Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 65(1) 2011 21-26

Multi-Response and Non-Traditional Experiments*

S.P. Mukherjee
Indian Association for Productivity, Quality and Reliability
AD 276, Salt Lake City, Kolkata

INTRODUCTION

Experimentation - in a broad sense - has been an age-old exercise by seekers of TRUTH. Simple and improvised experiments have led to wonderful innovations, though sophisticated and complicated experiments have become bad necessities in some recent contexts. While a wide array of experiments has existed over time, planning experiments to yield results that would provide a basis for inductive inferences, started in the context of Agriculture only in the early twentieth century. Since then, planning an experiment and analyzing the experimental results has become the key element in any scientific or technological investigation to-day.

Developments in Design and Analysis of Experiments were taking into account diverse problems encountered by experimenters even before the experiment could be designed. Such problems arose in identifying the appropriate factors and their levels along with constraints on these, in taking note of possible incompatibilities among levels of different factors and of prior information about the relative importance of a factor and its relation with other factors. Further, problems in identifying and subsequently quantifying the response variable(s) posed some problems. Theoretical developments were obviously based on certain assumptions that could not be always validated.

In a large variety of situations, the investigator is interested in the behaviour of more than one responses, which may be inter-related to varying extents, expressed in different units of measurement, and have different influences on the conclusions regarding the phenomenon or process under study. Multi-response experiments have attracted the attention of many research workers and remains a field with many more novel ideas to emerge.

The present article is essentially a review of some approaches to analysis of multi-response experiments that goes beyond the conventional approaches and recognizes sampling fluctuations in the response functions and offers a new formulation of the problem in terms of stochastic programming. It also portrays some typical problems with the design aspects as also with the responses, without — of course — suggesting any solutions. All these problems may arise in agricultural as well as industrial experiments.

THE PANORAMA OF EXPERIMENTS

An experiment, generically, means an exercise to gather (generate or compile) empirical evidences (not based on theory) and acquire knowledge relating to a phenomenon (occurring in Nature or Society or Economy) or a system or an operation. Experiments range from simple and quick observations, through design, development and administration of a test or a treatment, to complicated and even hazardous ventures into unknown tracts and entities. Simple laboratory experiments to identify and estimate properties of substances, bioassays to determine relative potency of a new biological preparation compared to a standard one, agricultural field trials with different manures or varieties or irrigation practices, clinical trials with

*Dr. V.G. Panse Memorial Lecture delivered at the 64th Annual Conference of the Indian Society of Agricultural Statistics held at BCKV, Kalyani during 3-5 December 2010.

E-mail address : mukherjeesp2000@yahoo.co.in

different drugs or treatment protocols, complicated and costly experiments in space research, etc. illustrate the panorama of experiments.

It is useful to distinguish between two types of experiments viz. absolute and comparative. An example of the first type could be an experiment to determine the electric charges of an electron or to find the mean root-length of a plant species treated with a specified level of an antibiotic compound or to find out if a given dose of a certain biological preparation does produce muscular contraction when applied to a mouse. It is for such experiments that the theory of errors was originally devised. Repeated outcomes of such an experiment do not agree exactly with one another and the problems considered were to obtain best estimates and associated measures of reliability from sets of outcomes.

On the other hand, in a comparative experiment two or more experimental conditions are compared with respect to their effects on a chosen characteristic of the population under study. Suppose, for example, a metallurgical engineer is interested in studying the effects of two different hardening processes, oil quenching and saltwater quenching, on some alloy steel. The engineer subjects a number of alloy specimens to each quenching medium and measures the hardness of each specimen after quenching.

It should be pointed out that an experiment, where only one factor or causal entity is considered at different levels and the corresponding effects are noted with the object of determining the relation between levels of the causal entity and the responses, involves comparison across levels, but can also be regarded as absolute or better exploratory.

Continuing with comparative experiments, one can proceed to have several conditions tried out to find which experimental condition is the best. Alternatively, one can try to find out which level or dose of only one preparation will yield the best result in terms of some physiological parameter or some change therein. These are sometimes branded as optimization experiments or response surface experiments.

THE "RESPONSE" PARADIGM

The properties or features or changes in such properties or features which follow the application of a treatment or treatment combination on each experimental unit which are ultimately analyzed to meet the objective(s) of the experiment and which are expected to vary from one unit to another are recognized as response variables. Most experiments involve only one response variable.

Responses of different units may be just observable (attribute) or countable (discrete variable) or measurable (continuous variable that appears to be discrete when recorded). Data arising from the experiment are accordingly recognized as categorical or discrete or continuous. In some cases of agricultural or biological experiments, the response could be quantal (all or nothing) or binary. We require different types and methods of analysis in these three different situations.

It may be important in some situations to distinguish between the 'yield' or 'response' that can be noted in respect of the experimental units and the variable of interest. This may happen if the variable of interest is difficult or time-consuming to note just at the end of an experiment and we tend to note some other variable as a proxy for the basic variable of interest or a variable that is related to the basic variable. The analysis of experimental data has to recognize three different situations viz.

(1) the variable of interest can be directly noted and is the same as the 'yield' or 'response',

(2) the variable of interest is not directly noted and is obtained through a transform (not necessarily monotonic) of the 'yield' or 'response' variable that is noted. For example, the variable of interest is reliability or survival probability but what is directly noted for each experimental unit is just time-to-failure or length of life. A second example could be the dry weight of a crop standing and harvested from a plot and this weight is a transform of the weight of the harvested produce,

(3) the response variable is the variable of interest but cannot be directly obtained, rather it is estimated from some data X which depend on the factor levels as also from data Y derived from some other source. Thus the yield of a crop is not directly noted in respect of some plots and is sometimes estimated from Remote Sensing data or from relations connecting crop yield with agro-climatic and agronomic variables.

In the last two situations, the analysis of data should take into account the nature of the transform and the relation connecting the variable of interest with the variables which have been noted.

All this relates to the case of a single variable of interest. Complications arise in dealing with more than one variables of interest.

THE FACTOR SPACE

The design of an experiment is essentially a plan to conduct the experiment so that valid and relevant conclusions can be reached efficiently and economically. Such a plan is in terms of decisions related to the choice of experimental units, the choice of factor levels or factor level combinations (treatments) and their associations with different experimental units (or different runs of the experiment, as the case may be) along with the measurements or other records to be kept for each experimental unit.

A good design should incorporate all prior information about the factors, factor levels, possible inter-relations among factors and should require only a few factor level (treatment) combinations to be tried out in an experiment (or an experimental run) as also only a few runs to reach the optimum level combinations.

In the traditional factorial experiments, all levels of a given factor are assumed to be compatible with each level of any other factor, so that the number of treatment combinations which can be accommodated in the framework of an experiment is simply the product of the numbers of levels of all the factors (which may or may not be the same for each factor). Of course, in some experiments some of these treatment combinations may not be included in each replicate or even not included at all, because of resource constraints or because of prior information justifying their exclusion for the purposes to be served by the experiment. It is also possible that although all or most of the treatment combinations were included in the design, response(s) could be missing for some of the experimental units. We can even think of situations where responses to one or more treatment combinations were completely missing. We have tools to estimate the missing response values and proceed with the analysis.

However, in practice, it is quite possible that some levels of certain factors cannot be combined with some levels of some other factors because of adverse or otherwise inadmissible joint effects on the experimental units, so that we have to think of designs which are structurally different from usual factorial experiments. Let us consider some simple examples of such a situation.

Let us take a 2×2 experiment where, suppose, the higher level of A cannot go with the lower level of B, so that we are left with only 3 treatment combinations e.g. (1) b and ab with the result that the 2 degrees of freedom associated with the treatment combinations can be used to estimate somehow effects A and B only in terms of simple effects viz. (ab)-(b) and (b)-(1) respectively. Going to the $2 \times 2 \times 2$ experiment there will be, depending on which treatment combinations are not compatible, problems in estimating some main effects and some interactions. At least, the second order interaction will not have the usual interpretation in terms of the difference between a first order interaction corresponding to the higher and that corresponding to the lower levels of the third factor. One can only expect more complications with general factorial experiments.

Mixture experiments are quite relevant in many industrial and agricultural situations. In most such experiments, the factors correspond to levels of some constituents in a combination that produces a single response (or several responses). If x_i denotes the level of constituent i (as a fraction of the over-all content of the mixture) with $x_i \geq 0$ and $\Sigma x_i = 1$, constraints that are known to be taken care of are like $x_i \leq (\geq) A_i$. However, several real-life situations require constraints like $x_i \geq (\leq) x_j$ to be satisfied. These await proper accommodation within the current framework of mixture experiments.

MULTI-RESPONSE EXPERIMENTS

An industrial engineer may want to study the influence of cutting speed and depth of cut on the life of a tool and the rate at which it loses metal. A food technologist may be interested in determining optimum combinations of the various ingredients of a product on the basis of acceptability, nutritional value, economics, and other considerations. A medical researcher studying the effects of complexing agents on the yield of a certain antibiotic may also be interested in the product

cost. Hill and Hunter (1966) cite several papers in which multiple responses are investigated.

Multi-response experiments are also quite relevant in the context of Agriculture. Materials about several such experiments were communicated to this author by V.K. Gupta (2010). These are briefly indicated in what follows.

An experiment was conducted during 2004-05 at Department of Agronomy, BCKV on integrated nutrient management on Rapeseed. Three different sources of sulphur were tested along with recommended dose of Nitrogen (N), Phosphorous (P) and Potassium (K) and Farmyard Manure (FYM). The experiment was laid out in a Randomized complete block (RCB) design in 3 replications. Data on the following response variables were collected :

I. Number of branches per plant (P1);
II. Number of siliqua per plant (P2);
III. Number of seed per siliqua (P3);
IV. Seed yield (g/plant) (P4);
V. Straw yield (g) (P5);
VI. HI (Havest Index) (P6);
VII. Test weight (P7);
VIII. Leaf area Index 45 days after sowing P(8);
IX. Leaf Area Index 90 days after sowing P(9).

An experiment was conducted on osmotic dehydration of banana to determine optimum combination of power levels, temperature and air velocity at Division of Agricultural Engineering, IARI, New Delhi. The data were collected on energy use efficiency (%) [Y_1], rehydration ratio [Y_2], total soluble solids (TSS) [Y_3], total sugars [Y_4] and total carbohydrates (mg/g dry matter) [Y_5]. The experimenter is interested in obtaining the optimum combination of the controllable factors that maximizes all the response variables simultaneously.

In case of complete multi-response situation, parameter estimation is not a problem at all. However, in case of incomplete multi-response experiments, the problem of parameter estimation and then obtaining the point of optimum response vector arises.

The response variables should not be investigated individually and independently of one another. Interrelationships that may exist among the responses can render such univariate investigation meaningless.

The goal of multi-response experiments is to find the setting of the design variables that achieve an optimal compromise of the response variables. By optimal compromise we mean finding the operating level of the design variables such that each response characteristic is as "close" as possible to its ideal value.

The multiresponse optimization technique initially proposed by Harrington (1965) and later modified by Derringer and Suich (1980) finds an optimal compromise of product characteristics using a distance metric called the desirability function. Values of desirability range from 0 to 1, with a number closer to 1 being more desirable.

Estimates of each response mean at treatment x are $\hat{Y}_i(x)$. These estimates are transformed into a desirability value $d_1(x)$ using a transformation function. There are two types of transformation function. The first is used when the optimal value of a response is a target value. The second type of transformation is used for a response value that is to be maximized or minimized. This method is attractive because it is intuitive and simple. The inputs are the mean response estimates, $\hat{Y}_i(x)$, the target values, τ_1, and the upper and lower acceptability bounds, ub_i and lb_i, respectively.

Pignatiello (1993), as extension of Taguchi's (1989) single-response loss function, presents a multirespone technique based on the criteria of minimizing deviation from taget and maximizing robustness to noise. If the product characteristics deviate from the target values a weighted penalty or loss is incurred. The distance metric denoted by $\hat{D}_p(x)$ consists of two components, a deviation-from-target component and a variance component.

An advantage of this method is the addition of a variance component into the distance metric because we would like to avoid those operating conditions that introduce variability into the process. A disadvantage is that we can not specify an acceptability region for the responses, as well, the distance metric is intended for product characteristics that have a specific target value, and therefore its appropriateness for characteristics to be minimized (maximized) is questionable.

A multiresponse technique similar to the previous approach, which was developed by Ribeio and Elsayed (1995), adds a third component to $\hat{D}_p(x)$, namely a term associated with fluctuations in the design variables. We call the distance metric $\hat{D}_{\mathrm{REO}}(x)$. Here, the design variables are random variables. The information captured by the third term of Eq. (*D*1) is an estimate of the variability induced on the response variables due to the fluctuations of the design variables. The third term will place an additional penalty on those operating conditions (i.e., settings of the *X*'s) in which design-variable variation has an effect on the product characteristics. When optimizing this distance metric, the analyst is led away from those operating conditions where the *X*'s cannot be set accurately.

A scenario in which this method might be useful is the case in which an experiment is conducted in a laboratory or pilot plant and the results are to be applied to a full-scale production system. Control variables that exhibit small fluctuations in the pilot system will often be inflated in full-scale production. If the optimization is performed on the pilot plant, the optimal obtained may not be the best choice for the full-scale system. Using this distance metric, if the design variable fluctuations in full-scale production were known (from control charts, for example), the analyst could incorporate this into the optimization analysis.

The last multiresponse technique investigated is a method proposed by Khuri and Conlon (1981), whose distance metric is $D_{\mathrm{KC}}(x)$. This distance metric uses the squared deviations of the product characteristics from their targets, but then normalizes these deviations by the variance of prediction of the response variables.

Khuri and Conlon propose a method for determining the tagets τ_i's. For response variables where smaller-is-best (bigger-is-best), τ_i is set equal to the minimum (maximum) of $\hat{Y}_i$ over the region of the experiment. [For example, consider the case of one design variable $0 < X < 1$ and one response variable where smaller-is-best. If $\hat{Y}(x) = 2 + 3X$, then τ is set equal to 2.]

DUAL RESPONSE SYSTEM

Myers and Carter proposed an approach associated with the exploration of two response functions – one recognized as primary and the other as secondary. Their objective was to find the operating settings which optimize a primary response $\hat{Y}_p$ subject to the condition that a secondary response $\hat{Y}_s$ takes on a desirable value, usually called a "target value" (denoted by T). Both $\hat{Y}_p$ and $\hat{Y}_s$ are assumed to be second-order polynomial regression functions. Thus, the problem considered is the non-convex quadratic programme

$$\text{Min. } \hat{Y}_p(X) = b_o + X^T b + X^T \mathbf{B} x$$

$$\text{s.t. } \hat{Y}_s(X) = c_o + X^T c + X^T \mathbf{C} x = T$$

$$\mathbf{X}^T x \le \rho^2,$$

where $\hat{Y}_p$ is the primary response, $\hat{Y}_s$ is the fitted secondary response, T is the target value $\hat{Y}_s$, and $X^{\mathrm{T}} = (x_1, x_2, \ldots, x_k)$ is the vector of control factors, **B**, b and b_o contain the estimated regression coefficients for $\hat{Y}_p$, similarly, **C**, c and c_o contain the estimated regression coefficients for $\hat{Y}_s$. That is

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12}/2 & \ldots & b_{1k}/2 \\ : & b_{22} & \ldots & b_{2k}/2 \\ : & : & \ldots & : \\ b_{1k}/2 & \ldots & \ldots & b_{kk} \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12}/2 & \ldots & c_{1k}/2 \\ : & c_{22} & \ldots & c_{2k}/2 \\ : & : & \ldots & : \\ c_{1k}/2 & \ldots & \ldots & c_{kk} \end{bmatrix}$$

$\mathrm{B}^{\mathrm{T}} = [b_1, b_2, \ldots, b_k]$; $c^{\mathrm{T}} = [c_1, c_2, \ldots, c_k]$.

Here, the search for optimal operating settings is confined to a ball of radius ρ.

It is assumed that all factors are scaled such that the designed expiment is centered at $X = 0$.

STOCHASTIC PROGRAMMING FORMULATIONS

However, the resulting solution is only a point estimate for the location of the global optimum and this neglects the sampling variability and bias of the fitted responses.

To address the inherent sampling error in DRS, a so-called "optimal region" where global optimum of the true system resides with a certain probability has been

considered by some authors. Herein, a Monte Carlo simulation study of DRS is performed; then for each simulated dual system of responses an algorithm is applied to compute the global optimal settings.

The issue of random (sampling) variations in the estimated responses can be taken care of in a stochastic programming formulation of the problem. This requires derivation of the somewhat complicated univariate normal distribution of each response using the result $\hat{B}_i \sim N(\beta, (X^T X)^{-1} \sigma^{ii})$ for $i = p, s$.

Here again we may have several alternative formulations by considering

(a) $E(\hat{Y}_p)$ for maximization w.r.t. X

(b) $Var(\hat{Y}_p)$ for minimization

(c) $CV(\hat{Y}_p)$ to be minimized

(d) $Pr\{\hat{Y}_p \geq y_o\}$ for a given y_o to be maximized

(e) $Y_o = G^{-1}(\alpha)$ for a given α to be maximized

where $G(t) = Pr\{\hat{Y}_p \leq t\}$ is the distribution function of $\hat{Y}_p$; all these may be subject of $Pr\{\hat{Y}_p \geq y_{po}\} \geq \beta$, β being a pre-assigned large value.

These are difficult chance-constrained programming problems, not yet fully explored and await any practical applications. While $E(\hat{Y}_p)$, $V(\hat{Y}_p)$ and $CV(\hat{Y}_p)$ may not be that difficult to obtain, the derivation of $G(t)$ will be quite complicated. Even assuming a normal distribution for $\hat{Y}_p$, the second formulation gets translated into a quadratic programming problem while the remaining three formulations call for solving fractional programming problems.

In situations where both the responses are of equal importance and the experimenter tries to achieve some target for each, one can formulate a goal programming problem. Even differences in importance can be accommodated in terms of weights while considering the weighted total of deviations in the estimated responses from the targets as the objective function. Not much is yet known about stochastic goal programming if we try to accommodate constraints like $Pr\{\hat{Y}_p \geq y_{po}\} \geq \beta$ and $Pr\{\hat{Y}_s \geq y_{so}\} \geq \nu$, β and ν being pre-assigned large values and we may have to take recourse to simulation for a search of the optimal solution X.

The goal-programming approach, even without chance constraints, can be extended to the case of any number of (estimated) responses. Of course, fixing targets for the different responses and associating appropriate priorities or weights can pose complex problems of choice.

SOME REFERENCES

Castillo, E. Del (1996). Multiresponse process optimization via constrained confidence regions. *J. Qual. Tech.,* **28**, 61-70.

Castillo, E. Del and Montogomery, D.C. (1993). A nonlinear programming solution to the dual response problem, *J. Qual. Tech.,* **25**, 199-204.

Castilo, E. Del, Fan, S.K. and Semple, J. (1997). Computation of global optima in dual response systems. *J. Qual. Tech.,* **29**, 347-353.

Derringer, G. and Suich, R. (1980). Simultaneous optimization of several response variables. *J. Qual. Technol.,* **12**, 214-219.

Harrington, E.C. Jr. (1965). The desirability function. *Ind. Qual. Control,* **21**, 494-498.

Khuri, A.I. and Conlon, M. (1981). Simultaneous optimization of multiple responses represented by polynomial regression functions. *Technometrics,* **23**, 363-375.

Myers, R.H. and Carter, W.H. Jr. (1973). Response surface techniques for dual response systems. *Technometrics*, **15**, 301-317.

Pignatiello, J.J. Jr. (1993). Stategies for robust multiresponse quality engineeing. *IIE Trans.,* **25**, 5-15.

Ribeiro, J.L. and Elsayed, E.A. (1995). A case study on process optimization using the gradient loss function. *Int. J. Product. Res.,* **33**, 3233-3248.

Roy, S.N. (1961). On the planning and interpretation of multifactor multiresponse experiments. *Bull. Inst. Internat. Statistics*, **38**, 59-72.

Roy, S.N. and Srivastava, J.N. (1964). Hierarchical and p-block multiresponse designs and their analysis. *Contributions to Statistics – Mahalanobis 70th Birthday Volume,* 419-428.

Taguchi, G., Elsayed, E.A., and Hsiang, T. (1989). *Quality Engineering in Production Systems.* McGraw-Hill, New York.

Vining, G.G. and Myers, R.H. (1990). Discussion of off line quality control, parameter design, and the Taguchi Method. *J. Qual. Tech.,* **17**, 198-206.