Cstat: Inference and Goodness-of-fit

Yang Chen

University of Michigan

August 2, 2022

Outline

Literature

Plan

Literature

Key Papers

- Cash, W., *Parameter estimation in astronomy through application of the likelihood ratio*, Astrophysical Journal, Part 1, vol. 228, Mar. 15, 1979, p. 939-947.
  - Inference: MLE & confidence intervals
  - Goodness-of-fit Test: χ^2 for difference of likelihood ratios – if there exists a hypothesized fixed subset of parameters?
- Kaastra, J. S. *On the use of C-stat in testing models for X-ray spectra*, Astronomy & Astrophysics 605 (2017): A51.
  - Goodness-of-fit Test: Approximate Gaussian
- Gaps and problems:
  - Numbers of bins larger than number of counts for faint sources
  - Goodness-of-fit tests are approximate
  - Approximate likelihood with discretization

Multiple Attempts

- Asymptotics for C-stat with small counts per bin and large bin count
- Asymptotic/conservative test of goodness
- Dynamic bin split and merge
- Practical implementation with discretized likelihood

Results: Asymptotic Normality

Theorem

Let $\hat{\theta}_n$ be the maximum likelihood estimate. Assume that $\{s_i(\theta_0)\}_{i\geq 1}$ is bounded from above and $\mathrm{rank}\left(\frac{\partial \boldsymbol{s}_{1:n}(\theta)}{\partial \theta}\Big|_{\theta=\theta_0}\right) = d$.

1. *Assume that for all $\theta \in \Theta$, $\sum_{i=1}^{n} [\log s_i(\theta)]^2 = O(n^{1-\alpha})$ for some $\alpha > 0$. Then $\hat{\theta}_n \to \theta_0$ almost surely as $n \to \infty$. Furthermore,*

$$I_n(\theta_0)^{-1}\left[-\frac{\partial^2 \log L_n(\theta)}{\partial\theta\partial\theta^\top}\bigg|_{\theta=\theta_0}\right] \xrightarrow{P} 1 \quad \text{as} \quad n \to \infty.$$

2. *Assume that (a) for any θ in an small neighborhood of θ_0, each $s_i(\theta)$ is second order continuously differentiable and $[\log s_i(\theta)]''$ is uniformly bounded by a finite constant, (b) $[\log n]^2/I_n(\theta_0) \to 0$ as $n \to \infty$; then $\sqrt{I_n(\theta_0)}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, I_d)$ as $n \to \infty$.*

Results: C-stat Property

Lemma

For any n, $-C_n(\hat{\theta}_n) + C_n(\theta_0) = \mathrm{LR}_n^*$*, where* LR_n^* *is given by*

$$\mathrm{LR}_n^* = -2\log\frac{L(s_1(\theta_0),\ldots,s_n(\theta_0)|N_1,\ldots,N_n)}{L(s_1(\hat{\theta}_n),\ldots,s_n(\hat{\theta}_n)|N_1,\ldots,N_n)}$$
$$= 2\sum_{i=1}^{n}\left[N_i\log s_i(\hat{\theta}_n) - N_i\log s_i(\theta_0) + s_i(\theta_0) - s_i(\hat{\theta}_n)\right],$$

which is the likelihood ratio statistics for testing the null hypothesis $H_0 : \theta = \theta_0$ *versus the alternative* $H_1 : \{s_i(\theta), 1 \le i \le n\} \in \mathcal{S}$*. As* $n \to \infty$, $\mathrm{LR}_n^* \xrightarrow{\mathcal{D}} \chi_d^2$*.*

Results: Binning Impacts

Theorem

Performing finer partitions does not decrease the Fisher information. In fact, with any finer partition, the Fisher information increases unless in the following situation.

- *there exists* $s_j^*(\theta), 1 \leq j \leq M$ *and a partition of* $\{1, \ldots, n\}$*, denoted by* $\{\sigma_1, \ldots, \sigma_M\}$*, such that for any* $1 \leq i \leq n$*, there exists* $1 \leq j \leq M$ *such that* $i \in \sigma_j$ *and* $s_i(\theta) = c_{ij} s_j^*(\theta)$ *for some constant* c_{ij}