVA analysis suggests the interaction effect A $\times$ C. Therefore we perform the statistical test for the null model A/B/C against the alternative model AC/B. As we have seen in Section 3, these hypotheses are equivalent to the models of 2^5 contingency tables, MABC/yABC/MyA/MyB/MyC (null model) and MABC/yABC/MyAC/MyB (alternative model). The fitted values for these models are given in Table 4.

The likelihood statistic is calculated as 13.06 with degree of freedom 1. Therefore we have a conclusion that the null model is rejected ($p = 0.000301$) and the interaction effect A $\times$ C is statistically significant. To show the efficacy of our procedure, we also give an imaginary data set of Table 3. The fitted values are given in Table 4 in this case, and the likelihood statistic is 11.56. Though this result also suggests the significance of the interaction effect A $\times$ C, the p value increases to $p = 0.000674$.

Table 5: Fitted values of Table 1 under the null model A/B/C (left) and the alternative model AC/B (right)

				null model				alternative model			
No.	A	B	C	n_{11}	n_{12}	n_{21}	n_{22}	n_{11}	n_{12}	n_{21}	n_{22}
1	1	1	1	27.7	12.3	10.3	9.7	26.7	13.3	11.3	8.7
2	1	1	2	24.1	15.9	3.9	16.1	25.2	14.8	2.8	17.2
3	1	2	1	19.0	21.0	6.0	14.0	17.3	22.7	7.7	12.3
4	1	2	2	13.2	26.8	3.8	16.2	14.8	25.2	2.2	17.8
5	2	1	1	39.0	1.0	13.0	7.0	40.0	0.0	12.0	8.0
6	2	1	2	39.2	0.8	8.8	11.2	38.1	1.9	9.9	10.1
7	2	2	1	38.3	1.7	12.7	7.3	40.0	0.0	11.0	9.0
8	2	2	2	38.5	1.5	8.5	11.5	36.9	3.1	10.1	9.9

Table 6: Fitted values of Table 3 under the null model A/B/C (left) and the alternative model AC/B (right)

				null model				alternative model			
No.	A	B	C	n_{11}	n_{12}	n_{21}	n_{22}	n_{11}	n_{12}	n_{21}	n_{22}
1	1	1	1	27.2	12.8	10.8	9.2	26.7	13.3	11.3	8.7
2	1	1	2	24.4	15.6	3.6	16.4	25.3	14.7	2.7	17.3
3	1	2	1	19.1	20.9	5.9	14.1	17.3	22.7	7.7	12.3
4	1	2	2	13.3	26.7	3.7	16.3	14.7	25.3	2.3	17.7
5	2	1	1	9.7	0.3	3.3	1.7	10.0	0.0	3.0	2.0
6	2	1	2	19.6	0.4	4.4	5.6	19.1	0.9	4.9	5.1
7	2	2	1	38.0	2.0	13.0	7.0	40.0	0.0	11.0	9.0
8	2	2	2	38.7	1.3	8.3	11.7	36.9	3.1	10.1	9.9

5 Discussion

In this paper, we give a general testing procedure to investigate the main and the interaction effects of the controllable factors in the two-valued input-output systems. For this type of data set, simple ANOVA model for the standardized SN ratio is widely used, which is a proposal of Taguchi method. However, since the standard SN ratio is only a function of the sample odds ratio, we cannot evaluate the influence of the sample size of the data appropriately in the simple ANOVA model. In contrast, since our method is based on the conditional distribution for the various null models, sample size is considered in p values.

Note that we do not deny the ANOVA model completely. It is unquestionable that the simple ANOVA model is useful in applications. However, it seems that the ANOVA model for the standard SN ratio is introduced heuristically. We believe that we have to investigate the theoretical validity for the Taguchi method carefully, and modify it as the need arises. We think that our procedure in this paper makes some contribution in this fields.

Though we only give an illustration of the likelihood ratio test based on the asymptotic χ^2 distribution in Section 4, various exact procedures or Monte Carlo procedures are also possible. See Agresti (1992) for the exact tests and Aoki and Takemura (2006) for the Markov chain Monte Carlo tests, for example.

References

[1] Agresti, A. (1992). A survey of exact inference for contingency tables. *Statistical Science*, **7**, pp. 131–177.

[2] Aoki, S. and Takemura, A. (2006). Markov chain Monte Carlo tests for designed experiments. Mathematical Engineering Technical Reports METR 2006-56, Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo. Submitted for publication.

[3] Hamada, M. and Nelder, J. A. (1997). Generalized linear models for quality-improvement experiments. *Journal of Quality Technology*, **29**, 292–304.

[4] McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, 2nd ed. Chapman & Hall, London.

[5] Miyakawa, M. (2006). *Advanced Course of The Design of Experiments*, Union of Japanese Scientists and Engineers, Tokyo. (in Japanese.)

[6] Taguchi, G. (1987). *System of Experimental Design*, White Plains, NY. Unipub/Kraus International Publications.

[7] Taguchi, G. (1991). *Taguchi Methods, Signal-to-Noise Ratio for Quality Evaluation*, Vol. 3, Dearborn, MI. American Supplier Institute Press.

[8] Wu, C. F. J. and Hamada, M. (2000). *Experiments: Planning, analysis, and parameter design optimization.* Wiley Series in Probability and Statistics: Texts and References Section. John Wiley & Sons Inc., New York. A Wiley-Interscience Publication.