Comparison of treatments via empirical likelihood

Ian McKeague
Department of Statistics
Florida State University
Tallahassee, FL 32306-4330, USA
<http://stat.fsu.edu/∼mckeague>

Partly based on joint work with Yichuan Zhao.

Outline

- Motivating examples
- Background on empirical likelihood (EL)
- EL methods for one-sample problems in survival analysis
- Two-sample problems
- Conclusions

Motivating examples

Mayo Clinic trial

Randomized clinical trial for primary biliary cirrhosis of the liver. Treatment and placebo groups to be compared.

Aim: find an EL-based simultaneous confidence band for the *relative survival function*:

$$\theta(t) = S_1(t)/S_2(t)$$

More suitable than the difference $S_1(t) - S_2(t)$ when the risk of failure is moderate.

Vaccine efficacy estimation

Halloran, Struchiner and Longini (1997)

VE is defined as 1 minus some measure of relative risk in the vaccinated group compared with the unvaccinated group:

$$\mathrm{VE}(t) = 1 - \frac{\alpha_{\mathrm{vaccine}}(t)}{\alpha_{\mathrm{placebo}}(t)}$$

$$\mathrm{VE}_c(t) = 1 - \frac{A_{\mathrm{vaccine}}(t)}{A_{\mathrm{placebo}}(t)}$$

Aim: find an EL-based simultaneous confidence band for $\mathrm{VE}_c(t)$

Background on empirical likelihood

- Thomas and Grunkemeier (1975) for survival function estimation. Owen (1988, 1990, . . . , 2001).
- First developed for finite-dimensional features $\theta = \theta(F)$ of a cdf (e.g., mean, median, cdf at a single point).

Nonparametric likelihood

$$L(F) = \prod_{i=1}^{n} (F(X_i) - F(X_i-)).$$

EL ratio

$$\tilde{R}(F) = \frac{L(F)}{L(F_n)} = \prod_{i=1}^{n} np_i$$

where (part of) the mass on X_i is $p_i \geq 0$, $\sum_{i=1}^{n} p_i \leq 1$, and F_n is the empirical cdf.

EL function

$$R(\theta_0) \;\; = \;\; \sup\{\tilde{R}(F) : \theta(F) = \theta_0\}$$

Equivalently

$$R(\theta_0) \quad = \quad \frac{\sup\{L(F) : \theta(F) = \theta_0\}}{\sup\{L(F)\}}$$

EL hypothesis tests

Accept $\theta(F) = \theta_0$ when $R(\theta_0) \geq r_0$ for some threshold r_0.

EL confidence regions

$$\{\theta \colon R(\theta) \geq r_0\}$$

with r_0 chosen via an EL analogue of Wilks's theorem.

EL for means

$\mu = E(X) \in \mathbb{R}^d$

$$R(\mu) = \max\left\{\prod_{i=1}^{n} np_i : \sum_{i=1}^{n} p_i X_i = \mu, p_i \geq 0, \sum_{i=1}^{n} p_i = 1\right\}$$

Theorem (ELT, Owen 1990) $X_1, \ldots, X_n$ iid with finite mean μ_0, finite covariance matrix of rank $q > 0$. Then

$$-2\log R(\mu_0) \xrightarrow{\mathcal{D}} \chi^2_q.$$

Extensions

- Smooth functions of means: $\theta = h(\mu)$
- Linear functionals of F: $\theta = E(h(X)) = \int h(x)\, dF(x)$.
- Implicitly defined parameters: $E(m(X, \theta)) = 0$ where $m(X, \theta)$ is the estimating function; e.g., median, $m(X, \theta) = 1\{X \leq \theta\} - .5$.

$$R(\theta) = \max \left\{ \prod_{i=1}^{n} np_i \colon \sum_{i=1}^{n} p_i m(X_i, \theta) = 0, p_i \geq 0, \sum_{i=1}^{n} p_i = 1 \right\}$$

Theorem (Owen 1990) Let $X_1, \ldots, X_n$ be iid, and suppose $m(X, \theta_0)$ has finite covariance matrix of rank $q > 0$. If $E(m(X, \theta_0)) = 0$, then

$$-2 \log R(\theta_0) \xrightarrow{\mathcal{D}} \chi^2_q.$$

Proof Immediate from ELT upon some changes in notation: X is replaced by $m(X, \theta)$, which has mean zero when $\theta = \theta_0$.

EL simultaneous band for F

Local EL function at $\theta_0 = F_0(t)$:

$$
\begin{aligned}
R(t) &= \frac{\sup\{L(F) : F(t) = F_0(t)\}}{\sup\{L(F)\}} \\
&= \left(\frac{\sqrt{n}(F_n(t) - F_0(t))}{\sqrt{F_0(t)(1 - F_0(t))}} \right)^2 + o_P(1)
\end{aligned}
$$

As a *process* in $t \in [a, b]$:

$$
\begin{aligned}
-2 \log R(t) &\xrightarrow{\mathcal{D}} \left(\frac{W^o(F_0(t))}{\sqrt{F_0(t)(1 - F_0(t))}} \right)^2 \\
&\stackrel{\mathcal{D}}{=} \left(\frac{W(\sigma^2(t))}{\sigma(t)} \right)^2 ,
\end{aligned}
$$

W^o standard tied-down Wiener process (Brownian bridge)
W standard Wiener process

$$\sigma^2(t) = \frac{F_0(t)}{1 - F_0(t)}.$$

Simultaneous confidence band for F over an interval $[a, b]$:

$$\{(t, F_0(t)) : -2 \log R(t) \leq C_\alpha, t \in [a, b]\}$$

C_α the upper α-quantile of

$$\sup_{t \in [\hat{\sigma}^2(a), \hat{\sigma}^2(b)]} \frac{W^2(t)}{t}.$$

Equal precision LR band. Narrower in tail than Hollander, McKeague, Yang (1997) band.

Survival analysis setting

Right-censored lifetime data

Observe n iid pairs (Z_i, δ_i)
$Z_i = \min(X_i, Y_i)$, $\delta_i = I\{X_i \leq Y_i\}$, X_i and Y_i independent.
F : cdf of X_i
G : cdf of Y_i
$S = 1 - F$: survival function, $S(0) = 1$
A: cumulative hazard function

$$A(t) = \int_{(0,t]} \frac{dF(s)}{1 - F(s-)}$$

Nonparametric likelihood

$$L(S) = L(F) = \prod_{i=1}^{n} (F(Z_i) - F(Z_i-))^{\delta_i} (1 - F(Z_i))^{1-\delta_i}.$$

To maximize $L(F)$, we only need consider F supported on the uncensored lifetimes.

EL function

$$R(\theta_0) \quad = \quad \frac{\sup\{L(S) : \theta(S) = \theta_0\}}{\sup\{L(S)\}}$$

EL suddenly becomes difficult because of the censoring!

Unless $\theta(S)$ has a particularly simple form, $R(\theta_0)$ may be intractable.

Known tractable forms of $\theta(S)$ or $\theta(A)$:

- $S(t_0)$
- $A(t_0)$
- quantiles
- linear functionals $\theta(F) = \int h(t)\,dF(t)$
- linear functionals $\theta(A) = \int h(t)\,dA(t)$

Thomas and Grunkemeier (1975), Li (1995), Murphy (1995), Li, Hollander, McKeague and Yang (1996), Murphy and van der Vaart (1997), Pan and Zhou (2002)

EL for the Cox model regression parameters

$$\alpha(t|z) = \alpha_0(t)\exp(\beta' z)$$

Estimating equation for β:

$$E(U(\beta_0)) = 0$$

where U is the partial likelihood score function.

Qin and Jing (2001): standard EL for this estimating equation.

Murphy and van der Vaart (1997): a profile EL for β for current status data.

EL for survival function at a fixed point

$p = S(t_0)$, with t_0 fixed, $0 < p < 1$.

Method of Lagrange multipliers is tractable.

Theorem If S is continuous, $0 < p = S(t_0) < 1$ and $G(t_0) < 1$, then

$$-2 \log R(p) \xrightarrow{\mathcal{D}} \chi_1^2$$

Thomas and Grunkemeier (1975), Li (1995), Murphy (1995)

EL simultaneous band for S

As a *process* in $t \in [a, b]$,

$$-2 \log R(t) \xrightarrow{\mathcal{D}} \left(\frac{W(\sigma^2(t))}{\sigma(t)} \right)^2,$$

Simultaneous confidence band for S over an interval $[a, b]$:

$$\{(t, S(t)) : -2 \log R(t) \leq C_\alpha, t \in [a, b]\}$$

C_α the upper α-quantile of

$$\sup_{t \in [\hat{\sigma}^2(a), \hat{\sigma}^2(b)]} \frac{W^2(t)}{t}$$

Equal precision LR band. Narrower in the tail than Hollander, McKeague, Yang (1997) band.

Li and Van Keilegom (2001): adjustment for a covariate effect (continuous one-dimensional covariate).

Two-sample problem

Comparison of treatment and placebo groups.

Notation
Index sample by j
Assume $n_j/n \to p_j > 0$
Total sample size $n = n_1 + n_2$
Nonparametric likelihood: $L(S_1, S_2) = L_1(S_1)L_2(S_2)$.

- Standard method: logrank test for $S_1 = S_2$.

- Wald-type comparison of S_1 and S_2 using some smooth functional $\varphi(S_1, S_2)$ and the functional delta method typically leads to intractable limiting distributions. Simulation needed.

Gaussian multiplier simulation technique

Martingale increments $dM_i(t)$ replaced by $G_i dN_i(t)$, where $G_i \sim N(0,1)$. (Lin, Wei and Ying, 1993)

Parzen, Wei and Ying (1997) constructed a Wald-type confidence band for $S_1(t) - S_2(t)$ using this technique.

Q-Q plot

$\{(F_1^{-1}(p), F_2^{-1}(p)) : 0 < p < 1\}$

Einmahl and McKeague (1999) constructed an EL confidence band for the Q-Q plot:

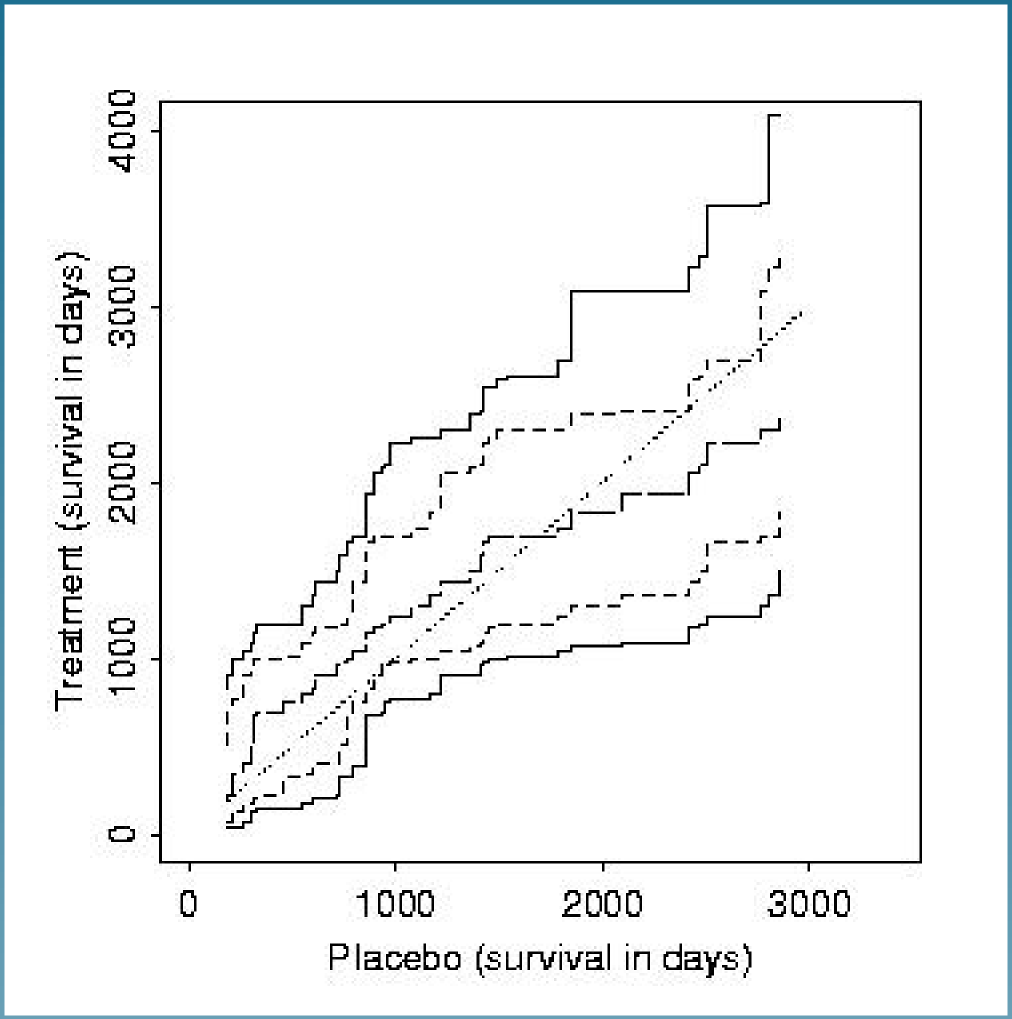
$$\{(t_1, t_2)\colon -2 \log R(t_1, t_2) \leq C_\alpha, t_1 \in [a, b]\}$$

where C_α uses

$$\sigma^2(t) = \sigma_1^2(t)/p_1 + \sigma_2^2(t')/p_2$$

and $t' = F_2^{-1}(F_1(t))$. Simulation not needed.

Mayo Clinic trial Q-Q plot



Relative survival

$\theta(t) = S_1(t)/S_2(t)$

McKeague and Zhao (2002) construct an EL simultaneous band:

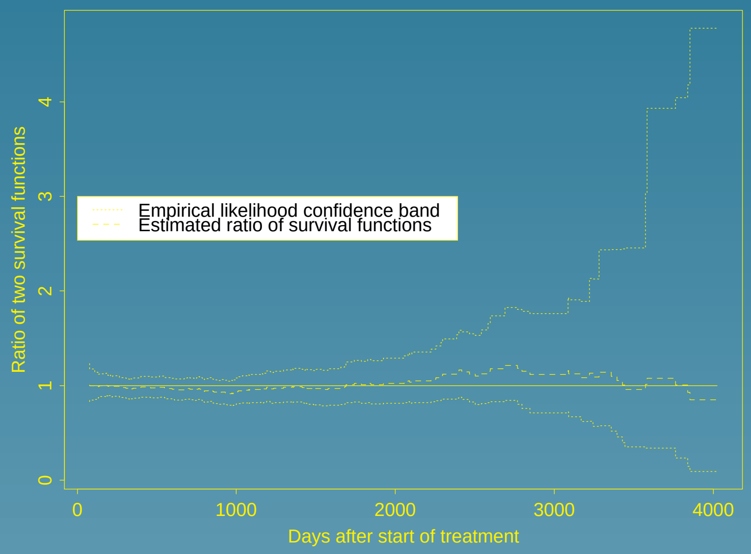
$$\{(t, \theta(t))\colon -2\log R(t) \leq C_\alpha, t \in [a, b]\}$$

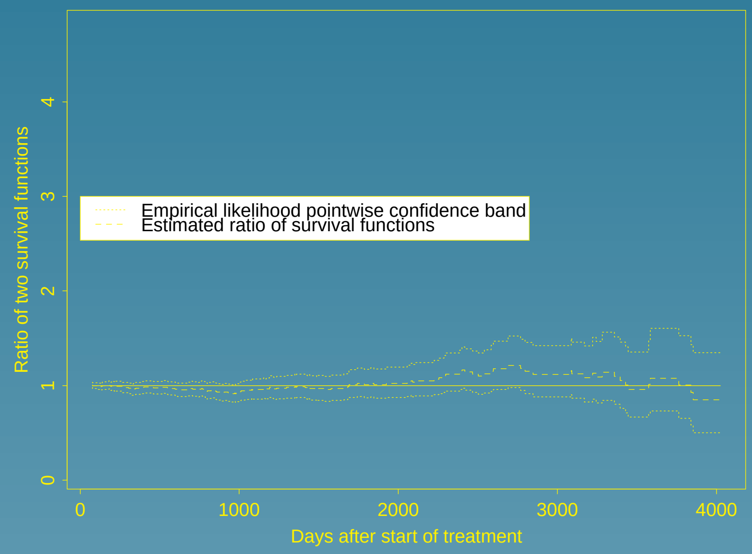
where C_α uses

$$\sigma^2(t) = \sigma_1^2(t)/p_1 + \sigma_2^2(t)/p_2.$$

Simulation not needed.

Mayo clinic trail: placebo/treatment relative survival





- Simultaneous band for differences in cumulative hazards:

$$A_1(t) - A_2(t) = -\log(S_1(t)/S_2(t))$$

EL works without simulation, McKeague and Zhao (2002).

- Simultaneous band for vaccine efficacy:

$$\mathrm{VE}_c(t) = 1 - \frac{A_{\mathrm{vaccine}}(t)}{A_{\mathrm{placebo}}(t)}$$

EL works, McKeague and Zhao (2002). Gaussian multiplier simulation needed. Reason EL works:

$$A_1(t)/A_2(t) = \log S_1(t)/\log S_2(t)$$

and the product-limit form of the survival function.

- Ratios of cdfs: $F_1(t)/F_2(t)$, EL intractable?

Conclusions

- EL techniques are tractable in the two-sample setting for simultaneous inference concerning *ratios* of survival functions and cumulative hazards.
- EL shows great promise for further development in more complex clinical trial settings.