BIOST/EPI 537
Survival data analysis for epidemiology

Chapter 4: Regression models in survival analysis: the accelerated failure time model

Marco Carone
Department of Biostatistics
School of Public Health, University of Washington

Winter 2020

LAST UPDATED ON FEB 20TH 2020

Contents of this chapter

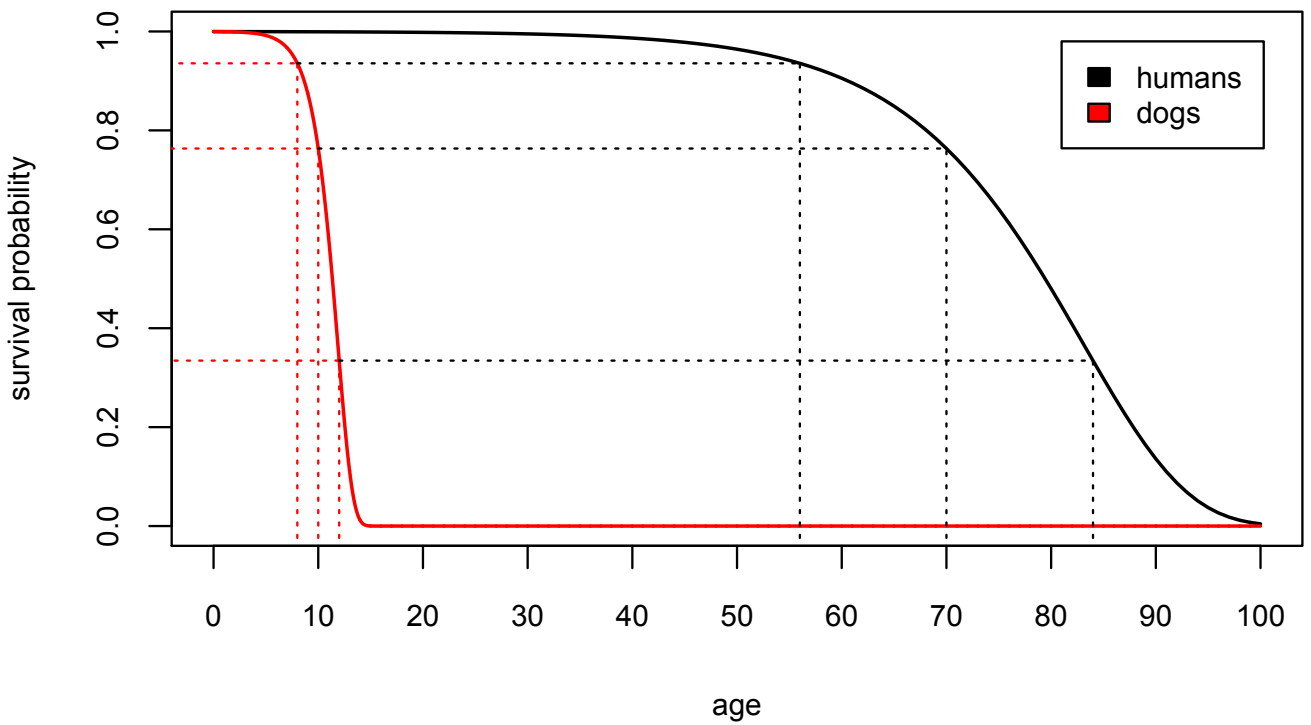
- Formulation, properties and parameter interpretation
- Parametric accelerated failure time model
- Diagnostic tools

Formulation, properties and parameter interpretation

Common belief: a **dog year** is equivalent to **seven human years**.

Denoting by S_0 and S_1 the survival functions of lifetime in humans and in dogs, respectively, we can write this as

$$S_1(t) = S_0(7t) .$$

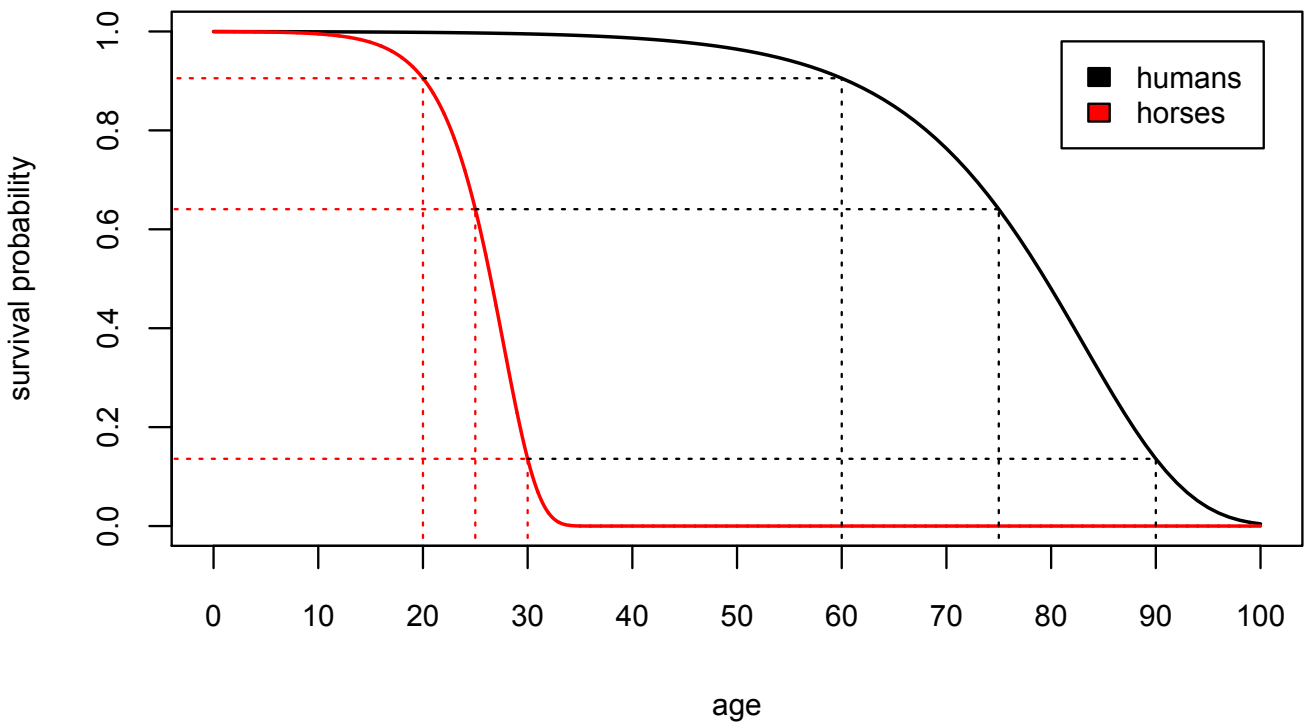


Formulation, properties and parameter interpretation

Less common belief: a **horse year** is often said to be equivalent to **three human years**.

Denoting by S_0 and S_1 the survival functions of lifetime in humans and in horses, respectively, we can write this as

$$S_1(t) = S_0(3t) \ .$$



Formulation, properties and parameter interpretation

Generic formulation of the model: (two subgroups, $Z = 0$ and $Z = 1$)

Denoting by S_0 and S_1 the subgroup-specific survival functions corresponding to $Z = 0$ and $Z = 1$, respectively, the **accelerated failure time (AFT) model** stipulates that

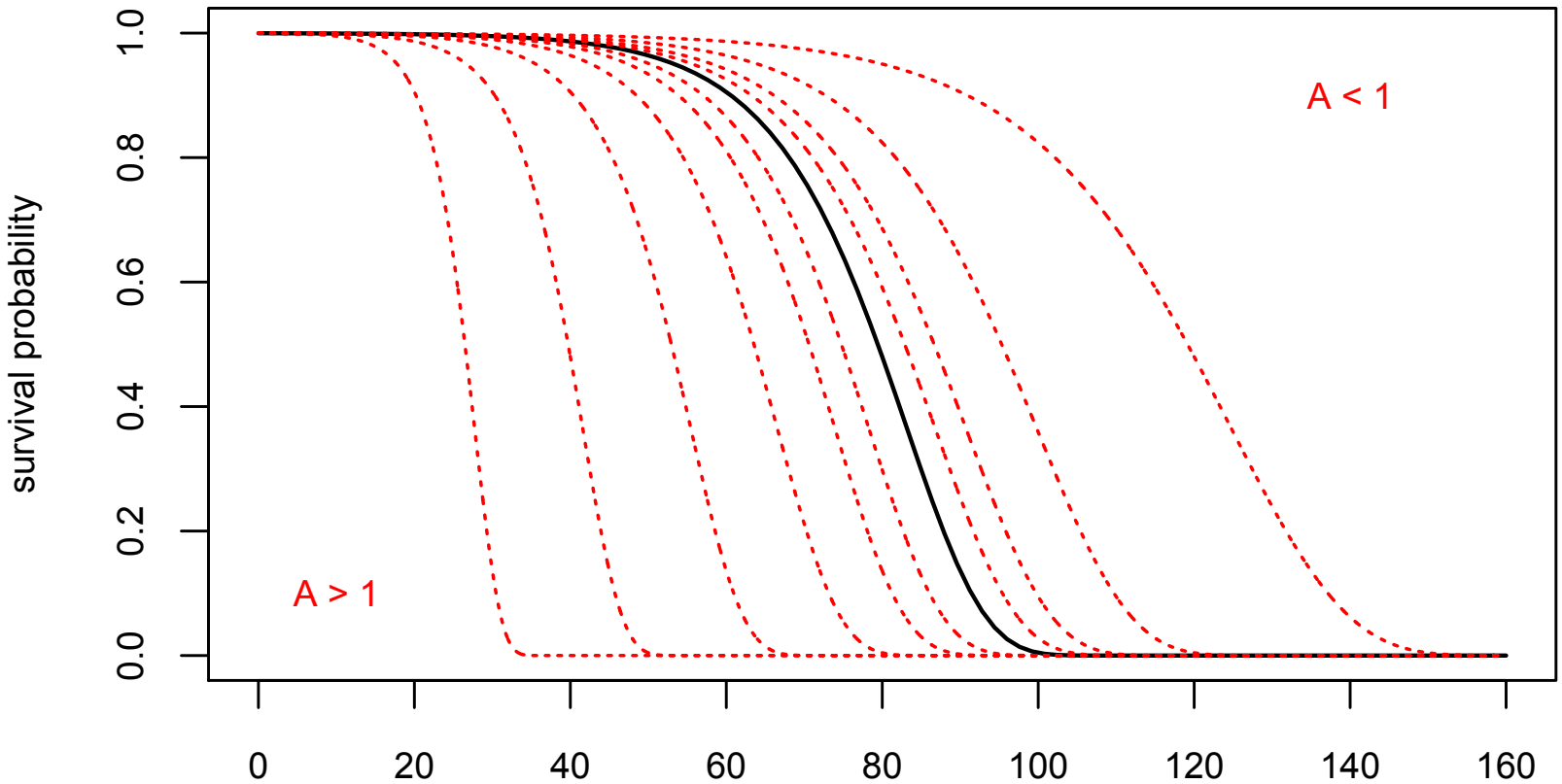
$$S_1(t) = S_0(At)$$

for some unknown **acceleration factor** or **time scaling factor** $A > 0$.

A few observations:

- Progress of time is accelerated when $A > 1$: the time scale is contracted.
- Progress of time is decelerated when $A < 1$: the time scale is stretched out.
- If the baseline survival function S_0 is known, the survival distribution for group $Z = 1$ is completely determined by the scalar acceleration factor A.

Formulation, properties and parameter interpretation



Formulation, properties and parameter interpretation

The AFT model implies important relationships.

$$E(T \mid Z = 1) = A^{-1} \cdot E(T \mid Z = 0)$$

$$\text{median of } T \text{ given } Z = 1 = A^{-1} \cdot \text{median of } T \text{ given } Z = 0$$

. . . and much more.

For example, if $A = 1.2$ and so $A^{-1} \approx 0.83$, we can conclude that. . .

the average (or median) survival time of individuals in subgroup $Z = 1$ is approximately 17% smaller than that of individuals in subgroup $Z = 0$.

If instead $A = 0.8$ and so $A^{-1} = 1.25$, we can conclude that. . .

the average (or median) survival time of individuals in subgroup $Z = 1$ is 25% greater than that of individuals in subgroup $Z = 0$.

Formulation, properties and parameter interpretation

Denoting by $A(z)$ the acceleration factor corresponding to subgroup $Z = z$ (relative to $Z = 0$), so that $A(0) = 1$ and $A(1) = A$, the model can be alternatively written as

$$S(t \mid z) = S_0(A(z)t)$$

for each t and $z = 0, 1$.

This allows us to **generalize the model** from a binary covariate **to arbitrary covariates**.

Suppose that $Z_1, Z_2, \ldots, Z_q$ are available covariates of interest. The general form of the **accelerated failure time model** states that

$$S(t \mid z_1, z_2, \ldots, z_q) = S_0(A(z_1, z_2, \ldots, z_q)t)$$

for each t and $(z_1, z_2, \ldots, z_q)$.

Formulation, properties and parameter interpretation

$$E\left(T \mid Z_1 = z_1, Z_2 = z_2, \ldots, Z_q = z_q\right)$$
$$= A(z_1, z_2, \ldots, z_q)^{-1} \cdot E\left(T \mid Z_1 = 0, Z_2 = 0, \ldots, Z_q = 0\right)$$

$$\text{median of } T \text{ given } Z_1 = z_1, Z_2 = z_2, \ldots, Z_q = z_q$$
$$= A(z_1, z_2, \ldots, z_q)^{-1} \cdot \text{median of } T \text{ given } Z_1 = 0, Z_2 = 0, \ldots, Z_q = 0$$

Most often, we will resort to a **loglinear model** for the acceleration factor:

$$A(z_1, z_2, \ldots, z_q) = e^{-(\varphi_1 z_1 + \varphi_2 z_2 + \ldots + \varphi_q z_q)}$$

with $\varphi := (\varphi_1, \varphi_2, \ldots, \varphi_q)$ a vector of unknown coefficients.

This is a particularly convenient model because it yields the following properties:

- the covariates have a multiplicative effect on the time scale;
- the acceleration factor is strictly positive in each subgroup;
- the acceleration factor is one for the baseline group (i.e., $z_1 = z_2 = \ldots = z_q = 0$).

Formulation, properties and parameter interpretation

Suppose that Z_1 is a binary covariate (e.g., sex: female=0 vs male=1).

$$\frac{E\left(T \mid Z_1 = 1, Z_2 = z_2, \ldots, Z_q = z_q\right)}{E\left(T \mid Z_1 = 0, Z_2 = z_2, \ldots, Z_q = z_q\right)} = \frac{A(1, z_2, \ldots, z_q)^{-1} \cdot E\left(T \mid Z_1 = 0, Z_2 = 0, \ldots, Z_q = 0\right)}{A(0, z_2, \ldots, z_q)^{-1} \cdot E\left(T \mid Z_1 = 0, Z_2 = 0, \ldots, Z_q = 0\right)}$$

$$= \frac{e^{\varphi_1 \cdot 1 + \varphi_2 z_2 + \ldots + \varphi_q z_q}}{e^{\varphi_1 \cdot 0 + \varphi_2 z_2 + \ldots + \varphi_q z_q}} = e^{\varphi_1}$$

$$\frac{\text{median of } T \text{ given } Z_1 = 1, Z_2 = z_2, \ldots, Z_q = z_q}{\text{median of } T \text{ given } Z_1 = 0, Z_2 = z_2, \ldots, Z_q = z_q} = e^{\varphi_1}$$

For subpopulations of equal $Z_2, Z_3, \ldots, Z_q$ but differing in sex, the mean survival time of individuals in the subpopulation corresponding to men is. . .

$(\varphi_1 > 0)$ $\quad$ $\ldots 100 \times [\exp(\varphi_1) - 1]\,\%$ greater. . .

$(\varphi_1 < 0)$ $\quad$ $\ldots 100 \times [1 - \exp(\varphi_1)]\,\%$ smaller. . .

. . . than the mean survival time in the subpopulation corresponding to women.

Similarly for median survival time (and any quantile of the failure time distribution).

Formulation, properties and parameter interpretation

Suppose that Z_1 is a continuous covariate (e.g., systolic blood pressure).

$$\frac{E\left(T \mid Z_1 = z_1 + 1, Z_2 = z_2, \ldots, Z_p = z_p\right)}{E\left(T \mid Z_1 = z_1, Z_2 = z_2, \ldots, Z_q = z_q\right)} = \frac{A(z_1 + 1, z_2, \ldots, z_q)^{-1}}{A(z_1, z_2, \ldots, z_q)^{-1}}$$
$$= \frac{e^{\varphi_1(z_1+1)+\varphi_2 z_2+\ldots+\varphi_q z_q}}{e^{\varphi_1 z_1+\varphi_2 z_2+\ldots+\varphi_q z_q}} = e^{\varphi_1}$$

$$\frac{\text{median of } T \text{ given } Z_1 = z_1 + 1, Z_2 = z_2, \ldots, Z_q = z_q}{\text{median of } T \text{ given } Z_1 = z_1, Z_2 = z_2, \ldots, Z_q = z_q} = e^{\varphi_1}$$

For subpopulations of equal $Z_2, Z_3, \ldots, Z_p$ but differing in SBP by one unit, the mean survival time of individuals in the subpopulation corresponding to higher SBP is. . .

$(\varphi_1 > 0)$ . . . $100 \times [\exp(\varphi_1) - 1]\,\%$ greater. . .

$(\varphi_1 < 0)$ . . . $100 \times [1 - \exp(\varphi_1)]\,\%$ smaller. . .

. . . than the mean survival time in the subpopulation corresponding to lower SBP.

Similarly for median survival time (and any quantile of the failure time distribution).

Formulation, properties and parameter interpretation

The **AFT model** can be equivalently restated via the equation

$$\log T = \varphi_0 + \varphi_1 Z_1 + \varphi_2 Z_2 + \ldots + \varphi_q Z_q + \epsilon \ .$$

The random error ϵ is assumed independent of $Z = (Z_1, Z_2, \ldots, Z_q)$ and of mean zero.

Thus, the AFT model is nothing more than a **linear model for log survival time**!

The baseline survival function S_0 is determined by the intercept φ_0 and the error distribution since

$$\begin{aligned} S_0(t) &= P(T > t \mid Z = 0) = P(\log T > \log t \mid Z = 0) \\ &= P(\varphi_0 + \varphi_1 Z_1 + \varphi_2 Z_2 + \ldots + \varphi_q Z_q + \epsilon > \log t \mid Z = 0) \\ &= P(\varphi_0 + \epsilon > \log t) = P(\epsilon > \log t - \varphi_0) \ . \end{aligned}$$

Causal interpretation of regression coefficient

Provided the set of confounders adjusted for is sufficiently rich, the regression coefficient in an AFT model may be interpreted as a population-averaged causal effect.

In such cases, the AFT model $\log T = \varphi_0 + \varphi_A A + \varphi_1 Z_1 + \varphi_2 Z_2 + \ldots + \varphi_q Z_q + \epsilon$, where A is a binary treatment, implies that

$$\begin{aligned} \varphi_A &= E[E[\log T \mid A = 1, Z_1, Z_2 \ldots, Z_q] - E[\log T \mid A = 0, Z_1, Z_2 \ldots, Z_q]] \\ &= \text{average treatment effect on log survival time .} \end{aligned}$$

In contrast, the hazard ratio cannot be interpreted as a population-averaged causal effect, and in fact, can have quite a problematic interpretation.

- the HR generally changes over time, so the estimand is a weighted average of period-specific HRs;
- because the relative composition of the treatment groups changes over time, the HR is subject to a selection effect that complicates the interpretation of the HR.

(For example, read *The hazards of hazard ratios* by Miguel Hernán.)

Interpretation of hazard ratios in the context of cures

Many times, the proportional hazards and AFT models are used for survival time distribution that include a so-called **cure fraction** – that is, a portion of the population never experiences the terminating event.

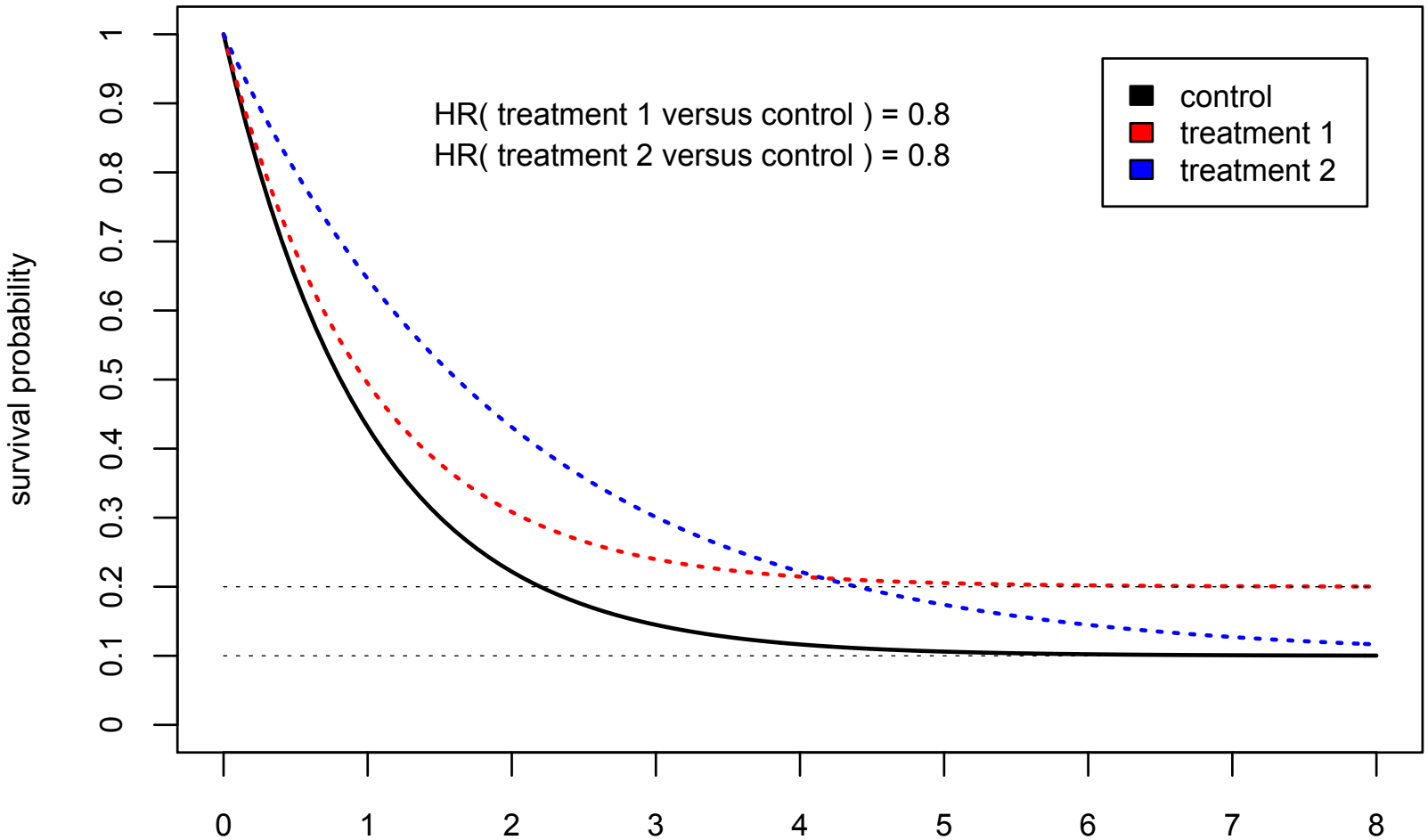
Whenever there is a cure fraction, the HR cannot simply be interpreted as suggesting that the terminating event occurs more or less rapidly, as is often thought.

In a study of treatment versus control, we may find HR < 1 because:

(a) the terminating event occurs less rapidly in the treated patients;
(b) a greater proportion of treated patients experience a cure, even though in those who do not, the survival distribution is identical;
(c) the cure fraction and the survival distribution in those not cured both vary.

In an AFT model, the cure fraction is not allowed to change between groups, so the regression parameter isolates the dilation/contraction of time. A simple extension of the AFT model allows separate modeling of changes in the cure fraction.

Formulation, properties and parameter interpretation



Parametric accelerated failure time model

Suppose we assume the AFT model $S(t \mid z) = S_0(te^{-(\varphi_1 z_1 + \varphi_2 z_2 + \ldots + \varphi_q z_q)})$.

Then what? How do we proceed?

- In practice, neither S_0 nor $\varphi := (\varphi_1, \varphi_2, \ldots, \varphi_q)$ are known, though it is possible to estimate them using the available data.
- However, without further assumptions, estimating S_0 can be difficult.
- In practice, we often assume a parametric model for S_0:
  - e.g., S_0 = exponential survival function with unknown parameter $\theta = \lambda$;
  - e.g., S_0 = Weibull survival function with unknown parameter $\theta = (\lambda, p)$;
  - e.g., S_0 = Gamma survival function with unknown parameter $\theta = (\lambda, \alpha)$;
  - e.g., S_0 = generalized Gamma survival function with unknown parameter $\theta = (\lambda, \alpha, p)$.
- This reduces the *unknowns* to θ and φ – we can use a usual maximum likelihood approach to find good estimators of these unknowns.

Parametric accelerated failure time model

Suppose that we assume a **Weibull AFT model** so that the baseline survival function is

$$S_0(t) = e^{-(\lambda t)^p}$$

for some unknown parameter values λ and p.

The subgroup-specific **survival function** is given by

$$S(t \mid z) = S_0(e^{-\varphi z} t) = \exp\left[-(\lambda e^{-\varphi z} t)^p\right] = \exp\left(-\lambda^p e^{-\varphi p z} t^p\right)$$

and its **density function** is $f(t \mid z) = \lambda^p e^{-\varphi p z} p t^{p-1} \exp\left(-\lambda^p e^{-\varphi p z} t^p\right)$.

As such, the likelihood contribution of an individual with data (y, δ, z) is

$$\left[\lambda^p e^{-\varphi z} p t^{p-1} \exp\left(-\lambda^p e^{-\varphi z} t^p\right)\right]^{\delta} \left[\exp\left(-\lambda^p e^{-\varphi z} t^p\right)\right]^{1-\delta} .$$

Parametric accelerated failure time model

```r
options(width=100)

herpes = read.csv("herpes.csv")
s.herpes = with(herpes, Surv(timetorec, event))
herpes$treat. = as.factor(herpes$treat)
herpes$type. = as.factor(herpes$type)

coxfit = coxph(s.herpes~treat.+type.+duration+age+male, data=herpes)
summary(coxfit)$coef
```

```
##                 coef exp(coef)    se(coef)          z     Pr(>|z|)
## treat.1   0.05794035 1.0596518 0.201157083  0.2880354 7.733197e-01
## treat.2   0.39637536 1.4864272 0.124801315  3.1760511 1.492946e-03
## treat.3  -0.24892000 0.7796423 0.326222789 -0.7630368 4.454414e-01
## type.2    1.23151989 3.4264334 0.182783530  6.7375867 1.610390e-11
## type.3    1.57730667 4.8418974 0.218852895  7.2071547 5.713208e-13
## duration  0.02356189 1.0238417 0.007004126  3.3640010 7.682127e-04
## age      -0.02326305 0.9770054 0.009266324 -2.5104942 1.205623e-02
## male      0.30606991 1.3580772 0.114391630  2.6756320 7.458849e-03
```

Parametric accelerated failure time model

```
flexsurvreg(s.herpes~treat.+type.+duration+age+male, data=herpes, dist="gengamma")
```

```
## Call:
## flexsurvreg(formula = s.herpes ~ treat. + type. + duration +
##     age + male, data = herpes, dist = "gengamma")
##
## Estimates:
##           data mean  est       L95%      U95%      se        exp(est)  L95%      U95%
## mu              NA    5.68776   4.83975   6.53577   0.43266        NA        NA        NA
## sigma           NA    1.43131   1.31571   1.55707   0.06150        NA        NA        NA
## Q               NA    0.12616  -0.18551   0.43783   0.15902        NA        NA        NA
## treat.1    0.08811   -0.00211  -0.51300   0.50879   0.26067   0.99790   0.59869   1.66328
## treat.2    0.28634   -0.49608  -0.81410  -0.17805   0.16226   0.60892   0.44304   0.83690
## treat.3    0.03524    0.45797  -0.34108   1.25703   0.40769   1.58087   0.71100   3.51495
## type.2     0.71806   -1.68352  -2.10763  -1.25941   0.21639   0.18572   0.12153   0.28382
## type.3     0.12996   -2.15261  -2.68838  -1.61684   0.27336   0.11618   0.06799   0.19852
## duration  19.56828   -0.03085  -0.04833  -0.01336   0.00892   0.96963   0.95282   0.98673
## age       25.44273    0.02947   0.00508   0.05386   0.01244   1.02991   1.00509   1.05534
## male       0.32819   -0.25339  -0.56214   0.05536   0.15753   0.77616   0.56999   1.05692
##
## N = 454,  Events: 371,  Censored: 83
## Total time at risk: 45953
## Log-likelihood = -2054.5, df = 11
## AIC = 4130.999
```

Parametric accelerated failure time model

If the distribution of T given Z follows a **Weibull AFT model** as in the last slide, the subgroup-specific hazard function is

$$h(t \mid z) = \frac{f(t \mid z)}{S(t \mid z)} = \frac{\lambda^p e^{-\varphi pz} p t^{p-1} \exp\left(-\lambda^p e^{-\varphi pz} t^p\right)}{\exp\left(-\lambda^p e^{-\varphi pz} t^p\right)} = \underbrace{\lambda^p p t^{p-1}}_{h_0(t)} e^{-\varphi pz}.$$

So, it is a **Cox model** with Weibull baseline and regression coefficient $\beta = -\varphi p$.

Similarly, if the distribution of T given Z follows a **Cox model** with Weibull baseline and regression coefficient β, it also follows an **AFT model** with Weibull baseline and acceleration parameter $\varphi = -\beta/p$.

The Weibull distribution is the only parametric family for which both the AFT and the Cox model formulations are satisfied simultaneously!

Parametric accelerated failure time model

```
flexsurvreg(s.herpes~treat.+type.+duration+age+male, data=herpes, dist="weibull")

## Call:
## flexsurvreg(formula = s.herpes ~ treat. + type. + duration +
##     age + male, data = herpes, dist = "weibull")
##
## Estimates:
##           data mean  est        L95%       U95%       se         exp(est)   L95%       U95%
## shape            NA   8.13e-01   7.50e-01   8.82e-01   3.36e-02         NA         NA         NA
## scale            NA   4.90e+02   2.19e+02   1.10e+03   2.02e+02         NA         NA         NA
## treat.1    8.81e-02  -1.13e-01  -5.99e-01   3.73e-01   2.48e-01   8.93e-01   5.50e-01   1.45e+00
## treat.2    2.86e-01  -5.36e-01  -8.34e-01  -2.37e-01   1.52e-01   5.85e-01   4.34e-01   7.89e-01
## treat.3    3.52e-02   2.95e-01  -4.91e-01   1.08e+00   4.01e-01   1.34e+00   6.12e-01   2.95e+00
## type.2     7.18e-01  -1.63e+00  -2.07e+00  -1.19e+00   2.24e-01   1.96e-01   1.26e-01   3.04e-01
## type.3     1.30e-01  -2.05e+00  -2.58e+00  -1.53e+00   2.68e-01   1.28e-01   7.59e-02   2.17e-01
## duration   1.96e+01  -3.02e-02  -4.69e-02  -1.34e-02   8.54e-03   9.70e-01   9.54e-01   9.87e-01
## age        2.54e+01   3.19e-02   9.66e-03   5.41e-02   1.13e-02   1.03e+00   1.01e+00   1.06e+00
## male       3.28e-01  -4.38e-01  -7.11e-01  -1.66e-01   1.39e-01   6.45e-01   4.91e-01   8.47e-01
##
## N = 454,  Events: 371,  Censored: 83
## Total time at risk: 45953
## Log-likelihood = -2068.979, df = 10
## AIC = 4157.959
```

Parametric accelerated failure time model

```
flexsurvreg(s.herpes~treat.+type.+duration+age+male, data=herpes, dist="weibullPH")

## Call:
## flexsurvreg(formula = s.herpes ~ treat. + type. + duration +
##     age + male, data = herpes, dist = "weibullPH")
##
## Estimates:
##           data mean  est       L95%      U95%      se        exp(est)  L95%      U95%
## shape           NA    0.81336   0.75015   0.88189   0.03357        NA        NA        NA
## scale           NA    0.00649   0.00305   0.01381   0.00250        NA        NA        NA
## treat.1    0.08811    0.09194  -0.30331   0.48719   0.20166   1.09630   0.73837   1.62773
## treat.2    0.28634    0.43574   0.19102   0.68047   0.12486   1.54611   1.21048   1.97481
## treat.3    0.03524   -0.24001  -0.87890   0.39889   0.32597   0.78662   0.41524   1.49016
## type.2     0.71806    1.32474   0.96712   1.68236   0.18246   3.76120   2.63035   5.37822
## type.3     0.12996    1.67006   1.24170   2.09842   0.21856   5.31249   3.46148   8.15330
## duration  19.56828    0.02454   0.01083   0.03825   0.00699   1.02484   1.01089   1.03899
## age       25.44273   -0.02592  -0.04406  -0.00778   0.00925   0.97441   0.95690   0.99225
## male       0.32819    0.35663   0.13315   0.58010   0.11402   1.42850   1.14243   1.78622
##
## N = 454,  Events: 371,  Censored: 83
## Total time at risk: 45953
## Log-likelihood = -2068.979, df = 10
## AIC = 4157.959
```

Model diagnostics

Many of the diagnostic tools introduced in Chapter 3 can also be used to assess the fit of the AFT model and to investigate potential outliers and influential observations.

These include Cox-Snell, martingale and deviance residuals as well as delta-betas.

An additional diagnostic tool tailored to the AFT model:

If the AFT model holds, then for each z and u it must be that

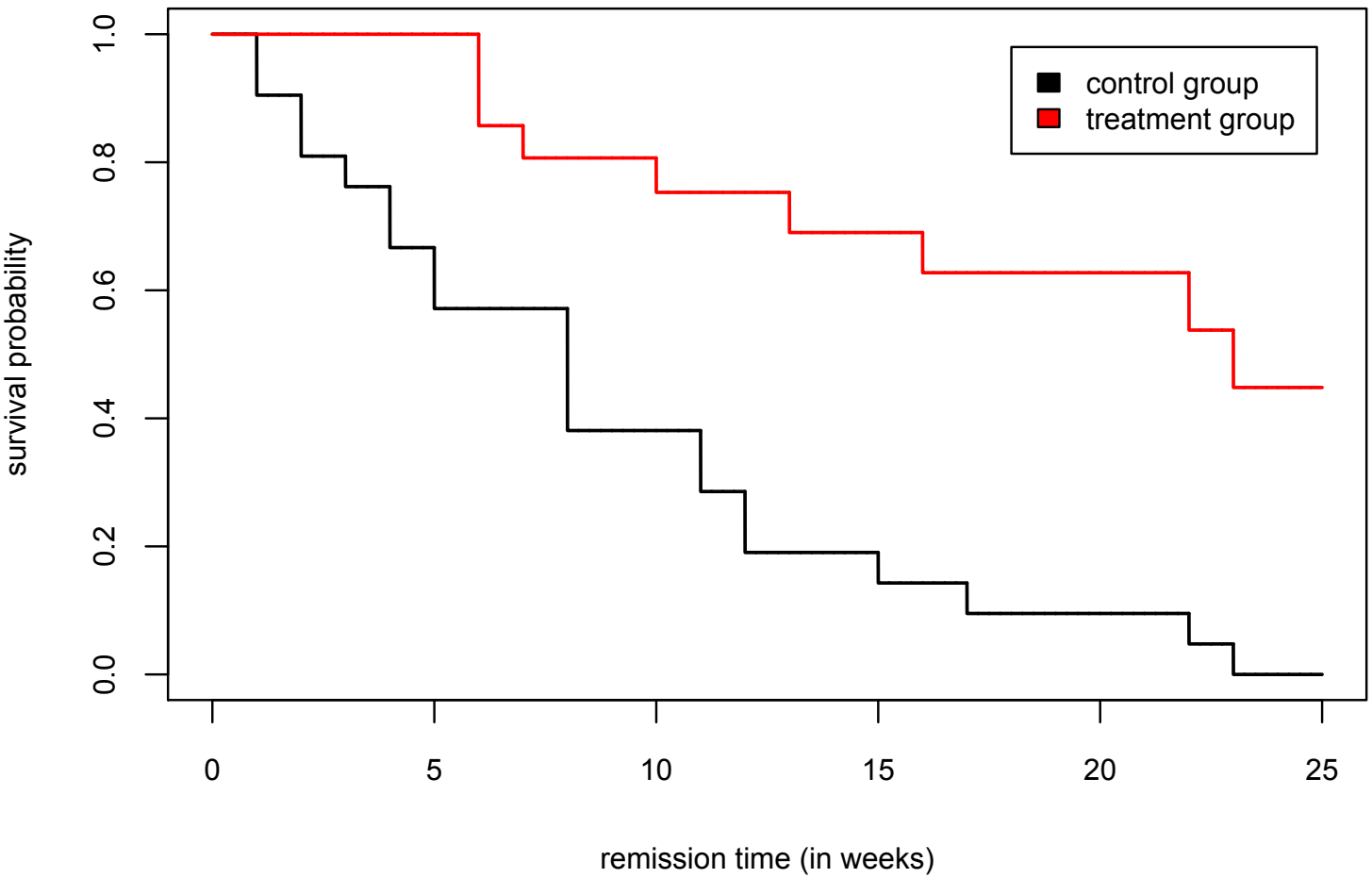
$$S(e^{\varphi z} u \mid z) = S_0(u) \ .$$

If Z is categorical, this suggests the following simple diagnostic scheme:

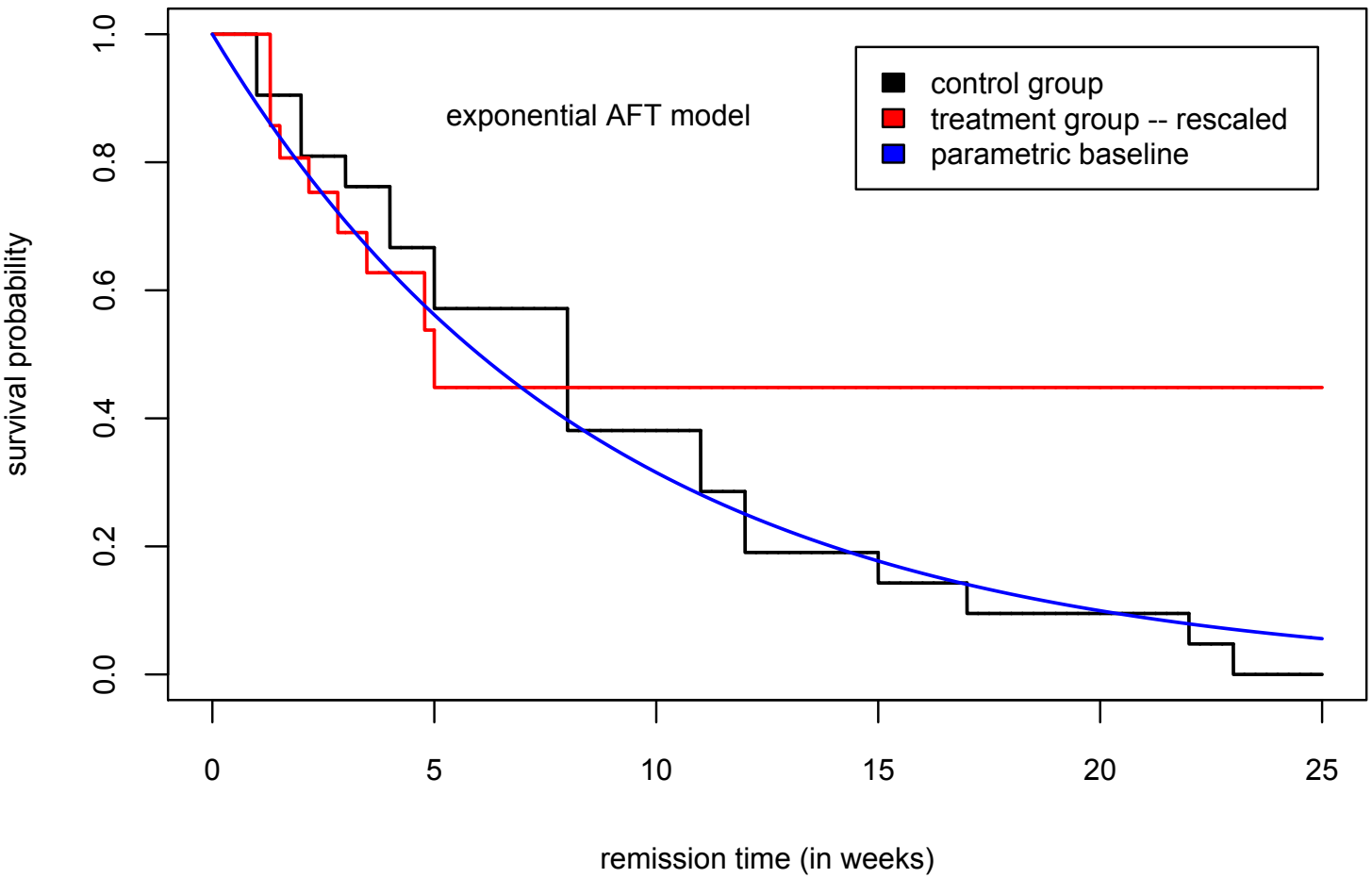
1. fit postulated AFT model to obtain estimate of φ and S_0;
2. estimate $S(t \mid z)$ using the KM estimator restricted to observations with $Z = z$;
3. on the same graph, plot versus u each of $\widehat{S}(e^{\hat{\varphi} z} u \mid z)$ (for each z) and $\widehat{S}_0(u)$.

If the AFT model is correctly specified, all curves should be approximately equal.

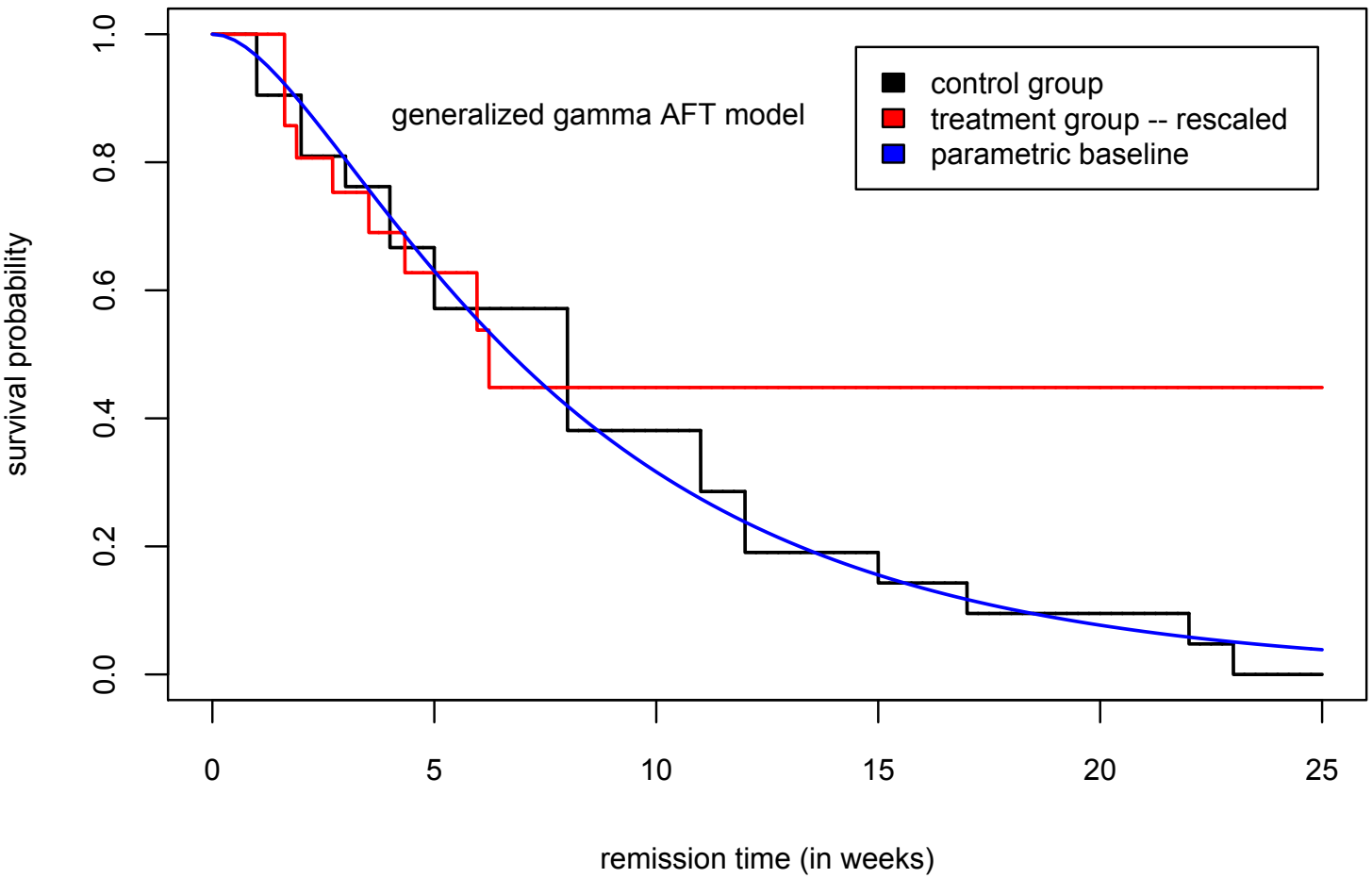
Model diagnostics



Model diagnostics



Model diagnostics



Additional reading and references

ADDITIONAL READING:

- Moore, DF. *Applied survival analysis using R.* Chapter 10.
- Hosmer, DW, Lemeshow, S, May, S. *Applied survival analysis.* (2nd ed.) Chapter 8.
- Kleinbaum, DG, Klein, M. *Survival analysis: a self-learning text.* (3rd ed.) Chapter 7.

- Hernán, M.A. 2010 *The hazards of hazard ratios.* Epidemiology.
- Struthers, C.A., Kalbfleisch, J.S. 1986 *Misspecified proportional hazard models.* Biometrika.
- Hutton, J.L., Monaghan, P.F. 2002 *Choice of parametric accelerated life and proportional hazards models for survival data: asymptotic results.* Lifetime Data Analysis.
- Keiding, N., Andersen, P.K., Klein, J.P. 1997 *The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates.* Statistics in Medicine.
- Yamaguchi, K. 1992 *Accelerated failure-time regression models with a regression model of surviving fraction: an application to the analysis of "permanent employment" in Japan.* Journal of the Americal Statistical Association.