STATISTICAL DEFINITIONS, DESIGNS, AND RESPONSE EQUATIONS FOR EXPERIMENTS ON FIXED-RATIO MIXTURES IN AGRICULTURE

Walter T. Federer and Anila Wijesinha
U.S.A. and Sri Lanka

1. INTRODUCTION

Research in and applications of statistical theory have been predominately for univariate responses and for a single type of treatment effect known as a direct effect. This is the effect of a treatment in the period when it is applied. In the real world of research investigations, responses are often multivariate in nature, and several types of treatment effects and side effects may be encountered. In tropical-zone agriculture, and to some extent in temperate-zone agriculture, the growing of mixtures of cultivars in specified proportions or of a specified succession of crops on a given unit of land have been common practices for centuries. The reasons are many and varied, some of which are (i) for increased yields, (ii) for disease control, (iii) for insect control, (iv) for erosion control, (v) for stabilizing annual yields, (vi) for spreading labor and saleable products over a calendar year, (vii) for decreasing use of commercial fertilizer, (viii) for providing shade or standards for vine and creeper crops, (ix) etc. Appropriate and/or correct statistical analyses for investigations on mixtures of cultivars has lagged far behind the needs. In fact, it is postulated that the most important statistical problem associated with tropical agriculture investigations at the present time is the statistical design, appropriate response model equations, statistical analyses (both nonsequential and sequential), and statistical inferences for such investigations. Some of the problems and possible solutions are considered in the present paper.

2. TYPES OF TREATMENT DESIGN

It has been found useful and enlightening to categorize types of treatments that may constitute a treatment design. One such classification follows:

1. Controls, standards, checks, placebos, or other items required as experimental reference points.
2. Discrete levels of the variables or factors under study in an experiment. (These are sometimes denoted as qualitative factors. The commonly known factorial design falls in this category.)
3. Continuous levels of factors or variables under study in an experiment. (These are sometimes denoted as quantitative factors. The so-called "response surface" designs fall in this category.)
4. Mixtures of k of v factors with the proportion of each factor in the mixture being specified by the experimenter, i.e., there is only one level per factor. (The commonly known diallel crossing

system in breeding, the matched pairs design in social studies, the round robin and other tournaments, etc. fall in this category. The two factors are in a 50:50 proportion.)

A treatment design consists of treatments from one or more of the above categories, or of combinations thereof.

3. TYPES OF MIXTURE DESIGNS IN CATEGORY 4 FOR AGRICULTURAL EXPERIMENTS

Treatment designs made up from treatments in category 4 are designated as fixed-ratio mixture designs. There are many types of these designs, and some of the more common ones in agricultural research are: (i) those involving competition between lines, cultivars, or species of plants; (ii) those involving interspersed plants or rows of lines, cultivars, or species; and (iii) those involving a succession of crops on the same experimental unit in a prearranged fashion such as, e.g., rotations.

There are several types of mixtures involved in short term multiple cropping situations; some of these are: (i) Two or more crops are randomly mixed within the same experimental unit; the proportions in the mixture are specified by the investigator. (ii) Individual plants are alternated according to some systemmatic plan. Plans of this type are often for the purpose of understanding the basic physiology of competition. (iii) Rows of members of a mixture are alternated for crops requiring the same growing season. (iv) For crops with a different length of growing season, shorter term crops may be planted in between the rows of the longer season plants; they are harvested early and the longer season plants then occupy the experimental unit as a single crop during the last part of the growing season. (v) Vines or creepers are planted in a stand of trees or bushes when the vines or creepers require some sort of standard for climbing. (vi) A sequence of crops on the same experimental unit in a specified period, say one year, is denoted as successive cropping.

It should be noted that investigations for short-term experiments are interested solely in the effect of the treatments in the periods when they are applied, that is, direct effects. When we come to long-term and/or repeated measures experiments, it is necessary to consider several other types of treatment effects such as the residual (effect of a treatment in periods beyond the period of treatment application), cumulative (the additional effect of application over no application), and permanent (direct plus residual). These additional effects in a linear model can considerably complicate the design, the analysis, and the inference structure.

4. SOME STATISTICAL MODELS RELATED TO CATEGORY 4 MIXTURE DESIGNS

When one considers a statistical analysis for mixtures, many problems arise. The first question is how to make use of the responses obtained. To illustrate, suppose that we have an experiment involving one crop of cassava per year (treatment one), two crops of corn per year (treatment two), two crops of soybeans per year (treatment three), and two crops of cowpeas per year (treatment four). The total yields of these four treatments are very different, have different error variances, different insect and disease problems, different nutritional values, and different economic values per kilograms. The point is, we wish to compare these four treatments instead of considering each crop individually as is usually done. What measure should be used? Some univariate contenders that come to mind are: (i) total calories, (ii) total yield in kilograms, (iii) total economic value or profit, (iv) land equivalent ratio or relative yield total (defined to be the relative land area for single crops to produce the amount obtained with a mixture), and (v) yield per calendar month of a year. Or, should one consider each treatment yield as a variate and use (i) multivariate procedures, (ii) linear programming procedures, (iii) stochastic programming procedures, or (iv) some other approach? The exact approach to be used for a particular study requires thought, investigation, and applications to actual experimental data. Perhaps several approaches will be necessary.

One particular univariate approach for mixtures has been put forth for a linear model for a mixture of k of v cultivars with general mixing, and n^{th}-specific mixing effects, $n=2,\cdots,k$ both for total yields of a mixture and for individual component yields of a mixture. The response model for experimental unit totals for k=3 is given below for a fixed-ratio mixture treatment design and a randomized complete block experiment design:

$$Y_{ghij} = \mu + \rho_g + [\tau_h+\delta_h+\tau_i+\delta_i+\tau_j+\delta_j]/3 + 2[\gamma_{hi}+\gamma_{hj}+\gamma_{ij}]/3 + \pi_{hij} + \epsilon_{ghij}, \tag{5.1}$$

where μ is an effect common to every observation, ρ_g is the g^{th} block effect, ϵ_{ghij} have mean zero and common variance σ^2_ϵ, $\tau_h (h,i,j=1,2,\cdots,v)$ is the h^{th} cultivar effect when grown in pure stand, δ_h is the general mixing effect for cultivar h when grown in a mixture, γ_{hi} is the bi-specific (interaction) effect of the pair of cultivars h and i when grown in a mixture, and π_{hij} is the tri-specific mixing effect of cultivar h, i, and j when grown in a mixture.

When individual yields of a mixture are available, a response model equation would be obtained for each member of the mixture. For k=3, the following equation would be:

$$Y_{gh(ij)} = (\mu+\rho_g+\tau_h+\delta_h+\gamma_{hi}+\gamma_{hj}+\pi_{hij})/3 + \epsilon_{gh(ij)}, \tag{5.2}$$

where the subscript h, i, or j not in parentheses indicates the yield for that cultivar in the presence of the cultivars in the parentheses. Thus, $Y_{gh(ij)}$ is the yield for cultivar h grown in the mixture hij. The effects μ, ρ_g, and π_{hij} as defined for equation (5.1) are split equally between cultivars h, i, and j. Likewise, a bi-specific mixing effect, say γ_{hi}, is split equally between the pair involved. This may not be a tenable model for some fixed-ratio mixture designs. In fact, most of the bi-specific mixing effects could be allocated to one of the cultivars involved. The same situation could prevail for the tri-specific mixing effect π_{hij}. One would need to alter the response model equation (5.2) as follows:

$$Y_{gh(ij)} = (\mu+\rho_g+\tau_h+\delta_h)/3 + \gamma_{h(i)} + \gamma_{h(j)} + \pi_{h(ij)} + \epsilon_{gh(ij)}, \qquad (5.3)$$

where $\gamma_{h(i)}$ is a bi-specific mixing effect of cultivars h and i grown in a mixture that is attributable to cultivar h and $\pi_{h(ij)}$ is a tri-specific mixing effect of cultivars h when grown in a mixture involving the three cultivars h, i, and j.

Minimal treatment designs to obtain unique solutions has received some discussion, but further work is required in this area. The class of symmetrical balanced incomplete block designs form minimal designs for general mixing effects. For general mixing effects plus bi-specific mixing effects of v items, it is necessary to have $v(v-1)/2$ combinations.

We are currently pursuing a number of aspects concerned with the statistical design and analysis of fixed-ratio mixture designs. One aspect would be to consider the yields in (5.2) as multivariates and use a multivariate analysis. Another is to consider alternate models such as (5.3). Still another aspect is to investigate further minimal treatment designs for this type of experiment. This paper is a condensation of a technical report (listed below) which contains an extensive list of references on this topic. Most of the references listed were published within the last ten years.

Some references of current work on the topic of fixed-ratio mixture designs by the authors are:

Federer, W. T. (1979). Statistical design and response models for mixtures of cultivars. Agronomy J. 71(5) (to appear).

Federer, W. T., Connigale, J. C., Rutger, J. N. and Wijesinha, A. (1979). Statistical analyses of yields from uniblends and biblends of eight dry bean cultivars. BU-670-M in the Mimeo Series of the Biometrics Unit, Cornell Univ., May.

Federer, W. T., Hedayat, A., Lowe, C. C. and Raghavarao, D. (1976). Applications of statistical design theory to crop estimation, with special reference to legumes and mixtures of cultivars. Agronomy J. 68, 914-919.

Federer, W. T. and Wijesinha, A. (1979). Statistical definitions, designs, response equations, and analyses for experiments on fixed-ratio mixtures. BU-677-M in the Mimeo Series of the Biometrics Unit, Cornell Univ., June.

Raghavarao, D. and Federer, W. T. (1979). Block total response as an alternative to the randomized response method in surveys. J. Royal Statist. Soc. B 41(1) (to appear).

Smith, L. L., Federer, W. T. and Raghavarao, D. (1974). A comparison of three techniques for eliciting truthful answers to sensitive questions. Proceedings, Soc. Statist. Sect., Amer. Statist. Assoc., 447-452.