行政院國家科學委員會專題研究計畫成果報告

什麼是無信息統計量及如何消去多餘參數
What is Noninformative Statistics and How to Eliminate the Nuisiance Parameter

計畫編號：NSC 87-2118-M-009-003
執行期限：86 年 8 月 1 日至 87 年 7 月 31 日
主持人：洪慧念　　國立交通大學統計學研究所

一、摘要

中文摘要

我們知道若一統計模型中有多餘參數時，往往會使問題困難許多。對於此問題當多餘參數只有一維時，過去我們以有了相當深入的探討，而且在實際的應用上也有不錯的結果。但若多餘參數為多維度時，到目前尚未有很好的結果。

因此本計畫是希望能找到解決此一問題的方法。首先我們知到 Jeffrey's 事前分佈在參數為單維度時具有非常好的性質，但在多維度時則不然。所以本計畫最重要的部份即在於找出一個在多維度的多餘參數空間上的一個好的事前分佈函數。此外我們還要在本計畫中探討參數的垂直性，我們的重點在於當其中有一參數為離散且多餘參數為多唯度的情形。我們相信此問題對於貝氏項中的廣義事前分佈中的常數決定非常重要。最後我們要研究的是如何衡量一組數據對某一個固定參數所提功的訊息，對此問題我們最關心的部份是何時不提供任何訊息。相信解決了這些問題後，我們必能找出消去概似函數中的多餘參數的最佳方法。

關鍵詞：多餘參數，垂直性，概似函數，群模型

Abstract

As we know, how to eliminate the effect of nuisance parameters in a statistical model is a difficult problem. For the single parameter case, we have many good results in our early work. But for the many nuisance parameter cases, this problem is not clear yet.

The purpose of this project is trying to solve this problem in several steps. It is well known that the Jeffrey's prior is a good prior and have many good properties when the parameter is single. However, if the parameter is multi dimensional, then the Jeffrey's prior is no longer good. So the first step in this project is to find a good prior on the high dimensional parameter space. In the second step, we will study the orthogonality of parameters. Up to now, people define the orthogonality of parameters only through mathematics formula, but we think that we should look at the orthogonality form the statistical point of view such that the orthogonality can be applied to the irregular or discrete situation. The last thing we want to do in this project is to discuss the situations in which the data does not contain any information about one parameter. We

believe that if we can solve the above problems, then we know how to eliminate the effect of the nuisance parameter.

Keywords: nuisance parameters, likelihood function, group transformation model, orthogonality.

二、結果與討論

Let a statistical model be parameterized as (,) where is the parameter of interest and is the nuisance parameter. It is quite possible that a single observation does not contain any information about . For example, if X is a random variable with normal distribution with mean and variance then it is reasonable to say that X does not contain any information about . If the data contain no information about , then we will hope that the likelihood function for is a constant function.

To give a definition of "no information" is a difficult problem. Here, we give a sufficient condition for a random variable to contain no information about as follows (This idea can also be found in Barndorff-Nielsen (1976) and Dawid (1975)): If X satisfies the condition (N) Hung and Wong (1996), then we can say that X contains no information about .

This definition can be justified by the idea of invariant test (Cox and Hinkley 1974), see Hung and Wong (1996).

From experience we know that if we can get the "right" likelihood function when we have only a single observation, then we will get the right likelihood function for all sample sizes. Therefore, we believe that one should pay more attention on the case when the sample size is unity. In this research, we will discuss the relationship between the average likelihood and the noninformative statistic in the group transformation models.

(1)

In a group transformation model, if the group G can be embedded into the real line such that the composition and inverse operators are both continuous functions, then there exists h( ) such that h( ) $=g^{-1}$ and , i.e., is orthogonal to g h( ).

(2)

Let X_1 and X_2 have densities p(x), q(x) respectively. Suppose h is a one-to-one transformation of X, and let p'(x), q'(x) denote the densities of $h(X_1)$ and $h(X_2)$ respectively, then

(3)

Let $_\theta$(dg) denote (g|)dg, then $_\theta$(dg) is a left invariant measure on G. Since all the left invariant measures on G are up to some constant and the choice of weighting function in average likelihood is also up to some constant, we can choose $_\theta$(dg) to be independent of , say (dg).

(4)

Let G be a unimodular group and Y_θ f(y; ,g) satisfies condition (N), and conditions in Lemma 2.1, then the average likelihood function of is a constant function.

(5)

Let $X_1, X_2, \ldots, X_n$ are i.i.d. f(x| , g), such that for each the family f(x| , g) satisfies the condition (N). Then the marginal distribution of (X_1^{-1} X_2, X_1^{-1} $_\theta$ X_3, ... , X_1^{-1} X_n) depends only on . And the conditional distribution of X_1 given (X_1^{-1}

$X_2, X_1^{-1} \theta X_3, \ldots, X_1^{-1} X_n$) satisfies the condition (N).

(6)

Assume $X_1 X_2, \ldots, X_n$ satisfies the conditions in 5. Then the average likelihood function of is proportion to the marginal density of ($X_1^{-1} X_2, X_1^{-1} \theta X_3, \ldots, X_1^{-1} X_n$).

(7)

In the group transformation model, we require that the group is unimodular. This is a reasonable assumption. In fact, the compact groups (e.g. rotation group), finite groups (e.g. finite permutation group), denumerable discrete groups (e.g. integer) and abelian groups (e.g. scale and location groups) are all unimodular.

三、參考文獻

Barndorff-Nielsen, O.(1973), "On Mancillary", Journal of Biometrika(1973), 60, 3, p.447-455.

Barndorff-Nielsen, O.(1976), "Nonformation", Journal of Biometrika(1976), 63, 3, p.567-571.

Fisher, R. A.(1935) "The logic of inductive inference", Journal of the Royal Statistic Society(1935), 98, p.39-54.

Hung, H.N. and Wong W.H. (1996) "Average Likelihood". Technical report, University of Chicago.

Sprott, D. A.(1975), "Marginal and conditional sufficiency", Journal of Biometrika, 62, 3, p.599-605.