Introduction to optband

Tom Chen & Sam Tracy

2016-08-26

Classical simultaneous confidence bands for survival functions (Hall-Wellner, Equal Precision, etc) are derived from transformations of the convergence to a Weiner process with a strictly-increasing variance function These transformations are often motivated by factors such as:

- Time intervals for which the analyst prefers to be thinner or wider
- Tractability of computing critical values of the asymptotic distribution in attaining the prescribed (nominal) coverage level

This package instead approaches the problem purely from an optimization perspective: given a certain coverage level, obtain bands such that the area between is minimized. While an exact solution cannot be obtained in closed form, `optband` provides an approximate solution based off local time arguments for both the survival and cumulative-hazard functions.

Usage

`optband` requires the `stats`, `survival`, `taRifx`, and `LambertW` packages. Of primary use is the `opt.ci` function with method `opt.ci(survi, conf.level = 0.95, fun = 'surv', tl = NA, tu = NA)`.

`opt.ci` takes a `survfit` object with the desired $1-\alpha$ coverage level, function of interest (either `'surv'` for the survival function or `'cumhaz'` for the cumulative-hazard function), and optional upper or lower bounds for data truncation. Defaults are $\alpha = 0.05$, `fun = 'surv'`, `tl = NA`, `tu = NA`.

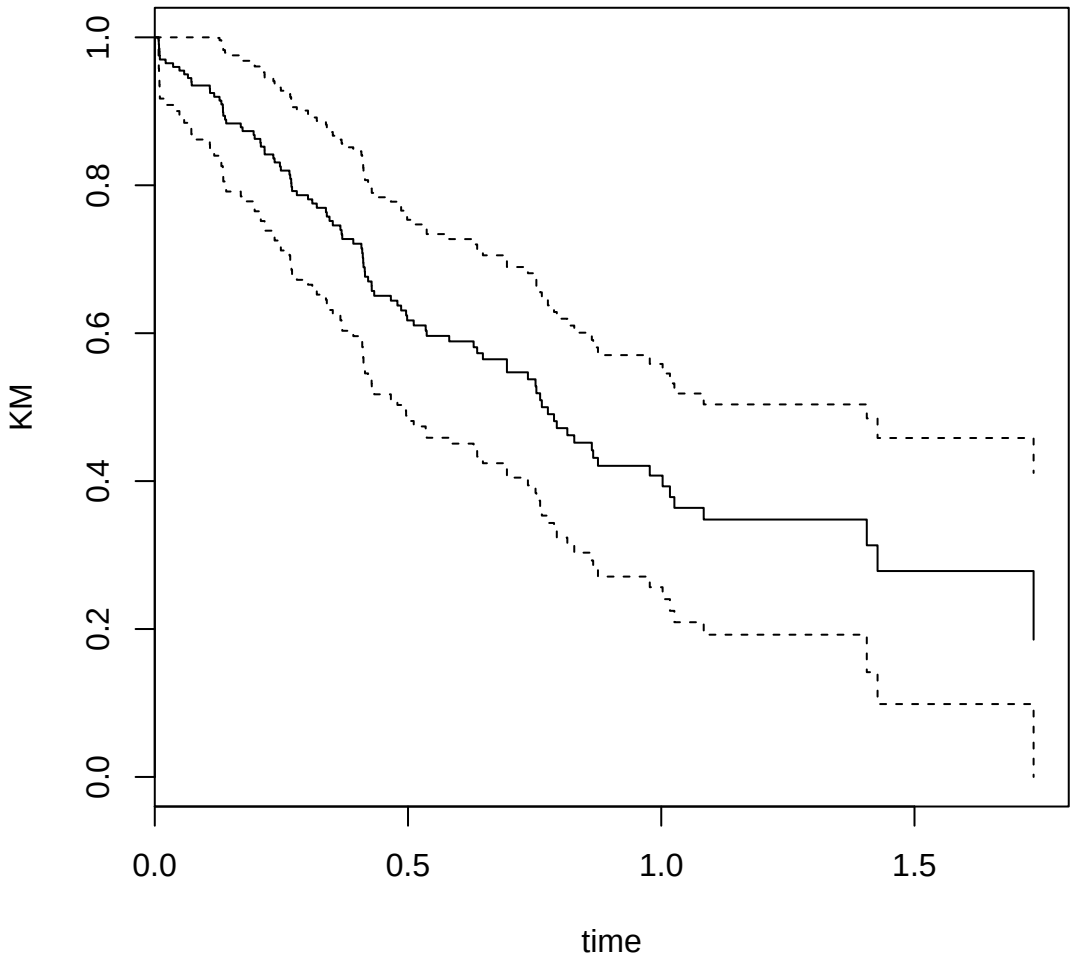
Example

First, let's generate some data and estimate the Kaplan-Meier curve:

```
N = 200
x1 <- rweibull(N, 1, 1)
x2 <- rweibull(N, 2, 1)
d <- x1 < x2
x <- pmin(x1,x2)
mydata = data.frame(stop = x, event = d)
S = survfit(Surv(x, d) ~ 1, type="kaplan-meier")
```

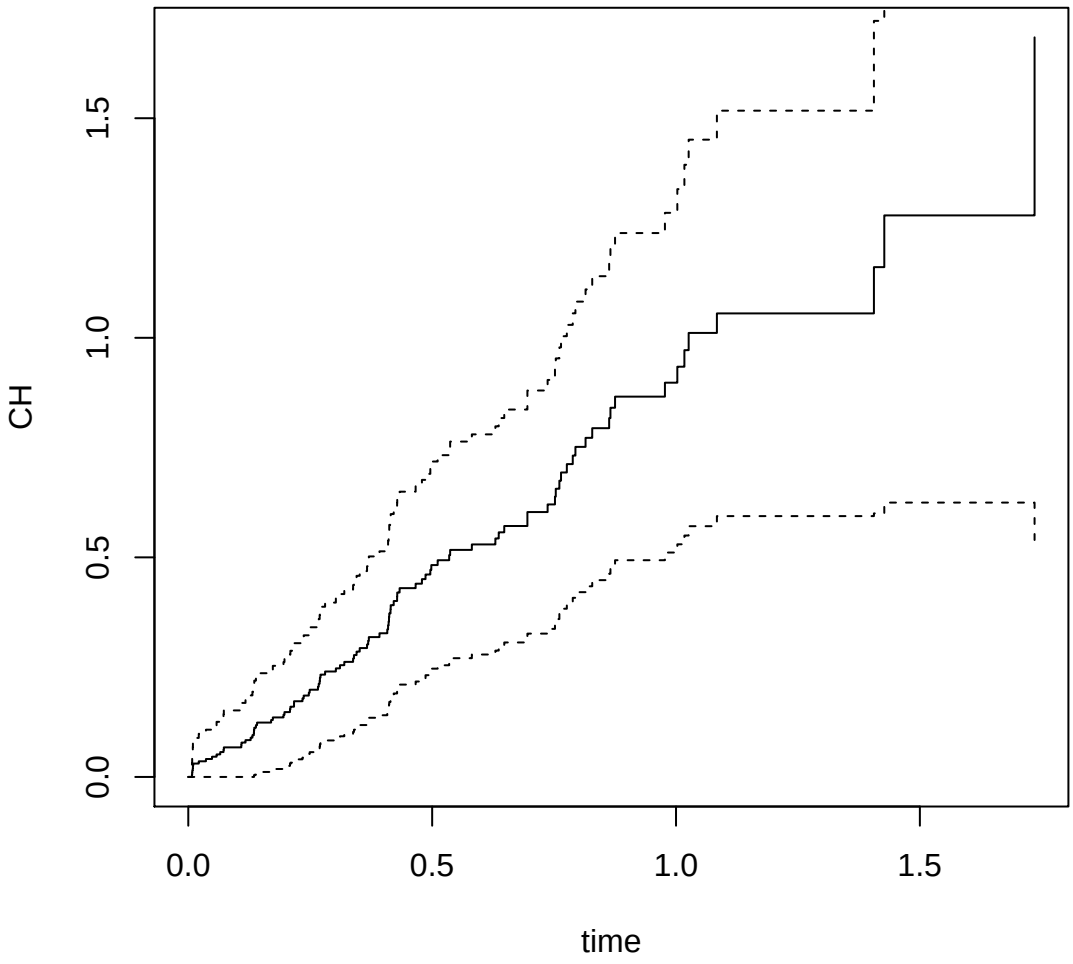
Now we can estimate the optimized confidence bands and plot the curve:

```
opt_S <- opt.ci(S, conf.level = 0.95, fun = "surv", tl = NA, tu = NA)
plot(opt_S, xlab="time", ylab="KM", mark.time=FALSE)
```



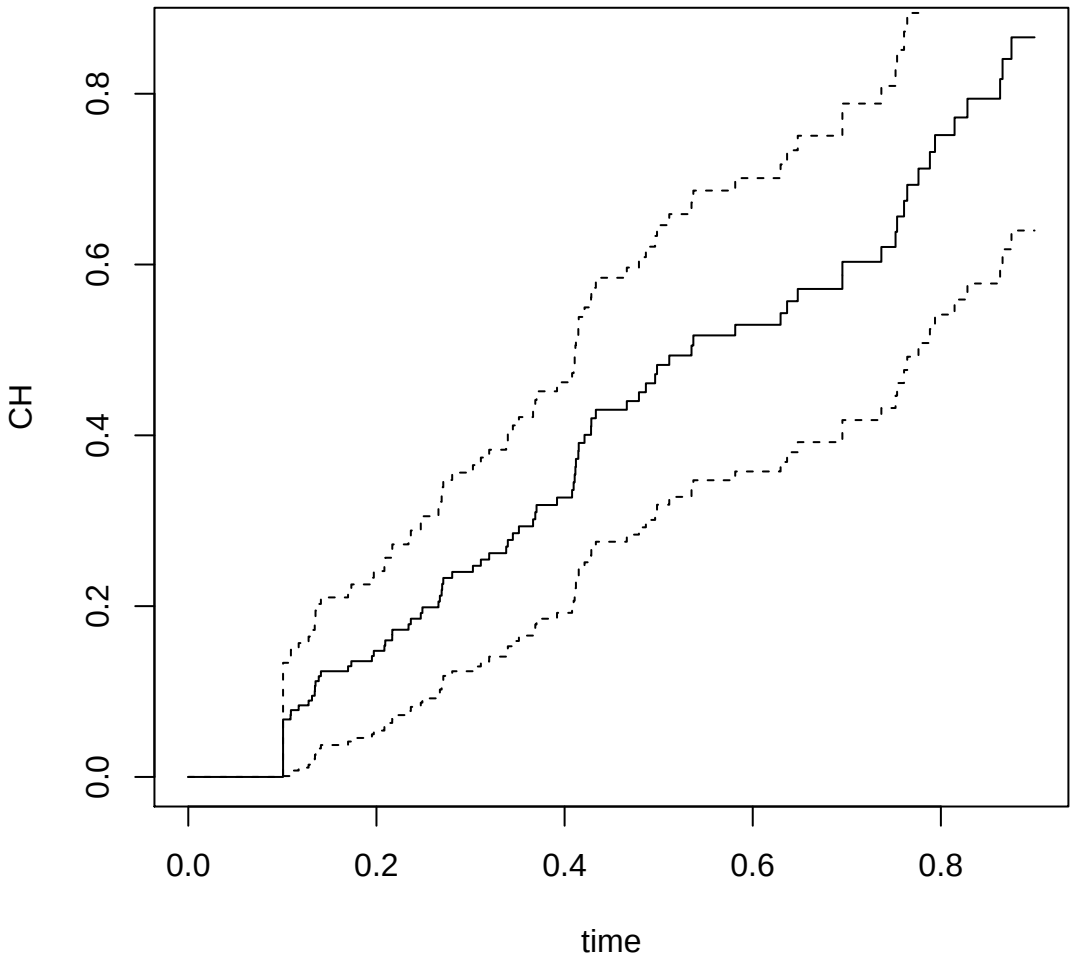
And we can do the same with the estimated cumultive-hazard function:

```
opt_H <- opt.ci(S, conf.level = 0.95, fun = "cumhaz", tl = NA, tu = NA)
plot(opt_H, fun="cumhaz", xlab="time", ylab="CH", mark.time=FALSE)
```



We can further play with the procedure by adjusting the α level and truncating the data:

```
opt_H <- opt.ci(S, conf.level = 0.90, fun = "cumhaz", tl = .1, tu = .9)
plot(opt_H, fun="cumhaz", xlab="time", ylab="CH", mark.time=FALSE)
```

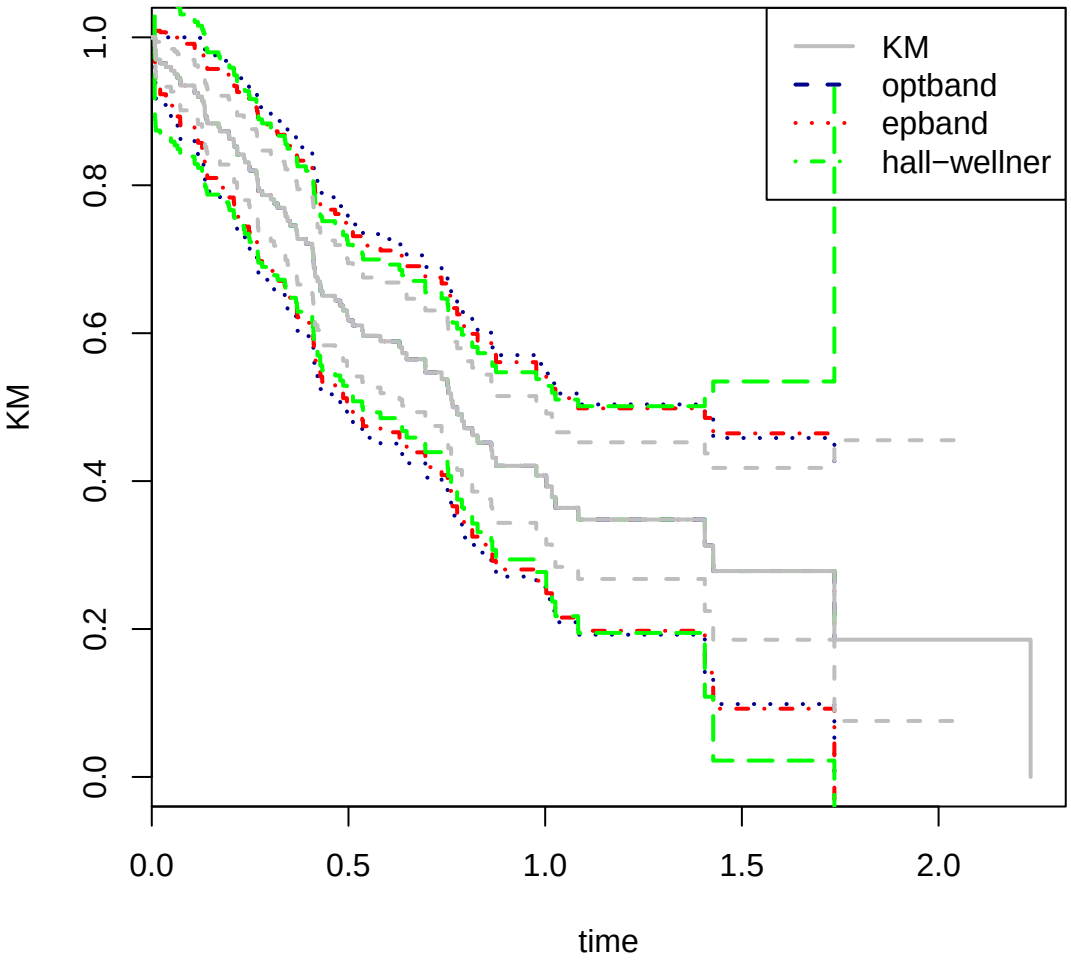


And compare the results of this band to the Equal Precision and Hall-Wellner bands (for this we will use the `km.ci` package):

```
library(km.ci)
color <- c("grey", "darkblue", "red", "green")
plot(S, mark.time=FALSE, xlab="time", ylab="KM", col = "white")

lines(opt_S, col=color[2], lty=2, lwd=2, mark.time=F)
e <- km.ci(S, conf.level = 0.95, method = "epband")
h <- km.ci(S, conf.level = 0.95, method = "hall-wellner")
lines(e, col=color[3], lty=3, lwd=2, mark.time=F)
lines(h, col=color[4], lty=4, lwd=2, mark.time=F)
lines(S,col="grey", lwd=2, mark.time=F)

legend("topright", c("KM", "optband", "epband", "hall-wellner"),
       lwd=2, lty=1:4, col=color)
```



References

Kendall, W.S., Marin, J.M. and Robert, C.P. (2007). Confidence bands for Brownian motion and applications to Monte Carlo simulation. Statistics and Computing, 17(1), pp.1-10.

Therneau T (2015). A Package for Survival Analysis in S. version 2.38, .

Terry M. Therneau and Patricia M. Grambsch (2000). Modeling Survival Data: Extending the Cox Model. Springer, New York. ISBN 0-387-98784-3.

Ralf Strobl (2009). km.ci: Confidence intervals for the Kaplan-Meier estimator. R package version 0.5-2. https://CRAN.R-project.org/package=km.ci

Ari B. Friedman (2014). taRifx: Collection of utility and convenience functions.. R package version 1.0.6. https://CRAN.R-project.org/package=taRifx

Georg M. Goerg (2016). LambertW: An R package for Lambert W x F Random Variables. R package version 0.6.4.

Georg M. Goerg (2011): Lambert W random variables - a new family of generalized skewed distributions with applications to risk estimation. Annals of Applied Statistics 3(5). 2197-2230.

Georg M. Goerg (2014): The Lambert Way to Gaussianize heavy-tailed data with the inverse of Tukey's h transformation as a special case. The Scientific World Journal.