DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN

Technical Report No. 106

March, 1967

ISN'T MY PROCESS TOO VARIABLE FOR EVOP?

by

G.E.P. Box and N.R. Draper

This research was supported in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFSOR Grant No. AF-AFOSR-1158-66 and by the United States Navy through the Office of Naval Research under Contract Nonr-1202(17), Project NR 042-222. Computer time was provided by the Wisconsin Alumni Research Foundation through the University Research Committee.

Isn't My Process Too Variable for EVOP?

by

George E. P. Box and Norman R. Draper
University of Wisconsin

SUMMARY

One often hears the question "Isn't my process too variable for EVOP to work properly?" In this paper this question is answered in the negative. First it is pointed out that large variability implies the existence of large effects waiting to be found. Second it is shown that an EVOP scheme which produces only a modest increase in the basic process variation can lead to detection of effects in a very few cycles.

1. INTRODUCTION

In this paper we intend to clarify a point about Evolutionary Operation (Box, 1957; Box and Hunter, 1959) which appears to be frequently misunderstood. The comment is often made that a particular process is "too variable" for EVOP techniques to be used and that the high intrinsic variation of the process will require "too many cycles" before effects can be determined. As we intend to show, this line of argument is not a valid one, for two reasons:

1. If the process is very variable, it is very likely that large effects are being concealed by the large error. In fact the size of effects waiting to be found can be expected to be of the order of the standard deviation of the process, whatever that standard deviation may be. This is called the "cutting the grass" argument (see Section 2).
2. When an EVOP program is run, additional response variation is inevitably introduced because we are deliberately varying the process conditions. It can be shown that with the deliberate variation in the process variable increasing the standard deviation of the response by only 20% or 30%, the effects of the variable can be detected in a few cycles with conservative probability guarantees (see Section 3).

2. "CUTTING THE GRASS"

Suppose a new plant has just been built. The ordinary variation in the operation of the process which occurs because of imperfect control and sometimes because of operating mistakes will produce differences in performance which in many cases could theoretically be turned to profitable account.

Suppose these latent improvements are measured on some suitable scale (e.g. the profitability of each) and are arranged in descending order of magnitude to produce a diagram like Figure 1.

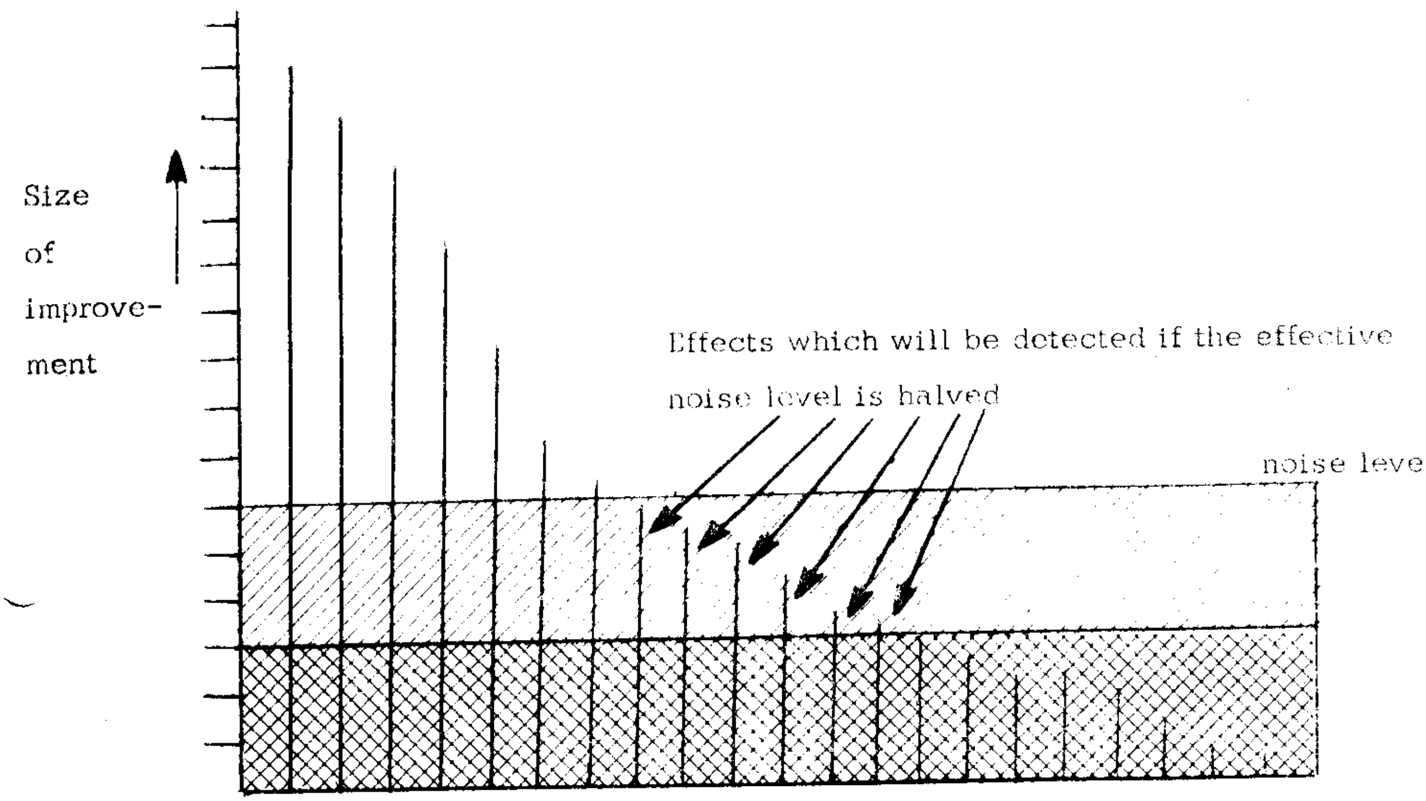

Figure 1. Effects and the "noise level"

The few very large effects to the left of the diagram will be obvious in start up and will lead very quickly to suitable adjustments in process running. The reason for rapid detection of these effects is that the "signal" they produce is large compared with the underlying level of variation of the process, sometimes called the noise level or the "height of the grass".

Noise arises from a variety of sources including variability of raw material, inability to precisely maintain input variables at set levels, as well as from instrument and measurement errors. The final variation in the measured output is a composite of all these.

At any given time in the life of a plant the size of the error variation, as measured by the standard deviation σ , has some definite magnitude which depends on the degree of process control which has been achieved up to that time. Some specified multiple of this standard deviation can be called the noise "level" or the "height of the grass". If a change in level of an input process variable produces an effect in the response which greatly exceeds the noise level, it "sticks out of the grass", and is detected easily. However if the effect produced is much smaller than the noise level, it is unlikely to be detected.

To discover effects buried in noise we must improve the signal to noise ratio. We must either decrease the effective noise level or increase the signal level. In EVOP we do both. The signal is increased by deliberately introducing changes of a carefully chosen kind in the variables under study. The effective noise level is reduced by repetition of the changes and averaging of the results. The diagram shows how a number of effects previously hidden will be shown up if we cut the noise to a half of its previous value. In general as the noise level is reduced by repeated cycles, i.e. as we "cut the grass lower," more and more effects emerge. This argument leads to the conclusion that no process can be said to be "too variable for EVOP". The very fact that the noise level of a process is high means that effects of a sizeable order are "lurking in the grass" waiting to be detected.

This raises another question. "Granted that effects may be present, but concealed in individual runs by the noise level, might it not take many cycles to detect them in a very variable process.

The answer is: -"No, not if it is percentage increase in standard deviation which governs the toleration of an EVOP scheme." Specifically, whether a process has a standard deviation $\sigma = 1$ or a standard deviation $\sigma = 10$, presumably the manufacturer and the customer have learnt to live with this. The result of applying Evolutionary Operation will be to slightly increase the standard deviation, but it seems reasonable to suppose that the slight inconvenience of running with, say a 20% increase in standard deviation (that is with $\sigma = 1.2$ for the first process and $\sigma = 12$ for the second) is equally tolerable. If this is so, it will take precisely the same number of cycles to detect effects in either case.

3. HOW MANY CYCLES ARE NECESSARY TO DETECT EFFECTS OF REASONABLE SIZE?

3.1 Results

In running an Evolutionary Operation scheme, one deliberately introduces additional variation by changing the variables it is desired to study. Thus Evolutionary Operation does not really get something for nothing. It obtains information from the process at the expense of slightly increasing its variability. Now as might be expected, it is the proportional increase in variability which determines the detectibility of the effects of the variables. It follows that we can obtain some idea of the number of cycles which might be needed for a given phase of an EVOP program by considering the proportional increase in the standard deviation of the response which would be acceptable. Normally, formal tests of significance at some fixed significance level are out of place in the routine running of Evolutionary Operation. Nevertheless, for the purposes of making calculations it becomes necessary to formalise our model and in this investigation we have done this in terms of the familiar Neyman-Pearson theory of testing of hypotheses.

In Table 1 the quantity k is the proportional increase in standard deviation produced by changing the variables in accordance with the factorial design of the EVOP program. The Table shows results for both

Design	α	β	k: 1.2	1.3	1.4	1.5	1.6	1.7	1.8
2^2 factorial	0.10	0.10	5.6	3.6	2.6	2.0	1.6	1.3	1.1
	0.10	0.05	6.8	4.4	3.1	2.4	1.9	1.6	1.3
	0.05	0.10	6.9	4.4	3.2	2.4	1.9	1.6	1.4
	0.05	0.05	8.2	5.2	3.8	2.9	2.3	1.9	1.6
2^3 factorial	0.10	0.10	3.2	2.1	1.5-	1.1	0.9	0.8	0.6
	0.10	0.05	3.9	2.5	1.8	1.4	1.1	0.9	0.8
	0.05	0.10	3.9	2.5	1.8	1.4	1.1	0.9	0.8
	0.05	0.05	4.6	3.0	2.1	1.6	1.3	1.1	0.9

Table 1 Number of cycles required to detect, with probability $(1-\beta)$, using an α level test, main effects which increase the standard deviation of a process from α to $k\sigma$, in two-factor and three-factor Evolutionary Operation schemes without center points.

the 2^2 and 2^3 factorial design. (For the sake of simplicity, center points which are often included in EVOP designs, are omitted.) The quantities α and β are the usual errors of the first and second kind, α representing the chance of wrongly detecting the effects of the variables when no effects exist and β the risk of failing to detect effects which do exist. In our analysis, again for simplicity, we have supposed that only main effects occur. The entries in the body of the table indicate the number of cycles necessary to achieve the stated values of α and β. For information, these numbers have been quoted (to one decimal place) in the manner in which they emerge from the appropriate calculations. Of course the number of cycles must be an integer, so that suitable rounding should be performed where necessary. The Table shows for example that, for a three variable EVOP, if changes in the variables could be tolerated which would produce a total increase of 30% in the standard deviation, and for 5% levels of α and β risks, three cycles would be sufficient to detect the effects.

The results indicate that where, as would often be the case, increases in standard deviation of 30% could be tolerated, then two or three cycles of three-variable EVOP programs should be enough to ensure reasonable values of α and β. Of course where more precise estimation is needed, a greater number of cycles would be required. Overall the table indicates that a rather modest number of cycles should normally be adequate.

3.2 Technical Derivation of Table 1

Suppose p variables are being examined by means of a 2^p factorial design in one phase of an EVOP scheme; usually p will equal two or three. Suppose, further, that the levels selected for these variables produce effects $E_1, E_2, \ldots, E_p$ on a single response variable and that there are no interactions between the variables. Thus, for example, a change from the lower to the upper level of the i-th variable produces an increase of E_i $(i=1, 2, \ldots, p)$ on the response, independently of the levels of other variables.

Let e_i denote the estimate of E_i obtained from the mean responses after n cycles in the usual manner and let σ^2_e denote the variance of any of these effects, since all have equal variance.

Consider the hypothesis H_0: all $E_i = 0$ against the alternative $H_1: E_i \neq 0$ for at least one i.

If the random errors attached to the observations are $N(0, \sigma^2)$ then each e_i will also be normally distributed, in fact $e_i \sim N(E_i, \sigma^2_e)$ where $\sigma^2_e = 4\sigma^2/n2^p$. Thus if H_o is <u>true</u>,

$$u = \sum_{i=1}^{p} e_i^2 \sim \chi^2_p \sigma^2_e \tag{3.2.1}$$

where χ^2_p denotes a chi-square variable with p degrees of freedom. However if H_1 is true,

$$u = \sum_{i=1}^{p} e_i^2 \sim \chi_p^2 (\Lambda) \sigma^2_e \tag{3.2.2}$$

where $\chi_p^2(\Lambda)$ denotes a non-central χ^2 variable with p degrees of freedom and non-centrality parameter

$$\Lambda = \sum_{i=1}^{p} E_i^2/\sigma_e^2 . \qquad (3.2.3)$$

(Keeping, 1962, page 397; Fix, 1949.) We shall now wish to select the value of n, the number of cycles, so that we can detect differences with a probability $1-\beta$ while using an α-level test. This means that we shall require that <u>if</u>

$$u/\sigma_e^2 \geq \chi^2_{p,\ 1-\alpha}$$

where the right hand side denotes the $1-\alpha$ point of the χ_p^2 distribution leaving area α in the right tail, then also

$$u/\sigma_e^2 \geq \chi^2_{p,\beta}(\Lambda)$$

where the right hand side is the β point of the $\chi_p^2(\Lambda)$ distribution leaving area β in the left tail. This can be achieved if n is chosen so that

$$\chi^2_{p,\ 1-\alpha} = \chi^2_{p,\beta}(\Lambda) \qquad (3.2.4)$$

In practice a non-integral value of n will be required to satisfy this equation, but the value can be suitably rounded. Larger values of n provide a higher certainty of detection at the same level α or the same certainty at a level smaller than α.

The non-centrality parameter

Suppose the variance of observations in the absence of effects is $V(y) = \sigma^2$. If a change in level of the i-th variable produces an effect E_i in the response, the overall variance increases by $E_i^2 / 4$. Overall, the variance increases to

$$\sigma^2 + \frac{1}{4} \sum_{i=1}^{p} E_i^2 \qquad (3.2.5)$$

which by (3.2.3) can be written as

$$\sigma^2 \{ 1 + \frac{1}{4} \Lambda \frac{\sigma_e^2}{\sigma^2} \} . \qquad (3.2.6)$$

If the variance is thus increased by a factor k^2 from σ^2 to $k^2\sigma^2$, then we obtain, setting (3.2.6) equal to $k^2\sigma^2$,

$$\Lambda = 4 \frac{\sigma^2}{\sigma_e^2} (k^2 - 1), \qquad (3.2.7)$$

or, since $\sigma_e^2 = 4\sigma^2 / n2^p$, it follows that

$$\Lambda = n2^p (k^2 - 1). \qquad (3.2.8)$$

Thus effects which increase the standard deviation of a process from σ to $k\sigma$ produce a non-centrality parameter of a size given by equation (3.2.8).

Exact solution of equation (3.2.4)

If we have tables of percentage points of the non-central χ^2 distribution we can proceed as follows:

1. For specified values of p, α, and β, look up, in non-central χ^2 tables, the value $\Lambda = \Lambda_0$ which satisfies equation (3.2.4)

2. Evaluate, for any selected value of k,

$$n = \Lambda_0 / 2^p (k^2 - 1),$$

using equation (3.2.8) and round off to the next highest integer (or to the next lowest if the difference is less than 0.1).

This method is simple, but possible only for $\alpha = 0.05$ and 0.01 and $\beta = 0.1(0.1)0.9$. The appropriate tables are given by Fix (1949). For other values of α and β an approximate method of calculation can be employed which appears to provide perfectly adequate accuracy.

Approximate solution of equation (3.2.8)

A useful (Patnaik, 1949) approximation to the β percentage point, $\chi^2_{p,\beta}(\Lambda)$, of the non-central χ^2 distribution with p degrees of freedom and non-centrality parameter Λ is given by $A\chi^2_{B,\beta}$ where

$$\begin{aligned} A &= 1 + \Lambda/(p+\Lambda) \\ B &= p + \Lambda^2/(p+2\Lambda) \end{aligned} \tag{3.2.9}$$

where Λ is given by equation (3.2.8). This approximation converts equation (3.2.4) into the (approximate) equality

$$\frac{p + n2^p (k^2 - 1)}{p + n2^{p+1} (k^2 - 1)} \chi^2_{p,1-\alpha} = \chi^2_{B,\beta} \, . \tag{3.2.10}$$

The simplest way to solve (3.2.10) for n when the values of α, β, p, and k are selected is to tabulate, for various (integer) values of n, both the left-hand side and the values of B which correspond. The values $\chi^2_{B,\beta}$ are then obtained by interpolating in the central χ^2 tables, since B will be non-integral. These calculations will then allow a non-integral value of n to be interpolated which satisfies equation (3.2.10). The results are as already given in Table 1.

ACKNOWLEDGEMENTS

We are grateful to David E. Tierney who performed the calculations. This research was supported partially by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFSOR Grant No. AF-AFOSR-1158-66, and partially by the United States Navy through the Office of Naval Research under Contract Nonr-1202(17), Project NR 042-222. Computer time was provided by the Wisconsin Alumni Research Foundation through the University Research Committee.

REFERENCES

Box, G.E.P. (1957). Evolutionary Operation: a method for increasing industrial productivity. Applied Statistics, 6, 3-23.

Box, G.E.P. and Hunter, J.S. (1959). Condensed calculations for Evolutionary Operation programs. Technometrics, 1, 77-95.

Fix, Evelyn (1949). Tables of non-central χ^2. University of California Publications in Statistics, 1, No. 2, 15-19.

Keeping, E.S. (1962). Statistical Inference, D. Van Nostrand Co., Inc., N.Y.

Patnaik, P.B. (1949). The non-central χ^2 - and F - distributions and their applications. Biometrika, 36, 202-232.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

1. Office of Aerospace Research and Office of Naval Research	2a. Unclassified
	2b. --
3. ISN'T MY PROCESS TOO VARIABLE FOR EVOP?	
4. --	
5. G. E. P. Box & N. R. Draper	
6. March, 1967	7a. 11 pp. 7b. 5 references
8a. AF-AFOSR-1158-66 Nonr-1202(17)	9a. Report No. 106
8b. NR 042-222	9b. --
10. Distribution of this document is unlimited.	
11. --	12. Office of Aerospace Research Arlington, Virginia Office of Naval Research Washington, D. C.

13. Abstract: One often hears the question "Isn't my process too variable for EVOP to work properly?" In this paper this question is answered in the negative. First it is pointed out that large variability implies the existence of large effects waiting to be found. Second it is shown that an EVOP scheme which produces only a modest increase in the basic process variation can lead to detection of effects in a very few cycles.