Biometrika (1975), **62**, 2, *p.* 251
Printed in Great Britain

Sufficiency and conditionality

BY JOHN D. KALBFLEISCH
Department of Statistics, University of Waterloo, Ontario

SUMMARY

Ancillary statistics are divided into two logically distinct types: those determined by the experimental design and those determined by the mathematical modelling of the problem. It is pointed out that, in the class of inference problems where our purpose is to gain information or insight into the nature of a chance set-up, a weakened conditionality principle when applied first removes the possibility of deriving the likelihood principle. Since to some extent a conditionality principle must be applied in experiment definition, it is argued that this is a necessary first step if full acceptance of the likelihood axiom is to be avoided.

Some key words: Ancillary statistics; Conditionality; Likelihood; Principles of inference; Sufficiency.

1. INTRODUCTION

Birnbaum (1962) considered a concept of 'the evidential meaning of an experiment E with outcome x' which is written $\mathrm{Ev}(E,x)$. On the basis of this undefined concept, he formulated principles of parametric inference. The main result in that paper was that the sufficiency principle, S, and the conditionality principle, C, that Birnbaum formulated were shown to imply the likelihood principle, L. This aroused considerable interest since many who find the likelihood principle unattractive do find appealing some type of sufficiency principle and some type of conditionality principle. Durbin (1970) has pointed out that if the conditionality principle is applied to the derived experiment in which only the minimal sufficient statistic is observed, the likelihood principle no longer follows. This approach would employ S to reduce to the minimal sufficient statistic and then apply a conditionality principle, C_S. This ordering has considerable appeal but suffers some drawbacks as pointed out by Savage (1970) and Birnbaum (1970). In this paper, an alternative approach is suggested which appears closer to the operational approaches of both the Fisherian and the Neyman–Pearson schools of statistical inference.

A simple example serves to illustrate the implications of the two approaches mentioned above. Suppose that E_1 and E_2 are two experiments characterized by negative binomial sampling to one success and by binomial sampling with $n = 5$. In both cases the purpose is to make inferential statements about the probability of success. If Y is the number of failures in E_1, and X the number of successes in E_2, then S and C together would imply that the outcome $Y = 4$ in E_1 and the outcome $X = 1$ in E_2 have the same evidential meaning; that is

$$\mathrm{Ev}(E_1, Y = 4) = \mathrm{Ev}(E_2, X = 1).$$

On the other hand, the modified sequence, S and C_S, would allow for no such conclusion. But if there were a chance mechanism which led to performing E_1 with probability α and

E_2 with probability $(1-\alpha)$ for known α $(0 < \alpha < 1)$, the experiment under consideration is now $E = \alpha E_1 + (1-\alpha) E_2$. Since in this experiment the two outcomes 'performing E_1 with result $Y = 4$' and 'performing E_2 with result $X = 1$' give identical likelihood functions, they lie in the same isostat of the minimal sufficient partition. It therefore follows from S and C_S that they have equivalent evidential meaning; that is,

$$\mathrm{Ev}\{E, (E_1, Y = 4)\} = \mathrm{Ev}\{E, (E_2, X = 1)\}.$$

In fact this conclusion follows from the sufficiency principle alone.

This connexion between the likelihood principle and the sufficiency principle was noted by Barnard, Jenkins & Winsten (1962) in their arguments favouring the likelihood principle. Fraser (1963) also notes this relationship and suggests that his arguments against the likelihood principle apply equally to the sufficiency principle. Criticisms of L are often based on examples similar to the one just considered. The critic defines two experiments E and E' which have possible outcomes x and x' yielding the same likelihood function for the parameters so that L implies $\mathrm{Ev}(E, x) = \mathrm{Ev}(E', x')$. He then argues against this conclusion on intuitive or operational grounds; see, for example, Fraser (1963) and Armitage (1963). But, since merely by defining a mixture experiment the sufficiency principle alone leads to similar conclusions, it would appear that intuitive and operational arguments against L can be made to apply to S with equal force. In order to reject L and accept S one must attach great importance to the possibility of choosing randomly between E and E'.

In what follows, we consider a subset of inference problems where our concern lies with the determination and description of a chance set-up M. Accordingly, we conduct an experiment which has as an integral part of it the collection of information about M. In §3, the notion of a chance set-up will be formalized, but for the present the intuitive meaning of the term will be sufficient. Within this subset of problems, a sequence of principles is constructed which leads to an approach very closely in line with the operational approach of Fisher to problems of this type.

2. AN ALTERNATIVE APPROACH

Birnbaum (1962) defines a mixture as follows:

> An experiment E is called a mixture (or a mixture experiment) with components $\{E_h\}$ if it is mathematically equivalent (under relabelling of the sample points) to a two stage experiment of the following form:
>
> (a) An observation h is taken on a random variable H having a fixed and known distribution G. (G does not depend on the unknown parameter.)
>
> (b) The corresponding component experiment E_h is carried out yielding an outcome x_h.

The idea that we wish to develop in this paper is a very simple one, though difficult to formalize. It is merely to recognize that there are two logically distinct types of mixtures: the first of these is a mixture as a result of the experimental design and the chance set-up under study and is called an experimental mixture; the second, a mathematical mixture, is a mixture experiment as a consequence of a class of models hypothesized for the chance mechanism, e.g. a parametric family. This distinction is very similar in spirit to the one made by Basu (1964) in considering real and conceptual, or performable and nonperformable, experiments. In §3, definitions of the two types of mixtures remove at least some of the subjectivity inherent in Basu's considerations. As a matter of terminology, an ancillary

H which determines an experimental mixture is called an experimental ancillary; correspondingly a mathematical mixture determines a mathematical ancillary.

An analogous situation arises in the distinction which can be made between the notion of independent experiments and the notion of independent events. The first of these is a result of the physical independence of two experiments which are performed. If A is an event whose occurrence or nonoccurrence is determined by the outcome of the first experiment and B an event whose occurrence or nonoccurrence is determined by the second experiment, then $\text{pr}(AB) = \text{pr}(A)\,\text{pr}(B)$ regardless of what probability measures are assigned to the two sample spaces. However, the independence or lack of independence of two events C and D in the same experiment will depend on the probability measure assigned.

A few examples will illustrate the distinction between the two types of mixtures or ancillaries.

Example 1. Suppose that X is an integer valued random variable with probability function $f(x)$ and an experiment is designed in order to gain information about the function $f(x)$. The experiment E consists of tossing an unbiased coin to decide between E_1 and E_2. The experiment E_1 involves taking two observations on the random variable X while E_2 involves selecting only one observation on X. Clearly the experiment E is an experimental mixture with components E_1 and E_2. The random variable

$$H = \begin{cases} 1 & \text{if the toss gives a head,} \\ 2 & \text{if the toss gives a tail,} \end{cases}$$

is an experimental ancillary; its distribution is fixed regardless of what assumptions are made about the form of $f(x)$. Suppose, however, that the parametric form of $f(x)$ is assumed to be

$$f(x;\mu) = \begin{cases} \dfrac{3}{\pi^2(x-\mu)^2} & (x \neq \mu), \\ 0 & (x = \mu), \end{cases} \tag{1}$$

where μ is some unknown integer. If X_1 and X_2 are the random variables associated with the first and second outcomes in E_1, then $X_1 - X_2 = A$ is an ancillary, its distribution being free of μ. Thus the experiment E_1 is a mathematical mixture of an infinite number of experiments indexed by the random variable A. Note, however, that A is a mathematical ancillary; E_1 is not an experimental mixture of these experiments. The mathematical mixture is a result purely of the model (1) for the probability function of X.

Example 2. Consider the following experiment to evaluate the effectiveness of cloud seeding in producing rain. On some specific day each cloud in the sky at twelve o'clock is to be seeded or not seeded with equal probabilities. It is recorded whether the cloud produces rain or not by 4.00 p.m.

Clearly, we are interested in the proportion of clouds which produce rain in each of the seeded and unseeded groups. The number of experimental units, N, is itself a random variable with probability distribution which is likely not known. Then N is an experimental ancillary, its distribution not being dependent on the probability measure assigned to the particular phenomenon under question. It is also apparent that given $N = n$, the n-tuple with ith entry 1 if the ith cloud is seeded and 0 if the ith cloud is not seeded is an experimental ancillary, the probability of any such sequence being 2^{-n}. Thus, the experiment is an experimental mixture of a large number of possible experiments, each of the sub-experiments being indexed by a number n and an n-tuple of zeros and ones.

Example 3. Suppose that an experiment is being performed to determine the regression of Y on X when neither variable is being controlled. Accordingly a sequence of n independent observations on X and Y are taken. In this case, the random variables $X_1, \ldots, X_n$ are experimental ancillaries and the experiment is an experimental mixture with components indexed by the various possible values of $X_1, \ldots, X_n$. If it is assumed that the regression of Y on X is linear with normal errors, the standardized residuals are mathematical ancillaries, their distribution being completely known as a result of the assumed normal parameterization.

This same experiment might be performed in order to gain information about the joint distribution of X and Y. In this case, however, the random variables $X_1, \ldots, X_n$ are not experimental ancillaries. Thus, what is and what is not an experimental ancillary will depend on the purpose of the experiment being performed. In order to identify what the experimental ancillaries are, one must be careful in defining the question that is being posed.

Example 4. For a final example, suppose we are interested in evaluating a physical constant β and that we have at our disposal a measuring intrument whose properties are completely known to us. The instrument is not exact but attaches $N(0,1)$ errors to β, so that we observe $x = \beta + \epsilon$, where $\epsilon \sim N(0,1)$. In this case, when n measurements $x_1, \ldots, x_n$ are taken, the residuals $(x_i - \bar{x})$ $(i = 1, \ldots, n)$ are experimental ancillaries, their distribution being determined by the choice of the measuring instrument independently of any assumptions we might make about β.

There is a superficial similarity between Example 1 and Example 4. In Example 1, the chance mechanism under study is that which generates x. The parametric family of probability functions was introduced as an assumption about the chance mechanism under study and consequently, the ancillaries are mathematical ancillaries. In Example 4, however, the ancillaries arise as a consequence of the choice of the measuring instrument and not from assumptions about any chance mechanism under study. As such, they are experimental ancillaries.

The measurement model and its generalization to the structural model has been considered by Fraser (1968, pp. 21, 49). The distinction made above between Examples 1 and 4 is closely allied with the distinction between the structural and the 'classical' model that Fraser makes.

In §3, we shall attempt to formalize the notion of an experimental ancillary and an experimental mixture. But as the above examples illustrate, in most practical situations the distinction between an experimental and mathematical ancillary is clear.

Since an experimental mixture is a result, not of the properties of the chance mechanism under study, but rather of impositions of the experimenter or consequences of random mechanisms external to the one under study, it seems reasonable that, in interpreting the data, one should first condition on the observed values of the experimental ancillary variates. Not to do so will, in fact, leave arbitrary the starting point of an experiment. In formalizing this idea, one arrives at the following principle of experimental conditionality.

PRINCIPLE OF EXPERIMENTAL CONDITIONALITY, C_E. *If an experiment E, designed to gain information on a chance set-up M, is an experimental mixture with component experiments $\{E_h\}$ and possible outcomes (E_h, x_h), then*

$$\mathrm{Ev}\{E, (E_h, x_h)\} = \mathrm{Ev}\,(E_h, x_h).$$

This principle implies that in interpreting evidential meaning, one should condition on the particular experiment physically performed. In the class of problems considered, this would imply that the random components of an experimental design play no direct role in interpreting the evidential meaning of a particular observed sample. It should be stressed, however, that this does not preclude the possibility that randomization incorporated in an experimental design may tend to safeguard the distributional assumptions made. Of course, the use of the randomization distribution for inference in problems where our purpose is not to identify a chance mechanism is not contraindicated by this principle. Fisher's (1966, pp. 10 *et seq.*) example of the lady tasting tea is a classic example of a problem falling outside the present framework.

If in a particular experiment one can obtain a maximal experimental ancillary, as is often the case, this principle would suggest that the first step is to reduce the reference set from the total experiment E to the experiment E_h which was actually physically performed. Then E_h is an experiment which is not an experimental mixture of any experiment except, trivially, of itself. Such an experiment will be called a minimal experiment. It should be noted that a minimal experiment is minimal with respect to the chance set-up under study; see Example 3. Further principles of inference could then be applied to the minimal experiment. This suggests the following revised principle of sufficiency.

The Principle of Sufficiency, S'. *Let E be any minimal experiment, with sample space $\{x\}$ and let $t(x)$ be any sufficient statistic. Let E' denote the derived experiment having the same parameter space, such that when any outcome x of E is observed the corresponding outcome $t = t(x)$ of E' is observed. Then for each x, $\mathrm{Ev}\,(E, x) = \mathrm{Ev}\,(E', t)$, where $t = t(x)$.*

As Durbin (1970) notes, this implies that the evidential meaning depends only on the observed value of the minimal sufficient statistic.

It is easily seen that because of the reduction to the minimal experiment suggested by C_E, the likelihood principle does not follow by the usual arguments from C_E and S'. In fact, Birnbaum's (1962, 1972) argument depends on the construction of an experiment E which is an experimental mixture of two experiments E' and E''. The revised sufficiency principle, S', would not apply to the mixture experiment E.

Following Durbin's (1970) approach further, one can define a modified principle.

Modified Principle of Mathematical Conditionality, C'_M. *If a minimal experiment E is mathematically equivalent to a mixture G of components E_h, with possible outcomes (x_h, E_h), where h depends only on the value of the minimal sufficient statistic, then*

$$\mathrm{Ev}\{E, (E_h, x_h)\} = \mathrm{Ev}\,(E_h, x_h).$$

The approach would then involve successive reductions and simplifications of the reference set and the data. We first reduce to the minimal experiment by conditioning on the maximal experimental ancillary, assuming that this exists. Given the minimal experiment one obtains from it the derived experiment which consists of observing the minimal sufficient statistic. The conditionality principle, C'_M, is then applied to this derived experiment.

The choice of an appropriate mathematical ancillary variate has been discussed by Barnard & Sprott (1971) and by Cox (1971). As yet, no completely satisfactory theory of choosing ancillary variates has been developed, since Cox's approach suffers from a degree of arbitrariness while that of Barnard & Sprott makes use of the unclear concept of shape. Nonetheless, in particular problems both of these approaches are practically quite sound and lead to the definition of sensible reference sets.

Since the classical approaches to problems of statistical inference involve the computing of probabilities over the sample space, they require a clear definition of the experiment being performed. It might be argued in this case that in defining any experiment, one is always conditioning on experimental ancillaries. The principle, C_E, says that in defining the experiment one should condition on all experimental ancillaries and state the experiment actually done. It therefore gives a starting point to which further principles of inference can be applied. The particular sequence of principles, C_E, S' and C'_M was chosen since this appears closely in conformity with the approach adopted by Fisher.

The continual emergence of the likelihood principle as the end product of principles of inference is quite possibly due to the fact that a principle like C_E has not first been applied. Without this principle, there will always be arbitrariness in the definition of an experiment and this arbitrariness itself would suggest that only a principle like L which makes no use of the experiment definition could be satisfactory.

3. Experimental mixtures

Basu (1964) makes a distinction between real or performable experiments and conceptual or nonperformable experiments. He leaves this distinction largely undefined but discusses examples and indicates his own opinion on the nature of various ancillary variates. Apparently he viewed the distinction as being largely subjective as is indicated by his discussion; see, for example, his footnote 2 on p. 13. In this section, an attempt is made to define in a formal manner the concept of an experimental mixture. This can to some extent be viewed as a formalization of Basu's original idea of performability; his terminology has not been used, however, since performability is not characteristic of the type of distinction we are making. The determination of the regression of Y on X in Example 3 takes $X_1, \ldots, X_n$ as experimental ancillaries, but these would not necessarily specify a 'performable' experiment.

The logical distinction between types of ancillary statistics is a difficult one to formalize completely. As with other basic concepts in statistics such as 'independent repetitions of an experiment', there are difficulties in defining experimental ancillaries in purely mathematical terms. From the examples above, the identification of experimental ancillaries requires a clear statement of the chance set-up being determined and of the experiment being performed.

Some progress can be made, however, in formalizing the distinction. We can view the chance set-up under study as being defined by a probability space $\{S, F, P\}$, where S is the sample space and F is the σ-field of events on which P is defined. We shall assume that S and F are completely specified and our problem reduces to determining P which would then completely identify the chance set-up. To begin with, we shall introduce no assumptions about P but merely assume that it is one of the class of probability measures P_F which can be defined on F.

In collecting information about the chance set-up an experiment E is performed which involves as an integral part of it the collection of data either functionally or stochastically related to outcomes of the chance set-up. Corresponding to E, there is a probability space $\{\mathscr{S}, \mathscr{F}, \mathscr{Q}\}$ and the probability measure $\mathscr{Q}$ is to some extent a function of the measure P of interest, but will in general depend also on the experimental design.

We shall call an observable random variate H an experimental ancillary in the experiment E designed to gain information on the chance set-up $\{S, F, P\}$ if H has the same distribution G for all choices of $P \in P_F$.

In a statistical analysis of the data, we typically introduce additional assumptions which restrict the class of measures from which P is obtained. For example we might assume $P \in \{P_\theta \mid \theta \in \Lambda\}$, where θ is a parameter, possibly of infinite dimension, which indexes the class of measures being considered. Such an assumption is made, for example, when the result of the chance mechanism is assumed to have a binomial distribution or when the class of probability measures is restricted to the class of continuous measures. We shall call the restricted class of probability measures P_θ.

A mathematical ancillary is defined with reference to the restricted class. An observable random variable H is a mathematical ancillary if it has a fixed distribution G for all choices of $P \in P_\theta$.

In general, the probability space associated with E will be quite complex in its makeup. The probability measure $\mathscr{Q}$ defined on the sample space of E will to some extent depend on the measure P of interest, but there will be other components which will depend on the experimental design. Some of these components will be at the experimenter's control, as for example the random choice between experiments in Examples 1 and 2 or the choice of measuring instrument in Example 4. Still other components of $\mathscr{Q}$ will arise from sources of variation which are not at the experimenter's control; this may for example determine the number of available experimental units as in Example 2.

From these considerations, we are led to the following definition of an experimental mixture.

An experiment E, designed to yield information about a chance set-up with probability space $\{S, F, P\}$, is an experimental mixture with components $\{E_h\}$ provided the following conditions hold. No matter what probability measure P in P_F is assumed, E is equivalent to a two-stage experiment:

(i) An observation is taken on a random variable H with a probability distribution G the same for all P, although possibly unknown.

(ii) The corresponding experiment E_h is carried out with outcome x_h.

It should be stressed that what is and what is not an experimental mixture is dependent on the experiment actually performed and on the purpose of that experiment. Accordingly C_E is not a universal principle like Birnbaum's S and C which may be applied automatically. The applicability of C_E must be judged anew with each new experimental situation.

One novel point of C_E is that the mixing distribution may be unknown. Such is the case in the cloud seeding example above. The definition requires, however, that the distribution of H must not be functionally dependent on the distribution P defined for the chance mechanism.

While the distinction between mathematical ancillaries and experimental ancillaries is clear in most examples, the measurement model is different and of some interest. In Example 4, the chance mechanism under study can be viewed as the process which determines the constant β. This can be formalized as a probability space $\{R, F, P\}$, where F is the Borel field on the real line R. The class P_F is the class of all measures that can be assigned to F. If n measurements are made of β using the instrument described, the residuals

$$(x_i - \bar{x}) \quad (i = 1, \ldots, n)$$

have a distribution which does not depend on the choice of $P \in P_F$. They are, therefore, experimental ancillaries. Subsequent statistical analysis might involve the choice of a family of priors for β or restriction of the class of measures to be considered to those measures which attach probability one to some $a \in R$.

4. MINIMAL EXPERIMENTS AND MAXIMAL EXPERIMENTAL ANCILLARIES

Many experiments involve choosing by some random procedure one of several subexperiments, each of which is composed of taking a fixed number of repeated observations on the chance set-up. By conditioning on experimental ancillary variates, one can in this situation always reduce to one of these experiments. We show here that this is a maximal reduction since such a subexperiment is minimal.

Let E' be an experiment which consists of observing a single realization of the chance set-up $\{S, F, P\}$.

LEMMA. *If E consists of a fixed number n of independent repetitions of the experiment E', then E is a minimal experiment.*

Proof. Let $X_1, \ldots, X_n$ be the independent outcomes arising from the repeated experiments E'. If $H(X_1, \ldots, X_n)$ is an experimental ancillary, its distribution by definition is the same for all choices of the measure P. But, if P is chosen to attach probability one to outcome x in S and zero to any other outcome, then H takes on the value $H(x, x, \ldots, x) = h$ with probability one. Thus H is trivial and E is a minimal experiment.

Other results on minimal experiments are difficult to obtain. One case which also arises very frequently is the case of repetitions of the experiment E' subject to some stopping rule which depends, at least to some extent, on the outcomes of the previous trials. Such experiments are often minimal, but there are situations where this is not the case. It is usually easy to verify that a particular experiment is minimal, but general results are difficult due to difficulties in defining general experimental situations.

5. OTHER PRINCIPLES

In §2, the sequence of principles C_E, S' and C'_M was suggested and it was pointed out that this was in close conformity with Fisher's approach to inferential problems. In this section, the possibility of incorporating a censoring principle into this scheme is examined. Again C_E is applied first and its introduction prohibits the deduction of the likelihood axiom as before.

The Principle of Irrelevant Censoring defined below was implicitly introduced by Pratt (1962) and formalized by Birnbaum in an unpublished 1964 technical report. In words this principle says that a censoring structure which might have affected, but in fact did not affect, the outcome of the experiment is irrelevant in the interpretation of the evidential meaning. This principle can be introduced after C_E as follows.

PRINCIPLE OF IRRELEVANT CENSORING, C_e. *Suppose E is a minimal experiment with sample space S and that E^* is a censored version of E with sample space S^* such that certain of the points in S cannot be observed. If x is an outcome in both S and S^*, then*

$$\mathrm{Ev}(E, x) = \mathrm{Ev}(E^*, x).$$

It is incidentally easily shown that if E is minimal then E^* is minimal.

This principle can be used to redefine the reference set of the minimal experiment so that irrelevant censoring is ignored in subsequent procedures. To incorporate such a principle, one should apply C_E, C_e, S' and C'_M in sequence. The domain of definition of S' requires alteration to the uncensored or less censored minimal experiment suggested by C_e.

A simple example serves to illustrate this sequence. Suppose that the experiment E^* consists of observing three survival times subject to censoring at 5, 10, 15 time units. The outcome observed is 4, 10^+, 12 where the + indicates a censored data point. Then E^* is minimal and C_e implies that the evidential meaning of this outcome in E^* is the same as if it had come from an experiment E censored only for the second survival time. If the survivals are assumed to be exponential with unknown failure rate θ, S' applied to this revised experiment E gives the sum of the first and third survival times along with an indicator variable for the second as minimally sufficient.

Of course many different extensions to less censored experiments could be made. In some sense, we want the extension to the simplest sample space that we can obtain. In the above example, E is an experiment which gives rise a minimal sufficient set for θ with easily deduced probability distributions.

I would like to thank Professor A. Birnbaum and J. G. Kalbfleisch for making very helpful suggestions on an earlier version and also Professors G. A. Barnard, V. P. Godambe, D. A. Sprott and M. E. Thompson for helpful comments and criticisms. This investigation was supported by the National Cancer Institute while the author was at the State University of New York at Buffalo.

References

Armitage, P. (1963). Sequential medical trials: Some comments on F. J. Anscombe's Paper. *J. Am. Statist. Assoc.* **58**, 384–7.

Barnard, G. A., Jenkins, G. M. & Winsten, C. B. (1962). Likelihood inference and time series. *J. R. Statist. Soc.* A **125**, 321–72.

Barnard, G. A. & Sprott, D. A. (1971). A note on Basu's examples of anomalous ancillary statistics. In *Waterloo Symposium on Foundations of Statistical Inference*, pp. 176–96. Toronto: Holt, Rhinehart and Winston.

Basu, D. (1964). Recovery of ancillary information. *Sankhyā* A **26**, 3–16.

Birnbaum, A. (1962). On the foundations of statistical inference. *J. Am. Statist. Assoc.* **57**, 269–306.

Birnbaum, A. (1970). On Durbin's modified principle of conditionality. *J. Am. Statist. Assoc.* **65**, 402–3.

Birnbaum, A. (1972). More on concepts of statistical evidence. *J. Am. Statist. Assoc.* **67**, 858–61.

Cox, D. R. (1971). The choice between alternative ancillary statistics. *J. R. Statist. Soc.* B **33**, 251–5.

Durbin, J. (1970). On Birnbaum's theorem on the relation between sufficiency, conditionality and likelihood. *J. Am. Statist. Assoc.* **65**, 395–8.

Fisher, R. A. (1966). *The Design of Experiments*, 8th edition. Edinburgh: Oliver and Boyd.

Fraser, D. A. S. (1963). On the sufficiency and likelihood principles. *J. Am. Statist. Assoc.* **58**, 641–7.

Fraser, D. A. S. (1968). *The Structure of Inference*. New York: Wiley.

Pratt, J. W. (1962). Comments on Birnbaum's paper. *J. Am. Statist. Assoc.* **57**, 314–6.

Savage, I. J. (1970). Comments on a weakened principle of conditionality. *J. Am. Statist. Assoc.* **65**, 399–401.

[*Received June* 1974. *Revised October* 1974]