

## Parameterization of Survival Functions in TreeAge Pro and STATA

We compare the parameterization of survival functions based on distributions available in TreeAge Pro and the parametric survival analysis package `streg` in STATA.

### 1. Reference table for the relationship between parameters:

Distribution	Parameterization used in		Relationship
	TreeAge Pro	streg in STATA	
Exponential	$\lambda$	$\lambda$	$\lambda_{\text{TreeAge}} = \lambda_{\text{Stata}}$
Weibull	$\lambda, k$	$\lambda, p$	$\lambda_{\text{TreeAge}} = \lambda_{\text{Stata}}, k_{\text{TreeAge}} = p_{\text{Stata}}$
Lognormal	$\mu, \sigma$	$\mu, \sigma$	$\mu_{\text{TreeAge}} = \mu_{\text{Stata}}, \sigma_{\text{TreeAge}} = \sigma_{\text{Stata}}$
Log-Logistic	$a, b$	$\lambda, \gamma$	$a_{\text{TreeAge}} = 1/\lambda_{\text{Stata}}, b_{\text{TreeAge}} = 1/\gamma_{\text{Stata}}$
Gompertz	$\lambda, \gamma$	$\lambda, \gamma$	$\lambda_{\text{TreeAge}} = \lambda_{\text{Stata}}, \gamma_{\text{TreeAge}} = \gamma_{\text{Stata}}$
Generalized Gamma	$c, \alpha, \beta$	$\kappa, \mu, \sigma$	?

### 2. Reference table for comparing parameterization of survival functions:

Distribution	Survival Function in	
	TreeAge Pro	streg in STATA
Exponential	$S(x) = e^{-\lambda x}$	$S(t) = e^{-\lambda t}$
Weibull	$S(x) = e^{-\lambda x^k}$	$S(t) = e^{-\lambda t^p}$
Lognormal	$S(x) = 1 - \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right)$	$S(t) = 1 - \Phi\left(\frac{\log(t) - \mu}{\sigma}\right)$
Log-Logistic	$S(x) = [1 + (x/a)^b]^{-1}$	$S(t) = [1 + (\lambda t)^{1/\gamma}]^{-1}$
Gompertz	$S(x) = e^{-(\lambda/\gamma)(e^{\gamma x} - 1)}$	$S(t) = e^{-(\lambda/\gamma)(e^{\gamma t} - 1)}$
Generalized Gamma	$S(x) = 1 - I(\alpha, (x/\beta)^c)$ , where $I(\alpha, x) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t} dt$	$S(t) = 1 - I(\kappa^{-2}, \kappa^{-2} e^{\kappa((\log(t)-\mu)/\sigma)})$ ( $\kappa > 0$ )

3. **Summary:** The parameterization for the Exponential, Weibull, Lognormal and Gompertz distribution is equivalent in TreeAge Pro and STATA, save for parameter names in the Weibull case. For the Log-logistic distribution, however, one needs to be careful as the parameters used by STATA are inversely related to those used by TreeAge Pro. The conversion between the parameters used for the generalized gamma in STATA and the parameters used in TreeAge have not yet been determined.

## Parameterization of Survival Functions in TreeAge Pro and SAS

We compare the parameterization of survival functions based on distributions available in TreeAge Pro and the procedure `lifereg` in SAS. For parameterization in SAS, please consult the [SAS help page](#) for the `lifereg` procedure and look for “Supported Distributions”. It tells you how the SAS output relates to the parameters of the distribution. For instance, “if there are no covariates in the model,  $\mu$  = Intercept in the output; otherwise,  $\mu = x'\beta$ .  $\sigma$  = Scale in the output.” Then, under each distribution, you find how  $\mu$  and  $\sigma$  are related to the parameters of the distribution. For instance, when the exponential distribution is chosen, what SAS reports for the intercept (when there are no covariates) in the model is related to  $\alpha_{SAS} = e^{-\text{Intercept}}$ . In turn, the SAS parameter  $\alpha_{SAS}$  for the exponential distribution in SAS is the same as the parameter called  $\lambda$  for the exponential distribution in TreeAge, i.e.,  $\lambda_{\text{TreeAge}} = \alpha_{SAS}$ . When the Weibull distribution is chosen, what SAS labels “scale” in its output refers to  $\sigma_{SAS}$ , which in turn is related to  $\gamma_{SAS}$  of the Weibull distribution in SAS through  $\gamma_{SAS} = 1/\sigma_{SAS}$ , and this is equivalent to the parameter  $k_{\text{TreeAge}}$  used in TreeAge’s Weibull distribution, i.e.,  $k_{\text{TreeAge}} = \gamma_{SAS} = 1/\sigma_{SAS}$ .

### 1. Reference table for the relationship between parameters:

Distribution	Parameterization used in		Relationship
	TreeAge Pro	lifereg in SAS	
Exponential	$\lambda$	$\alpha$	$\lambda_{\text{TreeAge}} = \alpha_{SAS}$
Weibull	$\lambda, k$	$\alpha, \gamma$	$\lambda_{\text{TreeAge}} = \alpha_{SAS}, k_{\text{TreeAge}} = \gamma_{SAS}$
Lognormal	$\mu, \sigma$	$\mu, \sigma$	$\mu_{\text{TreeAge}} = \mu_{SAS}, \sigma_{\text{TreeAge}} = \sigma_{SAS}$
Log-Logistic	$a, b$	$\alpha, \gamma$	$a_{\text{TreeAge}} = (\alpha_{SAS})^{-1/\gamma_{SAS}}, b = \gamma_{SAS}$

### 2. Reference table for comparing parameterization of survival functions:

Distribution	Survival Function in	
	TreeAge Pro	lifereg in SAS
Exponential	$S(x) = e^{-\lambda x}$	$G(t) = e^{-\alpha t}$
Weibull	$S(x) = e^{-\lambda x^k}$	$G(t) = e^{-\alpha t^\gamma}$
Lognormal	$S(x) = 1 - \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right)$	$G(t) = 1 - \Phi\left(\frac{\log(t) - \mu}{\sigma}\right)$
Log-Logistic	$S(x) = [1 + (x/a)^b]^{-1}$	$G(t) = [1 + \alpha t^\gamma]^{-1}$

3. **Summary:** The parameterization for the Exponential, Weibull and Lognormal distribution is equivalent in TreeAge Pro and SAS, except for parameter names in the Exponential and Weibull case. For the Log-logistic distribution, however, one needs to be careful as the parameters used by SAS are combinations of the parameters used by TreeAge Pro. (The Gompertz is not available in SAS.)

## Parameterization of Survival Functions in TreeAge Pro and R

We compare the parameterization of survival functions based on distributions available in TreeAge Pro and the library `survival` in R, which is the standard library to use for time to event modeling in R. Within that library, the command `survreg` fits parametric survival models.

Note: The help on `survreg` mentions that the parameterization of available distributions is based on Chapter 2.2 of Kalbfleisch and Prentice “The statistical analysis of failure time data,” Wiley, 2002. (The notation  $F(t)$  for the survival function listed there is used in the Table below, although the more common notation would be  $S(t)$ .) The R output when fitting data with these distributions refers to the distribution of the error term in the model for the mean of the log of the survival time. For instance, what R reports under scale when using the Weibull distribution is the estimate of what is called  $\sigma$  in Kalbfleisch and Prentice, which in turn is related to the parameter  $\gamma$  of the Weibull distribution that `survreg` uses through  $\gamma = 1/\sigma$ . In fact, for all distributions fitted with R, we have that  $\gamma = 1/\sigma$  and  $\lambda = e^{-\mu}$ , where  $\mu$  is the modeled mean of the log survival time. The Table below translates the parameters  $\lambda$  and  $\gamma$  to the corresponding parameters that TreeAge uses. The Gompertz and generalized Gamma distributions are only available via the `flexsurv` package (which, for the other distributions above, uses a different (!) parameterization than the `survreg` package). In addition, the generalized gamma distribution is available in two different parametrization.

### 1. Reference table for the relationship between parameters:

Distribution	Parameterization used in		Relationship
	TreeAge Pro	<code>survreg</code> in R	
Exponential	$\lambda$	$\lambda$	$\lambda_{\text{TreeAge}} = \lambda_R$
Weibull	$\lambda, k$	$\lambda, \gamma$	$\lambda_{\text{TreeAge}} = (\lambda_R)^{\gamma_R}, k_{\text{TreeAge}} = \gamma_R$
Lognormal	$\mu, \sigma$	$\lambda, \gamma$	$\mu_{\text{TreeAge}} = -\log(\lambda_R), \sigma_{\text{TreeAge}} = 1/\gamma_R$
Log-Logistic	$a, b$	$\lambda, \gamma$	$a_{\text{TreeAge}} = 1/\lambda_R, b_{\text{TreeAge}} = \gamma_R$
Gompertz	$\lambda, \gamma$	$a, b^*$	$\lambda_{\text{TreeAge}} = b_R, \gamma_{\text{TreeAge}} = a_R$
Generalized Gamma	$c, \alpha, \beta$	$a, b, k^{**}$	$c_{\text{TreeAge}} = b_R, \alpha_{\text{TreeAge}} = k_R, \beta_{\text{TreeAge}} = a_R$
		$\mu, \sigma, Q^{***}$	$c_{\text{TreeAge}} = Q_R/\sigma_R, \alpha_{\text{TreeAge}} = 1/Q_R^2,$ $\beta_{\text{TreeAge}} = e^{\mu_R + (\sigma_R \log(Q_R^2)/Q_R)}$

\* from `flexsurv`

\*\* from `flexsurv`, using `flexsurvreg(..., dist="gengamma.orig", ...)`

\*\*\* from `flexsurv`, using `flexsurvreg(..., dist="gengamma", ...)`. Conversion only works when  $Q > 0$ .

2. **Reference table for comparing parameterization of survival functions:**

Distribution	Survival Function in	
	TreeAge Pro	survreg in R
Exponential	$S(x) = e^{-\lambda x}$	$F(t) = e^{-\lambda t}$
Weibull	$S(x) = e^{-\lambda x^k}$	$F(t) = e^{-(\lambda t)^\gamma}$
Lognormal	$S(x) = 1 - \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right)$	$F(t) = 1 - \Phi(\gamma \log(\lambda t))$
Log-Logistic	$S(x) = [1 + (x/a)^b]^{-1}$	$F(t) = [1 + (\lambda t)^\gamma]^{-1}$
Gompertz	$S(x) = e^{-(\lambda/\gamma)(e^{\gamma x} - 1)}$	$F(t) = e^{-(b/a)(e^{at} - 1)}$ *
Generalized Gamma	$S(x) = 1 - I(\alpha, (x/\beta)^c)$ , where $I(\alpha, x) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t} dt$	$F(t) = 1 - I(k, (x/a)^b)$ **

\* from flexsurv

\*\* from flexsurv, using `flexsurvreg(..., dist="gengamma.orig", ...)`

3. **Summary:** Only the parameterization for the Exponential distribution is equivalent in TreeAge Pro and R's `survreg` function. For the Gompertz and generalized Gamma distribution in original form with the `flexsurv` package, the parameterization is the same, only parameter names differ. If the new parameterization for the generalized Gamma is used with `flexsurv`, these new parameters can be converted to the old parameters, and then to TreeAge parameters. For all others, one needs to be careful to enter parameters estimated with R into specifications for distributions in TreeAge Pro.

4. **Example:** The following R dataset shows survival times (in minutes), some of which are censored (Censor = 0 indicates that the observation was right censored):

> mydat					
	Minutes	Censor			
1	11	1	20	20	0
2	19	1	21	19	1
3	21	1	22	21	1
4	20	1	23	20	1
5	30	1	24	25	1
6	16	1	25	24	0
7	23	1	26	21	1
8	25	0	27	23	1
9	26	0	28	24	1
10	32	0	29	30	0
11	12	1	30	19	1
12	19	1	31	20	1
13	21	1	32	21	1
14	21	1	33	27	1
15	21	0	34	14	1
16	16	1	35	21	1
17	23	1	36	23	1
18	23	1	37	24	1
19	32	0	38	30	1

To fit a parametric Weibull model to these survival times, we use:

```
> fit.wei = survreg(Surv(Minutes, Censor) ~ 1, data=mydat, dist='weibull')
```

This gives the following (edited) output:

```
> summary(fit.wei)
```

	Value	Std. Error	z	p
(Intercept)	3.22	0.0423	76.1	0.00e+00
Log(scale)	-1.47	0.1432	-10.2	1.28e-24

Scale= 0.231

Hence, the fitted parameter  $\lambda_R$  appearing in R's parameterization of the Weibull survival function is  $\hat{\lambda}_R = e^{-\hat{\mu}} = e^{-3.22} = 0.04$  and the fitted parameter  $\gamma_R$  appearing in R's parameterization of the Weibull survival function is  $\hat{\gamma}_R = 1/\hat{\sigma} = 1/0.231 = 4.33$ . (With this, the fitted survival function in R is  $F(t) = e^{-(0.04t)^{4.33}}$ .)

To transfer these two parameter estimates to the Weibull distribution in TreeAge Pro, in the TreeAge Pro menu set  $\lambda_{\text{TreeAge}} = (0.04)^{4.33}$  and  $k_{\text{TreeAge}} = 4.33$ .