A NOTE ON BAHADUR'S TRANSITIVITY

by

E. Greenshtein
Purdue University

Technical Report #90-04

Department of Statistics
Purdue University

January 1990

◇

A Note on Bahadur's Transitivity

E. Greenshtein

Abstract

Let $X_1, X_2, \ldots$ be a sequence of random variables, $(X_1, \ldots, X_n) \sim F_\theta^n$ $\theta \in \Theta$. In a work by Bahadur [1], it was shown that for some sequential problems, an inference should be based on a sequence of sufficient and transitive statistics $S_n = S_n(X_1, \ldots, X_n)$. A simple criterion for transitivity is given in Theorem 1.

Introduction

Let $X_1, X_2, \ldots, X_m$ $m \leq \infty$ be a sequence of random variables, $(X_1, \ldots, X_n) \sim F_\theta^n$, $\theta \in \Theta$. Bahadur has shown in [1] that in a typical sequential decision problem it is enough to consider a sequence of sufficient and transitive statistics. It was shown that: The risk function of any sequential procedure Δ based on $X_1, X_2, \ldots,$ can be achieved by a procedure Δ' based on $S_1, S_2, \ldots$, if $S_1, S_2, \ldots$ is a sequence of transitive sufficient statistics.

Definition: The sequence S_n is transitive if under each of the measures F_θ^m, $\theta \in \Theta$ $S_1, S_2, \ldots$ is a Markov sequence.

An important work dealing with the concept of transitivity is [2]. Some criteria for transitivity are given there using invariance considerations. The simple criterion stated in the following Theorem 1 seems to have been overlooked so far.

Section 1: We will assume in the sequel that:

$$F_\theta^n \ll F_{\theta_0}^n \quad \text{for some} \quad \theta_0 \in \Theta, \quad n = 1, \ldots \tag{1}$$

Let S_n be a sequence of sufficient statistics. Then:

$$dF_\theta^n(s_1, \ldots, s_n) = f_\theta^n(s_n) dF_{\theta_0}^n(s_1, \ldots, s_n) \tag{2}$$

and it is obvious that $S_1, S_2, \ldots$ is a Markov chain under F_θ iff it is a Markov chain under θ_0.

AMS 1985 subject classification 62B05, 62L99, 60J10

Key words and phrases. Sufficiency, transitivity, Markov chain, sequential analysis

Denote $dF_\theta^{n+1}(s_1,\ldots,s_n,s_{n+1}|S_n=s_n)$ the conditional distribution of $S_1,\ldots,S_{n+1}$ conditional on $S_n=s_n$. Then:

$$dF_\theta^{n+1}(s_1,\ldots,s_{n+1}|S_n=s_n)=\frac{f_\theta^{n+1}(s_{n+1})}{f_\theta^n(s_n)}dF_{\theta_0}^{n+1}(s_1,\ldots,s_{n+1}|S_n=s_n) \tag{3}$$

Lemma 1: Suppose $S_1,S_2,\ldots$ is a sequence of complete sufficient statistics for the measures $F_\theta^n(s_1,\ldots,s_n)$ $\theta\in\Theta$. Then for every value s_n, S_{n+1} is complete and sufficient for $S_1,\ldots,S_{n+1}$ in the parametric family $F_\theta^{n+1}(s_1,\ldots,s_{n+1}|S_n=s_n)$ $\theta\in\Theta$.

Proof: Sufficiency is obvious from (3), we will prove completeness.

Notice that $F_{\theta_0}^{n+1}(ds_{n+1})\gg F_{\theta_0}^{n+1}(ds_{n+1}|s_n)$. Here $F_{\theta_0}^{n+1}(ds_{n+1})$ and $F_{\theta_0}^{n+1}$ $(ds_{n+1}|s_n)$ are respectively the marginal distribution of s_{n+1} and the conditional distribution of S_{n+1} conditional upon $S_n=s_n$, under $F_{\theta_0}(ds_1,\ldots,ds_m)$. Hence we can write $F_{\theta_0}^{n+1}(ds_{n+1}|s_n)=\varphi(s_{n+1})F_{\theta_0}^{n+1}(ds_{n+1})$.

Let $h(s_{n+1})$ be a real valued function, then:

$$\int h(s_{n+1})\frac{f_\theta^{n+1}(s_{n+1})}{f_\theta^n(s_n)}\,dF_{\theta_0}^{n+1}(s_{n+1}|s_n)=0$$

if and only if

$$\int h(s_{n+1})f_\theta^{n+1}(s_{n+1})\varphi(s_{n+1})dF_{\theta_0}^{n+1}(s_{n+1})=0.$$

By completeness of S_{n+1}, the latest is true for every θ implies $h(s_{n+1})\cdot\varphi(s_{n+1})\equiv 0$; this implies $h(s_{n+1})\equiv 0$ for s_{n+1} belongs to the support of $F_{\theta_0}^{n+1}(ds_{n+1}|s_n)$, the proof now follows.

Theorem 1: Suppose $S_1,S_2,\ldots$ is a sequence of complete sufficient statistics. Then $S_1,S_2,\ldots$ is transitive.

Proof: Consider the family $dF_\theta^{n+1}(s_1,\ldots,s_{n+1}|s_n)$ $\theta\in\Theta$. By sufficiency of S_n, $(S_1,\ldots,S_{n-1})$ is ancillary for this family. By Lemma 1 S_{n+1} is complete and sufficient. Hence by Basu Lemma [3] S_{n+1} and $(S_1,\ldots,S_{n-1})$ are independent conditional on $S_n=s_n$.

Section 2: Examples

Most of the examples in [2], can be derived by applying Theorem 1. We will consider two examples; the first is taken from [2].

Example 1: Let $\{X_i\}$ be i.i.d. $X_i \sim N(\mu, \sigma^2)$. Let $\theta = \frac{\mu}{\sigma}$. Let $S_n = \frac{\overline{X}_n}{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}$, then S_n is sufficient for $S_1, \ldots, S_n$ when the parameter of interest is θ, see [4] [2]. Presenting $N(\mu, \sigma^2)$ as an exponential family, with $(\sum^n X_i, \sum^n X_i^2)$ a minimal sufficient statistics for $(\frac{\mu}{\sigma}, \frac{1}{\sigma})$ we get $(\sum^n X_i, \sum^n X_i^2)$ is complete. Hence any function of $(\sum^n X_i, \sum^n X_i^2)$ is complete, in particular $S_n(\sum^n X_i, \sum^n X_i^2) = \frac{\overline{X}_n}{\sum^n (X_i - \overline{X}_n)^2}$. By Theorem 1 S_n is transitive.

Example 2: Let $(X_1, \ldots, X_m) \sim N(\theta, \Sigma)$, Σ known θ unknown. A sequence of complete minimal sufficient statistics exists, using the exponential family presentation. This sequence is transitive, by Theorem 1.

Acknowledgment

The observation stated as Theorem 1, was made thanks to a discussion with Professor L. D. Brown.

References

[1] Bahadur, R. R. (1954). Sufficiency and Statistical Decision functions. *Ann. Math. Stat.* 423–462.

[2] Ghosh, J. K. Hall, W. J., and Wijsman, R. A. (1965). The relationship between sufficiency and invariance, with applications in sequential analysis. *Ann. Math. Stat.* 575–614.

[3] Basu, D. (1958). "On Statistics independent of sufficient statistics", *Sankhya* **20** 223–226.

[4] E. L. Lehman, Testing Statistical Hypothesis (1959).

Professor E. Greenshtein
Department of Statistics
Purdue University
West Lafayette, IN 47906