Robustness measures derived from functionals of the empirical distribution function

William C. Wojciechowski

November 1, 1999

Empirical Distribution Function

- $F_n(x) = n^{-1} \sum_{i=1}^{n} \delta_{x_i}(x)$

- $T_n = T(X_1 \ldots X_n) = T(G_n)$

- $\lim_{n\to\infty} G_n = G$

- Mean: $n^{-1} \sum_{i=1}^{n} x_i$ $\int x dG(x)$

- Median: $x_{n/2}$, $\frac{(x_{n/2}+x_{n/2+1})}{2}$ $G^{-1}(1/2)$

Robust measures from functionals

- $T(G)$ is a function of G

- Continuity of $T(G)$

  Qualitative Robustness

- Derivatives of $T(G)$

  Quantitative robustness

M-Estimators

- $\min_{T_n} \sum_{i=1}^{n} \rho(x_i; T_n)$

- $\sum_{i=1}^{n} \psi(x_i; T_n) = 0$

- $\psi_{mle} = \frac{d \log f_\theta}{d\theta}$

- $\psi_{median} = sign(x)$

- $\psi_{mpde} = uf_\theta^\alpha - \int uf_\theta^{\alpha+1} \qquad u = \psi_{mle}$

Influence Function

- $IF(x) = lim_{\varepsilon \to 0} \frac{T((1-\varepsilon)F + \varepsilon\delta(x)) - T(F)}{\varepsilon}$

  $\frac{d}{d\epsilon}[T((1-\epsilon)F + \epsilon\delta)]_{\epsilon=0}$

- $IF(x)$ for M-Estimators

- $$\frac{\psi(x;T(F))}{-\int(\partial/\partial\theta)\psi(x;T(F))F(dx)}$$

- $\int \psi(x, T(G))dG(x) = 0$

Theoretical IF(z)

- $IF(z, mean, \Phi) = z$

- $IF(z, median, \Phi) = \frac{sign(z)}{2\phi(0)}$

- $IF(z, mpde, \Phi) = J^{-1}(u_\theta(z) f_\theta(z)^\alpha - \eta)$

  $J = \int u_\theta(z)^2 f_\theta(z)^{1+\alpha} dz$

Change in Variance Function

- $CVF(z,T,F) = \frac{d}{dt}V(T,(1-\epsilon)F+\epsilon\delta(z))_{\epsilon=0}$

- CVF for M-Estimators

$$CVF(z,\psi,F) = \frac{A(\psi)}{B(\psi)^2}(1+\frac{\psi(z)^2}{A(\psi)}-2\frac{\psi'(z)}{B(\psi)})$$

$$A(\psi) = \int \psi^2 dF$$

$$B(\psi) = \int \frac{d}{d\theta}\psi dF$$

Theoretical CVF(z)

- $CVF(z, mean, \Phi) = z^2 - 1$

- $CVF(z, median, \Phi) =$

  $\frac{2}{(2\phi(0))^2}(I_{\Re-(0)}(z) - \frac{1}{\phi(0)}\delta_0(z))$

- $CVF(z, mpde, \Phi) =$ very complicated!

Dataset

- Cushny and Peebles (1905)

- Difference between two drugs effect on prolongation of sleep

- Data used by...

  Student (1908)

  Fisher (1925)

  Anderson (1958)

- 0.0, 0.8, 1.0, 1.2, 1.3, 1.3, 1.4, 1.8, 2.4, 4.6