Specifying Models and Parameter Starting Values

David Borchers

November 2, 2015

concordance=TRUE

You have to pass function `est.hmltm` the name of the discrete detection hazard function $h(,y)$ to use (via the argument `FUN`) a model specification (via the argument `model`) and a vector of starting parameters (via the argument `pars`). The possible forms for $h(x,y)$ are shown in Table 1, together with their scale and shape parameters (each form has one or more shape paramters and one or more shape parameters).

FUN is a character variable containing the name of the detection hazard model $h(x,y)$. Valid names are shown in the first columnn of Table 1.

model is a list containing two character variables named `y` and `x` that specify the form of the dependence of the model's scale parameter(s) on covariates (or `NULL` if there are no covariates). The syntax of the model specification is given below (it is similar to `R`'s usual regression model syntax.)

pars is a numeric vector specifying starting values for the model parameters. The parameter starting values obviously depend on `model` and on which, if any, covariates you're using.

1 Covariates and parameter starting values

Covariates always affect the scale parameter, not the shape parameter. The reason for this is historical: line transect estimators tend to be implemented with covariates affecting the scale parameter. The only rationale I have ever found for this is a line transect study of beer cans reported in Pollock et al. (1990) in which having covariates affect the scale parameter of detection functions was found to be more effective than having them affect the shape parameter.

Deciding on starting parameter values involves some trial and error. A reasonable starting value for σ is to make it similar size to the average x and y (when the model does not include σ_x), average y (when the model includes σ_x). A reasonable value for σ_x is something around the size of the average x. Shape parameter values between about 0.5 and 2 have worked for cases I've looked at, so a value around 1 might be a reasonable starting value guess.

Table 1: Forms and parameters of the discrete hazard function, $h(x, y)$. Here x and y are perpendicular and forward distances, respectively, $\mathbf{X}$ and $\mathbf{X}_x$ are design matrices for covariates affecting the scale parameters and $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_x$ are parameter vectors associated with the covariates affecting the scale parameters.

Name	Form	Shape parameter(s)	Scale parameter(s) No covariates	With covariates
h.IP.0	$\sigma^\gamma / \left(\sigma^2 + x^2 + y^2\right)^{\gamma/2}$	γ	σ	$\sigma \exp(\mathbf{X}\boldsymbol{\beta})$
h.EP1.0	$\exp\left\{-\left(\frac{x}{\sigma}\right)^\gamma - \left(\frac{y}{\sigma}\right)^\gamma\right\}$	γ	σ	$\sigma \exp(\mathbf{X}\boldsymbol{\beta})$
h.EP2.0	$\exp\left\{-\left(\frac{x}{\sigma}\right)^{\gamma_x} - \left(\frac{y}{\sigma}\right)^{\gamma_y}\right\}$	γ_x, γ_y	σ	$\sigma \exp(\mathbf{X}\boldsymbol{\beta})$
h.EP1x.0	$\exp\left\{-\left(\frac{x}{\sigma_x}\right)^\gamma - \left(\frac{y}{\sigma}\right)^\gamma\right\}$	γ	σ, σ_x	$\sigma \exp(\mathbf{X}\boldsymbol{\beta}), \sigma_x \exp(\mathbf{X}_x\boldsymbol{\beta}_x)$
h.EP2x.0	$\exp\left\{-\left(\frac{x}{\sigma_x}\right)^{\gamma_x} - \left(\frac{y}{\sigma}\right)^\gamma\right\}$	γ, γ_x	σ, σ_x	$\sigma \exp(\mathbf{X}\boldsymbol{\beta}), \sigma_x \exp(\mathbf{X}_x\boldsymbol{\beta}_x)$

When you've got covariates you also need to specify starting values for the parameters associated with the covariates. You do this by specifying starting values for $\exp(\beta)$ for each β in the vector $\boldsymbol{\beta}$. Some guidance can be got from the facts that

$\exp(\beta) = 1$ implies the covariate has no effect,

$0 < \exp(\beta) < 1$ implies that increasing values of the covariate decrease σ,

$1 < \exp(\beta)$ implies that increasing values of the covariate increase σ.

In the absence of information about the effect of the covariate on σ a starting value of $\exp(\beta) = 1$ seems sensible.

If we let $e^{\beta_1}, \ldots, e^{\beta_m}$ be the starting values for the m covariates affecting σ and $e^{\beta_{x1}}, \ldots, e^{\beta_{xm_x}}$ be the starting values for the m_x covariates affecting σ_x, then the starting values are put in a vector (named `pars` here) in the order shown below, omitting values for any parameters that are not relevant for the hazard function and model being used:
`pars=c(`$\gamma, \gamma_x, \sigma, e^{\beta_1}, \ldots, e^{\beta_m}, \sigma_x, e^{\beta_{x1}}, \ldots, e^{\beta_{xm_x}}$`)`.

2 Some example starting values for covariates

Model h.IP.0 with no covariates and starting values $\gamma = 1$, $\sigma = 10$:

```
hfun="h.IP.0"
models=list(y=NULL,x=NULL)
```

```
pars=c(1,10)
```

Model h.IP.0 with covariates `bf` and `cu` affecting σ, and starting values $\gamma = 3.4$, $\sigma = 54$, $e^{\beta_{bf}} = 0.7$, $e^{\beta_{cu}} = 0.9$ (Note that when there is no σ_x parameter, the dependence of σ is specified in the `model` via element `y=....`):

```
hfun="h.IP.0"
models=list(y="bf+cu",x=NULL)
pars=c(3.4, 54, 0.7, 0.9)
```

Model h.EP2x.0 with no covariates and starting values $\gamma_x = 0.76$, $\gamma_y = 0.86$, $\sigma = 19.2$, $\sigma_x = 13.1$:

```
hfun="h.EP2x.0"
models=list(y=NULL,x=NULL)
pars=c(0.76, 0.86, 19.2, 13.1)
```

Model h.EP2x.0 with a single covariate (`bf`) affecting only σ_x, and starting values $\gamma_x = 0.62$, $\gamma_y = 0.96$, $\sigma = 24.1$, $\sigma_x = 17.5$, $e^{\beta_{x,bf}} = 0.65$:

```
hfun="h.EP2x.0"
models=list(y=NULL,x="~bf")
pars=c(0.62, 0.96, 24.1, 17.5, 0.65)
```

Model h.EP2x.0 with a single covariate (`bf`) affecting only σ, and starting values $\gamma_x = 1.1$, $\gamma_y = 0.8$, $\sigma = 30$, $e^{\beta_{bf}} = 0.7$, $\sigma_x = 27$:

```
hfun="h.EP2x.0"
models=list(y="~bf",x=NULL)
pars=c(1.1, 0.8, 30, 0.7, 27)
```

Model h.EP2x.0 with covariate `bf` affecting σ and covarites `bf` and `cu` affecting σ_x, and starting values $\gamma_x = 1.1$, $\gamma_y = 0.8$, $\sigma = 30$, $e^{\beta_{bf}} = 0.7$, $\sigma_x = 27$, $e^{\beta_{x,bf}} = 0.6$, $e^{\beta_{x_cu}} = 0.9$:

```
hfun="h.EP2x.0"
models=list(y="~bf",x="~bf+cu")
pars=c(1.1, 0.8, 30, 0.7, 27, 0.6, 0.9)
```

3 Some example starting values for factors

Note done this yet...

References

Pollock, K.H. and Otto, M.C. 1990. Size bias in line transect sampling: a field test. *Biometrics* **46**: 239-245.