*Journal of the Korean Statistical Society (2007),* **36**: *2, pp 201–208*

# NUMBER OF CYCLES IN EVOLUTIONARY OPERATION†

YONG B. LIM[1] AND SUNG H. PARK[2]

ABSTRACT

Evolutionary operation (EVOP) proposed by Box (1957) is a method for continuous monitoring and improvement of a full-scale manufacturing process with the objective of moving the operating conditions toward the better ones. EVOP consists of systematically making small changes in the levels of the two or three process variables under consideration. Data are collected on the response variable at each point of two level factorial design with the center point and a cycle is said to have been completed. The cycles are replicated sequentially until the decision is made on whether further cycle of experiments is needed to conclude the significance of any of main effects or interaction effects or the curvature. In this paper, an improved flow chart of EVOP is proposed and how to determine the number of cycles is studied based on the size of type II error. In order to reject the alternative hypothesis of interests with more confidence and conclude that we believe in the null hypothesis of no effects, we propose a counter measure $p^*$-value corresponding to the $p$-value. The relationship of $p^*$-value to the probability of type II error $\beta$ under the alternative hypothesis of interests is analogous to that of $p$-value to the probability of type I error $\alpha$. Also the implementation of EVOP with a mixture experiment is discussed.

*AMS 2000 subject classifications.* Primary 62K20; Secondary 62P30.
*Keywords.* Evolutionary operation, number of cycles, size of type II error.

## 1. INTRODUCTION

Evolutionary operation (EVOP) proposed by Box (1957) is a well-known method for continuous monitoring and improvement of a full-scale manufacturing process with the objective of moving the operating conditions toward the

Received September 2006; accepted December 2006.
†Lim's research was supported in part by Grant No. R06-2002-012-01002-0 for the Women's University from Korea Ministry of Science and Technology. Park's research was supported in part by Grant No. R01-2003-000-10220-0 from Korea Science and Engineering Foundation.
[1]Corresponding author. Department of Statistics, Ewha Womans University, Seoul 120-750, Korea (e-mail: yblim@ewha.ac.kr)
[2]Department of Statistics, Seoul National University, Seoul 151-747, Korea

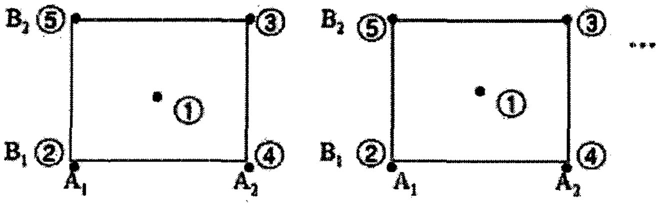


FIGURE 1.1 *Sequential $2^2$ designs augmented with the center point for EVOP.*

better ones. As Box said, the philosophy is that "a process should be run so as to produce not only product, but also information on how to improve the product". EVOP consists of systematically making small changes in the levels of the two or three process variables under consideration, which is necessary to avoid appreciable negative changes in the characteristics of the product.

Suppose that two process variables are screened for further consideration. Data are collected on the response variable at each point of two level factorial design with the center point and a cycle is said to have been completed. The cycles are replicated sequentially as it is shown in Figure 1.1 until the decision is made on whether further cycle of experiments is needed to conclude the significance of any of main effects or interaction effects or the curvature. Thus, cycles in EVOP can be treated as blocks. Steps in EVOP are summarized and visualized in Figure 1.2, the flow chart of EVOP.

An important statistical issues in EVOP is construction of the sequential decision-making procedure by which we accept the null hypothesis of no effects. The information on the power of the test at the detectable difference of the effects where the engineers are interested is helpful to make that decision.

By the pooling rules proposed by Lorenzen and Anderson (1993), a term in the model will be declared negligible if it is insignificant at the $\alpha$=.25 level. In addition to that, we propose to check whether we have enough power of the test such that we believe in the null hypothesis of no effects, since we want to be more conservative. In Section 2 how to determine the number of cycles in advance is studied based on the size of type II error at the standardized detectable difference of the effects. In order to reject the alternative hypothesis of interests with more confidence and conclude that we believe in the null hypothesis of no effects, we propose a counter measure $p^*$-value corresponding to the $p$-value. The relationship of $p^*$-value to the probability of type II error $\beta$ under the alternative

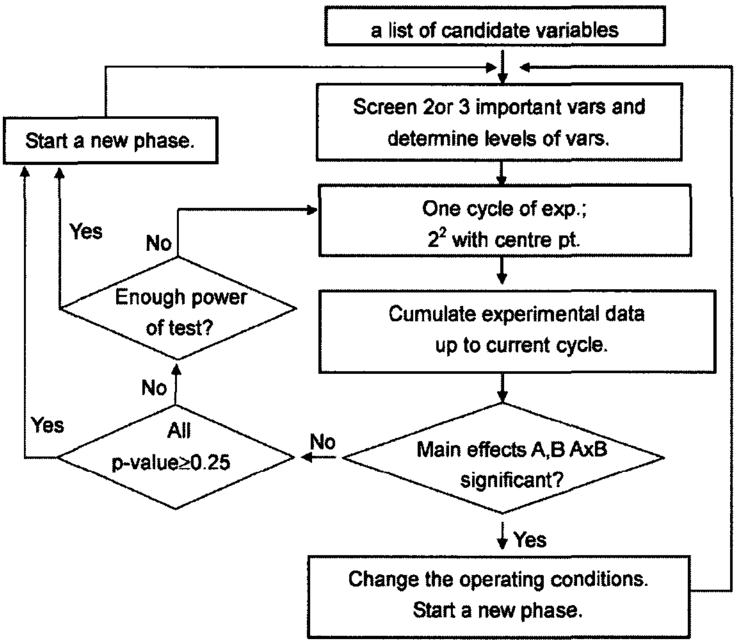


FIGURE 1.2 *Flowchart of EVOP.*

hypothesis of interests is analogous to that of $p$-value to the probability of type I error $\alpha$. Implementation of EVOP using MINITAB is discussed in Section 3. Also we discuss the implementation of EVOP with a mixture experiment in Section 4.

## 2. NUMBER OF CYCLES IN EVOP

Let us define the standardized detectable difference of effects $\Delta$ to be the ratio of the difference of mean responses between two groups to the standard deviation. A phase of EVOP is completed (or terminated) and a new phase is started when either any of main effects or interaction effects or the curvature effect is detected or none of them is declared to be significant after several cycles. In the latter case, the number of cycles $r$ depends on the power of the test, *i.e.*, the probability of type II errors at $\Delta$ where the engineers are interested, which is calculated from the distribution function of the noncentral $F$. Thus, when all the $p$-values are between .05 and .25, we have to check whether we have enough power of the test such that we believe in the null hypothesis of no effects.

It can be easily checked that

$$\beta = Pr[F_\lambda \le F_\alpha(1, 4(r-1))], \tag{2.1}$$

where $F_\lambda$ follows noncentral $F$ distribution with 1 and $4(r-1)$ degrees of freedom and noncentrality parameter $\lambda = r\Delta^2/2$ and $F_\alpha(1, 4(r-1))$ is the upper $\alpha^{th}$ percentile of $F$ distribution with 1 and $4(r-1)$ degrees of freedom. Noting that distribution function of the noncentral $F$ is a decreasing function of noncentrality parameter $\lambda$ which is proportional to $r\Delta^2/2$, the probability of type II error is a decreasing function of $\Delta$ and $r$. Given $\alpha$, $\beta$ and $r$, we find the solution $\lambda^*$ of equation (2.1) and then, the minimal standardized detectable difference $\Delta^*$. By implementing Norton's algorithm (1983) for the distribution function of noncentral $F$ and then, a simple MATLAB program is used to find $\Delta^*$ when $\alpha$, $\beta$ and $r$ are given. Table 2.1 lists the number of cycles needed with the level of significance $\alpha$=.05 and the probability of type II error $\beta$=0.1 at $\Delta^*$. In Six Sigma management, engineers are expected to detect at least the mean shift of $1.5\sigma$ of the process, *i.e.*, the process mean is shifted by 1.5 times standard deviation of the response variable. Table 2.1 suggests that engineers need to replicate at least six cycles with the level of significance $\alpha$=.05 and the probability of type II error $\beta$=0.1 at $\Delta^*$=1.5. Also, eleven cycles should be run to get at least the power of .9 at $\Delta^*$=1 with the level of significance $\alpha$=.05.

TABLE 2.1 *Number of cycles needed in EVOP with $\alpha = 0.05$ and $\beta = 0.10$ at $\Delta$*

| *Number of cycle* | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\Delta$ | | 3.11 | 2.14 | 1.77 | 1.55 | 1.39 | 1.28 | 1.19 | 1.11 | 1.05 |
| *Number of cycle* | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| $\Delta$ | 1.00 | .96 | .92 | .88 | .85 | .82 | .80 | .78 | .75 | .73 |

In the following we propose a counter measure $p^*$-value corresponding to the $p$-value. The relationship of $p^*$-value to the probability of type II error $\beta$ under $H_1 : \Delta = 1.5$ is analogous to that of $p$-value to the probability of type I error $\alpha$. Thus, the $p^*$-value is defind by the probability, under $H_1 : \Delta = 1.5$, of all the value of test statistics that are as normal in the direction (or in the acceptance region) of $H_0$ as the observed value. Recall that the $p$-value is a data dependent largest probability of type I error. Likewise, the $p^*$-value is a data dependent largest probability of type II error. As $p^*$-value gets smaller, we reject $H_1 : \Delta = 1.5$ with more confidence and conclude that we believe in the null hypothesis of no effects, since data supports rare chance of type II error being occurred. Here the $p^*$-value depends on the value $\Delta$ which will be determined by the interests of the engineers. In practice, EVOP data are analyzed sequentially at the end of each cycle $r$ and we calculated the probability, under $H_1 : \Delta = 1.5$, of all the value of test statistics that are as normal in the direction (or in the acceptance region) of $H_0$ as the observed value, conditional on having decided to continue at cycle $r - 1$. Since the decision to continue at cycle $r - 1$ will only be made if the $p^*$-value then was large, the $p^*$-value now, which is highly correlated with the previous one, is biasd upwards. Thus we propose the practical critical value 0.1 instead of the strict critical value of 0.05 for the significant $p^*$-value.

We set $\Delta$=1.5 and propose a rule for accepting the null hypothesis of no effects as follows:

1. When $r < 6$, the smallest $p$-value is greater than or equal to .25 and the corresponding $p^*$-value is less than or equals to .1.
2. When $r \geq 6$, the smallest $p$-value is greater than or equals to .25.

Setting $\Delta = 1$, we check the $p^*$-value to be small enough whenever the number of cycles is less then eleven.

EXAMPLE 2.1. Postulating that there are no significant effects, 30 random numbers from the standard normal distribution are generated using Minitab as follows:

TABLE 2.2 *EVOP data for six cycles which are randomly generated from the standard normal distribution*

| | *Conditions* | | | | |
|---|---|---|---|---|---|
| *cycle* | 2 | 5 | 1 | 4 | 3 |
| 1 | 0.363 | −0.275 | 1.673 | 0.596 | 0.737 |
| 2 | 1.024 | −1.852 | 0.233 | 0.039 | −1.450 |
| 3 | 0.689 | 1.229 | −0.278 | −1.747 | 0.405 |
| 4 | 0.400 | 0.066 | −0.236 | 0.082 | −1.172 |
| 5 | 0.370 | −1.457 | 1.208 | −0.039 | 0.077 |
| 6 | −0.722 | −0.445 | 0.372 | 0.470 | 0.663 |

At the end of each cycle, the experimental data are analyzed to give the following table of the smallest $p$-value of A, B and A×B, $p^*$-value at $\Delta = 1$ and $p^*$-value at $\Delta = 1.5$.

TABLE 2.3 *Values of p-value, $p^*$-value at $\Delta = 1$ and 1.5 at the end of each cycle*

| *cycle* | *p-value* | $p^*$*-value at* $\Delta$ =1 | $p^*$*-value at* $\Delta$ =1.5 |
|---|---|---|---|
| 2 | .093 (B) | .693 | .476 |
| 3 | .348 (A×B) | .224 | .056 |
| 4 | .300 (A) | .178 | .029 |
| 5 | .204 (B) | .182 | .025 |
| 6 | .280 (B) | .090 | .006 |

At the end of the third cycle, we declare none of A, B or A×B to be significant, and then start a new phase since the smallest $p$-value is .348 and the $p^*$-value at $\Delta = 1.5$ is .056

We get the same conclusion at the end of $6^{th}$ cycle by setting $\Delta = 1$.

## 3. Implementation of EVOP Using Minitab

At the end of each cycle, the analysis of experimental data had been done by entering responses in the EVOP calculation sheet. Since Minitab is available to most of companies adopting Six Sigma management, a more practical approach would be to implement EVOP using Minitab.

Note that cycles are treated as blocks, since cycles in EVOP are replicated sequentially. To generate a $2^2$ design augmented with center points in Minitab, click ▶ Stat ▶ DOE ▶ Factorial ▶ Create Factorial design, and then click ▶ Designs to choose number of center points to be 1.

For analysis of variance of EVOP data in Minitab, click ▶ Stat ▶ DOE ▶ Factorial ▶ Analyze Factorial Design and then, choose the response variable from the list of variables and click ▶ Terms to mark blocks and center points which are included in the model. To get the square plot of average responses at the end of a phase for the EVOP information board, click ▶ Stat ▶ DOE ▶ Factorial ▶ Factorial Plots and then, select cube plots and data means as type of means to use in plots.

## 4. Implementation of EVOP with a Mixture Experiment

We consider a mixture experiment in which the factors are the ingredients or components present at the mixture, and the response depends only on the proportions of the ingredients and not on the amount of a mixture. Suppose that the mixture consists of $q$ ingredients or components, and let $x_i$ represent the proportion of the $i^{th}$ ingredient in the mixture. Then the design space is the $q-1$ dimensional simplex,

$$x_1 + \cdots + x_q = 1, \qquad x_i \geq 0, \qquad i = 1, \ldots, q. \tag{4.1}$$

The constraint in equation (4.1) makes the levels of factors $x_i$ nonindependent, and this reduces the first-order model to the canonical form of the first-order mixture model which is given as follows:

$$E(y) = \sum_{i=1}^{q} \beta_i x_i.$$

Now, we study how to implement EVOP in a mixture experiment. Suppose that we screen three important components out of the 4 components mixture, say, $x_1, x_2, x_3$ and let $s = (s_1, s_2, s_3, s_4)$ be the current operating condition. We define the reference mixture to the current operating condition $s$ and 'Cox's direction' to be a line connecting $s$ to the original vertex $x_i = 1,\ i = 1, 2, 3$. Letting $\delta_i$ be a small change in the proportion of $x_i$, our proposed design in each cycle of EVOP consists of the current operating condition $s$ and three vertexes of the shrunken simplex along Cox's direction around $s$ corresponding to the three screened components as follows:

$$\begin{aligned}
&(s_1, s_2, s_3, s_4),\\
&\left(s_1 + \delta_1, s_2 - \delta_1 \times \frac{s_2}{s_2 + s_3}, s_3 - \delta_1 \times \frac{s_3}{s_2 + s_3}, s_4\right),\\
&\left(s_1 - \delta_2 \times \frac{s_1}{s_1 + s_3}, s_2 + \delta_2, s_3 - \delta_2 \times \frac{s_3}{s_1 + s_3}, s_4\right),\\
&\left(s_1 - \delta_3 \times \frac{s_1}{s_1 + s_2}, s_2 - \delta_3 \times \frac{s_2}{s_1 + s_2}, s_3 + \delta_3, s_4\right).
\end{aligned}$$

For analysis of variance of EVOP mixture data, do ANOVA for randomized block design to detect any treatment effects based on cumulative experimental data up to the current cycle. Treatment effects being significant, new operating condition could be determined as the design point with the largest sample mean.

## REFERENCES

BOX, G. E. P. (1957). "Evolutionary operation: a method for increasing industrial productivity", *Applied Statistics*, **6**, 81–101.

BOX, G. E. P., HUNTER, W. G. AND HUNTER, J. S. (1978). *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*, John Wiley & Sons, New Jersey.

CORNELL, J. A. (1990). *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data*, 2nd ed., John Wiley & Sons, New York.

LORENZEN, T. J. AND ANDERSON, V. L. (1993). *Design of Experiments: A No-name Approach*, Marcel Dekker, New York.

MYERS, R. H. AND MONTGOMERY, D. C. (1995). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, John Wiley & Sons, New York.

NORTON, V. (1983). "A simple algorithm for computing the non-central $F$ distribution", *Applied Statistics*, **32**, 84–85.