Comments on Paper by J. D. Kalbfleisch
Author(s): G. A. Barnard
Source: *Biometrika*, Vol. 62, No. 2 (Aug., 1975), pp. 260-261
Published by: Oxford University Press on behalf of Biometrika Trust
Stable URL: https://www.jstor.org/stable/2335357
Accessed: 30-11-2020 19:19 UTC

Comments on paper by J. D. Kalbfleisch

By G. A. BARNARD

Department of Mathematics, University of Essex, Colchester

1. Limitations of Birnbaum's model

Birnbaum's (1962) model of the statistical inference situation fails to represent all situations met in practice because it presupposes that the totality of possible hypotheses has been exactly specified; the probability function $f(x, \theta)$ is taken as known for every point x in the sample space S, and for every point θ in the parameter space Ω, and these are taken to exhaust all possibilities. In the absence of exact knowledge of f for all possible states of nature the applicability or otherwise of Birnbaum's principles M, mathematical equivalence; C, conditionality; and S, sufficiency, may become uncertain. In fact, M and S seem to be specially vulnerable to difficulties of this kind, and C less so.

Thus, for example, if we are estimating a location parameter θ with n independent observations x_i $(i = 1, \ldots, n)$ each supposed to have density function $\phi(x_i - \theta)$, we very often are unsure of the exact form of ϕ. But we can be sure that, whatever the form of ϕ, the sample configuration specified by the $n-1$ differences $x_i - x_1$ $(i \neq 1)$ has a joint distribution which is independent of θ and so is ancillary; and hence that the n observations are equivalent to a single observation $x = x_1$ from a distribution with density proportional to

$$\Pi\phi(x-\theta+d_i) = \psi(x-\theta),$$

where $d_i = x_i - x_1$ $(i = 1, \ldots, n)$. It may well turn out that the d_i are such that our uncertainty about the form of ϕ affects the form of ψ very little, in which case our inferences about θ will be robust; or it may turn out otherwise, in which case our inferences about θ must be expressed as being conditioned by the form of ϕ. Thus C may be relevant and applicable, but not M or S.

2. An extended model

To modify Birnbaum's model to take account of such possibilities we may suppose Ω to possess a 'penumbra', so that it is a subset of a larger class Ω^*, although our inferences may be expressed relative to Ω. Then it is evident that a function of the observations whose distribution is the same for all hypotheses in Ω need not retain this property when Ω is extended to Ω^*. The distinction between ancillaries which fail to remain ancillary when Ω is extended, and those which do remain ancillary, is one which corresponds to some extent, though not exactly, with the distinction Kalbfleisch (1975) makes between mathematical and experimental ancillaries, respectively.

To point up the distinction, we may call functions ancillary in Ω but not in Ω^* unstable, while those which remain ancillary in Ω^* could be called stable; though such terminology could well be objected to on the grounds that Ω^* will in practice rarely be completely well determined. For instance, in the location parameter example, we might know only that ϕ was 'not far from normal', so that the stability of an ancillary would be relative to the specified Ω^*. We might find it useful in a given context to imagine an increasing sequence of $\Omega^*, \Omega^{**}, \ldots$, and a corresponding sequence of ideas of stability.

The chance mechanisms used in the argument Kalbfleisch quotes from Barnard, Jenkins & Winsten (1962) are evidently both stable and experimental in Kalbfleisch's sense. On the other hand, if the possible alternative distributions in Kalbfleisch's Example 1 all have μ as a location parameter, then the difference $x_2 - x_1$ in his example will be a stable ancillary which is merely mathematical. The distinction between stable and unstable ancillaries would seem to have the same effect in preventing the derivation of the likelihood principle, in general, as Kalbfleisch's distinction between mathematical and experimental.

3. Vague alternatives

Another defect of Birnbaum's model, related to the first, is that it fails to represent the situation where we are testing a single hypothesis against vaguely specified alternatives. In this, the *locus classicus* of the simple test of significance, we may well have $f_0(x) = f_0(x')$, but x representing a much more discrepant result than x', for example where x lies out in the tail of the distribution, while x'

lies near the middle but happens to be a point with low probability function. But if the parameter space consists of a single point, Birnbaum's principle M would require that the two points x and x' would have the same evidential meaning. This suggests that in significance testing situations the likelihood principle fails to apply, which is, of course, what many statisticians think.

4. Minimal experiments

Kalbfleisch's concept of a minimal experiment is highly suggestive. While, perhaps necessarily, it remains a little vague, it points towards a deep-lying notion in the philosophy of natural science in general, and of statistical inference in particular. This is the notion of a repeatable experiment (Fisher, 1966, Chapter 2, §7): 'In order to assert that a natural phenomenon is experimentally demonstrable we need, not an isolated record, but a reliable method of procedure . . . etc.)'. What, minimally, must be repeated is a minimal experiment in Kalbfleisch's sense.

5. Applicability of Birnbaum's conclusions

Although these considerations may be taken as ways of avoiding the unpleasant consequence of the likelihood principle, and although the present writer does not subscribe to the general validity of this principle, it should perhaps be said that in those cases where Birnbaum's model really does apply, where there is no penumbra to the parameter space, where there is no ambiguity about the definition of the sample space and where Birnbaum's principle M applies, then the likelihood principle does appear to the present writer to be applicable. The growing practice, among geneticists and high-energy physicists, of reporting the likelihood function or its equivalents, for example, the score function, in cases of this kind, suggests that the principle accords with their intuition also, in such cases. It may be noted that conditionality arguments of the kind discussed in §1 above are consistent with the likelihood principle; the set of observations $(x_1, \ldots, x_n)$ and the equivalent single observation x give the same likelihood function, though the sample spaces are different.

References

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

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

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

Kalbfleisch, J. D. (1975). Sufficiency and conditionality. *Biometrika* **62**, 251–9.

[*Received December* 1974]

Comments on paper by J. D. Kalbfleisch

By OLE BARNDORFF-NIELSEN

Department of Theoretical Statistics, Aarhus University

The concept of experimental ancillaries is interesting, but Kalbfleisch (1975) has, in my view, not produced convincing reasons for being unable either to apply S before C_E, or to apply C_M before S. I am inclined to think that such reasons of principle cannot be found and that the resolution of the dilemma raised by Birnbaum's theorem lies, not in prescribing some fixed sequence of modified versions of C and S, but in accepting that statistical analyses, in general, will not and cannot lead to unequivocal conclusions, and, in particular, that various applications, to the same model and data, of the procedures of margining to a sufficient statistic or of conditioning on an ancillary statistic may lead to different views on the chance mechanism.

It is to be expected that reasoning from different angles about a subject matter on the basis of incomplete information will result in differing answers. Ambiguity of this kind is a difficult but not