New Results for Stepwise Tests in Orthogonal Saturated Designs

Weizhen Wang
Department of Mathematics & Statistics
Wright State University
Dayton, OH 45435
July, 2006

joint work with Dan Voss (Wright State University)

and Sam Wu (University of Florida)

partially supported by NSF through grant number DMS0308861.

Outline of the talk

1. **Introduction and Examples**

2. **Formulations and Motivations**

3. **The previous Results**

4. **The Step-down Tests**

5. **The Step-up Tests**

6. **Conclusion**

1 Introduction and Examples

- The saturated design is a design where the number of observations (or estimators) is equal to the number of parameters of interest, typically the contrasts of population means in 2^k (fractional) factorial design, and there is no degree of freedom left for the nuisance parameter, the population variance.

  - **screen many potentially important variables, factors or effects**
  - **used in early stages of experiments**
  - **used in screening experiments**

- For a saturated design, if the estimators of parameters of interest are independent, or equivalently, the contrasts are orthogonal under normality assumptions, then we have an orthogonal saturated design.

- Difficulty: The traditional estimator for the variance (MSE) is no longer available. It is challenging to make inferences on parameters of interest.
- The sparsity of effects principle: in most of systems, responses are driven largely by a limited number of main effects and lower-ordered interactions and only few of the effects are nonzero. Therefore, those estimators with mean zero can be used to estimate the variance.

- Examples

  - **Example 1.** (2^3 **full factorial design) Consider a three-way anova with 2 levels for each factor. There are 8 treatments and from each treatment one measurement is collected. Write the factor effect model in a regression setting,**

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \mu_{\dots} \\ \alpha_1 \\ \beta_1 \\ \gamma_1 \\ (\alpha\beta)_{11} \\ (\alpha\gamma)_{11} \\ (\beta\gamma)_{11} \\ (\alpha\beta\gamma)_{111} \end{bmatrix} + \begin{bmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{221} \\ \varepsilon_{222} \end{bmatrix}$$

$$= \underline{X}\underline{\beta} + \underline{\varepsilon}.$$

- **The parameters of interest:** $\alpha_1, \dots, (\alpha\beta\gamma)_{111}$, **seven of them. The effect sparsity says that most of them are zero. The nuisance parameters:** $\mu_{\dots}, \sigma^2$.

- **The data: the l.s.e.,** $\hat{\mu}_{\dots}, \dots, (\alpha\hat{\beta}\gamma)_{111}$, **eight of them, are independent since the covariance matrix is diagonal. No MSE for** σ^2.

- **Example 2. (The 12-run Plackett-Burman Design) Twelve observations are obtained to estimate the overall mean and the main effects of 11 factors each at two levels under a first-order model. The design matrix is as follows.**

$$\underline{X} = \begin{bmatrix} 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix} \tag{1}$$

2 Formulations and Motivation

- Consider a linear model

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i, \epsilon_i \sim \text{iid N}(0, \sigma^2) \qquad (2)$$

  for $1 \leq i \leq M$ for unknown $\beta_0, \ldots, \beta_k$ and σ_ϵ^2.

- The parameters of interest: $\beta_1, \ldots, \beta_k$

- Data: the independent l.s.e $\{\hat{\beta}_i\}_{i=1}^k$.

- Difficulties: M=k+1. Thus no degrees of freedom left for the nuisance parameter σ^2.

- Effect sparsity: Most of β_i's are zero. But do not know how many and which β_i's are zero.

- Motivation: Detect the first jump in the ordered list of $|\hat{\beta}_i|^2$.

- Goal: use simultaneous tests to identify active effects.
  - **Control the experimentwise error rates at α.**
  - **Use the data adaptively.**
- Two sets of null hypotheses:
  - **Step-down:**

$$\mathcal{A} = \{H_{0,I} = \{\beta_i = 0 : \forall i \in I\} : I \subset \{1, ..., k\}\}$$

  - **Step-up: let N be the number of $\beta_i' s$ being zero.**

$$\mathcal{B} = \{H_{0,m} = \{N \geq m\} : \nu + 1 \leq m \leq k\}$$

    - $H_{0,i} \subset H_{0,j}$ **if** $i \geq j$**.**
  - **Both sets are closed under the intersection.**

3 The Previous Results

- Step-down procedures
  - **Voss (1988) test for $\mathcal{A}$**
    - **control the experimentwise error rates, but not use data adaptively.**
  - **Ventel and Steel (1998) test for $\mathcal{B}$.**
    - **don't know whether to control the experimentwise error rates.**
    - **The zero effects are contaminated with the nonzero effects.**
  - **Voss and Wang (2006) test for $\mathcal{A}$**
    - **control the experimentwise error rates.**
    - **use data adaptively.**
    - **the estimator(s) with nonzero mean may be used to estimate the variance.**
  - **Holm, Mark and Adolfsson iterative step-down**

tests (2005)

  * **use data adaptively and control the error rates**
  * **use the data iteratively, i.e., if an effect is shown significantly large, it won't be used in the next step to estimate** σ**.**

- Step-up procedures
  - **Ventel and Steel (1998) test for** $\mathcal{B}$
    * **don't know whether to control the experimentwise error rates.**

4 The iterative step-down tests (Voss and Wang, 2005)

- The idea of iterative tests
  - **Let** $[1], \ldots, [k]$ **be the random indices such that** $|\hat{\beta}_{[1]}| < \cdots < |\hat{\beta}_{[k]}|$**. Let** $X_i = |\hat{\beta}_{[i]}|^2$**. Starting from the largest value** X_k**, if** X_k **is significantly large, declare** $\beta_{[k]}$ **to be active.**
  - **Now consider a data set of** $X_1, ..., X_{k-1}$**, the original data set excluding** X_k**, and see whether** $\beta_{[k-1]}$ **is active. Repeat this until we reach a zero effect.**

- Consider test statistics

$$T_i = \frac{X_i}{g_i(X_1, \ldots, X_i)}, \quad 2 \le i \le k, \qquad (3)$$

where each $g_i(t_1, \ldots, t_i)$ $(2 \le i \le \ldots, k)$ is a positive function with the following properties: i) scale invariant; ii) monotone. Obtain the critical value c_i $(i = 2, \ldots, k)$ as the upper-α quantile of the distribution of T_i at

$(\beta_1, ..., \beta_k) = (0, .., 0, \infty, ..., \infty)$ with i zeros.

- Iterative step-down test
  - **Step 1: If $T_k > c_k$ then assert $\beta_{[k]} \neq 0$ and continue; else stop.**
  - **Step 2: If $T_{k-1} > c_{k-1}$ then assert $\beta_{[k-1]} \neq 0$ and continue; else stop.**
  - **If...**

Theorem 1 *The above iterative step-down test strongly controls the probability of making any false assertions to be at most α under all parameter configurations $\beta = (\beta_1, \ldots, \beta_k)$ and for all σ.*

- For different choice of g_i, one may obtain different iterative step-down tests. For example, tests corresponding to Zahn (1975ab) and Venter and Steel (1998), Voss (1988), Voss and Wang (2006), Langsrud and Naes' (1998), and others.

5 The step-up tests, (Wu and Wang, 2006, AOS)

- Test for $\mathcal{B}$.
- For each hypothesis (in $\mathcal{B}$) $H_{0,m} : N \geq m$, let $\boldsymbol{\beta}_m = (\beta_1, ..., \beta_k) = (0, ..., 0, \infty, .., \infty)$ (m zeros).

Theorem 2 *For any $0 < \alpha < 1$, the following rejection region*

$$R_{m-1,m} = \{S_{m-1} < d_{m-1,m} X_m\}, \qquad (4)$$

defines a level-α test for $H_{0,m}$, where $S_{m-1} = \Sigma_{i=1}^{m-1} X_i$ and $d_{m-1,m}$ is given by

$$P_{\boldsymbol{\beta}_m}\left(\frac{S_{m-1}}{X_m} < d_{m-1,m}\right) = \alpha.$$

- There is a jump at X_m.
- Since $H_{0,m}$ decreases, we want $R_{m-1,m}$ increasing as m goes large.

To test all hypotheses in $\mathcal{B}$, let

$$R^*_{m-1,m} = \cup_{i=\nu+1}^{m}\{S_{i-1} < d^*_{i-1,i}X_i\} = \cup_{i=\nu+1}^{m}\{S_\nu < Q_i\} \tag{5}$$

where $Q_i = d^*_{i-1,i}X_i - (S_{i-1} - S_\nu)$ and $d^*_{i-1,i}$ is determined iteratively as follows:

- $d^*_{\nu,\nu+1} = d_{\nu,\nu+1}$.
- For $\nu + 2 \leq m \leq k-1$, suppose $d^*_{\nu,\nu+1}$ through $d^*_{m-2,m-1}$ are given. Then $d^*_{m-1,m}$ is determined by

$$\sum_{i=\nu+1}^{m} P_{\boldsymbol{\beta}_m}(max\{S_\nu, \{Q_j\}_{j=\nu}^{i-1}\} < Q_i) = \alpha. \tag{6}$$

- $d^*_{k-1,k}$ is given by

$$P_{\boldsymbol{\beta}_k}(R^*_{k-1,k}) = \alpha. \tag{7}$$

– $R^*_{m-1,m}$ increases and is level-α for a single $H_{0,m}$.

Theorem 3 *For testing all hypotheses in $\mathcal{B}$,*

$$\textit{assert not } H_{0,m} \textit{ if } R^*_{m-1,m} \textit{ is true.}$$

Then the experimentwise error rate is controlled at α.

The step-up tests procedure:

- Step 1: If $R^*_{\nu,\nu+1}$ is true, then conclude that $\beta_{[\nu+1]}$ through $\beta_{[k]}$ are the $k-\nu$ active effects ($= H \cap H_{A,\nu+1}$), and stop; otherwise go to step 2.
- Step 2: If $R^*_{\nu,\nu+2}$ is true, then conclude that $\beta_{[\nu+2]}$ through $\beta_{[k]}$ are the $k-\nu-1$ active effects ($= H_{0,\nu+1} \cap H_{A,\nu+2}$) and stop; o/w go to step 3.
- $\vdots$

- Step $k-\nu$: If $R^*_{\nu,k}$ is true, then conclude that $\beta_{[k]}$ is the only active effect ($= H_{0,k-1} \cap H_{0,k}$), and stop; o/w conclude no active effect and stop.

6 Discussion

–The stepwise tests are more powerful than the single-step tests

– The step-up tests seem more powerful to detect the first jump than the step-down tests.

– A better step-up tests procedure may exist by determining the constants $d^*_{m-1,m}$ as follows:

$$P_{\boldsymbol{\beta}_m}(R^*_{m-1,m}) = \alpha.$$

– A much more challenging problem is to make inferences in nonorthogonal designs.

7 References

- Alam, K. and Rizvi, M. H. (1966). Selection from multivariate normal populations. *Annals of the Institute of Statistical Mathematics* **18**, 307–318.

- Berk, K. N. and Picard, R. R. (1991). Significance tests for saturated orthogonal arrays. *Journal of Quality Technology* **23**, 79–89.

- Birnbaum, A. (1959). On the analysis of factorial experiments without replication, *Technometrics*, **1**, 343–357.

- Daniel, C. (1959). Use of half-normal plots in interpreting factorial two-level experiments, *Technometrics*, **1**, 311–341.

- Dudoit, S., Shaffer, J. P. and Boldrick, J. (2003). Multiple hypothesis testing in microarray experiments, *Statistical Science*, **18**, 71-103.

- Hamada, M. and Balakrishnan, N. (1998). Analyzing unreplicated factorial experiments: A review with some new proposals. *Statistica Sinica* **8**, 1–41.

- Hochberg, Y. and Tamhane, A. C. (1987). *Multiple Comparison Procedures*, John Wiley & Sons, New York.

- Holm, S., Mark, S., and Adolfsson, T. (2005). A step-down test for effects in unreplicated factorial designs. *Comm. Statist. Theory Methods* **34**, 405–416.

- Hsu, J. C. (1996). *Multiple comparisons.* Chapman & Hall/CRC.

- Hsu, J. C., Chang, J. Y., and Wang, T. (2003). Multiple comparisons of gene expressions from microarray data. *Preprint.*

- Kinateder, K. K. J., Voss, D. T., and Wang, W. (2000a). Exact confidence intervals in the analysis of nonorthogonal saturated factorial designs, *American Journal of Mathematical and Management Sciences*, **20** (2000), 71–84.

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

- Kunert, J. (1997). On the use of the factor-sparsity assumption to get an estimate of the variance in saturated designs. *Technometrics* **39**, 81–90.

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

- Mahamunulu, D. M. (1967). Some fixed-sample ranking and selection problems. *Annals of Mathematical Statistics* **38**, 1079–1091.
- Marcus, R., Peritz, E. and Gabriel, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika* **63**, 655–660.
- Venter, J. H. and Steel, S. J. (1998). Identifying active contrasts by stepwise testing. *Technometrics* **40**, 304–313.
- Voss, D. T. (1988). Generalized modulus-ratio tests for analysis of factorial designs with zero degrees of freedom for error. *Communications in Statistics: Theory and Methods* **17**, 3345–3359.
- Voss, D. T. (1999). Analysis of orthogonal saturated designs. *Journal of Statistical Planning and Inference* **78**, 111-130.
- Voss, D. T. and Wang, W. (1999). Simultaneous confidence intervals in the analysis of orthogonal saturated designs. *Journal of Statistical Planning and Inference* **81**, 383–392.
- Voss, D. T. and Wang, W. (2006). On Adaptive Testing in Orthogonal Saturated Designs. *Statistica Sinica*, **16**, 227-234.
- Voss, D. T. and Wang, W. (2005). Establishing Strong Control of Familywise Error Rate for Iterative Step-down Tests. Submitted.
- Wang, W. and Voss, D. T. (2001). Control of error rates in adaptive analysis of orthogonal saturated designs. *Annals of Statistics* **29**, 1058–1065.
- Wang, W. and Voss, D. T. (2003). On adaptive estimation in orthogonal saturated designs. *Statistica Sinica* **13**, 727–737.
- Wu, S. S. and Wang, W. (2006). Step-up Simultaneous Tests for Identifying Active Effects in Orthogonal Saturated Designs. To appear in *Annals of Statistics*.