Journal of the Nigerian Association of Mathematical Physics
***Volume* 19 (November, 2011), pp 115 – 118**
© *J. of NAMP*

Investigating the Validity of the Estimate, $\hat{E}(s^2)$ of $E(s^2)$-Values for Supersaturated Design (SSD) Using Mean Square Error (MSE)

[1]Mbegbu, J. I. and [2]Todo, C. O.

[1]Department of Mathematics,
University of Benin, Benin City.
[2]Department of Statistics,
Delta State Polytechnic, Otefe-Oghara, Nigeria.

Abstract

We investigated the validity of the estimate, $\hat{E}(s^2)$ of $E(s^2)$ for a supersaturated design (SSD) recently proposed by Todo and Mbegbu (2011). The investigation is with reference to Lower bound estimate (LBE) proposed by Nguyen and Cheng (1996). We achieved this task by comparing the mean square error of $\hat{E}(s^2)$ and mean square error of LBE. The result showed that the mean square error for the estimate $\hat{E}(s^2)$ is less than the mean square error for LBE. This indicates that the estimate $\hat{E}(s^2)$ is valid.

Keywords: Supersaturated, mean square error, designs, estimate, lower bound, $E(s^2)$-criterion.

1.0 Introduction

A supersaturated design (SSD) is a design in which the number of factors m is more than n-1, where n is the number of runs (Nguyen and Cheng, 2008). Supersaturated design is a fractional factorial design in which the number of potential effects is greater than the number of runs. Such designs are helpful when experimentation is expensive and the number of effects is large but only a few are significant (Minqian and Kaitai, 2006).

According to Xu and Wu (2005), the popular criterion in the literature as a measure of goodness or for comparing supersaturated designs is the $E(s^2)$ criterion. $E(s^2)$ criterion which was proposed by Booth and Cox (1962) measures the average correlation among the columns of the design matrix of an SSD.

In the design with n runs and n factors, each of the factors (columns) has two levels, and we require that $\frac{n}{2}$ of the entries in each column be +1 and the others -1 (see Mbegbu and Todo, 2010).

In the literature, most of the supersaturated designs have $E(s^2)$ values but Todo and Mbegbu (2011) proposed $\hat{E}(s^2)$ which is an estimate of $E(s^2)$. We shall investigate the validity of $\hat{E}(s^2)$ using the mean square error of $\hat{E}(s^2)$, $MSE\left[\hat{E}(s^2)\right]$.

2.0 The Estimate $\hat{E}\left(s^2\right)$ of $E\left(s^2\right)$ Value for Supersaturated Designs (SSD)

Let X denote the incidence matrix of the supersaturated design, SSD (n,m) with n runs and m factors.
We define

$$X = \left[a_{ij}\right]_{m\times n} \qquad (2.1)$$

[1]Corresponding authors: ***Mbegbu, J. I.*** E-mail:-, Tel. +2348020740989

and the transpose

$$X^T = \left[a_{ij}\right]_{n\times m} \qquad (2.2)$$

so that the design matrix becomes

$$\left[s_{ij}\right]_{m\times m} = X^T X \qquad (2.3)$$

according to Bulutoglu and Cheng (2004),

$$E\left(s^2\right) = \frac{1}{mC_n}\sum_{i\le j} s_{ij}^2 \qquad (2.4)$$

where $$mC_n = \frac{m!}{n!(m-n)!}$$

and s_{ij} is the value of the entry at the ith row and jth column of the design matrix, $X^T X$.
For two-level supersaturated design (SSD)

$$s_{ij} = \begin{cases} 1\text{, if factor } i \text{ occurs in the high level in run } i \\ -1\text{, if factor } j \text{ occurs in the low level in run } i \end{cases}$$

The estimation $\hat{E}\left(s^2\right)$ of $E\left(s^2\right)$ values for supersaturated design (SDD) had been proposed by Mbegbu and Todo, (2010). Obviously, for any supersaturated design, shuffling the elements of design matrix with respect to the factors always results to a new design. The number of SSDs that can result from shuffling the element of design matrix is $m!$, which have different $E\left(s^2\right)$ values though close to each other but within a lower bound proposed by Nguyen (1996), and Tang and Wu (1997).

According to Todo and Mbegbu (2011), the estimate of $E\left(s^2\right)$ for a family of supersaturated designs with n runs and m factors is

$$\hat{E}\left(s^2\right) = a + b\left(nm^2\right) \qquad (2.5)$$

where

$$a = \left[\overline{{}^*E(s^2)}\right] - b\left[\overline{nm^2}\right] \qquad (2.6)$$

and

$$b = \frac{\sum_{k=1}\left(\left[nm^2\right]_k - \left[\overline{nm^2}\right]\right)\left(\left[{}^*E\left(s^2\right)\right]_k - \left[\overline{{}^*E\left(s^2\right)}\right]\right)}{\sum_{k=1}\left(\left[nm^2\right]_k - \left[\overline{nm^2}\right]\right)^2} \qquad (2.7)$$

The estimate $\hat{E}(s^2)$ satisfies the lower bound proposed by Tang and Wu (1997). In equations (2.6) and (2.7) $\left[{}^*E\left(s^2\right)\right]_k$ are the values of $E(s^2)$ for known SSDs.

3.0: Materials and Method of Comparison

In line with the definition of mean square error of any estimator, we have

$$MSE\left[\hat{E}(s^2)\right] = \frac{1}{N}\sum_{k=1}^{N}\left(\left[E(s^2)\right]_k - \left[\hat{E}\left(s^2\right)\right]_k\right)^2 \qquad (3.1)$$

We use the mean square error as a measure of goodness of the estimate $\hat{E}\left(s^2\right)$ by comparing the estimate $\hat{E}\left(s^2\right)$ and the lower bound estimate (LBE) with $E(s^2)$ established by

(i) Bulutoglu and Cheng (2004)
(ii) Nguyen and Cheng (2008)

respectively.

According to Todo and Mbegbu (2011), Table 1 below depicts the $E(s^2), \hat{E}(s^2)$, and LBE for SSD(n,m) constructed by Bulutoglu and Cheng (2004), Todo and Mbegbu (2011) and Tang and Wu (1997) respectively.

Table 1: SSD(n,m) and its corresponding $E(s^2)$, $\hat{E}(s^2)$, LBE values

SSD (n,m)	$E(s^2)$: Bulutoglu and Cheng (2004)	$\hat{E}(s^2)$: Todo and Mbegbu (2011)	LBE: Tang and Wu (1997)
SSD (10,14)	5.0549	5.225768	4.046154
SSD (10,15)	5.5238	5.243084	4.535714
SSD (14,17)	4.9412	5.560321	3.611607
SSD (14,18)	5.6732	5.59425	4.277311
SSD (14,19)	6.0585	5.628189	4.869048

The lower bound estimate (LBE) is

$$E\left(s^2\right) \geq \frac{\left(m\left(n^2+n-1\right)-n^3\right)}{n\left(m-1\right)}, \quad \text{n is even.} \tag{3.2}$$

(see Tang and Wu, 1997)

We shall also consider the $E(s^2), \hat{E}(s^2)$ and LBE for SSD (n, m) constructed by Nguyen and Cheng (2008), Todo and Mbegbu (2011) and Tang and Wu (1997) respectively (see Table 2).

Table 2: SSD(n,m) and its Corresponding $E(s^2), \hat{E}(s^2)$, and LBE Values

SSD (n,m)	$E(s^2)$: Nguyen and Cheng (2008)	$\hat{E}(s^2)$: Todo and Mbegbu (2011)	LBE: Tang and Wu (1997)
SSD (10,14)	5.0549	4.82302472	4.046154
SSD (12,14)	4.7473	4.95532617	2.833333
SSD (12,18)	5.9608	5.47373187	5.205882
SSD (14,18)	5.6732	5.69243427	4.277311
SSD (14,22)	6.9091	6.44844258	6.306122
SSD(16,18)	5.0196	5.91113667	2.870000
SSD(16,22)	6.6494	6.77514617	5.553571
SSD (16,26)	7.8769	7.81195756	7.375000

4.0: Results of the Comparison Using MSE

Implementing equation (3.1) in Table 1 and 2 respectively yields the following Tables:

Table 3: Comparison of $E(s^2)$ constructed by Bulutoglu and Cheng (2004), and $\hat{E}(s^2)$ constructed by Todo and Mbegbu (2011).

SSD(n, m)	$E(s^2)$	$\hat{E}(s^2)$	$\left[E(s^2) - \hat{E}(s^2)\right]^2$	$MSE\left[\hat{E}(s^2)\right]$
SSD (10, 4)	5.0549	5.225768	0.029196	0.136539
SSD (10,15)	5.5238	5.243084	0.078801	
SSD (14,17)	4.9412	5.56031	0.383297	
SSD (14,18)	5.6732	5.59425	0.006233	
SSD (14,19)	6.0585	5.628189	0.185168	

Table 4: Comparison of $E(s^2)$ constructed by Bulutoglu and Cheng (2004), and LBE constructed by Tang and Wu (1997).

SSD(n, m)	$E(s^2)$	LBE	$\left[E(s^2) - (LBE)\right]^2$	$MSE[LBE]$
SSD (10,14)	5.0549	4.046154	1.017568	1.425000
SSD (10,15)	5.5238	4.535714	0.976314	
SSD (14,17)	4.9412	3.6166607	1.767818	
SSD (14,18)	5.6732	4.277311	1.948506	
SSD (14,19)	6.0585	4.869048	1.414796	

Table 5: Comparison of $E(s^2)$ constructed by Nguyen and Cheng (2008), and $\hat{E}(s^2)$ constructed by Todo and Mbegbu (2011).

SSD(n, m)	$E(s^2)$	$\hat{E}(s^2)$	$\left[E(s^2) - \hat{E}(s^2)\right]^2$	$MSE\left[\hat{E}(s^2)\right]$
SSD (10,14)	5.0549	4.82302472	0.053766	
SSD (12,14)	4.7473	4.95532617	0.043275	
SSD (12,18)	5.9608	5.47373187	0.237235	
SSD (14,18)	5.6732	5.69243427	0.000370	0.170215
SSD (14,22)	6.9091	6.44844258	0.212250	
SSD(16,18)	5.0196	5.91113667	0.794838	
SSD(16,22)	6.6494	6.77514617	0.015812	
SSD (16,26)	7.8769	7.81195756	0.004218	

Table 6: Comparison of $E(s^2)$ constructed by Nguyen and Cheng (2008), and LBE constructed by Tang and Wu (1997).

SSD(n, m)	$E(s^2)$	LBE	$\left[E(s^2) - LBE\right]^2$	$MSE[LBE]$
SSD (10,14)	5.0549	4.046154	1.017568	
SSD (12,14)	4.7473	2.833333	3.66327	
SSD (12,18)	5.9608	5.205882	0.569901	
SSD (14,18)	5.6732	4.277311	1.948506	1.70186
SSD (14,22)	6.9091	6.306122	0.363582	
SSD(16,18)	5.0196	2.875000	4.599309	
SSD(16,22)	6.6494	5.553571	1.200841	
SSD (16,26)	7.8769	7.375000	0.251904	

5.0 Discussion and Conclusion

A comparison of $MSE\left(\hat{E}(s^2)\right)$ and $MSE(LBE)$ in Tables 3 and 4 shows a large difference. $MSE\left(\hat{E}(s^2)\right)$ is comparatively far less then $MSE(LBE)$. Also, a comparison of $MSE\left(\hat{E}(s^2)\right)$ and $MSE(LBE)$ in Tables 5 and 6 shows a significant difference. $MSE\left(\hat{E}(s^2)\right)$ is far less than $MSE(LBE)$. This goes to demonstrate the goodness of $\hat{E}(s^2)$. $MSE\left(\hat{E}(s^2)\right)$ is only about 10% of $MSE(LBE)$.

Conclusively, the estimate $\hat{E}(s^2)$ has a minimum mean square error, 0.136539 compared to LBE. Hence, the estimate, $\hat{E}(s^2)$ is valid.

References

1. Booth, K. H. and Cox, D. R. (1962), Some Systematic Supersaturated Designs, Technometrics, Vol.4, pp.489-495; Vol. 39, series B pt.1, pp76-83.
2. Bulutoglu, D. A. and Cheng, C. (2004), Construction of E (s^2)-optimal Supersaturated Designs, Annual of stat., Vol. 32, No.4, August, pp1662-1678.
3. Mbegbu, J. I. and Todo, C.O. (2010), Explicit Construction of an E (s^2)-optimal Supersaturated Design with 20 Runs and 57 Factors, International Journal of Natural and Applied Sciences, 6(1): 109-115.
4. Todo, C.O. and Mbegbu, J. I. (2011), Regression Approach to Estimate the $E(s^2)$ Values for Supersaturated Design, International Journal of Natural and Applied Sciences, 7 (1): 54-57.
5. Minqian, L. and Kaitai, F. (2006), A Case Study in the Application of Supersaturated Design to Computer Experiments, Acta Mathematics Scienta, Vol. 65, pp. 152-172.
6. Nguyen, N. (1996), An Algorthmic Approach to Constructing Supersaturated Designs, Technometrics, Vol.38, pp.67-73.
7. Nguyen, N. and Cheng, C. (2008), New $E(s^2)$-optimal Supersaturated Designs Constructed from Incomplete Block Designs, Technometrics, Vol. 50, No.1, February, pp 26-31.
8 Tang, B. X. and Wu, C. E. J. (1997), A Method for Constructing Supersaturated Design and its E (s^2) optimality, Canadian Journal of Statistics, Vol.25, pp 191-201.
9. Xu, H. and Wu, C. F. I. (2005), Construction of Optimal Multi-level Supersaturated Designs, Annals Stat., Vol.33, No.6, pp 132-150.