On statistics which are almost sufficient from the viewpoint of the Fisher metrics

Kaori Yamaguchi*

August 14, 2023

This is joint work with Hiraku Nozawa (Ritsumeikan University).

The notion of sufficient statistics [4] is fundamental in statistics and information theory. A statistic on 2-integrable parametrized measure model is sufficient if the Fisher metric of the induced model coincides with that of the original model [3, 4]. We introduce and study a quantitatively weak version of sufficient statistics of which the Fisher metric of the induced model are bi-Lipschitz equivalent to the original one. Such statistics can be useful when there is no sufficient statistics. We show a Fisher-Neyman type result to characterize such statistic with a Lipschitz condition on the conditional probability or a certain decomposition of the density function.

In this article, we consider the following quantitatively weak version of sufficient statistics.

Definition 1. Let $(M, \Omega, \mathbf{p})$ be a 2-integrable parametrized measure model and $0 \leq \delta \leq 1$. Then a statistic $\kappa : \Omega \to \Omega'$ is called δ-*almost sufficient statistics* for $(M, \Omega, \mathbf{p})$ if we have

$$\delta^2 \mathfrak{g}(v, v) \leq \mathfrak{g}'(v, v) \tag{0.1}$$

for all $v \in T_\xi M$, where $\mathfrak{g}$ and $\mathfrak{g}'$ are the Fisher metrics on M given by $(M, \Omega, \mathbf{p})$ and the model induced by κ, respectively.

By the monotonicity theorem $\mathfrak{g}'(v, v) \leq \mathfrak{g}(v, v)$, the condition (0.1) means that the Fisher metric $\mathfrak{g}'$ of the induced model is bi-Lipschitz equivalent to the Fisher metric $\mathfrak{g}$ of the original model. A 1-almost sufficient statistic is a sufficient statistic. Almost sufficient statistics can be useful when there are no sufficient statistics.

The Fisher-Neyman characterization is as follows to which we include some related characterization results as well.

Theorem 2 (Fisher-Neyman characterization [4], see also [1, 2])**.** *Let $(M, \Omega, \mathbf{p})$ be a parametrized measure model of the form*

$$\mathbf{p}(\xi) = p(\omega; \xi)\mu_0$$

for a probability measure μ_0 such that p is positive. Let $\kappa : \Omega \to \Omega'$ be a statistic. Assume that M is connected. Then k is Fisher-Neyman sufficient for $(M, \Omega, \mathbf{p})$ if and only if there exist a function $s : \Omega' \times M \to \mathbb{R}$ and a function $t \in L^1(\Omega, \mu_0)$ such that for all $\xi \in M$ we have $s(\omega', \xi) \in L^1(\Omega', \kappa_* \mu_0)$ and

$$p(\omega; \xi) = s(\kappa(\omega); \xi)t(\omega), \mu_0\text{-}a.e..$$

Then the following are equivalent:

(i) *κ is Fisher-Neyman sufficient for $(M, \Omega, \mathbf{p})$.*

(ii) *We have $\partial_v \log p = \partial_v \kappa^* \log p'$ for every tangent vector v on M.*

(iii) *The map $\xi \mapsto \frac{p(\cdot;\xi)}{p'(\kappa(\cdot);\xi)}$ is a constant map.*

(iv) *There exist a function $s : \Omega' \times M \to \mathbb{R}$ and a function $t \in L^1(\Omega, \mu_0)$ such that we have $s(\cdot; \xi) \in L^1(\Omega', \kappa_* \mu_0)$ for all $\xi \in M$ and*

$$p(\omega; \xi) = s(\kappa(\omega); \xi)t(\omega), \qquad \mu_0\text{-}a.e.\ and\ \forall \xi \in M.$$

*Graduate School of Science and Engineering, Ritsumeikan University, Nojihigashi 1-1-1, Kusatsu, Shiga, 525-8577, Japan. E-mail-address: `ra0097vv@ed.ritsumei.ac.jp`

Our result is the following Fisher-Neyman type characterization of almost sufficient statistics:

Theorem 3. *Let $(M, \Omega, \mathbf{p})$ be a 2-integrable parametrized measure model of the form*

$$\mathbf{p}(\xi) = p(\omega; \xi)\mu_0$$

for a finite measure μ_0 on Ω such that p is positive. Assume that M equipped with the Fisher quadratic form is a finite dimensional Riemannian manifold of class C^2. Let $\kappa : \Omega \to \Omega'$ be a statistic and $0 < \delta \leq 1$. Then the following are equivalent:

(i) *κ is δ-almost sufficient for $(M, \Omega, \mathbf{p})$.*

(ii) *We have*

$$\left\| \partial_v \log \frac{p(\cdot; \xi)}{\kappa^* p'(\cdot; \xi)} \right\| \leq \sqrt{1 - \delta^2} \|\partial_v \log p(\cdot; \xi)\| \tag{0.2}$$

for every tangent vector v on M, where $\| \cdot \|$ denotes the L^2-norm.

(iii) *The map $M \to L^2(\Omega, \mu_0); \xi \mapsto \log \frac{p(\cdot;\xi)}{p'(\kappa(\cdot);\xi)}$ is locally $\sqrt{1-\delta^2}$-Lipschitz with respect to the distance defined by the Fisher quadratic form and the L^2-metric.*

(iv) *There exist measurable functions $s : \Omega' \times M \to \mathbb{R}$ and $t : \Omega \times M \to \mathbb{R}_{>0}$ such that $\log t(\cdot; \xi) \in L^2(\Omega, \mu_0)$ for $\forall \xi \in M$,*

- *$p(\omega; \xi) = s(\kappa(\omega); \xi)t(\omega; \xi)$ for μ_0-a.e. $\omega \in \Omega, \forall \xi \in M$ and*
- *the map $M \to L^2(\Omega, \mu_0); \xi \mapsto \log t(\cdot; \xi)$ is locally $\sqrt{1-\delta^2}$-Lipschitz with respect to the distance defined by the Fisher quadratic form and the L^2-metric.*

Main points of the proof are that we consider a geodesic $\xi(t)$ with unit speed on M with respect to the Fisher metric such that $\xi(t_0) = \xi_0$ and $\xi(t_1) = \xi_1$ (in the proof for the equivalence of condition (ii) and (iii)) and that, as it is well-known, we use the following fact (in the proof for the equivalence of condition (i) and (ii));

$$\|\kappa^*\phi'\|^2 = \|\phi'\|^2 = \int_{\Omega'} (\phi')^2 d\mu' = \int_{\Omega'} \phi' d\kappa_*(\phi\mu) = \int_{\Omega} (\kappa^*\phi')\phi d\mu = \langle \phi, \kappa^*\phi' \rangle.$$

Therefore, we have

$$\|\phi - \kappa^*\phi'\|^2 = \|\phi\|^2 - \|\kappa^*\phi'\|^2. \tag{0.3}$$

References

[1] Ay, N., Jost, J., Lê, H.V., Schwachhöfer, L., *Parametrized measure models*, Bernoulli **24**(3) (2015) 1692–1725. `https://doi.org/10.3150/16-BEJ910`

[2] Ay, N., Jost, J., Lê, H.V., Schwachhöfer, L., *Information Geometry*, Ergeb. Math. Grenzgeb. (3) Folge. A Series of Modern Surveys in Mathematics, 64. Springer, Cham, 2017. xi+407 pp. ISBN: 978-3-319-56477-7; 978-3-319-56478-4.

[3] Fisher, R.A., On the mathematical foundations of theoretical statistics *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* ontaining Papers of a Mathematical or Physical Character 222, 309–368. `http://doi.org/10.1098/rsta.1922.0009`

[4] Neyman, J., Sur un teorema concernente le considette statistiche sufficienti, *Giorn. Istit. Ital. Att.* **6** (1935) 320–334.

[5] Yamaguchi, K., Nozawa, H., *A Fisher-Neyman type characterization of statistics which are almost sufficient*, Preprint