Parametric Models

1 Parametric Models: The Intuition

As we saw early, a central component of duration analysis is the hazard rate. The hazard rate is the probability of experiencing an event at time t_i conditional on having survived to time t_i. The precise shape of the hazard rate – the way it changes with time – is likely to vary from one situation to the next. For example, the hazard rate might increase with time in some situations:

$$\frac{dh(t)}{dt} > 0 \tag{1}$$

This means that the risk of an event occurring increases with time. This type of situation exhibits what is often referred to as *positive duration dependence*. The hazard rate might decrease with time in other situations:

$$\frac{dh(t)}{dt} < 0 \tag{2}$$

This means that the risk of an event occurring decreases with time. This type of situation exhibits what is often referred to as *negative duration dependence*. The hazard rate might be constant across time in still other situations:

$$\frac{dh(t)}{dt} = 0 \tag{3}$$

This means that the risk of an event occurring is constant over time. We can also think of other situations with more complicated hazard rates that increase and decrease over time or that increase or decrease at faster or slower rates. Exactly how the hazard rate varies with time is generally referred to as *time dependency*.

The logic of parametric duration models is that they assume a particular shape for the hazard rate. Below are some of the common parametric models:

- Exponential
- Weibull
- Lognormal
- Log-Logistic
- Gompertz
- Generalized Gamma

Each of these different parametric models specifies a particular shape for the hazard rate i.e. the time dependency. For example, the exponential assumes a flat hazard; the weibull assumes a monotonic hazard; the log-normal and log-logistic assume a non-monotonic hazard. As I noted earlier, if the characterization of the underlying time-dependency is accurate - you pick the right distributional function – then parameter estimates will generally be more precise than estimates from semi-parametric and nonparametric models where the underlying time-dependency is left unspecified. So, there can be advantages to using parametric models. Problems arise, however, if you pick the wrong parametric form.

We will look a little later at how we might evaluate whether we have the ‘right’ shape for the time-dependency. I should just note here that people often provide poor justifications for the particular parametric model that they use.

- Some people look at the shape of the hazard function when there are no covariates – we did this when we looked at nonparametric models. They then look for the parametric model that matches this shape. The problem with this approach is that the shape of the hazard function is conditional on the covariates in the model. Once you start adding covariates, the shape of the underlying hazard function changes. As a result, the shape of the hazard function when there are no covariates may not be a good guide to the shape of the hazard function when there are covariates.

- Some people state that theory should determine the parametric model that they use. For example, they might argue that the risk that a government will collapse increases over time and that they should therefore use a parametric model that allows for monotonically increasing hazards such as the Weibull. The problem, though, arises once we recognize that time dependency is essentially what is left over after we have conditioned on some covariates. In other words, time dependency indicates how the hazard changes with time *after conditioning on covariates*. In effect, time dependency is what we cannot explain with our covariates. We might think that if we had all the right covariates in our model that there would be no time dependency – the hazard would be flat (the exponential model). Of course, we never have the 'right' model and so we have to try to capture the shape of the time dependency. The issue here, though, is that theory is unlikely to tell us about the shape of the time dependency once we've started adding covariates. Thus, theory should play a role in model selection but it should generally play a role only in helping us select what covariates to put in our model.

- Some people state that we should adopt flexible parametric models. There is certainly some truth to this, but we need to be careful about which parametric models are really more flexible than others. For example, the Weibull model is commonly used in political science. However, some scholars have argued that the Weibull is somewhat restrictive in that it assumes that the hazard increases or decreases monotonically. These scholars go on to claim that it is better to adopt the lognormal or log-logistic because this allows hazards to be nonmonotonic. The problem is that the Weibull, log-normal, or log-logistic are all two-parameter distributions – once the mean and variance are estimated, the shape is fixed.

As we will see in a moment, parametric models differ not only in terms of the assumption that they make about the shape of the hazard rate but also in terms of their specification and interpretation. More specifically, parametric models can be:

- proportional hazards (PH) models, or

- accelerated failure time (AFT) models.

PH and AFT classifications can be thought of as describing whole families of survival time distributions, where each member of a family shares a common set of features and properties (Jenkins 2008, 25). I'll describe the PH and AFT classifications in more detail in just a moment. First, though, let's look at how the different parametric models specify the hazard rate.

2 Specifying Different Parametric Models

Until now, we have been specifying the hazard rate simply as a function of time. However, we can now begin to think about how the hazard rate also changes with the characteristics of the individual units under consideration. Thus, the hazard rate is now:

$$h(t, X) = \frac{f(t)}{S(t)} = \lim_{\triangle t \rightarrow 0} \frac{Pr(t \leq T \leq t + \triangle t | T \geq t, X)}{\triangle t} \tag{4}$$

But how do the different parametric models specify the hazard rate?

2.1 Exponential Model

In the exponential model, the hazard rate is characterized as:

$$h(t, X) = \lambda \tag{5}$$

This implies that the conditional 'probability' of an event is constant over time (and that events occur according to a Poisson process). In other words, the risk of an event occurring is flat with respect to time.

Modelling the dependency of the hazard rate on covariates entails constructing a model that ensures a non-negative hazard rate (or non-negative expected duration time). One way to do this is simply to exponentiate the covariates such that:

$$h(t, X) = \lambda_i = e^{X_i \beta} \tag{6}$$

Given the way that the hazard rate is specified in the exponential model, the cumulative hazard can be written as:

$$H(t) = \lambda t \tag{7}$$

Recall from the earlier notes that $H(t) = -\ln[S(t)]$. As a result, we have:

$$\begin{aligned} S(t) &= e^{-H(t)} \\ &= e^{-\lambda t} \end{aligned} \tag{8}$$

This in turn means that the density is:

$$f(t) = h(t)S(t) = \lambda e^{-\lambda t} \tag{9}$$

This density means that the duration time T has an exponential distribution with mean $\frac{1}{\lambda} = E[t_i]$. In other words, the expected duration in an exponential model is:

$$E[t_i] = \frac{1}{\lambda} = \frac{1}{e^{X_i \beta}} = e^{-X_i \beta} \tag{10}$$

Having defined $h(t)$, $f(t)$, and $S(t)$, it is easy to construct the sample likelihood for the exponential model:

$$\mathcal{L} = \prod_{i=1}^{N} \left\{ \lambda e^{-\lambda t} \right\}^{d_i} \left\{ e^{-\lambda t} \right\}^{1-d_i} \tag{11}$$

2.2 Weibull

In the Weibull model, the hazard rate is characterized as:[1]

$$h(t, X) = \lambda p (\lambda t)^{p-1} \tag{12}$$

where

$$\lambda_i = e^{X_i \beta} \tag{13}$$

The Weibull model is more general and flexible than the exponential model and allows for hazard rates that are non-constant but monotonic. It is a two-parameter model (λ and p), where λ is the location parameter and p is the shape parameter because it determines whether the hazard is increasing, decreasing, or constant over time. The shape parameter works in the following way:

[1] You should be aware that different scholars often use different normalizations and notation for specifying the hazard rate in the Weibull model. For example, you may see the Weibull hazard written as $h(t, X) = pt^{p-1}\lambda$ or as $h(t, X) = \frac{1}{\sigma} t^{\frac{1}{\sigma}-1}\lambda$ where $\sigma \equiv \frac{1}{p}$. You will see why when we discuss the difference between PH and AFT models.

- If $\hat{\rho} < 1$, then the hazard is monotonically decreasing with time.
- If $\hat{\rho} > 1$, then the hazard is monotonically increasing with time.
- If $\hat{\rho} = 1$, then the hazard is flat and we have the exponential model i.e. the Weibull model nests the exponential model. This means that we can use the Weibull model to test to see if the exponential model is appropriate.

The survivor function for the Weibull is:

$$S(t) = e^{(-\lambda t)^p} \tag{14}$$

and the density function is:

$$f(t) = \lambda p(\lambda t)^{p-1} e^{(-\lambda t)^p} \tag{15}$$

Having defined $h(t)$, $f(t)$, and $S(t)$, it is easy to construct the likelihood for the Weibull model:

$$\mathcal{L} = \prod_{i=1}^{N} \left\{ \lambda p(\lambda t)^{p-1} e^{(-\lambda t)^p} \right\}^{d_i} \left\{ e^{(-\lambda t)^p} \right\}^{1-d_i} \tag{16}$$

The expected duration from a weibull model is

$$E(T) = \left(\frac{1}{\lambda}\right)^{\frac{1}{p}} \Gamma\left(1 + \frac{1}{p}\right) \tag{17}$$

where Γ denotes the gamma function.

2.3 Log-Logistic

In the log-logistic model, the hazard rate is characterized as:

$$h(t, X) = \frac{\lambda^{\frac{1}{\gamma}} t^{\left[\left(\frac{1}{\gamma}\right)-1\right]}}{\gamma\left[1 + (\lambda t)^{\left(\frac{1}{\gamma}\right)}\right]} \tag{18}$$

where

$$\lambda_i = e^{-(X_i \beta)} \tag{19}$$

As with the Weibull model, the log-logistic model has two parameters - λ is the location parameter and γ is the shape parameter. The log-logistic allows for non-monotonic unimodal hazards - in this case inverted U-shapes. The shape parameter works in the following way:

- If $\hat{\gamma} < 1$, then the conditional hazard first rises, then falls.
- If $\hat{\gamma} \geq 1$, then the hazard is declining.

Note that the hazard can never be monotonically rising in the log-logistic model.

The survivor function for the log-logistic is:

$$S(t) = \frac{1}{1 + (\lambda t)^{\left(\frac{1}{\gamma}\right)}} \tag{20}$$

and the density function is:

$$f(t) = \frac{\lambda^{\frac{1}{\gamma}} t^{\left[\left(\frac{1}{\gamma}\right)-1\right]}}{\left\{\gamma\left[1 + (\lambda t)^{\left(\frac{1}{\gamma}\right)}\right]\right\}^2} \tag{21}$$

and the integrated hazard function is:

$$H(t) = 1 + (\lambda t)^{\left(\frac{1}{\gamma}\right)} \tag{22}$$

Having defined $h(t)$, $f(t)$, and $S(t)$, it is easy to construct the likelihood for the log-logistic model:

$$\mathcal{L} = \prod_{i=1}^{N}\left\{\frac{\lambda^{\frac{1}{\gamma}} t^{\left[\left(\frac{1}{\gamma}\right)-1\right]}}{\left\{\gamma\left[1 + (\lambda t)^{\left(\frac{1}{\gamma}\right)}\right]\right\}^2} \times \frac{1}{1 + (\lambda t)^{\left(\frac{1}{\gamma}\right)}}\right\}^{d_i}\left\{\frac{1}{1 + (\lambda t)^{\left(\frac{1}{\gamma}\right)}}\right\}^{1-d_i} \tag{23}$$

The expected duration for the log-logistic only has a closed form solution when $\gamma < 1$. Using the STATA parameterization, this is:

$$E[t] = \frac{1}{\lambda} \times \frac{\gamma\pi}{\sin(\gamma\pi)} \tag{24}$$

2.4 Lognormal Model

In the lognormal model, the survivor function is:

$$S(t) = 1 - \Phi\left\{\frac{\ln(t) - \mu}{\sigma}\right\} \tag{25}$$

where Φ is the standard Normal cdf and $\mu = X\beta$. The density function is:

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}}\exp\left[\frac{-1}{2\sigma^2}\left\{\ln(t) - \mu\right\}^2\right] \tag{26}$$

and the hazard rate is:

$$h(t) = \frac{\frac{1}{t\sigma\sqrt{2\pi}}\exp\left[\frac{-1}{2\sigma^2}\left\{\ln(t) - \mu\right\}^2\right]}{1 - \Phi\left\{\frac{\ln(t)-\mu}{\sigma}\right\}} \tag{27}$$

The hazard rate is similar to that for the log-logistic for the case where $\gamma < 1$, i.e. it first rises and then falls.

2.5 Gompertz Model

In the Gompertz model, the hazard rate is:

$$h(t) = \lambda e^{\gamma t} \tag{28}$$

where $\lambda = e^{X\beta}$ and γ is a shape parameter. The survivor function is:

$$S(t) = e^{-\lambda\gamma^{-1}(e^{\gamma t}-1)} \tag{29}$$

The Gompertz model is useful for monotone hazard rates that either increase or decrease exponentially with time. The shape parameter works in the following way:

- If $\hat{\gamma} < 1$, then the hazard is monotonically decreasing with time.
- If $\hat{\gamma} > 1$, then the hazard is monotonically increasing with time.
- If $\hat{\gamma} = 1$, then the hazard is flat and we have the exponential model.

2.6 Generalized Gamma Model

The generalized gamma model has a quite complicated specification involving *two* shape parameters. The density of the generalized gamma distribution is:

$$f(t) = \frac{\lambda p(\lambda t)^{p\kappa-1}e^{-(\lambda t)^p}}{\Gamma(\kappa)} \tag{30}$$

where

$$\lambda_i = e^{-(X_i\beta)} \tag{31}$$

and p and κ are the two shape parameters.[2] The two shape parameters allow for quite a flexible hazard rate, including a U-shape. A nice characteristic of the generalized gamma model is that it nests several of the other parametric models as special cases: Weibull, exponential, log-normal, and the standard gamma. Thus, this model is good for adjudicating between (some) competing parametric models. The shape parameters work in the following way:

- If $\kappa = 1$, then the Weibull distribution is implied.
- If $\kappa = p = 1$, the exponential is implied.
- If $\kappa = 0$, the log-normal is implied.
- If $p = 1$, the gamma distribution is implied.

[2]Note that STATA refers to 'p' as 'σ' in its output.

2.7 Adjudicating Between Different Parametric Models

One of the limitations of parametric models is that you have to make an assumption about the shape of the hazard rate i.e. the shape of the time dependency. There is no way of knowing for sure whether you have chosen the correct parametric model. However, there are steps that you can make to limit the possibility of making an incorrect choice.

As we just saw, the generalized gamma model nests several other parametric models and can, therefore, be used to adjudicate between them. To estimate the generalized gamma model in STATA, you would type:

```
streg X, dist(gamma)
```

Once you have estimated the generalized gamma, you can then test to see if one of the nested models is appropriate. To see whether $\kappa = 0$, you can simply look at STATA's output because it reports whether κ is significantly different from zero. To see whether $p = 1$, you can again simply look at STATA's output because it reports whether ln$p$ (or in STATA's output ln σ) is significantly different from zero. To test whether $\kappa = 1$, you would type:

```
test [kappa]_b[_cons]=1
```

And to test whether $\kappa = p = 1$, you would type:

```
test [kappa]_b[_cons]=1, notest;
test [ln_sig]_b[_cons]=0, accum;
```

A limitation of the generalized gamma for adjudicating between parametric models is that it is only helpful for distinguishing between those parametric models that are nested within it. For example, you may find that the generalized gamma indicates that none of the parametric models that are nested within it are appropriate.

For non-nested parametric models, we might use Akaike's Information Criterion (AIC) to distinguish between different parametric models. Typically, we like model's whose log-likelihood is small. Akaike's method penalizes each model's log likelihood to reflect the number of parameters that are being estimated and then compares them:

$$\text{AIC} = -2\ln\mathcal{L} + 2(k + c) \tag{32}$$

where k is the number of model covariates and c is the number of model-specific distributional parameters. Essentially, you compare the AIC scores for different parametric models and then select the one with the smallest AIC score.

Table 1: Comparing Parametric Models

Distribution	Hazard Shape	c
Exponential	constant	1
Weibull	monotone	2
Gompertz	monotone	2
Lognormal	variable	2
Loglogistic	variable	2
Generalized gamma	variable	3

To get the AIC score in STATA, you should estimate a model and then type:

```
estimates stats
```

You can actually get the AIC for all models by typing:

```
foreach model in exponential weibull gompertz lognormal
                 loglogistic gamma {;
    quietly streg X, dist(‘model’);
    estimates store ‘model’;
    };
    estimates stats _all;
```

3 Proportional Hazard and Accelerated Failure Time

As I mentioned earlier, parametric models differ not only in terms of the assumption that they make about the shape of the hazard rate but also in terms of their interpretation.[3] Specifically, they differ in terms of whether they are proportional hazard (PH) or accelerated failure time (AFT) models.

3.1 The PH Specification

The hazard rate in all proportional hazard models can be written in the following way:

$$h(t, X) = h_0(t)e^{X_i\beta} = h_0(t)\lambda \tag{33}$$

where

- $h_0(t)$ is known as the baseline hazard and depends on t but not X. This part of the hazard rate indicates the pattern of time dependence that is assumed to be common to all units.
- $\lambda = e^{X_i\beta}$ is a unit-specific (non-negative) function of covariates (which does not depend on t) which scales the baseline hazard function common to all units up or down.

As you can see, you can write the hazard rate as a multiplicative function of the baseline hazard. In fact, this is the defining feature of all PH models. PH models are sometimes referred to as multiplicative hazard models for precisely this reason.

You will sometimes see the baseline hazard written slightly differently. For example, you might see the PH specification written as:

$$h(t, X) = h_0^*(t)\lambda^* \tag{34}$$

where $h_0^*(t) = h_0(t)e^{\beta_0}$ and $\lambda^* = e^{\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \beta}$.[4] In other words, you might see the coefficient on the constant term of your model placed in the baseline hazard as opposed to $e^{X_i\beta}$. The rationale behind this is that the intercept term is common to all units and should, therefore, not be included in the term summarizing unit heterogeneity in the hazard rate. The rationale for the original specification of the PH hazard rate is that the coefficient on the constant term is a regression parameter like any other element of β and should be included with them.

To estimate a PH parametric model in STATA and obtain coefficients, you type:

```
streg X, dist(model name) nohr
```

where MODEL NAME is the name of the parametric model you are using such as EXPONENTIAL, WEIBULL, GOMPERTZ, and NOHR means no hazard ratios, the meaning of which will become clearer in a moment.

[3] This section of the notes draws heavily on Jenkins (2008).

[4] We can do this because $e^{X_i\beta} = e^{\beta_0}e^{\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \beta}$.

3.2 Interpreting PH Models

As we just saw, the hazard rate in PH models is moved up or down as a function of the covariates associated with each unit. The PH property implies that absolute differences in X imply proportionate differences in the hazard rate at each t. For some $t = \bar{t}$, the ratio of hazard rates for two units i and j with vectors of characteristics X_i and X_j is:

$$\frac{h(\bar{t}, X_i)}{h(\bar{t}, X_j)} = e^{(X_i - X_j)\beta} \tag{35}$$

because the baseline hazards drop out. Eq. (35) indicates that the baseline hazard rate for unit i is $e^{(X_i - X_j)\beta}$ times different from that of unit j. Importantly, the right-hand side of Eq. (35) does not depend on time i.e. the proportional difference in the hazard rates of these two units is fixed across time. Put differently, the effect of the covariates in PH models are assumed to be fixed across time. We'll come back to this 'assumed' property and how we might test it a little later.

3.2.1 Factor Change in Baseline Hazard

We can use the PH property to help us substantively interpret our results. For example, we might be interested in how the hazard rate changes for a particular type of unit when we change one of the covariates. We can address this question by calculating the ratio of the hazard rates for this unit in a baseline scenario and a counterfactual scenario in which one of the covariates has been increased. Suppose we change the value of some covariate, X_k, by some amount δ. As Eq. (35) indicates, the ratio of the two hazards is:

$$\frac{h(\bar{t}, X_k + \delta)}{h(\bar{t}, X_k)} = e^{[X_k - (X_k + \delta)]\hat{\beta}_k} = e^{\hat{\beta}_k \delta} \tag{36}$$

If δ is just one unit, this simplifies to:

$$\frac{h(\bar{t}, X_k + 1)}{h(\bar{t}, X_k)} = e^{\hat{\beta}_k} \tag{37}$$

where $e^{\hat{\beta}_k}$ is known as the *hazard ratio*. How do you interpret this? Well, if the exponentiated coefficient, $e^{\hat{\beta}_k} = 1.10$, then we say that a one unit increase in X_k increases the baseline hazard by a factor of 1.1. In other words, the baseline hazard rate is 1.1 times larger, indicating that the expected duration will decrease. And if the exponentiated coefficient, $e^{\hat{\beta}_k} = 0.70$, then we say that a one unit increase in X_k increases the baseline hazard by a factor of 0.7. In other words, the baseline hazard rate is 0.7 times as large as in the baseline scenario, indicating a smaller hazard and, hence, an increase in the expected duration.

3.2.2 Percentage Change

We can also look at the percentage change in the baseline hazard associated with a change in the value of some covariate, X_k, by some amount δ:

$$\text{Percentage Change} = 100\left[e^{\hat{\beta}_k \delta} - 1\right] \tag{38}$$

If δ is just one unit, this simplifies to:

$$\text{Percentage Change} = 100\left[e^{\hat{\beta}_k} - 1\right] \tag{39}$$

Thus, if the exponentiated coefficient, $e^{\hat{\beta}_k} = 1.10$, then we say that a one unit increase in X_k increases the baseline hazard by 10%. And if the exponentiated coefficient, $e^{\hat{\beta}_k} = 0.70$, then we say that a one unit increase in X_k decreases the baseline hazard by 30%.

You can get STATA to report the exponentiated coefficients from a PH parametric model by simply typing:

```
streg X, dist(model name);
```

where MODEL NAME is the name of your parametric model. **NOTE** that STATA's default is to report exponentiated coefficients. Recall that if you want STATA to provide the actual coefficients, you have to add NOHR at the end of the command to indicate that you do not want hazard ratios i.e. that you do not want exponentiated coefficients. It is really important to remember exactly what STATA is reporting – exponentiated coefficients or just coefficients – because this obviously has a big effect on how you interpret the output.

3.3 The AFT Specification

The AFT model assumes a linear relationship between the log of (latent) survival time T and characteristics of the units, X:

$$\ln(T) = X\beta + z \tag{40}$$

where β is a vector of parameters and z is an error term. We can rewrite Eq. (43) as:

$$Y = \mu + \sigma u \tag{41}$$

or as:

$$\frac{Y - \mu}{\sigma} = u \tag{42}$$

where $Y \equiv \ln(T)$, $\mu \equiv X\beta$, and $u = \frac{z}{\sigma}$ is an error term with density $f(u)$ and σ is a scale factor which is related to the shape parameter of the hazard function as we will see. The AFT model is sometimes referred to as the log-linear model specification.[5]

As Table 2 indicates, distributional assumptions about u determine which sort of AFT model describes the distribution of the random variable T.

Table 2: Different Error Term Distributions and the Respective AFT Models

Distribution of u	Distribution of T
Extreme value (1 parameter)	Exponential
Extreme Value (2 parameter)	Weibull
Logistic	Log-Logistic
Normal	Lognormal
log Gamma (3 parameter Gamma)	Generalized Gamma

[5]Really they should be called semilog models because log-linear models typically involve having a logged dependent variable *and* logged independent variables.

- *Exponential*: In this model, the scale factor $\sigma = 1$ and $f(u) = e^{u-e^u}$.
- *Weibull*: In this model, σ is a free parameter and $\sigma = \frac{1}{p}$ to use the earlier notation. The density $f(u) = \frac{e^u}{1+e^u}$.
- *Log-logistic*: In this model, $f(u) = \frac{e^u}{1+e^u}$ and the free parameter $\sigma = \lambda$ to use the earlier notation.
- *Lognormal*: In this model, the error term is normally distributed and we have a specification of a (log)linear model similar to a standard linear model.

To estimate an AFT parametric model in STATA, you type:

```
streg X, dist(model name) time
```

where MODEL NAME is the name of the parametric model you are using such as EXPONENTIAL, WEIBULL, LNORMAL, LOGLOGISTIC, GAMMA, and TIME indicates that you want the AFT format in those cases (exponential and Weibull) where the models can be specified as either PH or AFT models.

3.4 Interpreting AFT Models

How can we interpret the coefficients from an AFT model? Recall that:

$$\ln(T) = X\beta + z \tag{43}$$

This is a standard semilog model that you probably came across in your introductory regression class. As a result, you can interpret it the same way as you would any other semilog model. For example, the marginal effect of X_k on ln$T$ is:

$$\frac{\partial \ln(T)}{\partial T} = \beta_k \tag{44}$$

Thus, a one unit increase in X_k leads to a β_k increase in the logged survival time. An alternative interpretation is that actual survival times increase at a rate of β_k or by $100 \times \beta_k$ per cent with a unit increase in X_k. None of these particular interpretations are that intuitive in the duration setting. There are more intuitive ways to interpret the results from AFT models, though.

3.4.1 Factor Change in Survival Times

From Eq. (43) we can see that:

$$T = e^{X\beta} e^{z} \tag{45}$$

Suppose we change the values of some covariate, X_k, by some amount δ. The ratio of the survival times is:

$$\frac{T(X_k + \delta)}{T(X_k)} = e^{[X_k - (X_k + \delta)]\hat{\beta}_k} = e^{\hat{\beta}_k \delta} \tag{46}$$

If δ is just one unit, this simplifies to:

$$\frac{T(X_k + 1)}{T(X_k)} = e^{\hat{\beta}_k} \tag{47}$$

where $e^{\hat{\beta}_k}$ is known as the *time ratio*. You can interpret this in a similar way to the hazard ratio in a PH model. For example, if the exponentiated coefficient, $e^{\hat{\beta}_k} = 1.10$, then we say that a one unit increase in X_k increases the survival time by a factor of 1.1. In other words, the survival time is 1.1 times longer. And if the exponentiated coefficient, $e^{\hat{\beta}_k} = 0.70$, then we say that a one unit increase in X_k increases the survival time by a factor of 0.7. In other words, the survival time is 0.7 times as large as in the baseline scenario.

3.4.2 Percentage Change

We can also look at the percentage change in the survival time associated with a change in the value of some covariate, X_k, by some amount δ:

$$\text{PERCENTAGE CHANGE} = 100\left[e^{\hat{\beta}_k\delta} - 1\right] \tag{48}$$

If δ is just one unit, this simplifies to:

$$\text{PERCENTAGE CHANGE} = 100\left[e^{\hat{\beta}_k} - 1\right] \tag{49}$$

Thus, if the exponentiated coefficient, $e^{\hat{\beta}_k} = 1.10$, then we say that a one unit increase in X_k increases the survival time by 10%. And if the exponentiated coefficient, $e^{\hat{\beta}_k} = 0.70$, then we say that a one unit increase in X_k decreases the survival time by 30%.

You can get STATA to report the exponentiated coefficients from an AFT model by tying:

```
streg X, dist(model name) time tr;
```

where MODEL NAME is the name of your AFT model, TIME indicates that you are using an AFT specification, and *tr* indicates that we want time ratios i.e. exponentiated coefficients.

3.5 Exponential and Weibull Models Revisited

It turns out that the exponential and Weibull models are the only parametric models that can be written with either a PH or AFT specification.

3.5.1 Exponential

PH Specification

According to the PH specification, which I used earlier without noting this, the exponential hazard rate is:

$$h(t, X) = \lambda_i = e^{X_i\beta} \tag{50}$$

and the expected duration is:

$$E[t_i] = \frac{1}{\lambda} = \frac{1}{e^{X_i\beta}} = e^{-X_i\beta} \tag{51}$$

AFT Specification

$$h(t, X) = \lambda_i = e^{-X_i\beta} \tag{52}$$

and the expected duration is:

$$E[t_i] = \frac{1}{\lambda} = e^{X_i\beta} \tag{53}$$

The key thing to note is that $\lambda_i = e^{X_i\beta}$ in the PH format and $\lambda_i = e^{-X_i\beta}$ in the AFT format. Notice the change in signs. This is important when calculating expected durations etc.. You'll find that the coefficients from the PH and AFT exponential models will have the same magnitude but have the opposite sign. The change in sign makes sense because the PH format uses covariates to model the hazard rate whereas the AFT format uses covariates to model the survival times. STATA uses the PH format by default and not the AFT format. To estimate the PH format, you type:

```
streg X, dist(exponential) nohr
```

and to estimate the AFT format, you type:

```
streg X, dist(exponential) time
```

3.5.2 Weibull

PH Specification

According to the PH specification, which I used earlier without noting this, the Weibull hazard rate is:

$$h(t, X) = \lambda p(\lambda t)^{p-1} \tag{54}$$

where $\lambda_i = e^{X_i\beta}$ and p is a shape parameter. The expected duration is:

$$E(T) = \left(\frac{1}{\lambda}\right)^{\frac{1}{p}} \Gamma\left(1 + \frac{1}{p}\right) \tag{55}$$

where Γ denotes the gamma function.

AFT Specification

According to the AFT specification, we have:

$$\ln(T) = X\beta + \sigma u \tag{56}$$

where u follows a Type-1 extreme value distribution and σ is a shape parameter that is equivalent to $\frac{1}{p}$ in the language of the PH framework shown above. Put differently, $p = \frac{1}{\sigma}$. It is possible to show that the coefficients from the PH and AFT formats are linked in the following manner i.e. $\hat{\beta}_{PH} = \frac{-\hat{\beta}_{AFT}}{\hat{\sigma}}$ or $\hat{\beta}_{AFT} = \frac{-\hat{\beta}_{PH}}{\hat{p}}$. Thus, the estimated coefficients are equivalent up to a scale factor equal to $\hat{\sigma}$.

It is also important to recognize that in the AFT format that:

$$\lambda_i = e^{-pX_i\beta} \tag{57}$$

and that the expected duration is:

$$E(T) = \frac{\Gamma\left(1 + \frac{1}{p}\right)}{\lambda} \tag{58}$$

It is easy to see that Eq. (58) is equivalent to Eq. (55) if $\lambda_i = e^{-pX_i\beta}$. For example, we can rewrite Eq. (58) as:

$$E(T) = [e^{-pX_i\beta}]^{-1}\Gamma\left(1 + \frac{1}{p}\right) \tag{59}$$

This in turn can be rewritten as:

$$E(T) = \left(\frac{1}{e^{X_i\beta}}\right)^{\frac{1}{p}} \Gamma\left(1+\frac{1}{p}\right) = \left(\frac{1}{\lambda}\right)^{\frac{1}{p}} \Gamma\left(1+\frac{1}{p}\right)$$

This is exactly the same as the expected duration in the PH format - they are equivalent.

You'll find that the coefficients from the PH and AFT Weibull models will again have the opposite signs and their magnitudes will be equivalent up to a scale factor of $\hat{\sigma}$ as just mentioned. STATA uses the PH format by default and not the AFT format. To estimate the PH format, you type:

```
streg X, dist(weibull) nohr
```

and to estimate the AFT format, you type:

```
streg X, dist(weibull) time
```

3.6 Summary of PH and AFT Specifications

In Table 3, I classify parametric models according to whether they can be interpreted in a PH or AFT format.

Table 3: PH and AFT Parametric Models

Distribution	PH	AFT
Exponential	Yes	Yes
Weibull	Yes	Yes
Gompertz	Yes	No
Lognormal	No	Yes
Loglogistic	No	Yes
Generalized gamma	No	Yes

4 Substantive Interpretation

There are a number of things you can do to interpret the results from parametric models:

1. Interpret the Sign and Statistical Significance of Coefficients

   In PH models, the sign of the coefficient indicates how a covariate affects the hazard rate. Thus, a positive coefficient increases the hazard rate and, therefore, reduces the expected duration. A negative coefficient decreases the hazard rate and, therefore, increases the expected duration. The statistical significance of the coefficient indicates whether these changes in the expected duration will be statistically significant or not.

   In AFT models, the sign of the coefficient indicates how a covariate affects the logged survival times. Thus, a positive coefficient increases the logged survival time and, hence, the expected duration. A negative coefficient decreases the logged survival time and, hence, the expected duration. The statistical significance of the coefficient indicates whether these changes in the expected duration will be statistically significant or not.

2. Calculate Factor and Percentage Changes in Hazard or Time Ratios

   In PH models, you can exponentiate the coefficients to obtain hazard ratios. You can then use these hazard ratios to calculate the factor change or percentage change in the baseline hazard associated with a one unit increase in a covariate.

   In AFT models, you can exponentiate the coefficients to obtain time ratios. You can then use these time ratios to calculate the factor change or percentage change in the expected survival time associated with a one unit increase in a covariate.

3. Calculate Expected Duration and Changes in Expected Duration

   In PH and AFT models, you can calculate the expected duration associated with different scenarios. You can then calculate the change in expected duration as you move from some baseline scenario to some counterfactual scenario. Using simulation methods to calculate these quantities of interest will allow you to easily obtain measures of uncertainty.

   You can get STATA to calculate the expected duration for the particular observations in your data set by typing:

   ```
   predict meantime, time mean
   ```

   after you have estimated your model. However, STATA will not provide a measure of uncertainty to go with the point estimate.

4. Plot Hazard and Survivor Functions

   After you have estimated your model, you can also get STATA to plot the hazard and survivor functions for different scenarios by typing:

   ```
   stcurve, survival
               at1(x1=0 x2=108.18 etc.)
               at2(x1=1 x2=108.18 etc.);
   ```

   To get the hazard function, simply switch SURVIVAL with HAZARD.

References

Jenkins, Stephen P. 2008. "Survival Analysis." Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester.