Empirical likelihood, II

Empirical (Nonparametric) likelihood:

For $X_1, \cdots, X_n$ iid $\sim F$.

$$EL(F) = \prod_{i=1}^{n} \Delta F(X_i)$$

EL ratio

$$ELR(F) = \frac{EL(F)}{EL(F_n)}$$

where F_n is the empirical distribution function. When no ties in the data, this is

$$= \prod_{i=1}^{n} np_i$$

where the mass of F on X_i is $p_i \geq 0$, and $\sum p_i = 1$.

ELR function for some finite dimensional parameters θ:

$$ELR(\theta) = \frac{\sup\{EL(F) \mid T(F) = \theta\}}{\sup EL(F)}$$

where $T(F)$ is some finite dimensional features of the F, like the mean of F: then $T(F) = \int t dF(t)$.

EL hypothesis tests:

Reject $H_0 : T(F) = \theta_0$ when $-2 \log ELR(\theta_0) > r_0$ for some threshold r_0, usually the 95% chi square quantile.

EL confidence regions:

$$\{\theta| \ -2 \log ELR(\theta) < r_0\}$$

EL for means $T(F) = \int x dF(x)$. For discrete F this is $T(F) = \sum t_i \Delta F(t_i)$ where t_i are the jump points of F. Let $\Delta F(t_i) = p_i$.

$$ELR(\mu) = \sup \left(\prod_{i=1}^{n} np_i \mid \int x dF(x) = \mu; p_i \geq 0; \sum_{i=1}^{n} p_i = 1 \right)$$

We now compute an explicite expression of $\log ELR(\mu)$. Maximize

$$\sum_{i=1}^{n} \log(np_i)$$

under the constraints:

$$n \sum_{i=1}^{n} p_i(X_i - \mu) = 0; \quad \sum_{i=1}^{n} p_i = 1$$

Write

$$G = \sum \log(np_i) - n\lambda \sum_{i=1}^{n} p_i(X_i - \mu) + \gamma[\sum_{i=1}^{n} p_i - 1]$$

λ and γ are Lagrange multipliers.

Taking derivatives, and set it to 0:

$$\frac{\partial G}{\partial p_i} = \frac{1}{p_i} - n\lambda(X_i - \mu) + \gamma = 0$$

We compute

$$0 = \sum p_i \frac{\partial G}{\partial p_i} = n + \gamma$$

giving $\gamma = -n$. Thus

$$p_i = \frac{1}{n} \frac{1}{1 + \lambda(X_i - \mu)}$$

where λ is solved from the following equation.

Plugging this back into the constraint to get λ:

$$g(\lambda) = \frac{1}{n} \sum_{i=1}^{n} \frac{X_i - \mu}{1 + \lambda(X_i - \mu)} = 0$$

This equation has a unique solution. (check)

Theorem 1 (ELT, Owen 1990) If X has finite mean μ_0 and finite covariance matrix of rank $q > 0$, ($q = \dim(X)$), then

$$-2 \log ELR(\mu_0) \to^D \chi_q^2$$

Proof for Case $q = 1$.

Note that $g(0, \mu) = \bar{X} - \mu$ Let $\hat{\sigma}^2 = \sum_{i=1}^{n} (X_i - \mu)^2$.

We now compute an expression for the solution of $g(\lambda) = 0$ when $\mu = \mu_0$. Taylor expanding g wrt λ at $\lambda = 0$, the equation becomes:

$$0 = g(\lambda) = g(0) + \lambda g'(0) + o_P(n^{-1/2}) = \bar{X} - \mu_0 - \lambda\hat{\sigma}^2 + o_P(n^{-1/2})$$

Thus

$$\lambda = (\bar{X} - \mu_0)/\hat{\sigma}^2 + o_P(n^{-1/2}) = O_P(n^{-1/2})$$

Recall

$$p_i = \frac{1}{n + n\lambda(X_i - \mu_0)}$$

so, using the Taylor expansion $\log(1+x) = x - x^2/2 + O(x^3)$,

$$
\begin{aligned}
-2\log ELR(\mu_0) = -2\sum_{i=1}^{n}\log(np_i) = 2\sum_{i=1}^{n}\log(1+\lambda(X_i-\mu_0)) \\
= 2n\lambda(\bar{X}-\mu_0) - n\lambda^2\hat{\sigma}^2 + o_P(1) \\
= 2n(\bar{X}-\mu_0)^2/\hat{\sigma}^2 - n(\bar{X}-\mu_0)^2/\hat{\sigma}^2 + o_P(1) \\
= n(\bar{X}-\mu_0)^2/\hat{\sigma}^2 + o_P(1) \\
\to^D \chi_1^2
\end{aligned}
$$

The proof can be modified to work with data that has ties.

EL with estimating equations:

Estimating function: $m(x;\theta)$ with $E(m(X;\theta_0)) = 0$; e.g., median, $m(X;\theta) = I_{[X\le\theta]} - 0.5$.

The EL function:

$$
R(\theta) = \sup\left\{\prod_{i=1}^{n} np_i \mid \sum_{i=1}^{n} p_i m(X_i;\theta) = 0; p_i \ge 0; \sum_{i=1}^{n} p_i = 1\right\}
$$

Theorem 2 (ELT) If $m(X,\theta_0)$ has a finite covariance matrix of rank $q > 0$, then

$$
-2\log ELR(\theta_0) \to^D \chi_q^2
$$

Proof: Immediate from basic ELT upon replacing X by $m(X;\theta_0)$.

What if the estimating function has some estimated parameters: The variance

$$
\sigma^2 = E(X-\mu)^2
$$

Since the mean is also unknown, the estimating function we actually use is

$$
\sum p_i(X_i-\bar{X})^2 - \sigma^2 = \sum p_i m_{ni} - \sigma^2
$$

where m_{ni} is not independent anymore. (homework: Show that $\frac{1}{\sqrt{n}}\sum m_{ni}$ have a asymptotic normal distribution under condition that $EX^4 < \infty$.)

Additional homework

For any CDF $F(t)$ we can define a cumulative hazard function:

$$\Lambda(t) = \int_{(-\infty,t]} \frac{dF(s)}{1 - F(s-)}$$

Notice this formular works for continuous and discrete F.

Show that the inverse formular is true:

$$1 - F(t) = \exp(-\Lambda(t))$$

works for continuous cdf F, and

$$1 - F(t) = \prod_{s \le t} 1 - \Delta\Lambda(s)$$

works for discrete F.