Wright State University

CORE Scholar

Mathematics and Statistics Faculty Publications

Mathematics and Statistics

8-2001

Control of Error Rates in Adaptive Analysis of Orthogonal Saturated Designs

Weizhen Wang
Wright State University - Main Campus, weizhen.wang@wright.edu

Daniel T. Voss

Follow this and additional works at: https://corescholar.libraries.wright.edu/math

🎨 Part of the Applied Mathematics Commons, Applied Statistics Commons, and the Mathematics Commons

Repository Citation

CONTROL OF ERROR RATES IN ADAPTIVE ANALYSIS OF ORTHOGONAL SATURATED DESIGNS

By Weizhen Wang and Daniel T. Voss

Wright State University

Individual and simultaneous confidence intervals using the data adaptively are constructed for the effects in orthogonal saturated designs under the assumption of effect sparsity. The minimum coverage probabilities of the intervals are equal to the nominal level $1 - \alpha$.

1. Introduction. Unreplicated factorial designs are extremely useful in industrial experimentation. Consequently, the analysis of saturated designs has received considerable attention in recent years.

A common scenario is as follows. An experiment is conducted using a single replicate or orthogonal fraction of a 2^k factorial design yielding observations $Y_1, \ldots, Y_n$, which are assumed to be independently normally distributed with homogeneous variance σ^2, and which are to be analyzed using a standard linear model. The design is said to be *saturated* if the factorial effect contrasts, $\mu_1, \ldots, \mu_p$ say, are estimable but $n = p + 1$ so there are no error degrees of freedom with which to independently estimate σ^2. Henceforth we refer to the factorial effect contrasts μ_i simply as "effects." Let X_i denote the least squares estimator of μ_i. The design is said to be *orthogonal* if the estimators $X_1, \ldots, X_p$ of the effects are uncorrelated. Thus, under normality the estimators X_i are independent. Furthermore, $X_i \sim N(\mu_i, a^2\sigma^2)$ for known constant a. Suppose the goal is to construct confidence intervals for the effects, $\mu_1, \ldots, \mu_p$. Lacking an independent variance estimate, the analysis is based solely on the estimators $X_1, \ldots, X_p$. This can be done assuming *effect sparsity*—namely, most of the effects μ_i are negligible. The difficulty is that we do not know how many or which of the effects are negligible.

More generally, factorial experiments may involve factors at other than two levels, and they may be asymmetric. Orthogonal polynomial contrasts may be used to accommodate factors at other than two levels. Then the estimator variances may differ, in which case $X_i \sim N(\mu_i, a_i^2\sigma^2)$ for known constant a_i.

Without loss of generality, we assume henceforth that $a_i^2 = 1$. Otherwise, use the X_i/a_i's instead to obtain confidence intervals for the μ_i/a_i's.

The problem of analysis of orthogonal saturated designs is not new. Hamada and Balakrishnan (1998) provided an extensive review, discussion, and empirical comparison of many methods. While many methods have been proposed and studied empirically, few are known to provide control of error rates under all parameter configurations, called *strong control of error rates* by Hochberg

and Tamhane [(1987), page 3]. Kinateder, Voss and Wang (2000) reviewed methods and problems concerning control of error rates in this context.

The most heuristically appealing methods of analysis utilize adaptive estimators of variability, so as to be more robust to the presence of a few large effects. Lenth (1989) proposed the first and most influential of these–a "quick and easy" method of analysis using an adaptive pseudo-standard error. Following Lenth, first obtain an initial estimate $\hat{\sigma}_o$ of σ as 1.5 times the median of the absolute estimates $|X_i|$. Secondly, set aside any absolute estimates which exceed $2.5\hat{\sigma}_o$, then compute $\hat{\sigma}$ as 1.5 times the median of the remaining absolute estimates. This is an *adaptive estimator* in the sense that, when viewed as a linear combination of the ordered absolute estimates, the coefficients are random, depending on the estimates. Variations on Lenth's approach were subsequently considered by Juan and Peña (1992), Dong (1993) and Haaland and O'Connell (1995). It has been an open problem to show that any such adaptive method of analysis of saturated designs provides strong control of error rates.

In this paper we provide a class of adaptive confidence intervals, both individual and simultaneous, which we show do provide strong control of error rates for the analysis of orthogonal saturated designs. The *confidence coefficient* of the interval or intervals, defined as the minimum or infimum over parameter configurations of the coverage probability of the interval or intervals, is obtained at the *null case*, that is, when all μ_i's are zero. It is common sense that one should use as many degrees of freedom as possible for estimating σ. In other words, one should use as many of those X_i's which have mean $\mu_i = 0$ as possible to estimate σ, though which and how many to use are unknown. In the next sections, we will obtain individual and simultaneous confidence intervals by carefully constructing an estimator G, defined in (5), for σ^2, which is continuous and monotone in each $|X_i|$ and uses the data adaptively.

The setting posed previously is that the estimators $X_1, \ldots, X_p$ are independently distributed $X_i \sim N(\mu_i, \sigma^2)$. However, for our results, the following more general conditions are sufficient and assumed henceforth to hold. Let $f_i(x)$ be the pdf of a continuous, unimodal distribution which is symmetric about zero with variance one, for $1 \leq i \leq p$. Assume independent estimators $X_1, \ldots, X_p$, where

$$(1) \qquad\qquad X_i \sim \frac{1}{\sigma} f_i\left(\frac{x_i - \mu_i}{\sigma}\right)$$

for unknown $\mu_1, \ldots, \mu_p$ and σ.

2. Individual confidence intervals. In this section, we discuss how to construct the individual confidence interval for each effect, μ_i. The method is the same for each, so consider μ_p. Denote the vector of effects by $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)$, with $\boldsymbol{\mu}_o = (0, \ldots, 0)$ representing the null case.

THEOREM 1. *Suppose* $G(x_1, \ldots, x_{p-1})$ *is a nonnegative function satisfying the following. For each* $1 \leq i \leq p-1$, *suppose* $G(x_1, \ldots, x_{p-1})$:

(i) *is symmetric about zero, that is,* $G(x_1, \ldots, x_{p-1}) = G(|x_1|, \ldots, |x_{p-1}|)$, *and nondecreasing in* $|x_i|$ *when the other variables* x_j ($j \neq i$) *are held fixed and*

(ii) *satisfies* $G(ax_1, \ldots, ax_{p-1}) = a^2 G(x_1, \ldots, x_{p-1})$ *for any* $a \geq 0$.

Then the probability $P_{\boldsymbol{\mu}, \sigma}(\frac{(X_p - \mu_p)^2}{G(X_1, \ldots, X_{p-1})} \geq d)$ *for any positive constant* d *depends on its parameters through* $\mu_1/\sigma, \ldots, \mu_{p-1}/\sigma$, *and is nonincreasing in each* $|\mu_i/\sigma|$ *when the others are fixed. Therefore,*

$$(2) \quad P_{\boldsymbol{\mu}_o, \sigma}\left(\frac{(X_p - \mu_p)^2}{G(X_1, \ldots, X_{p-1})} \geq d\right) = \sup_{\boldsymbol{\mu}, \sigma} P_{\boldsymbol{\mu}, \sigma}\left(\frac{(X_p - \mu_p)^2}{G(X_1, \ldots, X_{p-1})} \geq d\right)$$

and

$$(3) \quad X_p \pm \sqrt{d \, G(X_1, \ldots, X_{p-1})}$$

is an interval estimator for μ_p *with confidence coefficient* $1 - \alpha$, *where* α *is defined to be the left hand side of* (2).

PROOF. It is clear that the distribution of

$$Q = \frac{(X_p - \mu_p)^2}{G(|X_1|, \ldots, |X_{p-1}|)} = \frac{[(X_p - \mu_p)/\sigma]^2}{G(|X_1|/\sigma, \ldots, |X_{p-1}|/\sigma)}$$

depends on the parameters through $|\mu_1/\sigma|, \ldots, |\mu_{p-1}/\sigma|$ because of ii) and conditions on the f_i. Since $X_1, \ldots, X_p$ are independent, Q is nonincreasing as a function of $|x_i|$ for each $i < p$, and each $|X_i|/\sigma$ ($i < p$) is stochastically nondecreasing in $|\mu_i/\sigma|$, the distribution of Q is stochastically nonincreasing in each $|\mu_i/\sigma|$ [Alam and Rizvi (1966), Mahamunulu (1967) and Voss (1999)]. □

Now the remaining problem is to construct a function satisfying properties (i) and (ii) in Theorem 1. Let $|X|_{(i)}$ be the ith order statistic of $|X_1|, \ldots, |X_{p-1}|$, and let

$$(4) \qquad SS_i = \sum_{h=1}^{i} |X|_{(h)}^2$$

denote the sum of squares of the i smallest of these order statistics, with observed value $ss_i = \sum_{h=1}^{i} |x|_{(h)}^2$. Intuitively, it is more likely that the smaller order statistics $|X|_{(h)}$ will correspond to estimators X_i with negligible means. Thus, it is natural to use a multiple of ss_i to estimate σ^2, for reasonable choice of i. Consider how to choose i adaptively. If we believe a priori that at least ν of the means μ_i are negligible, then the sum in (4) should include at least ν terms–namely, we should use ss_i for some $i \geq \nu$. Also, if r of the means are negligible (where r is unknown) and the rest are large in magnitude, the procedure should adapt to this by using ss_i for i close to but not exceeding r.

To make the choice of i adaptive, we propose the following step-up approach. Start with $i = \nu$, to include (at least) the first ν terms. With i terms included, include the next term, namely $|x|^2_{(i+1)}$, as long as it is not too large relative to ss_i. Iteratively add terms in this way, iterating on i, until a term is too large to be added. In other words, we propose using as the variance estimator a multiple of ss_m, where m is the smallest value of $i \geq \nu$ for which $|x|^2_{(i+1)}/ss_i$ exceeds a specified value [see c_i in (7)], or $m = p - 1$ if this is never the case. While there is no guarantee that SS_m so obtained will be composed entirely from estimators X_i with negligible means, intuitively this is likely to be the case if ν is not chosen to be too large.

THEOREM 2. *Define*

(5) $$G(x_1, \ldots, x_{p-1}) = ss_m/k_m,$$

where

(6) $$m = \begin{cases} p - 1, & \text{if } |x|^2_{(i+1)} < c_i ss_i \\ & \forall\, i = \nu, \ldots, p - 2 \\ \min\{i : i \geq \nu,\ |x|^2_{(i+1)} \geq c_i ss_i\}, & \text{otherwise} \end{cases}$$

for

(7) $$c_i = c_\nu/[1 + (i - \nu)c_\nu]$$

and for c_ν a positive constant, and where

(8) $$k_i = 1 + (i - \nu)c_\nu.$$

Then the function G satisfies properties (i) *and* (ii) *in Theorem* 1.

PROOF.   It is easy to see that property (ii) is true for G and, for (i), that G is symmetric about zero. To show for i) that G is nondecreasing in $|x_i|$, we first prove the continuity of the function G.

For $\nu \leq i \leq p - 2$, let

$$B_i = \left\{(x_1, \ldots, x_{p-1}) : |x|^2_{(i+1)} < c_i ss_i\right\},$$

where c_i is as defined in (7), or equivalently, $c_{i+1} = c_i/(1 + c_i)$, and c_ν is a positive constant. Also, let

$$A_\nu = B_\nu^c, \quad A_i = \bigcap_{h=\nu}^{i-1} B_h \cap B_i^c \text{ for } \nu < i < p - 1 \quad \text{and} \quad A_{p-1} = \bigcap_{h=\nu}^{p-2} B_h.$$

It is clear that $\{A_i\}_{i=\nu}^{p-1}$ form a partition of the sample space. [Specifically, $(x_1, \ldots, x_{p-1}) \in A_m$, where m is as defined in (6).] On each A_i, G is continuous.

Consider G on the boundary between A_i and A_j for any $\nu \leq i < j \leq p - 1$. The equation

$$|x|^2_{(i+1)} = c_i ss_i$$

holds on the common boundary of A_i and A_j, so

$$ss_{i+1} = (1 + c_i)ss_i.$$

Also,

$$|x|^2_{(i+2)} \leq c_{i+1}ss_{i+1}$$

on the boundary of A_j. Therefore, on the common boundary of A_i and A_j,

$$(9) \qquad |x|^2_{(i+2)} \leq c_{i+1}ss_{i+1} = c_{i+1}(1 + c_i)ss_i = c_i ss_i = |x|^2_{(i+1)},$$

and so $|x|_{(i+2)} = |x|_{(i+1)}$. Similarly, $|x|_{(j)} = |x|_{(j-1)} = \ldots = |x|_{(i+1)}$, so

$$(10) \qquad \frac{ss_j}{k_j} = \frac{ss_i + (j - i)|x|^2_{(i+1)}}{k_j} = \frac{(1 + (j - i)c_i)ss_i}{k_j} = \frac{ss_i}{k_i},$$

so G is continuous on the common boundary of A_i and A_j. This implies the continuity of G on the entire sample space. Hence, G, as a function of x_i, is continuous for all $1 \leq i \leq p - 1$.

Now it suffices to prove that G, as a function of x_1, is nondecreasing on $x_1 > 0$. On each $A_j \cap \{x_1 > 0\}$ (now we use A_j as a set of x_1), the derivative of G with respect to x_1 is either 0 (if $x_1 > |x|_{(j)}$) or $2x_1/k_j$, which is positive. Therefore, G is nondecreasing on each $A_j \cap \{x_1 > 0\}$, and is nondecreasing on $x_1 > 0$ due to the continuity of G. $\square$

REMARK 1. From equations (9) and (10), one may see that the c_i and k_i for $i > \nu$ are uniquely determined given c_ν, k_ν, and the requirement that G be continuous, in order for property i) to be satisfied for a given c_ν. We have implicitly and without loss of generality defined k_ν to be one in equation (8).

REMARK 2. Our method is adaptive if and only if $c_\nu > 1/\nu$. In particular, $|x|^2_{(i+1)}/ss_i > 1/i$, so $(x_1, \ldots, x_{p-1}) \in A_j$ for $j > i$ is possible if and only if $c_i > 1/i$ for all $\nu \leq i < j$. However, $c_\nu > 1/\nu$ implies $c_i > 1/i$ for all $i > \nu$. In other words, for any value of $c_\nu > 1/\nu$, the method is adaptive and any value of $m \geq \nu$ is possible.

REMARK 3. The adaptive estimator of Lenth (1989)—his pseudo standard error—is not monotone in the absolute estimates $|X_i|$, so the confidence level of his interval cannot be established to be $1 - \alpha$ using the method of this paper. Likewise for the variations on Lenth's method considered by Juan and Peña (1992), Dong (1993) and Haaland and O'Connell (1995). To the best of our knowledge, none of the adaptive methods proposed in the literature satisfy the monotonicity condition of Theorem 1. Voss' (1999) individual confidence interval, which always uses SS_ν as denominator, that is, $G = SS_\nu$, is a special non-adaptive case of Theorem 2, obtained by choosing $c_\nu \leq 1/\nu$.

REMARK 4.   If effect sparsity is questionable, ν should be chosen to be small. Otherwise, one may use ν equal to the integer part of $(p+1)/2$ if one anticipates that at least half of effects are not active, or even larger depending on the knowledge of the effects under study. For a fixed ν, each c_i $(i > \nu)$ is increasing in c_ν, by equation (7). So the larger c_ν is, the more $|X|^2_{(i)}$'s that are likely to be included in the function G. If these observations are from populations with $\mu_i = 0$, the resulting confidence interval tends to be tighter. We recommend that c_ν be selected by solving the following equation:

$$(11) \quad P_{\boldsymbol{\mu}, \sigma}\left(\frac{|X|^2_{(\nu+1)}}{SS_\nu} \geq c_\nu \mid \mu_p = 0 \ \forall \ p \leq \nu + 1, \quad \mu_p = +\infty \ \forall \ p \geq \nu + 2\right) = \gamma,$$

using a small probability γ. This choice of c_ν is analogous to conducting a size γ test of the null hypothesis that $\nu + 1$ means are zero and the rest infinite against the alternative hypothesis that ν means are zero and the rest infinite. For example, assuming normality [i.e. assuming f_i in (1) is the pdf of a standard normal distribution], if $p = 15$, $\nu = 8$ and $\gamma = 0.05$, then $c_\nu = 1.765$, based on 500,000 simulations coded in Gauss. Further values of c_ν will be given later in Table 1.

REMARK 5.   By this approach, the value of m (i.e., the set A_m) is selected analogous to using a step-up testing procedure–namely, m is selected by stepping up from the value $i = \nu$ until obtaining a value $i = m$ for which $|x|^2_{(m+1)}/ss_m$ is sufficiently large. However, our inference procedure is not a stepwise procedure, as the choice of m does not imply an assertion that the effect μ_i corresponding to $|x|_{(m)}$ is nonzero, and our procedure does not control the probability of correctly choosing m in any sense. Some step-up tests in this context were considered by Loughin and Noble (1997), Venter and Steel (1996, 1998) and Langsrud and Naes (1998), though it remains open to show that their procedures provide control of error rates under all parameter configurations.

TABLE 1

Constants d_α and d'_α for $100(1-\alpha)\%$ individual and simultaneous confidence intervals for p effects, respectively, and c_ν for $\gamma = 0.05$

p	ν	c_ν	Individual confidence intervals			Simultaneous confidence intervals		
			$d_{0.10}$	$d_{0.05}$	$d_{0.01}$	$d'_{0.10}$	$d'_{0.05}$	$d'_{0.01}$
11	6	2.676	5.873	9.289	20.59	19.84	26.74	48.38
15	8	1.765	4.258	6.544	13.59	14.73	19.00	31.41
19	10	1.324	3.374	5.112	10.14	11.91	14.99	23.42
23	12	1.063	2.775	4.174	8.120	10.08	12.45	18.74
27	14	.8885	2.386	3.550	6.760	8.754	10.70	15.65
31	16	.7685	2.093	3.110	5.839	7.806	9.429	13.49

3. Simultaneous confidence intervals.   To construct simultaneous confidence intervals for $\{\mu_1, \ldots, \mu_p\}$, we follow the method of Voss and Wang (1999) but use the function G in Theorem 2. Let

$$(12) \qquad \hat{\underline{x}}_i = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p),$$

for $1 \le i \le p$. Note that $(x_1, \ldots, x_{p-1}) = \hat{\underline{x}}_p$ and $G(X_1, \ldots, X_{p-1}) = G(\hat{\underline{X}}_p)$.

THEOREM 3.   *Define*

$$(13) \qquad V_i^2 = \frac{(X_i - \mu_i)^2}{G(\hat{\underline{X}}_i)}$$

for $1 \le i \le p$, where G is defined in (5), *and let*

$$(14) \qquad W^2 = \max_{1 \le i \le p} V_i^2.$$

Then

$$(15) \qquad P_{\boldsymbol{\mu}_o, \sigma}(W^2 \ge d') = \sup_{\boldsymbol{\mu}, \sigma} P_{\boldsymbol{\mu}, \sigma}(W^2 \ge d')$$

where d' is a constant. Therefore,

$$(16) \qquad X_i \pm \sqrt{d' G(\hat{\underline{X}}_i)}$$

for $1 \le i \le p$ are simultaneous interval estimators for $\mu_1, \ldots, \mu_p$ with simultaneous confidence coefficient $1 - \alpha$, where α is defined to be the left hand side of (15).

PROOF.   Analogous to the proof of Theorem 1 in Voss and Wang (1999). $\square$

REMARK 6.   Suppose one guesses correctly (though unknowingly) that ν of the effects are zero, and suppose the nonzero effects are all large in magnitude. The confidence intervals will tend to be tight for the nonzero means. However, if X_i is one of those estimators with zero mean, then the corresponding error estimate will necessarily include the estimate of a nonzero effect, lengthening the interval considerably. This may not be considered a problem, since the focus is generally on detection of nonzero effects and corresponding directional inference. If it is considered a problem, one might intentionally start at $\nu - 1$ rather than ν.

REMARK 7.   Analogously, one can construct simultaneous confidence intervals for any subset $\{\mu_{i_1}, \ldots, \mu_{i_j}\}$ of $\{\mu_1, \ldots, \mu_p\}$ by consideration of

$$W^2 = \max_{1 \le h \le j} V_{i_h}^2.$$

REMARK 8.   The same approach as presented here for constructing individual and simultaneous confidence intervals can be used to obtain hypothesis tests of specified size.

REMARK 9.   Implementation of the methods requires availability of the constant d of equation (3) for individual confidence intervals or the constant d' of equation (16) for all simultaneous confidence intervals, as well as the constant c_ν of equation (7). These constants are given in Table 1 for common values of p assuming normality.

REFERENCES

ALAM, K. and RIZVI, M. H. (1966). Selection from multivariate normal populations. *Ann. Inst. Statist. Math.* **18** 307–318.

DONG, F. (1993). On the identification of active contrasts in unreplicated fractional factorials. *Statist. Sinica* **3** 209–217.

HAALAND, P. D. and O'CONNELL, M. A. (1995). Inference for effect-saturated fractional factorials. *Technometrics* **37** 82–93.

HAMADA, M. and BALAKRISHNAN, N. (1998). Analyzing unreplicated factorial experiments: A review with some new proposals. *Statist. Sinica* **8** 1–41.

HOCHBERG, Y. and TAMHANE, A. C. (1987). *Multiple comparison procedures*. Wiley, New York.

JUAN, J. and PEÑA, D. (1992). A simple method to identify significant effects in unreplicated two-level factorial designs. *Comm. Statist. Theory Methods* **21** 1383–1403.

KINATEDER, K. K. J., VOSS, D. T. and WANG, W. (2000). Analysis of saturated and super-saturated factorial designs: A review. In *Proceedings of the Indian International Statistical Association 1998 International Conference* (N. Balakrishnan, ed.) 325–347. Gordon and Breach, New Jersey.

LANGSRUD, O. and NAES, T. (1998). A unified framework for significance testing in fractional factorials. *Comput. Statist. Data Anal.* **28** 413–431.

LENTH, R. V. (1989). Quick and easy analysis of unreplicated factorials. *Technometrics* **31** 469–473.

LOUGHIN, T. M. and NOBLE, W. (1997). A permutation test for effects in an unreplicated factorial design. *Technometrics* **39** 180–190.

MAHAMUNULU, D. M. (1967). Some fixed-sample ranking and selection problems. *Ann. Math. Statist.* **38** 1079–1091.

VENTER, J. H. and STEEL, S. J. (1996). A hypothesis-testing approach toward identifying active contrasts. *Technometrics* **38** 161–169.

VENTER, J. H. and STEEL, S. J. (1998). Identifying active contrasts by stepwise testing. *Technometrics* **40** 304–313.

VOSS, D. T. (1999). Analysis of orthogonal saturated designs. *J. Statist. Plann. Inference* **78** 111–130.

VOSS, D. T. and WANG, W. (1999). Simultaneous confidence intervals in the analysis of orthogonal saturated designs. *J. Statist. Plann. Inference* **81** 383–392.

DEPARTMENT OF MATHEMATICS AND STATISTICS
WRIGHT STATE UNIVERSITY
DAYTON, OHIO 45435
E-MAIL: wwang@math.wright.edu
            dan.voss@wright.edu