

flexsurv: flexible parametric survival modelling in R. Supplementary examples

Christopher H. Jackson
MRC Biostatistics Unit, Cambridge, UK
chris.jackson@mrc-bsu.cam.ac.uk

Abstract

This vignette of examples supplements the main **flexsurv** user guide.

Keywords: ~survival.

1. Examples of custom distributions

1.1. Proportional hazards generalized gamma model

Crowther and Lambert (2013) discuss using the **stgenreg** Stata package to construct a proportional hazards parameterisation of the three-parameter generalised gamma distribution. A similar trick can be used in **flexsurv**. A four-parameter custom distribution is created by defining its hazard (and cumulative hazard) functions. These are obtained by multiplying the built-in functions **hgengamma** and **Hgengamma** by an extra dummy parameter, which is used as the location parameter of the new distribution. The intercept of this parameter is fixed at 1 when calling **flexsurvreg**, so that the new model is no more complex than the generalized gamma AFT model **fs3**, but covariate effects on the dummy parameter are now interpreted as hazard ratios.

```
> library(flexsurv)
> hgengammaPH <- function(x, dummy, mu=0, sigma=1, Q){
+   dummy * hgengamma(x=x, mu=mu, sigma=sigma, Q=Q)
+ }
> HgengammaPH <- function(x, dummy, mu=0, sigma=1, Q){
+   dummy * Hgengamma(x=x, mu=mu, sigma=sigma, Q=Q)
+ }
> custom.gengammaPH <- list(name="gengammaPH",
+   pars=c("dummy","mu","sigma","Q"), location="dummy",
+   transforms=c(log, identity, log, identity),
+   inv.transforms=c(exp, identity, exp, identity),
+   inits=function(t){
+     lt <- log(t[t>0])
+     c(1, mean(lt), sd(lt), 0)
+   })
```

```

+           })
> fs7 <- flexsurvreg(Surv(recyrs, censrec) ~ group, data=bc,
+                   dist=custom.gengammaPH, fixedpars=1)

```

2. Examples of custom model summaries

2.1. Plotting a hazard ratio against time

The following code plots the hazard ratio (Medium versus Good prognostic group) against time for both the proportional hazards model `fs7` and the better-fitting accelerated failure time model `fs2`. It illustrates the use of the following functions.

`summary.flexsurvreg` for generating the estimated hazard at a series of times, for particular covariate categories.

`normboot.flexsurvreg` for generating a bootstrap-style sample from the sampling distribution of the parameter estimates, for particular covariate categories.

`do.call` for constructing a function call by supplying a list containing the function's arguments. This is used throughout the source of **flexsurv**.

```

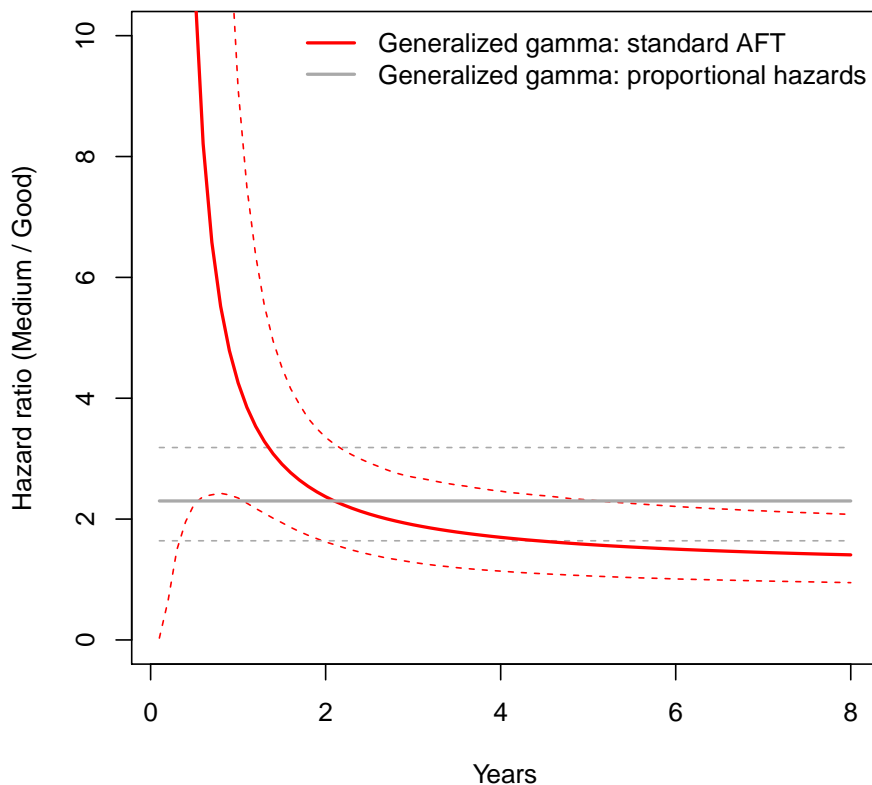
> fs2 <- flexsurvreg(Surv(recyrs, censrec) ~ group + sigma(group),
+                   data=bc, dist="gengamma")
> B <- 5000
> t <- seq(0.1, 8, by=0.1)
> hrAFT.est <-
+   summary.flexsurvreg(fs2, t=t, type="hazard",
+                       newdata=data.frame(group="Medium"),ci=FALSE)[[1]][,"est"] /
+   summary.flexsurvreg(fs2, t=t, type="hazard",
+                       newdata=data.frame(group="Good"),ci=FALSE)[[1]][,"est"]
> pars <- normboot.flexsurvreg(fs2, B=B, newdata=data.frame(group=c("Good","Medium")))
> hrAFT <- matrix(nrow=B, ncol=length(t))
> for (i in seq_along(t)){
+   haz.medium.rep <- do.call(hgengamma, c(list(t[i]), as.data.frame(pars[[2]])))
+   haz.good.rep <- do.call(hgengamma, c(list(t[i]), as.data.frame(pars[[1]])))
+   hrAFT[,i] <- haz.medium.rep / haz.good.rep
+ }
> hrAFT <- apply(hrAFT, 2, quantile, c(0.025, 0.975))
> hrPH.est <-
+   summary.flexsurvreg(fs7, t=t, type="hazard",
+                       newdata=data.frame(group="Medium"),ci=FALSE)[[1]][,"est"] /
+   summary.flexsurvreg(fs7, t=t, type="hazard",
+                       newdata=data.frame(group="Good"),ci=FALSE)[[1]][,"est"]
> pars <- normboot.flexsurvreg(fs7, B=B, newdata=data.frame(group=c("Good","Medium")))
> hrPH <- matrix(nrow=B, ncol=length(t))

```

```

> for (i in seq_along(t)){
+   haz.medium.rep <- do.call(hgengammaPH, c(list(t[i]), as.data.frame(pars[[2]])))
+   haz.good.rep <- do.call(hgengammaPH, c(list(t[i]), as.data.frame(pars[[1]])))
+   hrPH[,i] <- haz.medium.rep / haz.good.rep
+ }
> hrPH <- apply(hrPH, 2, quantile, c(0.025, 0.975))
> plot(t, hrAFT[1,], type="l", ylim=c(0, 10), col="red", xlab="Years",
+   ylab="Hazard ratio (Medium / Good)", lwd=1, lty=2)
> lines(t, hrAFT[2,], col="red", lwd=1, lty=2)
> lines(t, hrPH[1,], col="darkgray", lwd=1, lty=2)
> lines(t, hrPH[2,], col="darkgray", lwd=1, lty=2)
> lines(t, hrAFT.est, col="red", lwd=2)
> lines(t, hrPH.est, col="darkgray", lwd=2)
> legend("topright", lwd=c(2,2), col=c("red","darkgray"), bty="n",
+   c("Generalized gamma: standard AFT", "Generalized gamma: proportional hazards"))

```



References

Crowther MJ, Lambert PC (2013). “**stgenreg**: A Stata package for general parametric survival analysis.” *Journal of Statistical Software*, **53**, 1–17.