Quantile Regression

Ruibin Xi

Inference about quantiles (1)

- For the τth quantile $\hat{\xi}_\tau$,we have shown that under certain regularity conditions

$$\sqrt{n}(\hat{\xi}_\tau - \xi_\tau) \rightsquigarrow \mathcal{N}(0, \omega^2)$$

where $\omega^2 = \tau(1-\tau)/f^2(\xi_\tau)$

- More generally, for m different quantiles,

$$\hat{\zeta}_n = (\hat{\xi}_{\tau_1}, \ldots, \hat{\xi}_{\tau_m})$$

we have $\sqrt{n}(\hat{\zeta}_n - \zeta) \rightsquigarrow \mathcal{N}(0, \Omega)$

where

$$(\omega_{ij}) = (\tau_i \wedge \tau_j - \tau_i \tau_j)/(f(F^{-1}(\tau_i))f(F^{-1}(\tau_j)))$$

Inference about quantiles (2)

- Consider the hypothesis testing problem

$$H_0 : \xi(\tau) = \xi_0(\tau)$$

- We may construct a test using the asymptotic result, but this involves the unknown quantity $f^2(\xi_\tau)$

Inference about quantiles (3)

- By differentiating the identity

$$F(F^{-1}(t)) = t$$

- we get

$$\frac{d}{dt}F(F^{-1}(t)) = f(F^{-1}(t))\frac{d}{dt}F^{-1}(t) = 1$$

- and

$$\frac{d}{dt}F^{-1}(t) = \frac{1}{f(F^{-1}(t))}$$

Inference about quantiles (4)

- Thus, we may estimate $f^{-1}(\xi_\tau)$ by

$$
\begin{aligned}
f^{-1}(\xi_\tau) &= \frac{1}{f(F^{-1}(\tau))} \\
&\approx \frac{F^{-1}(\tau + h_n) - F^{-1}(\tau - h_n)}{2h_n} \\
&\approx \frac{F_n^{-1}(\tau + h_n) - F_n^{-1}(\tau - h_n)}{2h_n}
\end{aligned}
$$

Inference about quantiles (5)

- Alternatively, we may consider

$$Z_n = \sum_{i=1}^{n} I(Y_i - \xi_0(\tau))$$

which is binomial $B(n, \tau)$

- We reject H_0 if $T_n = n^{-1} Z_n - \tau$ is too large in absolute value

Inference about quantiles (6)

- Confidence interval

$$C_\alpha = \{\xi : T_n(\xi) \text{ does not reject at level } \alpha\}.$$

since $T_n(\xi)$ is piecewise constant and monotone non-decreasing

$$C_\alpha = [Y_{(L)}, Y_{(U)}]$$

- In Large sample

$$\{L, U\} = n\tau \pm c_{\alpha/2}\sqrt{n\tau(1-\tau)}$$

where $c_{\alpha/2} = \Phi^{-1}(1 - \alpha/2)$

Quantile Regression Asymptotic (1)

- In linear regression $y_i = x_i^\top \beta + u_i$

with iid error $\{u_i\}$.

- Assume $\{u_i\}$ have the same distribution F
- Then, the joint asymptotic distribution of

$$\hat{\zeta}_n = (\hat{\beta}_n(\tau_1)^\top, \ldots, \hat{\beta}_n(\tau_m)^\top)^\top$$

is $\sqrt{n}(\hat{\zeta}_n - \zeta) = (\sqrt{n}(\hat{\beta}_n(\tau_j) - \beta(\tau_j)))_{j=1}^m \rightsquigarrow \mathcal{N}\big(0, \Omega \otimes Q_0^{-1}\big)$

$n^{-1} \sum x_i x_i^\top \equiv Q_n$ converges to Q_0

$(\omega_{ij}) = (\tau_i \wedge \tau_j - \tau_i \tau_j)/(f(F^{-1}(\tau_i))f(F^{-1}(\tau_j)))$

Quantile Regression Asymptotic (2)

- If $\{u_i\}$ are not i.i.d.

$$\sqrt{n}(\hat{\beta}(\tau) - \beta(\tau)) \rightsquigarrow \mathcal{N}\left(0, \tau(1-\tau)H_n^{-1} J_n H_n^{-1}\right)$$

$$J_n(\tau) = n^{-1} \sum_{i=1}^{n} x_i x_i^\top$$

$$H_n(\tau) = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} x_i x_i^\top f_i(\xi_i(\tau))$$

- At different quantiles, the covariance matrix has the blocks

$$\begin{aligned} &\text{Acov}(\sqrt{n}(\hat{\beta}(\tau_i) - \beta(\tau_i), \sqrt{n}(\hat{\beta}(\tau_j) - \beta(\tau_j))) \\ &= [\tau_i \wedge \tau_j - \tau_i \tau_j] H_n(\tau_i)^{-1} J_n H_n(\tau_j)^{-1}, \end{aligned}$$

Quantile regression inference (1)

- Consider the two sample model

$$Y_i = \alpha_1 + \alpha_2 x_i + u_i$$

where $x_i = 0$ for n_1 observations from the 1st sample, and $x_i = 1$ for n_2 observations from the 2nd sample

Quantile regression inference (2)

- Testing from equality of the slope parameter across two quantiles τ_1 and τ_2 is equivalent to test

$$
\begin{aligned}
\alpha_2(\tau_2) - \alpha_2(\tau_1) &= (Q_2(\tau_2) - Q_1(\tau_2)) - (Q_2(\tau_1) - Q_1(\tau_1)) \\
&= (Q_2(\tau_2) - Q_2(\tau_1)) - (Q_1(\tau_2) - Q_1(\tau_1)) \\
&= 0,
\end{aligned}
$$

Quantile regression inference (3)

- The aymptotic variance of

$$\hat{\alpha}_2(\tau_2) - \hat{\alpha}_2(\tau_1)$$

is given by

$$\sigma^2(\tau_1, \tau_2) = \left[\frac{\tau_1(1-\tau_1)}{f^2(\xi_1)} - 2\frac{\tau_1(1-\tau_2)}{f(\xi_1)f(\xi_2)} + \frac{\tau_2(1-\tau_2)}{f^2(\xi_2)}\right]\left[\frac{n}{nn_2 - n_2^2}\right]$$

where $\xi_i = F^{-1}(\tau_i)$

- A test for the null hypothesis can be given based on the asymptotic normality of

$$T_n = (\hat{\alpha}_2(\tau_2) - \hat{\alpha}_2(\tau_1))/\hat{\sigma}(\tau_1, \tau_2)$$

Quantile regression inference (4)

- A test for the null hypothesis can be given based on the asymptotic normality of

$$T_n = (\hat{\alpha}_2(\tau_2) - \hat{\alpha}_2(\tau_1))/\hat{\sigma}(\tau_1, \tau_2)$$

where $\hat{\sigma}(\tau_1, \tau_2)$ is a consistent estimator of $\sigma^2(\tau_1, \tau_2)$

Quantile regression inference (5)

- General linear hypothesis on

$$\zeta = (\beta(\tau_1)^\top, \ldots, \beta(\tau_m)^\top)^\top$$

of the form

$$H_0: \; R\zeta = r$$

- The test statistic

$$T_n = n(R\hat{\xi} - r)^T [RV_n^{-1}R]^{-1}(R\hat{\xi} - r)$$

where V_n is the $mp \times mp$ matrix with ijth block

$$V_n(\tau_i, \tau_j) = [\tau_i \wedge \tau_j - \tau_i \tau_j] H_n(\tau_i)^{-1} J_n(\tau_i, \tau_j) H_n(\tau_j)^{-1}$$

Quantile regression inference (6)

The statistic T_n is asymptotically χ_q^2 under H_0 where q is the rank of the matrix R.

Estimation of covariance matrices(1)

- Need to estimate the sparsity function

$$s(\tau) = [f(F^{-1}(\tau))]^{-1}$$

which may be estimated by

$$\hat{s}_n(t) = [\hat{F}_n^{-1}(t + h_n) - \hat{F}_n^{-1}(t - h_n)]/2h_n$$

- In case of linear quantile regression

$$\hat{f}_i(x_i^\top \hat{\beta}(\tau)) = 2h_n / (x_i^\top (\hat{\beta}(\tau + h_n) - \hat{\beta}(\tau - h_n))$$

- To avoid negative values

$$\tilde{f}_i(x_i^\top \hat{\beta}(\tau)) = \max\{0, \hat{f}_i(x_i^\top \hat{\beta}(\tau))\}$$

Estimation of covariance matrices(2)

Bandwidth selection

- Bofinger (1975) showed that the optimal bandwidth is

$$h_n = n^{1/5}[4.5s^2(\tau)/(s''(\tau))^2]^{1/5}$$

- For normal distribution

$$h_n = n^{-1/5}\left[\frac{4,5\phi^4(\Phi^{-1}(t))}{(2\Phi^{-1}(t)^2+1)^2}\right]^{1/5}$$

Estimation of covariance matrices(3)

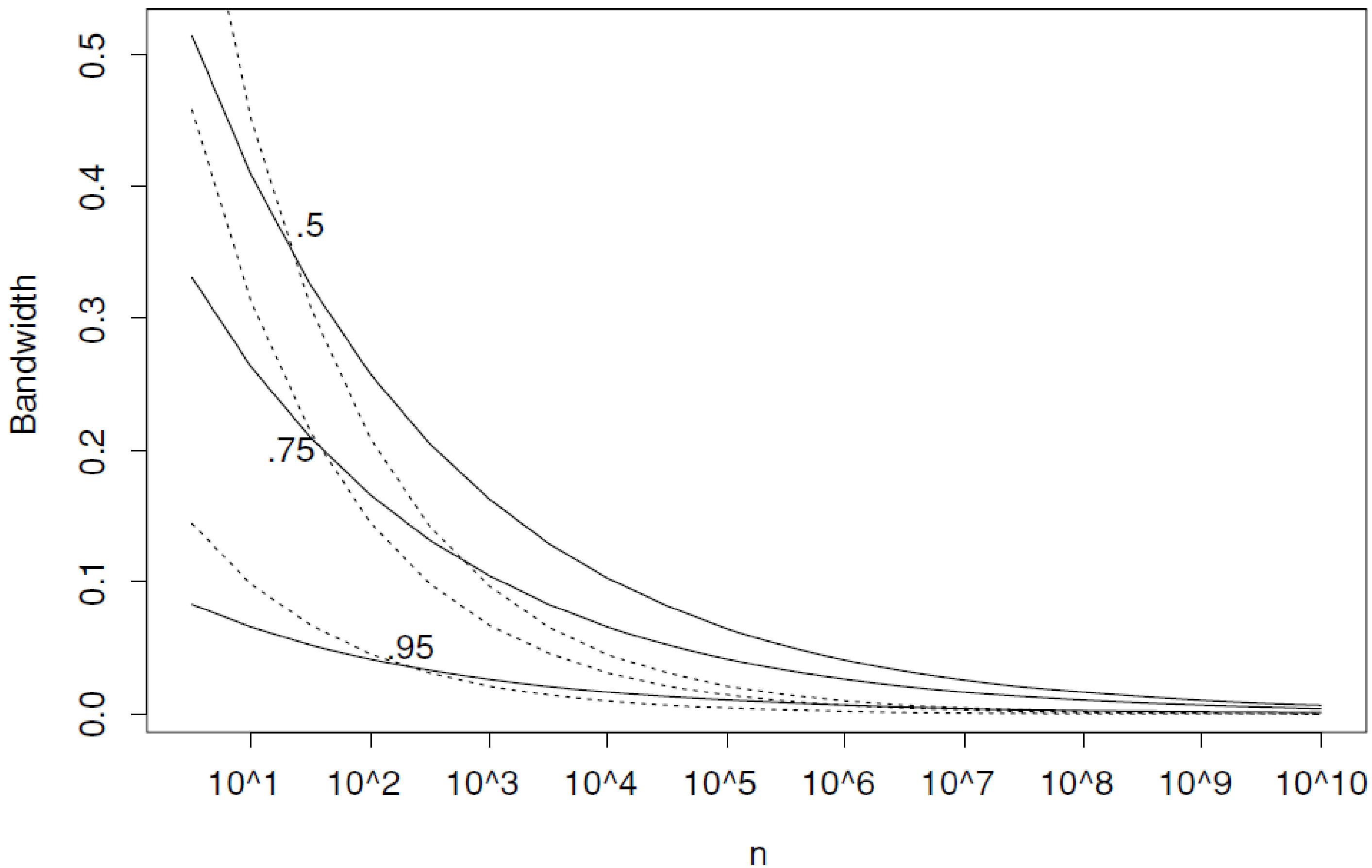
Bandwidth selection

- Hall and Sheather (1988) suggested the bandwidth

$$h_n = n^{-1/3} z_\alpha^{2/3} [1.5 s(t)/s''(t)]^{1/3}$$

where z_α satisfies $\Phi(z_\alpha) = 1 - \alpha/2$

Estimation of covariance matrices(4)

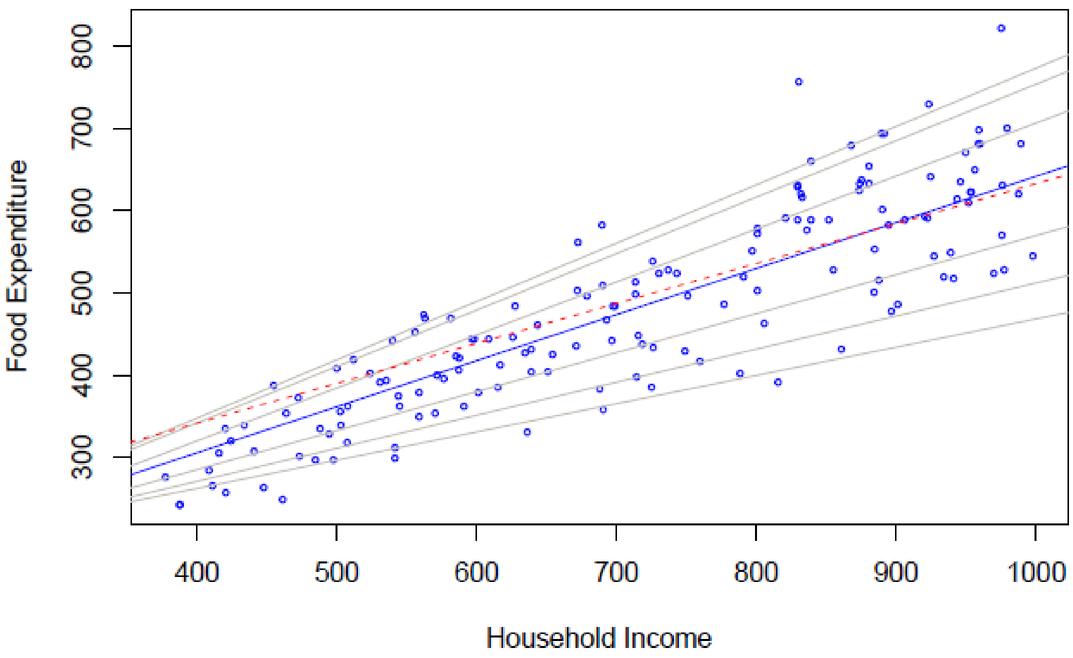


Bofinger: solid lines

Hall and Sheather: dotted lines

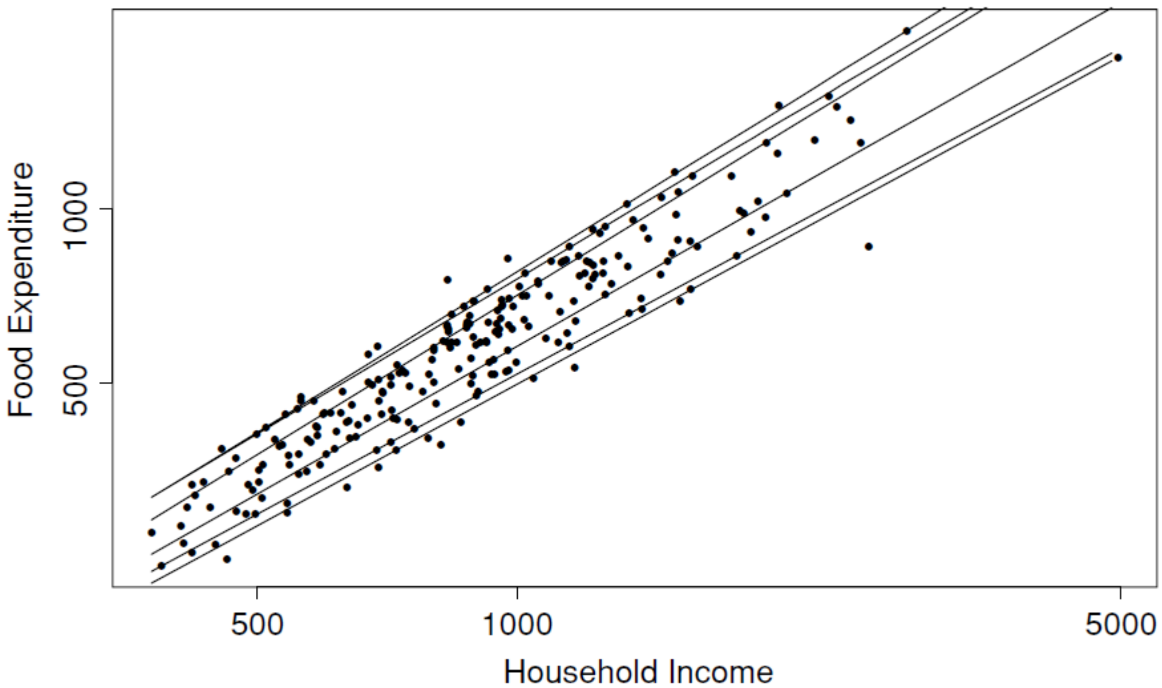
Engel's Food Expenditure Data (1)

- Food Expenditure VS Household Income



Engel's Food Expenditure Data (2)

- Food Expenditure VS Household Income (log scale)



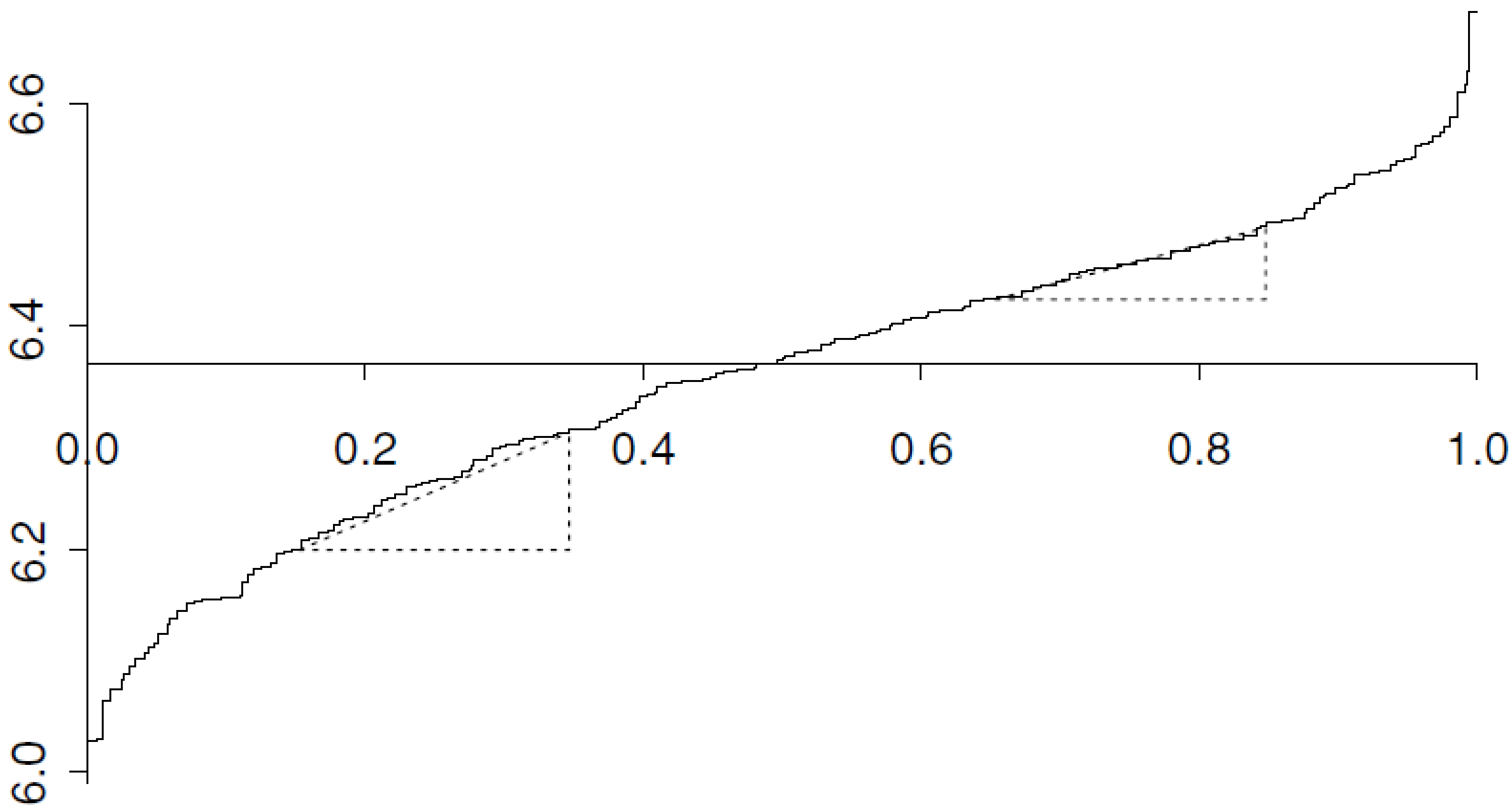
$$\hat{s}(1/4) = 0.543$$

$$\hat{s}(3/4) = 0.330$$

p-value 0.03

$$\hat{\beta}_2(3/4) - \hat{\beta}_2(1/4) = 0.915 - 0.849 = 0.0661$$

Engel's Food Expenditure Data (3)



This figure plots $\hat{Q}_Y(\tau|\bar{x}) = \bar{x}'\hat{\beta}(\tau)$

Bootstrap estimation (1)

Bootstrapping the residuals

Let $\hat{\beta}(\tau) = \operatorname{argmin} \sum \rho_\tau \left(y_i - x_i^\top b\right)$

$\hat{u}_i = y_i - x_i \hat{\beta}(\tau)$

The empirical distribution of the errors

$$\hat{F}_n(u) = n^{-1} \sum_{i=1}^{n} I(\hat{u}_i < u)$$

drawing bootstrap samples $u_i^*, \ldots, u_n^*$ from $\hat{F}_n(u)$

set $y_i^* = x_i \hat{\beta}(\tau) + u_i^*$

$\beta_n^*(\tau) = \operatorname{argmin} \sum \rho_\tau \left(y_i^* - x_i^\top b\right)$

Bootstrap estimation (1)

drawing bootstrap samples $u_i^*, \ldots, u_n^*$ from $\hat{F}_n(u)$

set $y_i^* = x_i \hat{\beta}(\tau) + u_i^*$

$\beta_n^*(\tau) = \operatorname{argmin} \sum \rho_\tau \left(y_i^* - x_i^\top b \right)$

DeAngelis et al. showed that

$$\hat{G}(z) = P(\sqrt{n}(\beta_{nj}^*(\tau) - \hat{\beta}_{nj}(\tau)) \le z_j, \ \ j = 1, \ldots, p | \mathcal{X})$$

converges to the limiting distribution of $\sqrt{n}(\hat{\beta}_n(\tau) - \beta(\tau))$

Bootstrap estimation (2)

Bootstrapping the observation

Draw (x_i^*, y_i^*) with replacement from the n pairs

$$\{(x_i, y_i) : i = 1, \dots, n\}$$

each with probability 1/n

$$\beta_n^* = \text{argmin}_b \sum \rho_\tau (y_i^* - b^T x_i^*)$$

Jackknife (1)

Suppose $\hat{\theta}_n$ is the median of the sample $\{X_1, \ldots, X_n\}$

$\hat{\theta}_{(i)}$ denote the median with the ith observation deleted

the jacknife estimate of the variance of the median is

$$v_n = \frac{n-1}{n} \sum_{i=1}^{n} \left(\hat{\theta}_{(i)} - \hat{\theta}_{(\cdot)}\right)^2$$

with $\hat{\theta}_{(\cdot)} = n^{-1} \sum \hat{\theta}_{(i)}$

Jackknife (2)

for $n = 2m$ $\quad v_n = \frac{n-1}{4}(x_{(m+1)} - x_{(m)})^2$

$$nv_n \rightsquigarrow \frac{1}{4f^2(F^{-1}(1/2))}\left(\frac{\chi_2^2}{2}\right)^2$$

$(\chi_2^2/2)^2$ is a random variable with mean 2 and variance 20

the asymptotic variance of $\sqrt{n}(\hat{\theta}_n - F^{-1}(1/2))$

should equal $1/(2f^2(F^{-1}(1/2)))$