CHAPTER 14

Some Special Problems

14.1 INTRODUCTION

In this chapter we deal briefly with some miscellaneous problems which do not fit conveniently into the previous chapters. In § 14.2 some designs are described for use when the uncontrolled variation is expected to consist predominantly of a smooth trend. Section 14.3 mentions an important class of problems in which a theoretical calculation is possible to find which of a number of systems of observations will lead to estimates of maximum precision. The next section contains an account of designs for finding optimum conditions. The final section deals with the special problems of assays, in particular of bioassays.

14.2 TREND-FREE SYSTEMATIC DESIGNS

Sometimes we have a small number of experimental units and a substantial part of the uncontrolled variation is expected to consist of a smooth trend, for example a trend in space or time. In the ordinary way this situation is dealt with by one of the methods of Chapters 3 and 4, either (*a*) by randomizing the allocation of treatments and adjusting the treatment means for the trend by the method of § 4.3, or (*b*) by using the method of randomized blocks to deal with variations other than the trend, then calculating adjusted treatment means to correct for the trend, or (*c*) by using the method of randomized blocks to control most of the effect of the trend, putting in the first block the units occurring first in space or time, and so on. The first two methods allow the form of the trend to be estimated directly.

If the experiment is a moderate-sized one with, say, more than four replicates of each treatment, these methods will usually be satisfactory. In smaller experiments, however (particularly if the trend is curved), the first two methods tend to give imprecise results. This is because the randomization is quite likely to throw up an arrangement which is markedly unbalanced with respect to the trend. The third method is

free from this defect, but is not satisfactory if it is desired to estimate as precisely as possible the form of the trend.

We need therefore a new sort of design for such situations that will enable both the trend and the treatment effects to be estimated simply with maximum precision and will take advantage of our knowledge about the expected form of the uncontrolled variation.

Example 14.1. Consider an experiment to investigate the effect on a textile process of changing the relative humidity and suppose that three relative humidities, 50, 60, and 70 per cent, are to be used. To obtain uniform experimental units a suitable quantity of raw material would be taken and thoroughly mixed and then divided into say nine batches to form nine experimental units. The first batch would be processed at one relative humidity in the first period, the second batch at a different relative humidity in the second period, and so on. Superimposed on any treatment effects and on the random variations remaining, is likely to be a smooth trend due to the ageing of the material. It would often be of interest to estimate this trend explicitly, as well as to set up the experiment so that the trend has little or no influence on the estimates of the treatment effects.

It can be shown that if the treatments are used in the systematic order

$$T_{60} \quad T_{50} \quad T_{70} \quad T_{70} \quad T_{60} \quad T_{50} \quad T_{50} \quad T_{70} \quad T_{60}$$

then

(*a*) any linear trend in time has no effect on the treatment comparisons;

(*b*) any curvature (i.e., second-degree component) in the trend has no effect on the linear part of the treatment effect, that is, on the comparison of 70% relative humidity with 50% relative humidity;

(*c*) there is some mixing of the estimate of the curvature of the trend and of the curvature of the treatment response curve, i.e., of the mean response to 50% and 70% relative humidity minus that for 60% relative humidity. Statistical analysis is necessary to sort these two out.

To see what is meant by the first property imagine that the observations consisted of a pure linear trend, in which values $1, 2, 3, \ldots, 9$ are obtained. Then the mean observation on T_{50} is $\frac{1}{3}(2 + 6 + 7) = 5$, the mean observation on T_{60} is $\frac{1}{3}(1 + 5 + 9) = 5$, and the mean observation on T_{70} also is 5. That is, this particular linear trend, and in fact any linear trend, leaves the estimates of the treatment differences unaltered.

It can be shown that designs chosen to have these properties of balance with respect to the trend simplify, and maximize the efficiency of, the estimates of both trend and treatment effects.

Cox (1951, 1952) has set out the method for selecting a design and for analyzing it. A few examples are given in Table 14.1. Box and Hay (1953) have described an ingenious and flexible method of dealing with similar experiments in which the treatments correspond to a set of at least two quantitative factors. In such cases there is enough freedom in the choice of factor levels to allow sufficient randomization to be brought into the design. In the simpler situations considered by Cox, such as

that of Example 14.1, there is however no randomization, other than possibly in naming the treatments. This would be a defect in a moderate or large sized experiment, in which randomization can be relied on to ensure freedom from systematic error and a correct estimate of the residual error. In a single very small experiment, absence of randomization is much less serious, since we are in any case bound to rely to some extent

TABLE 14.1

SOME TREND-FREE DESIGNS

No. of Treatments	Units per Treatment	Degree of Trend	Design
2	3	2	$T_1\,T_2\,T_2\,T_1\,T_1\,T_2$
2	4	2	$T_1\,T_2\,T_2\,T_1\,T_2\,T_1\,T_1\,T_2$
3	3	2	(*a*)* $T_2\,T_1\,T_3\,T_3\,T_2\,T_1\,T_1\,T_3\,T_2$
			(*b*)* $T_2\,T_3\,T_1\,T_2\,T_1\,T_3\,T_2\,T_3\,T_1$
3	4	3	$T_1\,T_2\,T_3\,T_3\,T_2\,T_1\,T_1\,T_2\,T_3\,T_3\,T_2\,T_1$
4	3	2	$T_2\,T_3\,T_4\,T_1\,T_1\,T_4\,T_3\,T_2\,T_2\,T_3\,T_4\,T_1$
4	4	3	$T_1\,T_2\,T_3\,T_4\,T_4\,T_3\,T_2\,T_1\,T_1\,T_2\,T_3\,T_4\,T_4\,T_3\,T_2\,T_1$

* The design (*a*) is the one mentioned in Example 14.1 and should be used when the comparison of T_3 with T_1 is of particular interest, as when three equally spaced levels of a quantitative factor are involved. In other cases design (*b*) should be used.

on our prior knowledge of the uncontrolled variation when the number of experimental units is small. Provided that the assumption that the uncontrolled variation is formed from a trend plus random variation is sensible, and that the design adopted is unlikely to correspond to a pattern in the uncontrolled variation, there can be no reasonable objection to the absence of randomization.

Similar principles can be used to pick out systematic Latin squares that are, for example, such that the treatment effects are uninfluenced by particular patterns of diagonal variation across the square. This is useful when the rows and columns of the square correspond to quantitative factors between which there may be a linear × linear component of interaction.

14.3 OPTIMUM ALLOCATION

In this section an outline is given of recent work on a class of problems which, although not all concerned with comparative experiments of the

sort we have been discussing up to now, do concern the most efficient distribution of experimental effort. The situations can be described like this: observations may be made on a number of experimental set-ups and the quantities so obtained are known to be expressed statistically in terms of certain unknown parameters. Which set-ups should be observed in order to estimate the parameters with maximum precision?

Example 14.2. In certain experiments on alloys cast from high-purity metals, and in other fields too, the following situation arises. The experimental units are arranged in runs, corresponding to production runs, and within each run they are ordered in time. Thus with each run corresponding to a 20-lb melt, it might be possible to cast, in sequence, four 5-lb ingots. Suppose that the treatments for comparison are different concentrations of the alloying elements and that when an ingot has been cast, the concentration of an alloying element in the remaining melt can be increased by the addition of fresh material but cannot be decreased.

In other words, if we think of just one alloying element, the factor level cannot decrease as we go from unit to unit within a run. This is a severe restriction on the type of design that can be used, since it excludes, for example, any Latin square. For if T_1, T_2, T_3, and T_4 denote successively increasing concentrations of the alloying element, any design which secures that each concentration occurs equally frequently in each run and in each period is bound to involve reversals of order within some runs.

Hence, if it is desired to set up a design in which run and period effects are eliminated from the error of the treatment comparisons, something less direct than a Latin square has to be used. Suppose for simplicity that there is just one alloying element occurring at two levels A_{-1} and A_1. Then with four periods per run, a design must consist of a mixture of sequences of the following five types:

Type I	A_{-1}	A_{-1}	A_{-1}	A_{-1}
II	A_{-1}	A_{-1}	A_{-1}	A_1
III	A_{-1}	A_{-1}	A_1	A_1
IV	A_{-1}	A_1	A_1	A_1
V	A_1	A_1	A_1	A_1

The difference between treatments is mixed up both with run and with period differences, unless

(*a*) all the sequences are of Type III, in which case the difference between treatments is completely identified with differences between periods, or

(*b*) one-half the sequences are of Type I and one-half of Type V, when the difference between treatments is completely identified with differences between runs.

Therefore if it is desired to eliminate variations between periods and between runs, a mixture of sequences of the five types must be used.

We can now formulate the basic problem. For any design formed from a mixture of sequences, we can, in theory, adjust the estimated difference between treatments for the lack of balance with runs and periods, and we can also find the standard error of the resulting adjusted estimate. For a given total number of observations, which of the mixtures of sequences minimizes the standard

error, i.e., leads to an estimate of the treatment effect of maximum precision? This is a mathematical problem and the solution can be shown to be to have collections of eight production runs, and within each collection one run of Types I and V and two each of the other types, i.e.,

A_{-1}	A_{-1}	A_{-1}	A_{-1}
A_{-1}	A_{-1}	A_{-1}	A_1
A_{-1}	A_{-1}	A_{-1}	A_1
A_{-1}	A_{-1}	A_1	A_1
A_{-1}	A_{-1}	A_1	A_1
A_{-1}	A_1	A_1	A_1
A_{-1}	A_1	A_1	A_1
A_1	A_1	A_1	A_1

The mathematical discussion and the extension to designs with several factors have been given by Cox (1954).

The general comment to be made on this example is that if practical considerations severely restrict the arrangement of treatments, it will, if it is important to make the most economical use of the experimental material, be necessary to evaluate theoretically the standard error of the estimated treatment effects for all admissible arrangements and to choose the arrangement which minimizes the standard error. Of course in particular cases it may well be better to adopt a simpler, but less efficient, procedure. Thus in Example 14.2 it would be possible to hold the treatment constant within each run, i.e., to use a whole run as an experimental unit. This would avoid the complication of using a special design, but elimination of the effect of variation from run to run would, of course, no longer be possible.

The next two examples deal with a rather different type of situation where, although there are no alternative treatments under comparison, a theoretical calculation of precision is possible for different experimental set-ups.

Example 14.3. There are a number of situations in which it is desired to estimate the density of particles which are distributed randomly in some medium, for example dust particles in space, bacteria or blood cells in suspension, etc. The direct method is the counting of particles in known small volumes of medium, but this is often tedious. A neat method of avoiding direct counting is based on the following fact. Suppose that the particles are distributed completely randomly through the medium: this requires for example that the proportion of the volume of medium occupied by particles should be negligible, so that there is no "crowding," and that the particles should be uncharged, so that there the particles have no electrostatic effect on one another. Then the mean number of particles in a volume v of original medium is

$$-2.303 \log_{10} \text{ (proportion of a large number of volumes } v \text{ that contain no particles).} \qquad (1)$$

Thus if we take a large number of volumes v and observe for each whether or not it contains particles, the density of particles can be estimated. In some cases the determination of whether or not a particular volume contains particles is easy; with bacteria, however, it will be necessary to assume that growth occurs on any plate that contains at least one bacterium.

The problem of design is to decide what volume v of medium should be used in each trial, e.g., what dilution of the original suspension should be taken. If rather a large volume v is taken, nearly all the trial volumes will contain particles, and equation (1) will lead to an estimate with very low precision; the precision will also be low if the trial volume v is too small. It can be shown that maximum precision of the volume v is chosen to contain on the average 1.6 particles (Fisher, 1951, p. 219; Finney, 1952, p. 573).

This result is no use as it stands, since if we knew what volume contains on the average 1.6 particles, we should know the concentration of particles and the estimation that we are considering would be pointless. However it may be shown that very nearly maximum precision is attained if the mean number of particles in the test volume is between 1 and $2\frac{1}{2}$ and that reasonably high precision is retained if the number is between $\frac{1}{2}$ and $3\frac{1}{2}$. Thus if a prior estimate is available correct to within a factor of 2 or 3, the method can be used.

If no such prior value is known, two procedures are available. The one widely used is to make observations at each of a series of volumes, say $v_0, 2v_0, 4v_0, 8v_0, \ldots$ (series of volumes progressing by factors of 4 or 10 are also sometimes used), chosen to be sure of covering the optimum value. The disadvantage of this is that it will usually happen that the observations at several of the volumes give little information about the density under estimate. The second procedure is to do a small preliminary trial with a range of volumes in order to estimate the single volume at which the main series of observations should be made. Obviously this cannot be done if the whole experiment must be laid out at one time.

Example 14.4. Andrews and Chernoff (1955) have discussed the following problem connected with the estimation of the virulence of a strain of bacteria. There are available 30 test animals and 10 ml of material containing this strain of bacteria in suspension. It is thought that the concentration of bacterial organisms in this suspension is about four organisms per ml and that the probability that a dose of one organism applied to a test animal leads to a response is about 1/5. This latter probability is to be estimated more accurately.

Part of the suspension must be used for a plate count to estimate the concentration and the remainder allocated among the test animals to determine virulence.

To deal with this problem it is again necessary to set up a statistical model to represent the observations. Briefly this is that there is a certain unknown chance α that one organism administered to a test animal will lead to a positive response, whereas if several organisms are administered these act independently and no response is obtained if and only if no response would have been obtained with each organism separately. Further it is assumed that the numbers of organisms in a certain nominal dose have a particular random distribution called the Poisson distribution. From these assumptions it can be shown that the probability that a nominal dose of D organisms produces a negative response is $e^{-\alpha D}$, where e is the base of natural logarithms.

Now if the concentration of the suspension of organisms were known and if there were an unlimited supply of test animals and organisms, we should have a situation mathematically the same as that of Example 14.3 and the optimum dose would be $1.6/\alpha$, where the best initial estimate, 0.2, would be used, for α, indicating an optimum dose of 8 organisms. We have, however, the additional features described above, and these appreciably complicate the mathematical solution. Andrews and Chernoff show that optimum procedure is approximately the following:

(*a*) Every test animal receives the same dose.

(*b*) The fraction of the suspension given to the test animals is approximately the smaller of $1/(1 + \sqrt{\alpha})$ and $1.6s/(\alpha\lambda)$, where α is the best initial estimate of the probability that a single organism will lead to a positive response, s is the number of test animals, and λ is the best initial estimate of the total number of organisms in the available suspension. The initial estimates of α and λ do not have to be very precise.

For example, with the values of 0.2 and 40 for α and λ, and with $s = 30$, $1/(1 + \sqrt{\alpha}) \simeq 0.69$ and $1.6s/(\alpha\lambda) \simeq 6$, so that 69% of the suspension should be divided into 30 equal portions for application to the test animals and 31% of the suspension used for a plate count. The estimation of α from the resulting observations is straightforward.

Equivalent problems occur in estimating the unknown parameters in theoretical relationships. For example, in a diffusion problem we might be able to measure the concentration of diffusing solute at different distances x into the diffusing medium and after different times t from the start of the diffusion process. Theory gives a relation between concentration and x, t involving as unknown parameters the diffusion coefficient and a constant defining boundary conditions. The problem is to decide at what values of x and t to observe concentration in order to get estimates of the unknown parameters of maximum precision.

It is characteristic of these problems that a statistical model has to be assumed to represent the situation and that the conclusions about the optimum arrangement would be wrong if the model were seriously wrong. For instance, in Example 14.4 the idea that all animals should receive the same nominal dose would certainly not be acceptable if there were a serious doubt about the formula $e^{-\alpha D}$ for the chance that a nominal dose D produces a negative response. We can sometimes cover this possibility, however, by including additional parameters in the model to represent departure from the form first assumed. A second characteristic feature is that, except in especially fortunate cases, initial estimates have to be available for one or more of the quantities in the experiment. In the example just discussed such estimates are needed for the quantities α and λ. It is a matter for investigation in each case how precise these initial estimates need to be; if values of sufficient precision are not available, a small part of the experimental material may be used to obtain rough

estimates and from these a suitable method of using the remaining material then determined. Elfving (1952) and Chernoff (1953) have given general mathematical discussions of these problems, assuming that any necessary prior estimates are available. The details of special cases are liable to be complicated.

14.4 SEARCH FOR OPTIMUM CONDITIONS

In the designs discussed in previous chapters it has been assumed that the object is the estimation of the differences between alternative treatments. Quite often, however, particularly in technological experiments, the ultimate object is the selection of the treatment or set of treatments which are in some sense the best. It may, however, still be necessary to estimate all relevant treatment effects, both in order to get added understanding of the system being investigated and also because the criterion determining the optimum conditions may be rather imprecisely defined. Thus if there are several treatments which differ only slightly with respect to, say, total cost per unit yield, we may decide to use the process which is best by some other standard, e.g., estimated long-run reliability (assuming that no allowance for this has been included in the calculation of cost). To be able to do this satisfactorily, it will be necessary to estimate, not just which is the best treatment according to the different criteria, but also the amounts by which other treatments depart from the optimum. Even in such cases, however, the need does arise for designs that will select a group of treatments, or range of experimental conditions, for fuller investigation.

Therefore it is of interest to examine procedures where the emphasis is on the estimation of optimum conditions, rather than on the estimation of treatment differences. First suppose that there is one quantitative factor v that can be varied. To determine the value of v for which a response y is maximized (or minimized) it will usually be best to proceed in two stages. In the first, a rough determination is made of the region within which the maximum lies. This can be done either by setting out an experiment with, say, about six equally spaced levels covering the range of interest or by proceeding in steps. In one form of the latter method, observations are taken at two levels v_0 and $v_0 + \Delta$; if the first level gives higher response than the second, observations are continued at a level $v_0 - \Delta$; if the second gives the higher response observations are taken at $v_0 + \Delta$ whereas if both levels give about the same response, further observations are taken at both $v_0 - \Delta$ and $v_0 + \Delta$. This procedure is continued, the aim being at each step to shift the treatment into a region in which higher response is obtained. There are clearly many ways in which such a procedure can be formalized, but it is probably best

to leave scope for special considerations, such as the ease with which the treatment may be changed, the amount of random variation relative to the slope of the response curve, and so on.

When the general form of the response curve has been established, the second stage of the experiment consists of a three-level (or possibly four-level) experiment centered on the suspected position of the maximum, the lower and upper levels being chosen as far apart as possible, subject to the response curve being reasonably parabolic in the region covered by the three levels. A second-degree equation (§ 6.8) is then fitted to the results of the second stage and any relevant results from the first stage, and the maximum of the fitted equation determined by plotting or by differentiation. A more detailed discussion is given by Hotelling (1941).

It is assumed here that the response curve is approximately of a special mathematical form within the range investigated in the second stage. Kiefer and Wolfowitz (1952) have discussed a very interesting procedure which requires only weak assumptions about the response curve; however in most practical cases it seems likely that more precise estimates of the position of the maximum can be obtained by judicious use of the parabolic approximation.

Box and Wilson (1951), see also Davies (1954), have suggested a procedure for use when there are several quantitative factors, $v_1, v_2, \ldots,$ that can be varied independently. Their idea is first to determine, by a two-level factorial experiment, the direction near the starting point in which the response surface rises most steeply. For example in Fig. 14.1,

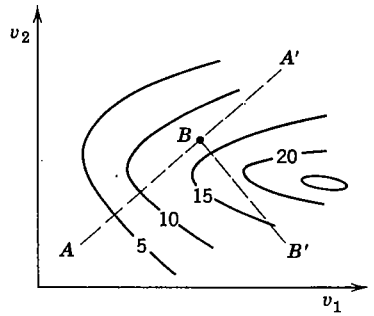

Fig. 14.1. Paths of steepest ascent.

if it is decided to start near A, an experiment will show that changes in the direction AA' produce maximum increments in yield. Since this line is at about 45° to the axes representing the two factors, equal changes in v_1 and in v_2 should be made, and this is done in the next stage, the optimum

position B along the line AA' being determined. A further two-level experiment centered at B determines the line BB' of steepest ascent from B and so on. When the optimum has been determined approximately, a three-level experiment is used to investigate the shape of the response surface near the optimum. A difficulty is that the direction of the line of steepest ascent depends on the units in which v_1 and v_2 are measured. A full account, with examples, of this method is given by Box and Wilson and by Davies.

When the different treatments are qualitative, the problems of design are different, although stage-by-stage investigation is again likely to be useful. Apart from the particular problems of selection in genetics (Cochran, 1951), little work seems to have been done on experimental designs for such situations, although there have been a number of theoretical investigations of so-called decision rules for selecting best treatments.

14.5 ASSAYS

An assay is a system of observations intended to give a number measuring a particular property of some experimental material, for example the potency of a drug, the strength of an insecticide, the mean fiber diameter of a consignment of wool, and so on. All such measurement involves in the last analysis comparison with some sort of standard. We can, however, for practical purposes distinguish between assays in which (*a*) a standard material similar to the experimental material is used and the observations on it compared with the observations on the experimental material and (*b*) no such standard material is used directly. We have already touched on this distinction in the discussion of Example 1.4.

For instance, in early attempts at standardizing insulin, potency was measured directly by the amount necessary to produce a certain response in mice, giving "animal units" of potency. This is the second type of assay and was found to be unsatisfactory because of variations between different batches of mice and between different times. In particular the comparison of different workers' results was unreliable. The introduction of international standard preparations against which experimental preparations could be tested enabled this difficulty to be avoided; in each trial, some mice receive the standard, others the experimental drug. The set-up then has the form of a comparative experiment of the type we have been discussing in this book, the object being to assess the unknown by direct comparison with the standard. Variations between batches of mice affect unknown and standard almost equally, and reproducible measures of potency are thereby obtained.

As another example, consider the measurement of the mean fiber

diameter of wool. One quick method is to form, in a controlled way, a plug of fibers and to measure the rate of flow of air through the plug when a fixed pressure drop is applied across it. From the rate of flow, the pressure drop, and the mass of wool in the plug, a quantity Q can be calculated that is closely correlated with the mean fiber diameter. In practice a value for estimated mean fiber diameter is derived from Q by the use of a calibration curve, obtained by testing plugs made of fibers of known mean diameter, determined by a more laborious optical method. This is an example of the second type of assay with no direct use of standards. That is, the standards, the fibers of known mean diameter, are used to construct a calibration curve, but this curve is assumed to remain fixed and the standards are not introduced directly in each individual determination. The calibration curve has in fact been shown to be quite constant and this method is satisfactory.

If this were not so and the relation between Q and mean fiber diameter tended to shift in time, the first type of assay could be used. A series of standard lots of wool would be taken each of known mean fiber diameter, the series covering a range of fiber diameters. To test a new batch, a group of two or three standards would be selected whose diameters are likely to straddle the diameter of the new batch. The experimental quantity Q would be determined for all, and the diameter of the experimental batch then estimated by an obvious graphical or statistical technique.

These examples are typical of a wide range of procedures used both in the physical and the biological sciences. Physical applications tend on the whole to be like the second example, where a fixed relation can be assumed between the quantity measured experimentally and the quantity it is required to estimate, or where the observation obtained can be used directly to measure the property of interest. The use of standards is restricted to initial calibration and occasional recalibrations. Biological applications tend on the whole to be like the first example, where it is desirable to introduce a standard explicitly into the determination. A very thorough and authoritative account of the statistical problems of design and analysis connected with this sort of assay has been given by Finney (1952), and only a few salient points will be mentioned here.

The assay of insulin mentioned above is *direct* in that the observation for each experimental unit is the quantity of drug, unknown or standard, necessary to produce a certain response. No new problems of design are involved in such an assay, the methods for the comparison of two or more alternative treatments being used. Most bioassays, however, are *indirect*, in that for one reason or another it is necessary to apply a

predetermined dose to each experimental unit and to observe the resulting response, quantitative or qualitative, not to increase gradually the dose until a fixed response is obtained. This raises some new problems.

The simplest design for indirect assays is the so-called symmetrical four-point assay. Two concentrations of the standard are used, one, say, λ times the other, and two concentrations of the unknown, the one also λ times the other. The concentrations should be adjusted so that the observations on corresponding doses of standard and unknown are expected to be about the same; prior knowledge is, of course, necessary to do this. The ratio λ between the two concentrations should be chosen to be as large as possible, subject to the requirement that the relation between observation and log dose should be linear over the whole range of concentrations of both drugs in the assay. (Considerable prior investigation of the dose-response is assumed to have been done in setting up the method in the first place.)

An experiment is now arranged to compare these four treatments, if possible increasing precision by some of the techniques discussed in previous chapters. This experiment will in effect be a 2×2 factorial experiment, one factor being the type of drug, the other the level of dose, high or low. From the results the relative potencies of unknown and standard can be estimated quite easily.

Notice that we are interested not in the treatment effects in their own right, but in using the treatment effects to estimate a special type of relation existing between the treatments. The assumptions on which the assay is based are two. First the relation between response and log-dosage must be linear over the range used; this cannot be checked from the data. Secondly the response curves for the two preparations must be parallel if the hypothesis that the test preparation is equivalent to an unknown dilution of the standard is to be maintained. This can be checked from the data.

If little reliable information is available about the shape of the response curve, a six-point assay will be suitable, in which each preparation occurs at three levels, equally spaced on the log-dosage scale. From the results, tests of both linearity and parallelism are possible; this corresponds to an ordinary factorial experiment with a quantitative factor, in which three levels are necessary to get an estimate of the curvature of the response curve. More than three levels will not normally be advisable in routine work.

Since the final experiment is set out as a comparative trial, all the methods discussed in earlier chapters, incomplete block techniques, confounding, cross-over designs, etc. are from time to time found useful in designing assays. Finney's book should be consulted for details and examples, as well as for careful discussion of the various special complications that arise.

SUMMARY

A number of unrelated special topics have been discussed briefly:

(*a*) Trend-free systematic designs are available that are sometimes suitable for small experiments in which a smooth trend is superimposed on the treatment effects;

(*b*) the optimum allocation of observations can sometimes be determined when a theoretical calculation of precision is possible for each of a number of possible set-ups, any of which may be observed;

(*c*) special methods are available for determining optimum conditions, i.e., the experimental conditions under which a suitable quantity is maximized;

(*d*) a few of the problems connected with assays, particularly bioassays, have been outlined.

REFERENCES

Andrews, F. C., and H. Chernoff. (1955). A large-sample bioassay design with random doses and uncertain concentration. *Biometrika*, **42**, 307.

Box, G. E. P., and W. A. Hay. (1953). A statistical design for the efficient removal of trends occurring in a comparative experiment, with an application in biological assay. *Biometrics*, **9**, 304.

——— and K. B. Wilson. (1951). On the experimental attainment of optimum conditions. *J. R. Statist. Soc.* B, **13**, 1.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Statist.*, **24**, 586.

Cochran, W. G. (1951). Improvement by selection. *Proc. 2nd. Berkeley Symp. on Math. Statist. and Prob.*, 449. Berkeley: University of California Press.

Cox, D. R. (1951). Some systematic experimental designs. *Biometrika*, **38**, 312.

——— (1952). Some recent work on systematic experimental designs. *J. R. Statist. Soc.* B, **14**, 211.

——— (1954). The design of an experiment in which certain treatment arrangements are inadmissible. *Biometrika*, **41**, 287.

Davies, O. L. (editor) (1954). *Design and analysis of industrial experiments.* Edinburgh: Oliver and Boyd.

Elfving, G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Statist.*, **23**, 255.

Finney, D. J. (1952). *Statistical method in biological assay.* London: Griffin.

Fisher, R. A. (1951). *Design of experiments.* 6th ed. Edinburgh: Oliver and Boyd.

Hotelling, H. (1941). Experimental determination of the maximum of a function. *Ann. Math. Statist.*, **12**, 20.

Kiefer, J., and J. Wolfowitz. (1952). Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.*, **23**, 462.