

91

## **On Models and theories of inference; structural or pivotal analysis\***

**D. Brenner, D. A. S. Fraser, G. Monette**

*Recent material is surveyed concerning the foundations of statistical models, the definition and logic of statistical inference, and the additives used in various applications of theories of inference. This background is then used to compare structural/structured models [Fraser 1966, 1968, 1979] and pivotal models [Barnard 1974, 1981]; the models are found to be identical except for nomenclature and basic justification which supports the structural format. An outline is given of the development of the structural/structured model and of the corresponding mode of inference. The recent pivotal methods of analysis are assessed with this development as background.*

### 0. Introduction

At Toronto, since the mid 1950's, we have been involved with detailed investigations of certain inference methods, of questions concerning the formation of statistical models, and more recently of comparisons of various inference theories. As part of this continuing process, I'd like to give a survey of our recent work on the formation of statistical models and the comparison of inference theories, and give also an outline of the development of structural methods. Then, more specifically, I'll make comparisons of pivotal methods with the structural methods that we've been developing since 1966.

Before this, though, let me say generally that I find it reassuring that periodically we do have symposia on statistical inference. I recall particularly the symposium here at Waterloo in 1970, the recent one at Oberwolfach in December 1980, and of course, the one we are involved with now. I think having periodic symposia on inference is fundamentally good. It represents a concern for the fundamentals, for the basics, for the meaning contained in the whole statistical process. It is also fundamentally good, because out there in the profession (certainly not here) there's a lot of diffuse anti-intellectualism and parochial development – not to say that a little bit of parochial development may not be good, as otherwise we would all be doing decision theory. The extremes out there go in many directions but one of these seems to stand almost in stark opposition to ours; it involves a CRT display, a few standard deviations, and no model at all. There is of course, richness in variety when a subject is developing, but an extreme that isolates itself from the rest can, in fact, be a negative contribution.

---

\*Prepared for and presented at a symposium on inference methods held at the University of Waterloo, on August 6, 1981.

## 1. Inference

What is statistical inference? The term didn't appear too much before maybe thirty years ago, it didn't really exist as something separable and distinct in statistics. In fact, it arose as an alternative to the formalities imposed on the profession by decision theory, which in turn came largely from acceptance sampling.

What is inference? Here is an attempt to put it in the form of a definition. *Inference is the theory and methods concerned with the way the background information, the available or given information concerning the process, together with current data make implications concerning the unknowns in the system under investigation.* I emphasize *implications* – in the true mathematical, logical sense, not just that there are steps that use logic or use mathematics but that the whole process is a precise logical deduction. Some discussion of this may be found in Brenner, Fraser and Monette [1981] and Fraser [1982].

If we dwell for a moment on the definition, it says, I think, that we're talking about *all* of central statistics and not just inference as conventionally packaged. But it wouldn't cover decision theory or control theory; both of these have a real place but that is certainly not what we are talking about here. What we do have under the heading inference is a lot of specialized areas – confidence theory, hypothesis testing, estimation theory, and others including for example the derivation of posterior distributions by various methods. However, I think these areas exist separately only to the degree that basic statistical inference is delinquent in its overall role as indicated by the definition.

At the meeting in Oberwolfach, Jose Bernardo, a rather committed Bayesian raised questions as to why I hadn't mentioned experimental design or why experimental design didn't seem to be included in the definition. Questions of design are of course fundamental and very important, and in fact really at the core of things, and I think that if you examine the definition openly, then design *is* included. Experimental design is concerned with what makes or strengthens the logical process. In a formal sense, this is concerned with the *domain* of the inference process, to see how it affects the *range*. This comment then is a rather belated response to Jose, and I won't dwell on it further.

## 2. Logic of Inference

What is the logic of inference? If we examine the statistical literature in search of answers, the kindest thing we can say is that the literature is pretty diffuse on this topic. You see special terms that leave me personally feeling very uneasy. The term "a model for inference". If you think of the inference process, that is one thing. But what does it mean to talk about a *model* for inference itself? or even talk about "theories of inference". If we think of the given material and what is implied by that, then why the plural, *theories* of inference; or a more extreme term "new ways of reasoning towards...". What do these terms actually mean? Just taking them as they stand you get the impression of some supra-rational process – new ways of obtaining things that are not con-

tained logically in the given material we're working with. They suggest thing magical, maybe mystical, a modern version of the crystal ball that you look into and the answer pops out; maybe it comes out of the CRT display. Or what about the term "evidential meaning"? The term is attributable to Alan Birnbaum, I think. It is a strange term. You get the idea that evidential meaning is something that's going to appear at the end of some process by behaving or manipulating or doing things in some *right* or *conventional* manner. "Right"; maybe that's acceptable. But "conventional" gets me very worried. Is inference really going to be "professional consensus"? There is a sentence in Professor Barnard's work that I find particularly disturbing: "It is more important for us statisticians to agree on our practical recommendations than on the reasoning underlying these". I hope he didn't mean it as it appears to me. In a sense it's saying that as long as we agree out there, then it doesn't matter whether we know what we're doing, what the implications are from the model and the data! That concerns me. I believe we ought to know what we're doing; I think we ought to be very clear about what is assumed known and what the implications are from the given material.

In another direction we frequently hear the term "structural inference". I don't know where the term comes from, but I do know that many who use it haven't read material on structural models and their analyses.

What is structural? Our concern, with a background of structural, is fundamentally with the analysis of statistical models and data. What kind of models? The traditional models, the structured models, the structural models and, indeed probability models. Probability models? If you have an objective prior, then your full model includes that prior and becomes a probability model, not a statistical model. There are Bayesians seeking respectability and non-Bayesians hanging on Bayesian coat-tails who talk of "empirical Bayes". Well, if there is an empirical *prior* then it's *part* of the statistical-probability modelling process, it is not an inference method. Maybe the prior has different, possibly theoretical, origins and maybe we want to set it temporarily aside from other experimental or observational information but it is still part of the basic model, hence the reference to a probability model for something that has improperly been classified as an inference method. For some discussion, see Fraser [1976, Section 9.6].

When there is structural information on the model we tend to put emphasis on certain things. We're concerned with probing the *substance*, the true *given* in any particular problem, or —, a heavily used term, the *structure*. What we're really talking about is *what's there* — or in another word the *non-arbitrary*.

*Statistical inference* — we view it as a formal deductive process. If you are going to talk about something that's formal, then the ingredients have to be formal; you have to have a formal starting point. The formally given material essentially separates into two parts: the *model*  $M$  which we view as the background information concerning the unknowns of the investigation; and the *data*  $D$ , the observed values for the variables in the model. Now examining questions of inference in the literature we see that this isn't all — there are almost always other things; there are *additives*, other introduced elements — the shifting sands of the discipline.

The underlying *inference base*, then, has composition  $(M, D)$ , the model and the data; or, if we want to embrace the qualified situations,  $(M, D, A)$ , which gives the model and the data together with specified additives. The inference

process, then, is a logical deductive analysis of (M,D) or (M,D,A) to obtain a direct expression of the information concerning the unknowns or the unknowns qualified by A.

### 3. The Given

#### A. The Context and the Background Information

Consider further this notion of the *given*, the given material that is the formal starting point for the inference process. There is of course some physical *context* that gives rise to the given. In a very general sense this context involves some *system that is subject to investigation*.

We use the term background information for all the direct and indirect information concerning the system.

#### B. Model for the System

The background information, when presented in a formal manner, becomes the *model for the system: a formal mathematical presentation of the set of possibilities for the system*. The model is NOT an *arbitrary* thing. The elements in the model must have correspondents in the system! Thus, the model is *descriptive*. And elements in the system must have correspondents back in the model, thus the model is *exhaustive*, exhaustive in the sense of coverage.

So let me emphasize: the model is *real*, not *arbitrary*. There should be, in an essential sense, a one-one correspondence between elements in the model and elements in the system under investigation. This agrees with what others say in general concerning models; the classic example of a real mouse and a tin mouse; the tin mouse can be a model for the real mouse; and indeed, the real mouse can be a model for the tin mouse.

In practice, however, when models are given there are often *extras* inserted, such as, for example, *internal variables* without direct reference to the system under investigation. Of course, there can be internal variables that do have objective support, in which case they would properly be included in the model. But for internal variables without evidence or objective support then we're getting into a dangerous area, we're getting into the uncertain elements that go with the *additives* I mentioned earlier.

All of the preceding discussion of models is in the sense of – *to a reasonable approximation*. Perhaps at times we may want to generalize the model to gain a closer approximation.

The term *true model*. That's a term that gets tossed around a lot. Is your model true? Such a question is often motivated by a mere desire to subvert the best intentions of the expression. It invites the answer No, with a suggested invalidity for all that surrounds the model. A calm answer to the question is, of course, *No* in a literal sense, but certainly *Yes* in the sense of giving a reasonable approximation, given the properties we've just been discussing.

### C. Model for the Investigation

If we consider a particular investigation, then we are concerned with the model for the system made specific to the investigation. *The model  $M$  for the investigation*, then, is the set of descriptions covering the unknowns of the investigation.

The model for the system is made specific in various ways, like specifying the number of repetitions, the values of input variables, and other particular characteristics of the investigation. This is a rather abbreviated definition but it's been elaborated on elsewhere, for example, Fraser [1979].

Basically, this model records the set of descriptions covering the unknowns of the investigation. There are some obvious criteria for it. For example: *Do not treat a known as an unknown; Do not treat an unknown as a known*. Why would these need mentioning? Well, you all know the example where an investigator tosses a coin and if *heads* he picks instrument I, if *tails* he picks instrument II, he makes his measurements with the chosen instrument and then builds the 50–50 probability into his model. He is of course modelling the system not the particular investigation. And indeed, he obtains various peculiar decision theoretic, testing, and confidence interval results; for related comments, see Fraser [1979, Section 5.2.2], Feuerverger and Fraser [1980]. But clearly he knows what instrument he used and it came by an objective probability mechanism that has nothing to do with the parameters being measured. He knows which way the coin fell and he knows which instrument he used. He needs to present the model for the investigation, as opposed to acting as if a known is an unknown.

### D. The Data

The data  $D$  are the observed values for variables in the model for the investigation. Again this has been elaborated on elsewhere [Fraser 1979].

### E. Remarks

Let me pick up again the phrase “No model is true” and the obvious answer “of course not, just to a reasonable approximation”. What do people mean when they use the phrase: I think it's a kind of shock treatment, an attempt to crumble the very starting point of an argument – it's absurd reduction: No model is true, therefore anything goes.

An investigator may explore a range of models. There would then be a corresponding range of implications or inferences.

For our purposes, there is a clear need to know what the *given* is, what purportedly is describing the investigation.

There is also a clear need to isolate the extras, the *additives* that are appended to the model and data, or are added “quietly” during subsequent analysis. These include ideosyncratic elements, the easy extras, the additional assumptions announced by phrases such as “it seems natural to...”. You should have your guard up when you hear that. There's a good chance something is being slipped in too easily; at the very least, I'd like to see something specific in the way of documentation.

## 4. The Additives

What are the additives? The additives are all the extras in the inference process that are not covered by the model  $M$  and the data  $D$ .

What are some examples of these additives? They include all the common *principles* widely used: *sufficiency*, *ancillarity*, *invariance*, *unbiasedness*, *likelihood*, and others. There is the *choice of the function* for use in an analysis. Or more generally, there is the *choice of the procedure* for an analysis. There is the insertion of *intermediate variables* in the model and in the analysis. There are the *criteria* such as *least squares*, *minimum length*, and *maximum power*. There are also notational elements and these are ones that can really slip by. Once things are in notation or are labelled, they acquire a kind of absolute status. But things that differ only in label and otherwise are just elements of a set are just notational and we shouldn't act subsequently as if they're not and introduce elaborate principles or methods to get rid of them. You acknowledge that they're labels. Other examples of additives can be found in the literature after disarming phrases such as "it is useful to", "it is natural to".

Indeed, often, it is not clear what the additives are in any particular analysis. They come in at various points throughout an analysis, often in an ad hoc manner, sometimes without any indication that something arbitrary is added.

The purpose of inference is to get from the model and the data, from the base  $(M, D)$ , a clear presentation of the information concerning the unknowns. We have just acknowledged that typically we have the model and data with additives, the extended  $(M, D, A)$ , and inference, then, is to get a clear presentation of the "information" concerning the unknowns, with quotation marks to suggest the qualification that comes from *additives* as partial input.

The preceding raises a very serious question. How do the results depend on the additives  $A$ ? The additives aren't objectively describing things, they're put in because they're nice, or "helpful", or "useful". I think there are perhaps two ways of approaching this question, neither one easy. The obvious first approach is *sensitivity analysis*. This is an *operational approach*. Sensitivity analysis means you start changing the  $A$ 's and see what happens at the implications' end of the inference process. In effect, you're acknowledging that you don't really know what's going on, so you change the elements of  $A$  and see what happens. However, I don't see how this can answer the question of what happens *without*  $A$ .

The other approach, which has particular interest for us, is to recast the additives in the form of an extended model, an enlarged frame *with actual references* for its elements; thus making the additives in effect *descriptive*. For this we are in effect bringing the additives back to the beginning to see what they can reasonably describe in the physical world. We're sort of saying: *Here's my model and my data and I have juxtaposed with them a collection of other things describing a larger context, and in that larger context I obtain some particular presentation of information*. To follow through on this approach represents a substantial and important challenge. It represents a concern for *meaning*, a counterpoint to the operational approach.

## 5. Pivotal

### A. Main Background

A *pivotal quantity* is a function of the response variable and principal parameter that has a fixed (or separately parameterized) distribution. A pivotal quantity is commonly called just a *pivotal*.

Their recent presence in Statistics dates back to Fisher's 1930 paper on Inverse Probability. In that paper a pivotal quantity, although not named as such, was used to derive a fiducial distribution for the correlation coefficient. In subsequent papers pivots became central (both in substance and in name) to the derivation of fiducial distributions. In most cases the pivotal quantity was constructed from the response variable as reduced to the minimal sufficient statistic. Many analyses appear in Fisher [1956].

Pivotal quantities also enter prominently in the presentation of confidence intervals and regions. Their role there is primarily one of convenience giving a mechanical device for deriving confidence intervals and regions; for some early examples see Fraser [1958].

### B. General Presence in Statistics

In a general sense pivots have been present in Statistics since at least the time of Gauss. Equations exist there in which response variables are expressed in terms of error variables that have a fixed distribution representing the measurement error. As Statistics was not codified in any detailed way we don't find there the formality of the present definition but some mathematical essence exists in those equations.

### C. Invariant Fiducial

My own concern with pivots goes back to the mid 1950's and the search for meaning in the fiducial method proposed by Fisher. This developed largely into the material in two papers [Fraser 1961a, b] in which a frequency interpretation and sampling consistency were shown to focus exactly on model parameters exhibiting group invariance, thus on transformation parameter models. Fisher wrote me at that time endorsing essentially all the material in those papers save my omission of the correlation coefficient as a valid group example. Subsequently the correlation coefficient was shown to have a *marginal* fiducial distribution in a special group framework using the positive lower triangular matrices; some discussion of this appears in Fraser [1964].

What came out of that fiducial analysis period was that if you're going to use pivots, even for the limited objective of deriving posterior distributions, the transformation group property for the parameter variable relationship was *fundamental* and *essential*.

#### D. Pivotal versus Expression Functions

Early in the 1960's as part of our continuing concern for fundamentals in inference we changed our notational emphasis at Toronto from the *error/pivotal* form:

$$\begin{pmatrix} \text{Error Variable} \\ \text{Pivotal Variable} \end{pmatrix} = f_1 \left( \begin{array}{c} \text{Response} \\ \text{Observable Variable} \end{array} \mid \theta \right) \quad (1)$$

to the *expression/structural* form

$$\begin{pmatrix} \text{Response} \\ \text{Observable Variable} \end{pmatrix} = f_2 \left( \begin{array}{c} \text{Error Variable} \\ \text{Pivotal Variable} \end{array} \mid \theta \right) \quad (2)$$

The functions  $f_1$  and  $f_2$  in the two cases are just inverses one of the other.

For purposes of modelling, the two equations (1) and (2) can be identical. They differ only in the choice of which direction we write down a one-one mapping. Or in the abstract sense of a function as a set of ordered pairs, it's just the notational choice of whether we had the pairs from left to right or right to left.

Mathematically, the equations or representations (1) and (2) are the same. Of course, in certain contexts it may be more convenient or fruitful to think of the pivotal form (1) and in others to think of the expression form (2). The two differ only in the names attached to the components.

Well – Why the shift from the pivotal (1) to the structural (2) equations? It seemed clear to us at Toronto in the early 1960's that there were a lot of applications that have a clear *objective error variable*. We felt the model should reflect this reality; thus include directly in the model a variable for the error: Error was real, so model it; don't pretend to "recover it" by some pivotal function. The pivotal functions were trying to compensate for a modelling mistake – *recovering from many distributions (in the ordinary model) the single distribution that was there originally*.

Many of the measurement error applications have very clearly an objective error. But there are other not so obvious cases such as the dilution and bio-assay problems that emerged much later, where, in retrospect, the error is so objective that it's sitting on the laboratory table [Fraser and Prentice 1971, Fraser 1979].

Why the notational shift from the pivotal function to the structural equation? You can think of it as a notational question, a matter of taste. Or, more fundamentally, you can think of it as a reorganization in terms of *objective components* – which goes back to my concern at the beginning – that models must involve objective components, otherwise we implicitly have additives.

Equations (1) and (2) can each be examined in a more general context that loses the bijective quality implicit as they stand above. An error/pivotal variable can be generated which is a reduction on the response variable, see Fraser [1979]. Also a response variable can be a reduction from a "larger" error variable; see the structured models in Fraser [1971] but recall the concerns just mentioned concerning objectivity for error.

## 6. Structural/Structured Models and Pivotal Models

In Toronto in the early 1960's, as just mentioned, we shifted from the pivotal equation to the structural equation. The *Biometrika* paper [Fraser 1966] examined a simple case and subsequent papers led to the detailed examination of structural models as summarized in *The Structure of Inference* [Fraser 1968]. This latter emphasized, unfortunately perhaps, posterior distributions but covered also estimation, testing, and confidence methods. It also restricted attention predominantly to the case where the expression functions formed a group. This was done quite deliberately out of concern for the interpretation and consistency as mentioned earlier.

The structured model appeared in University of Toronto theses in the late 1960's; it involved the more general case without the group restriction. I discussed aspects of the structured model at the symposium here eleven years ago; this appeared in the proceedings as *Events, Information Processing and the Structured Model* [Fraser 1971]. That paper examined the response as some function of the error and investigated requirements on the class of possible functions to validate *conditioning*, to have the needed partition for conditional probability. It turned out to be fundamental that the class have *closure properties*, which in turn imply group properties. This was subsequently examined in a very formal frame [Brenner and Fraser 1979].

The need for a group is not an idle thing; it is real and fundamental. You don't have to look far for examples. McGilchrist [1973] examined some structured/pivotal models and derived tests. In the discussion [Fraser 1973] of the paper I gave an example where the group property was absent and nominally calculated probabilities can be far from any actual, repetitive type probabilities. Of course in certain contexts you can be close to having a group, so that calculated probabilities can approximate real probabilities – a case perhaps of a model being a reasonable approximation.

But how are we to choose between two models: the structured/structural model and the pivotal model? The two models are the same except for nomenclature and the order in which elements are presented.

At Toronto we come down very firmly in favour of the structured/structural model. And as I mentioned earlier this represented a switch in the early 1960's from the pivotal version to the structural version. Our preference for the structural at that time was based on *organizing the model in such a way that the components could be clearly identified with objective elements in the application*. A consequence of this was that we were in a firmer position to check various related questions of validity and obtain extended models without losing prime properties.

From other points of view one could argue that it's a matter of personal taste how you lay out components of the model. In the pivotal form one can view certain things in an exploratory way and seek functions that pull off or exhibit characteristics in the basic traditional model. This avenue, however, has also been explored in a structurally oriented direction in Brenner and Fraser [1981].

The validity questions have been extensively explored in the structural/structured format and the group property found to be essential. Writing the model in pivotal form doesn't modify this fact in any way, but the avenues to explore validity may largely be dissipated.

The generality and extensiveness of published material is relevant. In the structural format, the monograph *The Structure of Inference* [Fraser 1968] examined tests, estimates, confidence intervals, and posterior probabilities for a wide range of structural models: location, scale, location-scale, regression, multivariate, multivariate regression, dilution, bioassay, circle and sphere. The emphasis, however, was heavily on the posterior distributions and the monograph was generally judged as being “all posterior distributions”. The numerical implementation for the conditional analyses with these models involved integration and was illustrated only for the simplest location model.

After the publication of that monograph it was clear that the tests, estimates and confidence intervals needed emphasis and that the conditional analysis of the location scale model should be given extensive attention. In Fraser [1976] the location-scale analysis was examined numerically and marginal likelihood used for estimating error distribution shape; some adaptive methods were also discussed. In Fraser [1979] structural/structured models were examined for tests, estimates, and confidence regions (posterior distributions were ignored) and the computer-based conditional analysis for the location-scale model was given major emphasis. This latter analysis applies whether the model is in classical, pivotal, or structural format, only the reasons or grounds for the analysis differ. These different grounds were discussed in context. Significant in this is the *automatic conditioning that occurs with the structural format*; there is no need to refer to the controversial ancillary principal as with the classical and pivotal formats.

It is of interest to mention here the location examples that appear in Gauss’s writing. Error variables were used, and indeed the potential was there for the *automatic conditioning* that in fact evolved much later with the structural model.

In the pivotal format the published material seems to date from Dempster [1966] who examined a pivotal version of the location model. Beran [1971, 1972] used a model that in a sense was a mathematical compromise between the structural and pivotal and pursued posterior probabilities in the pattern followed by Dempster. More recently, Barnard [for example, 1974] has considered inference for pivotal models, particularly, the location-scale case but the importance of likelihood for the shape parameter has been overlooked.

In summary we note the identity of the structured/structural model and the pivotal model; emphasize the objectivity available with the structural format; recall the extensive examination of models in the structural format, the development of the computer programs for conditional analysis that accompanied the surveys of structural applications; cite an exploratory role for pivotal models; and note the very fundamental questions that open ahead concerning the foundations of our subject.

## 7. Details: Structural and Pivotal Models

For the location scale model we can allow an additional parameter for the error or pivotal variable. Various terms have been used such as “second level param-

eter”, “model adjustment parameter”, and “discrepancy parameter”. For the structural model we prefer the terms “shape parameter” or “form parameter” to refer to unknown characteristics of the error distribution.

With various possibilities for the error distribution it becomes advantageous to introduce a standardization with respect to the transformation parameters such as location and scale. This standardization is also needed as part of the actual definition of the primary parameters, for example, the location and scale parameters. We have found the standardization with 68.26% probability in  $(-1, +1)$  to give substantial stability to the key error distributions that are used for the estimates and confidence limits; see Fraser [1979, Section 1.2.2].

In Barnard [1981] the ratio-of-means example involving the Darwin data is used to suggest that the group property is not needed and that a concept such as an “approximate sufficiency” can be invoked as needed. The example, however, is not so compliant. A marginal distribution is approximately independent of the parameter; this argues for “approximate ancillarity” and a conditional analysis. However, a marginal distribution is calculated for a t-statistic, and this would require an approximate sufficiency, which was not demonstrated or considered. The subsequent calculations then used an approximate linearity, thus an approximate group. Even without the shift from alleged ancillarity to alleged sufficiency *one cannot argue that calculations for an approximate group thereby illustrates that the group condition is not needed.*

Ancillarity as a principle has fundamental difficulties. Clearly, it works sometimes and it fails sometimes [for an early example, see Fraser 1973]. For recent discussion see Fraser [1979, Sections 4.2, 3.2]. It seems natural, then, to delineate or isolate the special aspects of the successful cases. A solution in this direction is formed in the objective component probability spaces used with the automatic conditioning of the structural model; see Fraser [1979, Section 3.2].

Barnard [1981] cites a “modus ponens probabilitates”:  $\Pr(A \text{ given } B) = p$ ,  $B$  is known to be true, therefore  $\Pr(A) = p$ . Unfortunately, as a theorem or assertion this is false. It is essential to know the origins of the information “ $B$  is known to be true”. For examples, see Fraser [1976, Section 4.1] and Freund [1965] where the “implication”  $\Pr(A) = p$  clearly does not obtain; see also Fraser [1971] where some general questions concerning information are addressed.

The questions surrounding conditional and long run confidence levels are serious; the same applies for conditional and long run tests of size  $\alpha$ . Welch [1939], of course, was arguing in favour of procedures that amounted to varying the conditional level, varying it to achieve minimum average confidence interval length or maximum power; this was in the Neyman Pearson tradition; a reassessment was given by Buehler [1959] and Wallace [1959] following the viewpoint expressed by Fisher [1956]. Some recent comment may be found in Fraser and McDunnough [1980] following the approach initiated in Fraser [1979].

I'll start from the premise that the conditional analysis consequent to objective modelling and automatic conditioning is the correct analysis for certain simple location, location-scale and regression models. This acceptance could be based on (i) the structural model without additives (ii) the pivotal model with the ancillarity principle (an additive) or on (iii) the traditional model with additives involving a reduction procedure and an ancillarity principle.

I see no way in which we could advertise a certain overall confidence level and then allow the conditional level for any presented intervals to fluctuate up and down whether to achieve some minimum mean length or to achieve say maximum power. To do so amounts to fraud in the face of the premise just mentioned. What would be really happening is that when a client comes in with precise measurements you give him a long confidence interval (high conditional confidence) and then later when a loose scientist comes in from another department, (he works with low precision and his intervals might reasonably be pretty long) you give him a short interval. You may advertise yourself as a 95% confidence-interval producer and qualify that with some reference to conditional confidence; but what it amounts to is *buying your mean length where its cheap*, a long interval for the precise scientist and a short interval for the loose scientist all to achieve some overall minimum mean length at 95% confidence. To me it is a serious misrepresentation. The thing that's basic is the *conditional level*, a consequence of the objective modelling with its automatic conditioning.

I've expressed the preceding in terms of different scientists arriving with data. But it applies equally if we consider it "within" scientists or "within" types of analysis, or even "within" particular problems like the uniform distribution example in the early Welch paper.

To contemplate overall confidence with fluctuating conditional confidence levels is to open oneself to very serious charges of misrepresentation.

## References

- BARNARD, G.A. (1974): Pivotal and robust estimation. In: Proceedings of the Conference in Foundation Question in Statistical Inference. (Eds.: O. Barndorff-Nielsen, P. Blaesild, G. Schow), Aarhus; 61–80.
- BARNARD, G.A. (1981): A coherent view of statistical inference. Symposium on Statistical Inference and Application. Waterloo.
- BERAN, R.J. (1971): On distribution-free statistical inference with upper and lower probabilities. *Annals of Mathematical Statistics* 42, 157–168.
- BERAN, R.J. (1972): Upper and lower risks and minimum procedures. In: Proceedings of the Sixth Berkeley Symposium, 1, 1–16.
- BRENNER, D., D.A.S. FRASER (1979): On foundations for conditional probability with structural models, when is a class of functions a function. *Statistische Hefte* 20, 147–160.
- BRENNER, D., D.A.S. FRASER (1981): A simplification of the traditional statistical model in the presence of symmetry. *Statistische Hefte* 22, 195–206.
- BRENNER, D., D.A.S. FRASER, G. MONETTE (1981): Theories of inference of simple additives. *Statistische Hefte* 22, 231–233.
- BUEHLER, R.J. (1959): Some validity criteria for statistical inferences. *Annals of Mathematical Statistics* 30, 845–863.
- DEMPSTER, A.P. (1966): New methods of reasoning toward posterior distributions based on sample data. *Annals of Mathematical Statistics* 37, 355–374.
- FEUERVERGER, A., D.A.S. FRASER (1980): Categorical information and the singular linear model. *Canadian Journal of Statistics* 8, 41–45.
- FISHER, R.A. (1930): Inverse probability. In: Proceedings of the Cambridge Philosophical Society 26, 528–535, (Reprinted as Paper 22 in FISHER (1950)).
- FISHER, R.A. (1950): *Contributions to Mathematical Statistics*. New York.
- FISHER, R.A. (1956): *Statistical Methods and Scientific Inference*. Edinburgh.

- FRASER, D.A.S. (1958): *Statistics. An Introduction.* New York.
- FRASER, D.A.S. (1961a): On fiducial inference. *Annals of Mathematical Statistics* 32, 661–676.
- FRASER, D.A.S. (1961b): The fiducial method and invariance. *Biometrika* 48, 261–280.
- FRASER, D.A.S. (1964): On the definition of fiducial probability. *Bulletin of the International Statistical Institute* 40, 842–856.
- FRASER, D.A.S. (1966): Structural probability and a generalization. *Biometrika* 53, 1–9.
- FRASER, D.A.S. (1968): *The Structure of Inference.* New York.
- FRASER, D.A.S. (1971): Events, information processing and the structured model. In: *Proceedings Symposium on the Foundations of Statistical Inference* (Eds.: V.P. Godambe, D.A. Sprott), Toronto.
- FRASER, D.A.S. (1973): The elusive ancillary. In: *Multivariate Statistical Inference* (Eds.: D.G. Kabe, R.P. Gupta), London.
- FRASER, D.A.S. (1973): Comments on: Post-data two sample tests of location. *Journal of the American Statistical Association* 68, 101–140.
- FRASER, D.A.S. (1976): *Probability and Statistics. Theory and Application.* Toronto.
- FRASER, D.A.S. (1979): *Inference and Linear Models.* New York.
- FRASER, D.A.S. (1982): Statistical Inference. In: *Encyclopedia of Statistical Sciences* (Eds.: S. Kotz, N.L. Johnson), New York.
- FRASER, D.A.S., P. McDUNNOUGH (1980): Some remarks in conditional and unconditional inference for location-scale models. *Statistische Hefte* 21, 224–229.
- FRASER, D.A.S., R.L. PRENTICE (1971): Randomized models and the dilution and bio-assay problems. *Annals of Mathematical Statistics* 42, 141–146.
- FREUND, J. (1965): Puzzle or paradox. *American Statistician* 19, 29–44.
- McGILCHRIST, C.A. (1973): Post-data two sample tests of location. *Journal of the American Statistical Association* 68, 97–101.
- WALLACE, D.L. (1959): Conditional confidence level properties. *Annals of Mathematical Statistics* 30, 864–876.
- WELCH, B.L. (1939): On confidence limits and sufficiency with particular reference to parameters of location. *Annals of Mathematical Statistics* 10, 58–69.

Prof. Dr. David Brenner, Prof. Donald A.S. Fraser, G. Monette, University of Toronto, Dept. of Statistics, Toronto ONT M5S 1A1, Canada.