A NOTE ON TRUNCATION AND SUFFICIENT STATISTICS[1]

BY WALTER L. SMITH

University of North Carolina

1. Introduction and summary. Generalizing earlier observations by Fisher and Hotelling, Tukey [1] showed that if a family of distributions admits a set of sufficient statistics, then the family obtained by truncation to a fixed set, or by a fixed selection, also admits the same set of sufficient statistics (this wording is Tukey's; we give a precise mathematical statement later). Tukey's proof assumed the relevant family of probability measures to be dominated by a fixed measure function and made use of the factorization theorem concerning sufficient statistics in this case. In the present short note we shall first re-prove Tukey's result without assuming domination (and, hence, without appealing to the factorization theorem). Then we shall show that, under general conditions, if a sufficient statistic has one or more of the properties of completeness, bounded completeness, or minimality, before truncation, then it preserves such after truncation.

The treatment is on the lines of the abstract discussion of sufficient statistics given by Halmos and Savage [2]. We shall assume familiarity with the results given in this latter paper. For definitions of completeness, bounded completeness, and minimality, and for a discussion of the significance of these concepts we refer to Lehmann and Scheffé [3].

2. On ϕ-truncation. Let X be an abstract space of elements x, and let $\mathfrak{F}_x$ be a (Borel) field of subsets of X. We write $\{\mu_\theta ; \theta \,\varepsilon\, \Omega\}$ for a family of probability measures on $(X, \mathfrak{F}_x)$, where Ω is an abstract parameter space. The statistic t is a mapping of X onto another abstract space T, that is to say, we suppose for simplicity, with no loss of generality, that T is precisely the range $\{t(x) : x \,\varepsilon\, X\}$ of the mapping t. If $B \subseteq T$, we write $t^{-1}B = \{x : t(x) \,\varepsilon\, B\}$ for the *origin* of B. The class of all $B \subseteq T$ such that $t^{-1}B \,\varepsilon\, \mathfrak{F}_x$ is written $\mathfrak{F}_t$; it is easy to show that $\mathfrak{F}_t$ is a (Borel) field.

We shall write $\mathcal{E}_\theta$ for expectation based on μ_θ. If $f(x)$ is any $(\mathfrak{F}_x)$-measurable function such that $\mathcal{E}_\theta|f(x)| < \infty$, for all θ in some set $\Lambda \subseteq \Omega$, we shall say $f(x)$ is Λ-integrable. If $f(x)$ is Ω-integrable, the conditional expectation $\mathcal{E}_\theta(f(x) \mid t)$ is given by the Radon-Nikodym derivative

$$\mathcal{E}_\theta(f(x) \mid t) = \frac{d\nu_\theta t^{-1}}{d\mu_\theta t^{-1}}, \tag{1}$$

where the measure ν_θ is defined by

$$d\nu_\theta = f(x)\, d\mu_\theta . \tag{2}$$

Received December 19, 1955.

[1] This research was supported in part by the United States Air Force through the Office of Scientific Research of the Air Research and Development Command.

For a proof of this assertion, see [2]. Note that the derivative in (1) is arbitrary on a set of $(\mu_\theta t^{-1})$-measure zero. If it transpires that for each Ω-integrable $f(x)$ the conditional expectation $\mathcal{E}_\theta(f(x) \mid t)$ may be taken as independent of $\theta \in \Omega$, then t is *sufficient* for $\{\mu_\theta\,;\,\theta \in \Omega\}$.

Suppose that $\phi(x)$ is a non-negative Ω-integrable function; define $\Omega_\phi = \{\theta : \mathcal{E}_\theta\phi(x) > 0\}$; and define a new family $\{\mu_\theta^\phi\,;\,\theta \in \Omega_\phi\}$ of probability measures on $(X, \mathfrak{F}_x)$ by the equation

$$d\mu_\theta^\phi = \frac{\phi(x)}{\mathcal{E}_\theta\,\phi(x)}\,d\mu_\theta\,. \tag{3}$$

We call $\{\mu_\theta^\phi\,;\,\theta \in \Omega_\phi\}$ the ϕ-truncation of $\{\mu_\theta\,;\,\theta \in \Omega\}$. When $\phi(x)$ is the characteristic function of some set $A \in \mathfrak{F}_x$, then ϕ-truncation corresponds to truncation to a fixed set, in the usual sense. When $\phi(x)$ is bounded above it is easy to see that it may be assumed to be bounded above by unity, and then ϕ-truncation corresponds to what Tukey [1] has called fixed selection. However, $\phi(x)$ may not be bounded above.

We sometimes write, for brevity, $\{\mu_\theta\}$ for $\{\mu_\theta\,;\,\theta \in \Omega\}$; $\{\mu_\theta^\phi\}$ for $\{\mu_\theta^\phi\,;\,\theta \in \Omega_\phi\}$; and $\{\mu_\theta\}_\phi$ for $\{\mu_\theta\,;\,\theta \in \Omega_\phi\}$.

We shall also write $\mathcal{E}_\theta^\phi$ for expectations based on μ_θ^ϕ; and if $f(x)$ is any $(\mathfrak{F}_x)$-measurable function such that $\mathcal{E}_\theta^\phi|f(x)| < \infty$, for all θ in some set $\Lambda \subseteq \Omega$, we shall say $f(x)$ is Λ^ϕ-integrable. If a statement is followed by an expression like $[\mathfrak{M}]$, where $\mathfrak{M}$ represents a family of probability measures, this will mean that the statement is false, at most, on a set of probability zero for each measure of $\mathfrak{M}$. In this connection let us notice that, since $\{\mu_\theta\}_\phi$ dominates $\{\mu_\theta^\phi\}$, we can always replace $[\{\mu_\theta\}_\phi]$ by $[\{\mu_\theta^\phi\}]$, $[\{\mu_\theta t^{-1}\}_\phi]$ by $[\{\mu_\theta^\phi t^{-1}\}]$.

LEMMA. *If $f(x)$ is Ω_θ^ϕ-integrable, then $\phi(x)f(x)$ is Ω_ϕ-integrable.*

PROOF. If $\theta \in \Omega_\phi$, and we define $X_\phi = \{x : \phi(x) > 0\}$,

$$\begin{aligned}
\mathcal{E}_\theta \mid \phi(x)f(x) \mid &= \int_{X_\phi} \phi(x)\,|f(x)|\,d\mu_\theta \\
&= \{\mathcal{E}_\theta\phi(x)\} \int_{X_\phi} |f(x)|\,d\mu_\theta^\phi \\
&= \{\mathcal{E}_\theta\phi(x)\}\{\mathcal{E}_\theta^\phi\,|f(x)|\} \\
&< \infty.
\end{aligned}$$

THEOREM. (i) *If t is sufficient for $\{\mu_\theta\,;\,\theta \in \Omega\}$, then t is sufficient for $\{\mu_\theta^\phi\,;\,\theta \in \Omega_\phi\}$;* (ii) *if, in addition, t is complete, minimal, for $\{\mu_\theta\,;\,\theta \in \Omega_\phi\}$, then t is complete, minimal, for $\{\mu_\theta^\phi\,;\,\theta \in \Omega_\phi\}$; and a similar remark applies if t is boundedly complete—provided, in this case, that $\phi(x)$ is bounded above.*

(Notice that we require t to be complete, etc., for the subfamily $\{\mu_\theta\,;\,\theta \in \Omega_\phi\}$. It is possible that t be complete, etc., for $\{\mu_\theta\,;\,\theta \in \Omega\}$ and not so for $\{\mu_\theta\,;\,\theta \in \Omega_\phi\}$.)

PROOF. We observe first that by (1), (2), and (3), for $\theta \in \Omega_\phi$,

$$\frac{d\mu_\theta^\phi t^{-1}}{d\mu_\theta\, t^{-1}} = \frac{\mathcal{E}(\phi(x) \mid t)}{\mathcal{E}_\theta(x)}\,, \tag{4}$$

where we may omit the suffix θ in $\mathcal{E}_\theta(\phi(x) \mid t)$ because t is sufficient for $\{\mu_\theta\,;\,\theta \in \Omega\}$.

Let $f(x)$ be any Ω_ϕ^ϕ-integrable function. Then by the lemma, $\phi(x)f(x)$ is Ω_ϕ-integrable, and by the sufficiency of t for $\{\mu_\theta ; \theta \varepsilon \Omega\}$, and hence for $\{\mu_\theta ; \theta \varepsilon \Omega_\phi \subseteq \Omega\}$, we may write $\mathcal{E}\{f(x)\phi(x) \mid t\}$ independent of $\theta \varepsilon \Omega_\phi$. Thus for any $A \varepsilon \mathfrak{F}_t$ we have for all $\theta \varepsilon \Omega_\phi$,

$$\int_A \mathcal{E}(f(x)\phi(x) \mid t) \, d\mu_\theta t^{-1} = \int_{t^{-1}A} f(x)\phi(x) \, d\mu_\theta$$

$$= \{\mathcal{E}_\theta \phi(x)\} \int_{t^{-1}A} f(x) \, d\mu_\theta^\phi, \text{ by (2.3)},$$

$$= \{\mathcal{E}_\theta \phi(x)\} \int_A \mathcal{E}_\theta^\phi(f(x) \mid t) \, d\mu_\theta^\phi t^{-1},$$

$$= \int_A \mathcal{E}_\theta^\phi(f(x) \mid t)\mathcal{E}(\phi(x) \mid t) \, d\mu_\theta t^{-1},$$

by (4). Since the last equation holds for every $A \varepsilon \mathfrak{F}_t$, we deduce from the Radon-Nikodym theorem that

$$(5) \qquad \mathcal{E}_\theta^\phi(f(x) \mid t)\mathcal{E}(\phi(x) \mid t) = \mathcal{E}(f(x)\phi(x) \mid t), \quad [\{\mu_\theta t^{-1}\}_\phi].$$

The function $\phi(x)$ is non-negative on X, from which it follows that $\mathcal{E}(\phi(x) \mid t) \geqq 0$, $[\{\mu_\theta t^{-1}\}_\phi]$. Write $T_\phi = \{t: \mathcal{E}(\phi(x) \mid t) > 0\}$; then for $t \varepsilon T - T_\phi$, we have $\mathcal{E}(\phi(x) \mid t) = 0$, $[\{\mu_\theta t^{-1}\}_\phi]$. Thus, if $\theta \varepsilon \Omega_\phi$,

$$\int_{T-T_\phi} d\mu_\theta^\phi t^{-1} = \int_{T-T_\phi} \frac{d\mu_\theta^\phi t^{-1}}{d\mu_\theta t^{-1}} \, d\mu_\theta t^{-1},$$

$$= \int_{T-T_\varphi} \frac{\mathcal{E}(\phi(x) \mid t)}{\mathcal{E}_\theta \theta(x)} \, d\mu_\theta t^{-1}, \qquad \text{by (4)},$$

$$= 0.$$

We have therefore shown that $\mathcal{E}(\phi(x) \mid t) > 0$, $[\{\mu_\theta^\phi t^{-1}\}]$, and may deduce from (5) that

$$\mathcal{E}_\theta^\phi(f(x) \mid t) = \frac{\mathcal{E}(f(x)\phi(x) \mid t)}{\mathcal{E}(\phi(x) \mid t)}, \qquad [\{\mu_\theta^\phi t^{-1}\}].$$

Hence for any Ω_ϕ^ϕ-integrable function $f(x)$ there exists a version of $\mathcal{E}_\theta^\phi(f(x) \mid t)$ which is independent of $\theta \varepsilon \Omega_\phi$. This is enough to prove that t is sufficient for $\{\mu_\theta^\phi ; \theta \varepsilon \Omega_\phi\}$.

Next suppose that t is a complete sufficient statistic for $\{\mu_\theta ; \theta \varepsilon \Omega_\phi\}$. To prove that t is complete for $\{\mu_\theta^\phi ; \theta \varepsilon \Omega_\phi\}$, we must show that if $\psi(t)$ is an arbitrary $\mathfrak{F}_t$-measurable function such that $\mathcal{E}_\theta^\phi \psi(t(x)) = 0$ for all $\theta \varepsilon \Omega_\phi$, then $\psi(t) = 0$, $[\{\mu_\theta^\phi t^{-1}\}]$. However, if

$$\int_T \psi(t) \, d\mu_\theta^\phi t^{-1} = 0, \qquad \text{all } \theta \varepsilon \Omega_\phi,$$

it follows from (4) that

$$\int_T \psi(t)\mathcal{E}(\phi(x) \mid t)\, d\mu_\theta t^{-1} = 0, \qquad \text{all } \theta \in \Omega_\phi .$$

But t is complete for $\{\mu_\theta ; \theta \in \Omega_\phi\}$, and $\psi(t)\mathcal{E}(\phi(x) \mid t)$ is an $(\mathfrak{F}_t)$-measurable function of t. Hence

$$\psi(t)\mathcal{E}(\phi(x) \mid t) = 0, \qquad [\{\mu_\theta t^{-1}\}_\phi].$$

Since we have already seen that $\mathcal{E}(\phi(x) \mid t) > 0$, $[\{\mu_\theta t^{-1}\}_\phi]$, the proof that t is complete for $\{\mu_\theta^\phi ; \theta \in \Omega_\phi\}$ is thus evident.

When t is boundedly complete we can employ precisely the same argument, assuming both $\psi(t)$ and $\phi(x)$ to be bounded so as to ensure, as is easily checked, that $\psi(t)\mathcal{E}(\phi(x) \mid t)$ is bounded $[\{\mu_\theta t^{-1}\}_\phi]$.

Lastly we deal with the minimality question. Suppose that $s(x)$ is any statistic defined on X which is sufficient for $\{\mu_\theta^\phi ; \theta \in \Omega_\phi\}$. Write $S = \{s(x) : x \in X\}$ for the abstract space on which s maps X. Then to prove that t is minimal we must demonstrate the existence of a mapping h of S on T such that $t(x) = h(s(x))$, $[\{\mu_\theta^\phi t^{-1}\}]$.

Recall $X_\phi = \{x : \phi(x) > 0\}$, and notice that since $\theta \in \Omega_\phi$ implies $\mathcal{E}_\theta\phi(x) > 0$, it also implies that $\mu_\theta(X_\phi) > 0$. Plainly, $\mu_\theta^\phi(X - X_\phi) = 0$ for all $\theta \in \Omega_\phi$; thus it will be enough to prove the relation $t(x) = h(s(x))$, $[\{\mu_\theta^\phi t^{-1}\}]$, merely on X_ϕ. To this end, let us define a new statistic

$$\begin{aligned} s_1(x) &= s(x) && \text{if} \quad x \in X_\phi , \\ &= x && \text{if} \quad x \in X - X_\phi . \end{aligned}$$

Since $s(x)$ is obviously a one-valued function of $s_1(x)$, it follows that $s_1(x)$ is also a sufficient statistic for $\{\mu_\theta^\phi ; \theta \in \Omega_\phi\}$ (by Theorem 6.4 of Bahadur [4]). We write $S_1 = \{s_1(x); x \in X\}$ and $\mathfrak{F}_{s_1}$ for the (Borel) field of all subsets $B \subseteq S_1$ such that $s_1^{-1}B \in \mathfrak{F}_x$.

If we set

$$\begin{aligned} \psi(x) &= \frac{1}{\phi(x)}, && x \in X_\phi , \\ &= 0, && x \in X - X_\phi . \end{aligned}$$

then $\psi(x)$ is a non-negative $(\mathfrak{F}_x)$-measurable function on X; and for $\theta \in \Omega_\phi$,

$$\begin{aligned} \mathcal{E}_\theta^\phi\psi(x) &= \int_{X_\phi} \frac{1}{\phi(x)}\, d\mu_\theta^\phi , \\ &= \int_{X_\phi} \frac{1}{\phi(x)} \frac{\phi(x)}{\mathcal{E}_\theta\,\phi(x)}\, d\mu_\theta , \qquad \text{by (3)}, \\ &= \frac{\mu_\theta(X_0)}{\mathcal{E}_\theta\,\phi(x)} < \infty . \end{aligned} \tag{6}$$

It follows from (6) that $\psi(x)$ is Ω_ϕ^ϕ-integrable, and so we may consider the ψ-truncation of $\{\mu_\theta^\phi ; \theta \in \Omega_\phi\}$. We shall employ obvious extensions of our notation, and observe that by (6), $\Omega_{\phi\psi}^\phi = \{\theta; \theta \in \Omega_\phi , \mathcal{E}_\theta^\phi\psi(x) > 0\} = \Omega_\phi$. Hence, for all $\theta \in \Omega_\phi$,

$$
\begin{aligned}
(7)\qquad d\mu_\theta^{\phi\psi} &= 0, && \text{if } x \,\varepsilon\, X - X_\phi ,\\
&= \frac{\psi(x)}{\mathcal{E}_\theta^\phi \psi(x)}\, d\mu_\theta^\phi ,\\
&= \frac{\mathcal{E}_\theta\, \phi(x)}{\phi(x)\mu_\theta(X_\phi)}\, \frac{\phi(x)}{\mathcal{E}_\theta \phi(x)}\, d\mu_\theta , && \text{by (3) and (6)},\\
(8)\qquad &= \frac{d\mu_\theta}{\mu_\theta(X_\phi)} , && \text{if } x \,\varepsilon\, X_\phi .
\end{aligned}
$$

Because s_1 is sufficient for $\{\mu_\theta^\phi ;\ \theta \,\varepsilon\, \Omega_\phi\}$, and because $\Omega_{\phi\psi}^\phi = \Omega_\phi$, it follows from the first part of our theorem that s_1 is also sufficient for the ψ-truncation $\{\mu_\theta^{\phi\psi} ;\ \theta \,\varepsilon\, \Omega_\phi\}$. Let $f(x)$ be any Ω_ϕ-integrable function. Then we notice that for $\theta \,\varepsilon\, \Omega_\phi$, by (7) and (8),

$$\mathcal{E}_\theta^{\phi\psi}\, |f(x)| = \frac{1}{\mu_\theta(X_\phi)} \int_{X_\phi} |f(x)|\, d\mu_\theta < \infty,$$

i.e., $f(x)$ is $\Omega_\theta^{\phi\psi}$-integrable; and by the sufficiency of s_1 there must exist a function $g(s_1) = \mathcal{E}^{\phi\psi}(f(x) \mid s_1)$ which is independent of $\theta \,\varepsilon\, \Omega_\phi$, is such that $g(s_1(x))$ is $(\mathfrak{F}_x)$-measurable, and is such that for any $B \,\varepsilon\, \mathfrak{F}_{s_1}$ and all $\theta \,\varepsilon\, \Omega_\phi$,

$$\int_{s_1^{-1}B} f(x)\, d\mu_\theta^{\phi\psi} = \int_{s_1^{-1}B} g(s_1(x))\, d\mu_\theta^{\phi\psi}.$$

This implies, by (7), that

$$\int_{X_\phi \cap s_1^{-1}B} f(x)\, d\mu_\theta^{\phi\psi} = \int_{X_\phi \cap s_1^{-1}B} g(s_1(x))\, d\mu_\theta^{\phi\psi},$$

and so, by (8), that

$$(9)\qquad \int_{X_\phi \cap s_1^{-1}B} f(x)\, d\mu_\theta = \int_{X_\phi \cap s_1^{-1}B} g(s_1(x))\, d\mu_\theta .$$

Finally, define an $(\mathfrak{F}_x)$-measurable function

$$
\begin{aligned}
g^*(s_1(x)) &= g(s_1(x)) && \text{if } s_1 \,\varepsilon\, S, \text{ i.e., if } x \,\varepsilon\, X_\phi\\
&= g(s_1(x))\ (= f(x)) && \text{if } s_1 \,\varepsilon\, X - X_\phi , \text{ i.e., if } x \,\varepsilon\, X - X_\phi .
\end{aligned}
$$

Thus

$$
\begin{aligned}
\int_{s_1^{-1}B} f(x)\, d\mu_\theta &= \int_{X_\phi \cap s_1^{-1}B} f(x)\, d\mu_\theta + \int_{(X - X_\phi)\cap s_1^{-1}B} f(x)\, d\mu_\theta\\
&= \int_{X_\phi \cap s_1^{-1}B} g(s_1(x))\, d\mu_\theta + \int_{(X - X_\phi)\cap s_1^{-1}B} g^*(s_1(x))\, d\mu_\theta\\
&= \int_{s_1^{-1}B} g^*(s_1(x))\, d\mu_\theta ,
\end{aligned}
$$

the last equality following from (9). Since $f(x)$ is an arbitrary Ω_ϕ-integrable function, this last equation, being true for all $B \varepsilon \mathfrak{F}_{s_1}$, shows that s_1 is a sufficient statistic for $\{\mu_\theta ; \theta \varepsilon \Omega_\phi\}$. But we are given that t is minimal sufficient for $\{\mu_\theta ; \theta \varepsilon \Omega_\phi\}$. Hence there is a mapping h of S_1 onto T such that $t(x) = h(s_1(x))$, $[\{\mu_\theta t^{-1}\}_\phi]$. If we now restrict x to X_ϕ it is evident that $t(x) = h(s(x))$, $[\{\mu_\theta^\phi t^{-1}\}]$, as was tc be proved.

REFERENCES

[1] J. W. TUKEY, "Sufficiency, truncation, and selection," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 309–311.
[2] P. R. HALMOS AND L. J. SAVAGE, "Application of the Radon-Nikodym theorem to the theory of sufficient statistics," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 225–241.
[3] E. L. LEHMANN AND H. SCHEFFÉ, "Completeness, similar regions and unbiased estimation, Part I," *Sankhya*, Vol. 10 (1950), pp. 305–340.
[4] R. R. BAHADUR, "Sufficiency and statistical decision functions, " *Ann. Math. Stat.* Vol. 25 (1954), pp. 423–462.

A CENTRAL LIMIT THEOREM FOR MULTILINEAR STOCHASTIC PROCESSES

BY EMANUEL PARZEN

Stanford University

1. Introduction. Let the random sequence $X(t)$ be observed for $t = 1, 2, \cdots$, and let $S(n) = X(1) + \cdots + X(n)$ be its consecutive sums. The random sequence may be said to obey the *classical* central limit theorem if, for any real number a,

$$\lim_{n\to\infty} \text{Prob}\left\{\frac{S(n) - ES(n)}{\sigma[S(n)]} < a\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-1/2x^2}\,dx. \tag{1.1}$$

Because of the importance of the central limit theorem in establishing the properties of statistical tests and estimates, it would appear that in order to develop a satisfactory theory of statistical inference for stochastic processes which are random sequences of dependent random variables, it is necessary to establish a central limit theorem for such processes. Diananda [2] has proved a central limit theorem for discrete parameter stochastic processes which are *linear* processes. We here introduce a class of stochastic processes which we call *multilinear* processes, for which we prove a central limit theorem. The results are capable of extension to the continuous parameter case, but we do not do so here.

Received December 19, 1955; revised August 16, 1956.