Insufficiency and the preservation of Fisher information

David Pollard

Yale University

April 2012
NESS meeting

Based on: arXiv:1107.3797 [math.ST]

Slides for this talk at: www.stat.yale.edu/˜pollard/Talks

1. Fisher information (for one observation)

- Family of densities $\mathcal{P} = \{p_\theta(x) : \theta \in \mathbb{R}\}$
- Score function

$$\Delta_\theta(x) = \frac{\partial}{\partial t} \log p_{\theta+t}(x) = 2\frac{\partial}{\partial t}\sqrt{\frac{p_{\theta+t}(x)}{p_\theta(x)}} \qquad \text{at } t = 0$$

- $\mathbb{I}_{\mathcal{P}}(\theta) = \mathbb{E}_\theta \Delta_\theta(x)^2$
- Distribution of statistic T under p_θ has density q_θ
- $\mathcal{Q} = \{q_\theta : \theta \in \mathbb{R}\}$
- $\mathbb{I}_{\mathcal{Q}}(\theta) \leq \mathbb{I}_{\mathcal{P}}(\theta)$ with equality if T is a sufficient statistic

2. Preservation of Fisher information

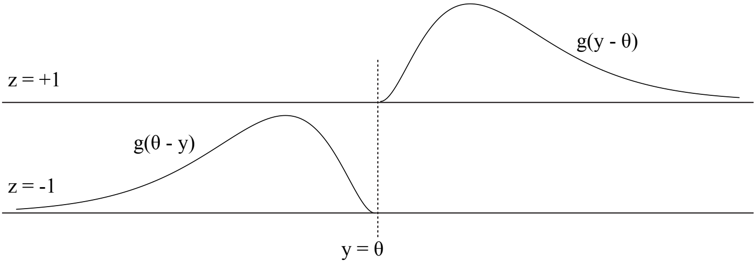
- If $\mathbb{I}_{\mathcal{Q}}(\theta) = \mathbb{I}_{\mathcal{P}}(\theta)$ then is T a sufficient statistic?
- Not always: Kagan and Shepp (2005)
- Counterexample: $x = (y, z) \in \mathbb{R} \times \{-1, +1\}$ with

$$\mathbb{P}\{z = -1\} = 1/2 = \mathbb{P}\{z = +1\} \qquad \text{(ancillary)}$$

$$f_\theta(y \mid z) = \begin{cases} g(y - \theta) & \text{if } z = +1 \\ g(\theta - y) & \text{if } z = -1 \end{cases}$$

where $g(w) = \frac{1}{2} w^2 e^{-w} \{w > 0\}$

3. Counterexample ($\approx$ K&S)



- $\mathbb{P}\{z = -1\} = 1/2 = \mathbb{P}\{z = +1\}$
- y is not sufficient: $P_\theta(z = 1 \mid y) = \{y > \theta\}$ a.s.
- Marginal: $q_\theta(y) = \frac{1}{2}g(y - \theta) + \frac{1}{2}g(\theta - y)$
- $\mathbb{I}_{\mathcal{Q}}(\theta) = \mathbb{I}_{\mathcal{P}}(\theta)$. Why?

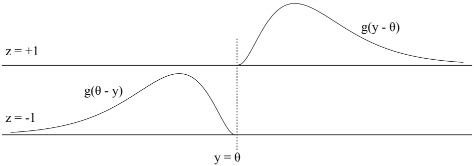
4. True NSC for no loss of Fisher info, in general

- Score function for $\mathcal{Q}$ is (use DQM?):

$$\widetilde{\Delta}_\theta(y) = \mathbb{E}_\theta(\Delta_\theta(x) \mid T(x) = y)$$

- No loss of Fisher information at θ iff $\Delta_\theta(x) = \widetilde{\Delta}_\theta(Tx)$ a.s. [$\approx$ Pitman (1979, pages 19–21)]
- Back to counterexample:

$$\Delta_\theta(x) = -\{z = +1\}\frac{\dot{g}(y-\theta)}{g(y-\theta)}\{y > \theta\} + \{z = -1\}\frac{\dot{g}(\theta - y)}{g(\theta - y)}\{y < \theta\}$$



5. Sample of size n (local asymptotic normality)

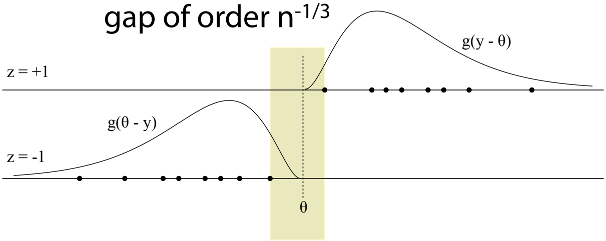
$x_i = (y_i, z_i)$ for $i = 1, \dots, n$

$$\log \prod_{i \le n} \frac{p_{\theta + t/\sqrt{n}}(x_i)}{p_\theta(x_i)} \approx \frac{t}{\sqrt{n}} \sum_{i \le n} \Delta_\theta(x_i) - \frac{t^2}{2} \mathbb{I}_{\mathcal{P}}(\theta)$$

$$\log \prod_{i \le n} \frac{q_{\theta + t/\sqrt{n}}(y_i)}{q_\theta(y_i)} \approx \frac{t}{\sqrt{n}} \sum_{i \le n} \widetilde{\Delta}_\theta(y_i) - \frac{t^2}{2} \mathbb{I}_{\mathcal{Q}}(\theta)$$

Both models are locally asymptotically like the $\{N(t, 1/\mathbb{I}(\theta)) : t \in \mathbb{R}\}$ model, in Le Cam's weak sense.

6. What more do the z_i's tell us about θ?



Root-n consistent estimator $\widehat{\theta}_n(y_1, \ldots, y_n)$.

$$z_i^* := \begin{cases} +1 & \text{if } y_i > \widehat{\theta}_n \\ -1 & \text{if } y_i < \widehat{\theta}_n \end{cases} \qquad x_i^* = (y_i, z_i^*)$$

$$\mathbb{P}_\theta\{z_i = z_i^* \text{ for } i = 1, \ldots, n\} \to 1 \qquad \text{rapidly}$$

7. Approximation in strong Le Cam sense

- Root-n consistent estimator $\widehat{\theta}_n(y_1, \ldots, y_n)$.

$$z_i^* := \begin{cases} +1 & \text{if } y_i > \widehat{\theta}_n \\ -1 & \text{if } y_i < \widehat{\theta}_n \end{cases} \qquad x_i^* = (y_i, z_i^*)$$

$\mathbb{P}_\theta\{z_i = z_i^* \text{ for } i = 1, \ldots, n\} \to 1$ rapidly

  - Only need y_i's to construct x_i^*'s and

$$\mathbb{P}_\theta\{x_i = x_i^* \text{ for } i = 1, \ldots, n\} \to 1 \qquad \text{rapidly}$$

- Inferences based on $x_1^*, \ldots, x_n^*$ are
(with prob tending to one)
the same as inferences based on $x_1, \ldots, x_n$.

References

Kagan, A. and L. A. Shepp (2005). A sufficiency paradox: an insufficient statistic preserving the Fisher information. *The American Statistician 59*, 54–56.

Pitman, E. J. G. (1979). *Some Basic Theory for Statistical Inference*. Monographs on applied probability and statistics. Chapman and Hall.

Pollard, D. (2012). A note on insufficiency and the preservation of Fisher information. In *[Volume in honor of Jon Wellner's 65th birthday]*. Institute of Mathematical Statistics. {arXiv:1107.3797v1 [math.ST]}.