SUFFICIENCY AND MODEL-PRESERVING TRANSFORMATIONS

by

D. Basu

University of North Carolina
and
Indian Statistical Institute

Institute of Statistics Mimeo Series No. 420

January 1965

This research was supported by the National Science Foundation Grant No. GP 16-60.

DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
Chapel Hill, N. C.

SUFFICIENCY AND MODEL-PRESERVING TRANSFORMATIONS

by

D. Basu

University of North Carolina and Indian Statistical Institute

1. Introduction and Summary

Given a statistical model $(\underline{X}, \underline{A}, \underline{P})$, where $\underline{P}$ is a family of probability measures on the σ-field $\underline{A}$ of sub-sets of the sample space $\underline{X}$, a mapping T of $\underline{X}$ onto itself will be called model-preserving if

i) T is a one-to-one bimeasurable map of the measurable space $(\underline{X}, \underline{A})$ onto itself.

ii) $PT^{-1}(A) \equiv P(A)$ for all $A \epsilon \underline{A}$ and $P \epsilon \underline{P}$.

Let $\underline{T} = \{T\}$ be the family of all model-preserving transformations. Since the identity map is always model-preserving, the family T is never vacuous. It is easy to check that $\underline{T}$ constitutes a group with the group operation taken as the composition of maps.

For example, consider a sample $x_1, x_2, \ldots x_n$ of n independent observations on a normal variable with mean zero and unknown variance σ^2. In this case, the family $\underline{T}$ includes all linear orthogonal transformations and also includes other non-linear transformations. For instance, if we define

$$y_i = \varphi(x_i)\ |x_i|\ , \qquad (i = 1,2,\ldots n)\ ,$$

where φ is an arbitrary skew-symmetric* function on the real line taking only the two values -1 and +1, then it is not difficult to see that $y_1, y_2, \ldots y_n$ are independent normal variables with means zero and variances σ^2. Again, any orthogonal transformation of the y_i's will leave the model invariant.

---

*$\varphi(x) = -\varphi(-x)$ for all x

Two measurable sets A and B are said to be $\underline{P}$ - equivalent if

$$P(A \Delta B) \equiv 0 \quad \text{for all } P \in \underline{P} ,$$

where Δ is the symmetric difference operator.

Two real valued measurable functions f and g defined on $(\underline{X}, \underline{A})$ are said to be $\underline{P}$-equivalent if the set

$$\{ x | f(x) \neq g(x) \}$$

is $\underline{P}$-equivalent to the null-set.

The set $A \in \underline{A}$ will be called T-invariant if A and $T^{-1}A$ are $\underline{P}$-equivalent. Let $\underline{A}_T$ be the class of T-invariant sets. Since the set-function T^{-1} preserves all set-operations, it follows that $\underline{A}_T$ is a sub-field* of $\underline{A}$. Let $\underline{A}^*$ be the intersection of the family $\{\underline{A}_T\}$ of sub-fields corresponding to different $T \in \underline{T}$. Thus, $\underline{A}^*$ is the sub-field of all measurable sets that are T-invariant for each model-preserving transformation T. The sub-field $\underline{A}^*$ corresponds to the maximal invariant statistic generated by the group $\underline{T}$ of model-preserving transformations.

Since the statistician is in the same situation (vis a vis the model) whether he observes x or Tx, the principle of invariance demands that the decision that he takes should be invariant with respect to each model preserving transformation T. This leads to the invariance reduction of the data by the substitution of the model $(\underline{X}, \underline{A}, \underline{P})$ with the simpler model

$$(\underline{X}, \underline{A}^*, \underline{P}) \quad .$$

On the other hand, in the presence of a sufficient sub-field $\underline{S}$, the principle of sufficiency leads us to the sufficiency reduction

$$(\underline{X}, \underline{S}, \underline{P}) \quad .$$

---

*abbreviation for sub-σ-field.

If $\underline{S}$ happens to be the minimal sufficient statistic then the above will be the maximal sufficiency reduction.

In many situations the maximal sufficiency reduction of the data is the same as the invariance reduction. For instance, consider the example of n independent observations on a normal variable with mean zero and unknown variance. This paper attempts to prove that the maximal sufficiency reduction cannot be less extensive than the invariance reduction. We prove the following two theorems

<u>Theorem 1:</u> If $\underline{S}$ be a boundedly complete (and hence minimal) sufficient sub-field then

$$\underline{S} \subset \underline{A}^* .$$

<u>Theorem 2:</u> If the family $\underline{P}$ of probability measures be dominated and $\underline{M}$ be the minimal sufficient sub-field then

$$\underline{M} \subset \underline{A}^* .$$

2. Proof of Theorem 1

Let $\underline{S}$ be a boundedly complete sufficient sub-field. We need the following well-known lemma.

<u>Lemma 1:</u> If z be a bounded $\underline{A}$-measurable function such that

$$E(z|P) \equiv 0 \quad \text{for all } P \in \underline{P} \quad ,$$

then for any bounded $\underline{S}$-measurable function f

$$E(zf|P) \equiv 0 \quad \text{for all} \quad P \in \underline{P} \quad .$$

We omit the proof of the lemma.

Now, let S and T be arbitrary but fixed members of $\underline{S}$ and $\underline{T}$ respectively, and let

$$S_o = T^{-1}S \quad .$$

Since T leaves the model $(\underline{X}, \underline{A}, \underline{P})$ invariant, we have

$$P(S) \equiv PT^{-1}(S) \equiv P(S_o) \quad \text{for all } P \in \underline{P} \quad ,$$

i.e. $\quad E(I_S - I_{S_o} \mid P) \equiv 0, \quad$ for all $P \in \underline{P}$,

where I_A stands for the indicator of the set A.

Noting that the two functions $I_S - I_{S_o}$ and I_S satisfy the conditions of Lemma 1, we at once have

$$E[(I_S - I_{S_o})I_S | P] \equiv 0 \quad \text{for all } P \in \underline{P}$$

or $\quad P(S) \equiv P(SS_o) \quad$ for all $P \in \underline{P}$,

But $\quad P(S) \equiv P(S_o) \quad$ for all $P \in \underline{P}$.

$\therefore \quad P(S \Delta S_o) \equiv 0 \quad$ for all $P \in \underline{P}$.

In other words, S is T-invariant.

Since S and T were arbitrary members of $\underline{S}$ and $\underline{T}$ respectively, we finally have

$$\underline{S} \subset \underline{A}^* \quad .$$

3. Proof of Theorem 2

Let us assume that $\underline{P}$ is dominated by a σ-finite measure. For the proof of Theorem 2 we need the following two lemmas.

<u>Lemma 2</u>: There exists a countable selection $P_1, P_2, \ldots.$ of measures from $\underline{P}$ such that the measure

$$Q = \Sigma\, c_i P_i \quad ,$$

where $0 < c_i < 1$ and $\Sigma\, c_i = 1$, dominates $\underline{P}$.

Lemma 3: If $f_P = \frac{dP}{dQ}$ and if $\underline{M}$ be the smallest sub-field of $\underline{A}$ such that f_P is $\underline{M}$-measurable for every $P \in \underline{P}$, then $\underline{M}$ is the minmal sufficient sub-field.

We omit the proofs of the above two well-known lemmas.

Let T be an arbitrary but fixed member of $\underline{T}$. Since T leaves each member of $\underline{P}$ invariant it follows at once that T leaves Q invariant also, i.e. the two measures

$$Q \text{ and } QT^{-1}$$

are identical.

Now, for any $A \in \underline{A}$ and $P \in \underline{P}$,

$$P(A) = \int_A f_p(x)dQ \quad (\because f_P = \frac{dP}{dQ})$$

$$= \int_A f_p(x)dQT^{-1} \quad (\because Q = QT^{-1})$$

$$= \int_{T^{-1}A} f_P T(x)dQ \ ,$$

the final step being a standard result in integration theory.

Again, $P(T^{-1}A) = \int_{T^{-1}A} f_P(x)dQ$.

Since $P(A) = P(T^{-1}A)$, we now have the identity

$$\int_{T^{-1}A} f_P(x)dQ \equiv \int_{T^{-1}A} f_P T(x)dQ, \text{ for all } A \in \underline{A} \ .$$

Remembering that T is one-to-one and bimeasurable we can re-write the last identity as

$$\int_A f_P(x)dQ \equiv \int_A f_P T(x)dQ, \quad \text{for all } A \in \underline{A} \ .$$

Thus, for each $P \in \underline{P}$, the two functions $f_P(x)$ and $f_P T(x)$ must be Q-equivalent and hence they must be $\underline{P}$-equivalent.

Let $\underline{M}_P$ be the sub-field generated by f_P, i.e., $\underline{M}_P$ is the smallest sub-field of $\underline{A}$ such that f_P is $\underline{M}_P$-measurable. If M be a typical member of $\underline{M}_P$, then there exists a Borel set B on the real line such that

$$M = f_P^{-1}(B) \quad .$$

Since f_P and $f_P T$ are $\underline{P}$- equivalent it follows that two sets

$$f_P^{-1}(B) \quad \text{and} \quad (f_P T)^{-1}(B) = T^{-1} f_P^{-1}(B)$$

are $\underline{P}$-equivalent.

That is, M is $\underline{P}$-equivalent to $T^{-1}M$. In other words, M is T-invariant. Since M and T were arbitrary members of $\underline{M}_P$ and $\underline{T}$ respectively, we have established that

$$\underline{M}_P \subset \underline{A}^* \qquad \text{for all} \quad P \in \underline{P} \quad .$$

Remembering that $\underline{M}$ is the smallest sub-field that contains all the $\underline{M}_P$'s we now have

$$\underline{M} \subset \underline{A}^* \quad .$$

References

1. Bahadur, R. R. (1954), Sufficiency and Statistical decision functions, Ann. Math. Statist. 25, 425-462.

2. Hall, W. J., Wijsman, R. Q., and Ghosh, J. K. (1964), The relationship between sufficiency and invariance, Institute of Statistics Mimeo Series No. 403.

3. Lehmann, E. L. (1959), Testing Statistical Hypotheses, Wiley, New York.