The Annals of Statistics
1993, Vol. 21, No. 4, 2163–2167

A NOTE ON BAHADUR'S TRANSITIVITY

BY EITAN GREENSHTEIN

Hebrew University and Ben Gurion University

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 it was shown that, for some sequential problems, an inference may be based on a sequence of sufficient and transitive statistics $S_n = S_n(X_1, \ldots, X_n)$ without any loss in statistical performance. A simple criterion for transitivity is given in Theorem 1.

1. Introduction. Let $X_1, \ldots, X_m$, $m \le \infty$, be a sequence of random variables, $(X_1, \ldots, X_n) \sim F_\theta^n$, $\theta \in \Theta$. Bahadur [1] has shown that in a typical sequential decision problem it is enough to consider a sequence of sufficient statistics which is transitive under every $\theta \in \Theta$. The reason is that the risk function of any sequential procedure Δ whose decision at the nth stage is a function of $X_1, \ldots, X_n$ can be achieved by a procedure Δ' whose decision at the nth stage is a function of $S_n(X_1, \ldots, X_n)$, if and only if $S_1, S_2, \ldots$ is a sequence of sufficient statistics which is transitive under every $\theta \in \Theta$.

DEFINITION 1. The sequence $\{S_n\}$ is transitive if the conditional distribution of S_{n+1}, conditional upon $X_1, \ldots, X_n$, is a function of $X_1, \ldots, X_n$ only through S_n.

A closely related concept is that the sequence $S_1, S_2, \ldots$ is a Markov sequence. When there is a one-to-one map from $S_1, \ldots, S_n$ to $X_1, \ldots, X_n$, the sequence $S_1, S_2, \ldots$ is Markovian if and only if it is transitive.

An important work dealing with the concept of transitivity is by Ghosh, Hall and Wijsman [3]. Some criteria for transitivity are given there, using invariance considerations.

In Theorem 1 we give a criterion for a sequence of sufficient statistics to be transitive. The same proof can be used to show that such a sequence is a Markov sequence.

2. Main result. We assume that $F_\theta^n \ll F_{\theta_0}^n$ for every n and for some $\theta_0 \in \Theta$, and that $X_1, \ldots, X_n$ is Euclidean for $n = 1, 2, \ldots$. Hence conditional distributions are well defined.

Denote by $F_\theta^{n+1}(x_{(n)}, s_{n+1} | s_n)$ the conditional distribution of $(X_{(n)}, S_{n+1})$ conditional on $S_n = s_n$, where $X_{(n)} = X_1, \ldots, X_n$. Let

$$f_\theta^n(s_n) = \frac{d\bar{F}_\theta^n(s_n)}{d\bar{F}_{\theta_0}^n(s_n)},$$

Received February 1990; revised September 1992.
AMS 1991 *subject classifications*. Primary 62L10; secondary 62B99.
Key words and phrases. Sufficiency, transitivity.

where $\bar{F}_\theta^n(s_n)$ is the distribution of S_n under θ. Finally, we denote by $\hat{F}_\theta^{n+1}(s_{n+1}|s_n)$ the conditional distribution of S_{n+1} conditional on $S_n = s_n$, and by $\tilde{F}_\theta^n(x_{(n)}, s_n, s_{n+1})$ the joint distribution of $X_{(n)}$, S_n and S_{n+1}.

LEMMA 1. *Let $S_n(X_1, \dots, X_n)$, $n = 1, 2, \dots$, be a sequence of sufficient statistics for $F_\theta^n(x_1, \dots, x_n)$, $\theta \in \Theta$. Then the following hold*:

$$\text{(i)}\quad dF_\theta^{n+1}\big(x_{(n)}, s_{n+1}|s_n\big) = \frac{f_\theta^{n+1}(s_{n+1})}{f_\theta^n(s_n)}\, dF_{\theta_0}^{n+1}\big(x_{(n)}, s_{n+1}|s_n\big)\quad a.e.\ \big(\bar{F}_{\theta_0}^n\big).$$

$$\text{(ii)}\quad d\hat{F}_\theta^{n+1}\big(s_{n+1}|s_n\big) = \frac{f_\theta^{n+1}(s_{n+1})}{f_\theta^n(s_n)}\, d\hat{F}_{\theta_0}^{n+1}\big(s_{n+1}|s_n\big)\quad a.e.\ \big(\bar{F}_{\theta_0}^n\big).$$

PROOF. (i) $d\tilde{F}_\theta^{n+1}(x_{(n)}, s_n, s_{n+1}) = f_\theta^{n+1}(s_{n+1})\, d\tilde{F}_{\theta_0}^{n+1}(x_{(n)}, s_n, s_{n+1})$ by sufficiency. Hence,

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

$$= \frac{f_\theta^{n+1}(s_{n+1})}{f_\theta^n(s_n)}\, dF_{\theta_0}^{n+1}\big(x_{(n)}, s_{n+1}|s_n\big).$$

(ii) Immediate from (i). □

COROLLARY 1. *S_{n+1} is sufficient for θ with respect to the family $F_\theta^{n+1}(x_{(n)}, s_{n+1}|s_n)$, $\theta \in \Theta$.*

THEOREM 1. *Let $S_n(X_1, \dots, X_n)$, $n = 1, 2, \dots$, be a sequence of sufficient statistics for $F_\theta^n(x_1, \dots, x_n)$, $\theta \in \Theta$. Suppose that, for $n = 1, 2, \dots$ and for every value of S_n, S_{n+1} is complete with respect to the family $\hat{F}_\theta^{n+1}(s_{n+1}|s_n)$, $\theta \in \Theta$. Then $\{S_n\}$ is transitive under every $\theta \in \Theta$.*

PROOF. It suffices to show that S_{n+1} and $X_{(n)}$ are independent conditional upon $S_n = s_n$.

By Corollary 1, S_{n+1} is sufficient for θ with respect to $F_\theta^{n+1}(x_{(n)}, s_{n+1}|s_n)$, $\theta \in \Theta$, and by assumption it is complete. By sufficiency of S_n, $X_{(n)}$ is ancillary. Hence by Basu's lemma ([5], page 191), S_{n+1} and $X_{(n)}$ are independent conditional upon $S_n = s_n$. □

In the following two propositions, we will consider the completeness assumption.

PROPOSITION 1. *Suppose $\{S_n\}$ is a sequence of complete and sufficient statistics for $F_\theta^n(x_1, \dots, x_n)$. Suppose $\hat{F}_{\theta_0}^{n+1}(s_{n+1}|s_n) \ll \bar{F}_{\theta_0}^{n+1}(s_{n+1})$ for every* *value of s_n. Then S_{n+1} is complete with respect to $\hat{F}_\theta^{n+1}(s_{n+1}|s_n)$, $\theta \in \Theta$ for every value of s_n.*

PROOF. Let $\varphi_{s_n}^{n+1}$ be such that

$$d\hat{F}_{\theta_0}^{n+1}(s_{n+1}|s_n) = \varphi_{s_n}^{n+1}(s_{n+1})\, d\bar{F}_{\theta_0}^{n+1}(s_{n+1}).$$

Let $h(s_{n+1})$ be any integrable real-valued function. Then

$$\int h(s_{n+1})\, d\hat{F}_\theta^{n+1}(s_{n+1}|s_n) = \int h(s_{n+1}) \frac{f_\theta^{n+1}(s_{n+1})}{f_\theta^n(s_n)}\, d\hat{F}_{\theta_0}^{n+1}(s_{n+1}|s_n).$$

The equality is shown by using Lemma 1(ii). The last expression equals zero for every θ if and only if

$$\int h(s_{n+1})\, f_\theta^{n+1}(s_{n+1}) \varphi_{s_n}^{n+1}(s_{n+1})\, d\bar{F}_{\theta_0}^{n+1}(s_{n+1}) = 0 \quad \text{for every } \theta.$$

By the completeness of S_{n+1} this implies $h\varphi_{s_n}^{n+1} = 0$ a.e. $(\bar{F}_{\theta_0}^{n+1})$, which implies $h = 0$ a.e. $(\hat{F}_{\theta_0}^{n+1}(\cdot|s_n))$, which by using Lemma 1(ii) implies $h = 0$ a.e. $(\hat{F}_\theta^{n+1}(\cdot|s_n))$ for all θ. The proof now follows. □

PROPOSITION 2. *Let $S_n(X_1, \dots, X_n)$ be a sequence of sufficient statistics. Suppose that, for every n, $\bar{F}_\theta^n(s_n)$, $\theta \in \Theta \subseteq \mathbb{R}^k$, is a k-dimensional exponential family with a canonical parameter θ and a canonical observation S_n (for a definition see* [2], *page* 1). *Suppose Θ has a nonvoid interior. Then S_{n+1} is complete with respect to $\hat{F}_\theta^{n+1}(s_{n+1}|s_n)$.*

PROOF. Let $\theta_0 = 0$, w.l.o.g. Then $d\bar{F}_\theta^n(s_n) = \exp(\theta \cdot s_n - \psi_n(\theta))\, d\bar{F}_0^n(s_n)$ and

$$d\hat{F}_\theta^{n+1}(s_{n+1}|s_n) = \frac{\exp(\theta \cdot s_{n+1} - \psi_{n+1}(\theta))\, d\hat{F}_0^{n+1}(s_{n+1}|s_n)}{\exp(\theta \cdot s_n - \psi_n(\theta))} \quad \text{a.e. } (\bar{F}_0^n),$$

by Lemma 1(ii).

Using this presentation we see that $\hat{F}_\theta^{n+1}(s_{n+1}|s_n)$, $\theta \in \Theta$, is an exponential family, where Θ is a canonical (nonvoid) parameter set and S_{n+1} is a canonical observation; hence S_{n+1} is complete. □

3. Examples. Most of the examples in [3] can be derived by applying Theorem 1. We will consider two examples; the first is taken from [3].

EXAMPLE 1. Let $\{X_i\}$ be i.i.d. $X_1 \sim N(\mu, \sigma^2)$. Let $\theta = \mu/\sigma$. Let

$$S_n = \bar{X}_n / \sqrt{\Sigma^n (X_i - \bar{X}_n)^2}.$$

Then S_n is sufficient for $S_1, \dots, S_n$ when the parameter of interest is θ (see

[4], Exercise 9, page 250, and [3]). Presenting $N(\mu, \sigma^2)$ as an exponential family with $(\Sigma^n X_i, \Sigma^n X_i^2)$ a minimal sufficient statistic for $(\mu/\sigma^2, -1/(2\sigma^2))$, we get that $(\Sigma^n X_i, \Sigma^n X_i^2)$ is complete. Hence, any function of $(\Sigma^n X_i, \Sigma^n X_i^2)$ is complete. In particular,

$$S_n\left(\sum^n X_i, \sum^n X_i^2\right) = \bar{X}_n / \sqrt{\Sigma^n (X_i - \bar{X}_n)^2}.$$

By Proposition 1 and Theorem 1, S_n is transitive.

In order to verify the dominance assumption of Proposition 1 note that

$$d\hat{F}_\theta^{n+1}(s_{n+1}|s_n) = \int dF_{(\mu,\sigma)}^{n+1}(s_{n+1}|s_n, \bar{x}_n)\, dG_{(\mu,\sigma)}^n(\bar{x}_n|s_n),$$

where $F_{(\mu,\sigma)}^{n+1}(s_{n+1}|s_n, \bar{x}_n)$ is the conditional distribution of S_{n+1}, conditional on S_n and $\bar{X}_n$; and where $G_{(\mu,\sigma)}^n(\bar{x}_n|s_n)$ is the conditional distribution of $\bar{X}_n$ conditional on S_n. The measure $F_{(\mu,\sigma)}^{n+1}(s_{n+1}|s_n, \bar{x}_n)$ is dominated by the Lebesgue measure (using some algebra and the fact that X_{n+1} is normal); hence the same is true for $\hat{F}_\theta^{n+1}(s_{n+1}|s_n)$. The conclusion follows because of the equivalency of the Lebesgue measure and $\bar{F}_{\theta_0}^{n+1}(s_{n+1})$.

EXAMPLE 2. Let $(X_1, \dots, X_m) \sim N(\theta \cdot 1_m, a\Sigma)$, where $\theta \cdot 1'_m = \theta \cdot (1, \dots, 1) = (\theta, \dots, \theta)$, a is a scalar and Σ is an $m \times m$ covariance matrix. Consider the following cases:

(i) θ unknown, $a\Sigma$ known;
(ii) a unknown, θ and Σ known;
(iii) a and θ unknown, Σ known.

In all three cases, a sequence of sufficient statistics that can be presented as an exponential family exists. Using Proposition 2 and Theorem 1, those sequences are transitive.

We will elaborate on part (i) of the example.

Let $\Sigma_{(n)}$ be the covariance matrix of $(X_1, \dots, X_n) = X'_{(n)}$, $n \le m$, and $\theta \cdot 1_{(n)}$ its expectation vector. Then

$$\begin{aligned} dF_\theta^n(x_1, \dots, x_n) &= \exp\left[-\tfrac{1}{2}(x_{(n)} - \theta \cdot 1_{(n)})' \cdot a^{-1} \cdot \Sigma_{(n)}^{-1}(x_{(n)} - \theta \cdot 1_{(n)})\right] d\mu_n \\ &= \exp\left[\theta \cdot 1'_{(n)} \cdot a^{-1} \cdot \Sigma_{(n)}^{-1} \cdot x_{(n)} - \psi_n(\theta)\right] d\nu_n, \end{aligned}$$

where μ_n, ν_n and ψ_n are implicitly defined.

We see that $S_n = 1'_{(n)} \cdot a^{-1} \cdot \Sigma_{(n)}^{-1} \cdot X_{(n)}$ is a canonical observation from a one-dimensional exponential family, with a canonical parameter θ.

Acknowledgments. The observation stated as Theorem 1 was made thanks to a discussion with Professor L. D. Brown. The author would also like to thank Professor R. A. Wijsman for pointing out errors in an earlier version and for a helpful discussion.

REFERENCES

[1] BAHADUR, R. R. (1954). Sufficiency and statistical decision functions. *Ann. Math. Statist.* **25** 423–462.

[2] BROWN, L. D. (1989). *Fundamentals of Statistical Exponential Families* IMS, Hayward, CA.

[3] GHOSH, J. K., HALL, W. J. and WIJSMAN, R. A. (1965). The relationship between sufficiency and invariance, with application in sequential analysis. *Ann. Math. Statist.* **36** 575–614.

[4] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[5] LEHMANN, E. L. (1986). *Testing Statistical Hypotheses*, 2nd ed. Wiley, New York.

DEPARTMENT OF INDUSTRIAL ENGINEERING
AND MANAGEMENT
BEN GURION UNIVERSITY
P.O. BOX 653, BEER SHEVA 84120
ISRAEL