

# Parametric Survival Models

Christoph Dätwyler and Timon Stucki

9. May 2011

# Contents

## Introduction

## Parametric Model

- Distributional Assumption

- Weibull Model

- Accelerated Failure Time Assumption

- A More General Form of the AFT Model

- Weibull AFT Model

- Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Contents

## Introduction

## Parametric Model

Distributional Assumption

Weibull Model

Accelerated Failure Time Assumption

A More General Form of the AFT Model

Weibull AFT Model

Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Parametric Survival Model

## Basic Idea

The survival time follows a distribution.

## Goal

Use data to estimate parameters of this distribution

⇒ completely specified model

⇒ prediction of time-quantiles

# Parametric Survival Model vs. Cox PH Model

## Parametric Survival Model

- + Completely specified  $h(t)$  and  $S(t)$
- + More consistent with theoretical  $S(t)$
- + time-quantile prediction possible
- Assumption on underlying distribution

## Cox PH Model

- distribution of survival time unknown
- Less consistent with theoretical  $S(t)$  (typically step function)
- + Does not rely on distributional assumptions
- + Baseline hazard not necessary for estimation of hazard ratio

# Contents

## Introduction

## Parametric Model

Distributional Assumption

Weibull Model

Accelerated Failure Time Assumption

A More General Form of the AFT Model

Weibull AFT Model

Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Contents

Introduction

Parametric Model

**Distributional Assumption**

Weibull Model

Accelerated Failure Time Assumption

A More General Form of the AFT Model

Weibull AFT Model

Log-Logistic Model

Other Parametric Models

The Parametric Likelihood

Frailty Models

Summary

# Probability Density, Hazard and Survival Function

## Main Assumption

The survival time  $T$  is assumed to follow a distribution with density function  $f(t)$ .

$$\Rightarrow S(t) = P(T > t) = \int_t^{\infty} f(u) du$$

Recall:

$$h(t) = -\frac{\frac{d}{dt}S(t)}{S(t)}, \quad S(t) = \exp\left(-\int_0^t h(u) du\right)$$



# Density Function in Relation to the Hazard and Survival Function

## Remark

$$f(t) = h(t)S(t)$$

Proof:

$$h(t)S(t) = -\frac{\frac{d}{dt}S(t)}{S(t)}S(t) = \frac{d}{dt} \int_{\infty}^t f(u)du = f(t)$$

## Key Point

Specifying one of the three functions  $f(t)$ ,  $S(t)$  or  $h(t)$  specifies the other two functions.

## Commonly Used Distributions and Parameters

Distribution	$f(t)$	$S(t)$	$h(t)$
Exponential	$\lambda \exp(-\lambda t)$	$\exp(-\lambda t)$	$\lambda$
Weibull	$\lambda p t^{p-1} \exp(-\lambda t^p)$	$\exp(-\lambda t^p)$	$\lambda p t^{p-1}$
Log-logistic	$\frac{\lambda p t^{p-1}}{(1+\lambda t^p)^2}$	$\frac{1}{1+\lambda t^p}$	$\frac{\lambda p t^{p-1}}{1+\lambda t^p}$

Modeling of the parameters:

- ▶  $\lambda$  is reparameterized in terms of predictor variables and regression parameters.
- ▶ Typically for parametric models, the shape parameters  $p$  is held fixed.

# Contents

## Introduction

## Parametric Model

Distributional Assumption

### **Weibull Model**

Accelerated Failure Time Assumption

A More General Form of the AFT Model

Weibull AFT Model

Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Weibull Model

Assuming  $T \sim \text{Weibull}(\lambda, p)$  with probability density function

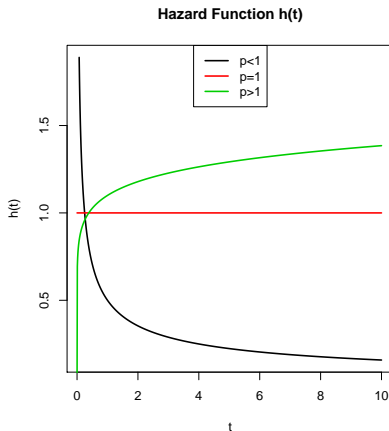
$$f(t) = \lambda p t^{p-1} \exp(-\lambda t^p), \text{ where } p > 0 \text{ and } \lambda > 0,$$

the hazard function is given by

$$h(t) = \lambda p t^{p-1}.$$

$p$  is called **shape parameter**:

- ▶ If  $p > 1$  the hazard increases
- ▶ If  $p = 1$  the hazard is constant (exponential model)
- ▶ If  $p < 1$  the hazard decreases



# Graphical Evaluation of Weibull Assumption

## Property of Weibull Model

The  $\log(-\log(S(t)))$  is linear with the log of time.

$$\begin{aligned} S(t) &= \exp(-\lambda t^p) \\ \Rightarrow -\log(S(t)) &= \lambda t^p \\ \Rightarrow \log(-\log(S(t))) &= \log(\lambda) + p \log(t) \end{aligned}$$

This property allows a graphical evaluation of the appropriateness of a Weibull model by plotting

$$\log(-\log(\hat{S}(t))) \text{ vs. } \log(t),$$

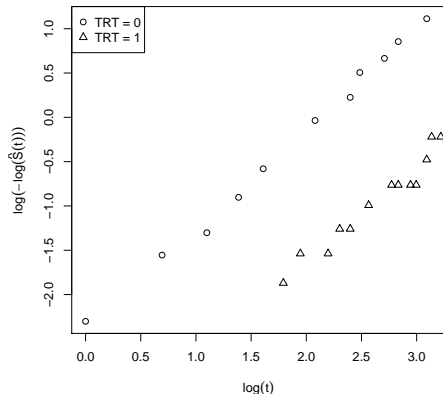
where  $\hat{S}(t)$  is Kaplan-Meier survival estimate.

## Example: Remission Data

We consider the remission data of 42 leukemia patients.

- ▶ 21 patients given treatment (TRT = 1)
- ▶ 21 patients given placebo (TRT = 0)

Note: The survival time (time to event) is the time until a patient went out of remission, which means that the patient relapsed.



- ▶ straight lines  $\Rightarrow$  Weibull
- ▶ same slope  $\Rightarrow$  PH

## Weibull PH Model

Recall:  $h(t) = \lambda pt^{p-1}$

Weibull PH model:

- ▶ Reparameterize  $\lambda$  with

$$\lambda = \exp(\beta_0 + \beta_1 TRT).$$

- ▶ Then the hazard ratio ( $TRT = 1$  vs.  $TRT = 0$ ) is

$$HR = \frac{\exp(\beta_0 + \beta_1)pt^{p-1}}{\exp(\beta_0)pt^{p-1}} = \exp(\beta_1),$$

which indicates that the PH assumption is satisfied.

Note: This result depends on  $p$  having the same value for  $TRT = 1$  and  $TRT = 0$  (otherwise time would not cancel out).

# Exponential PH Model

The exponential distribution is a special case of the Weibull distribution with  $p = 1$ .

Weibull density function:

$$f(t) = \underbrace{\exp(-\lambda t^p)}_{S(t)} \underbrace{\lambda p t^{p-1}}_{h(t)}$$

Setting  $p = 1$  gives the density function of an exponential distribution

$$f(t) = \underbrace{\exp(-\lambda t)}_{S(t)} \underbrace{\lambda}_{h(t)}.$$



## Exponential PH Model

Running the exponential model leads to the following output:

_t	Coef.	Std. Err.	z	p >  z
trt	-1.527	.398	-3.83	0.00
_cons	-2.159	.218	-9.90	0.00

The estimated hazard ratio is obtained from the estimated coefficient  $\hat{\beta}_1$  by

$$\widehat{HR}(TRT = 1 \text{ vs. } TRT = 0) = \exp(\hat{\beta}_1) = \exp(-1.527) = 0.22.$$

### Interpretation

This means that the hazard for the group with  $TRT = 0$  is bigger than the one for the group with  $TRT = 1$  because  $0.22 < 1$ , indicating a positive effect of the treatment.

# Contents

## Introduction

## Parametric Model

Distributional Assumption

Weibull Model

**Accelerated Failure Time Assumption**

A More General Form of the AFT Model

Weibull AFT Model

Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Accelerated Failure Time Assumption

First example: Humans vs. dogs

- ▶ Let  $S_H(t)$  and  $S_D(t)$  denote the survival functions for humans and dogs respectively.
- ▶ Known: In general dogs grow older seven times faster than humans.
- ▶ In AFT terminology: The probability of a dog surviving past  $t$  years is equal to the one of a human surviving past  $7t$  years.
- ▶ This means:

$$S_D(t) = S_H(7t)$$

- ▶ In other words, dogs accelerate through life about seven times faster than humans.

# AFT Models

## AFT Models

AFT Models describe stretching out or contraction of survival time as a function of predictor variables.

# Illustration of AFT Assumption

Second example: Smokers vs. nonsmokers

- ▶ Let  $S_S(t)$  and  $S_{NS}(t)$  denote the survival functions for smokers and nonsmokers respectively.

AFT assumption

$$S_{NS}(t) = S_S(\gamma t) \text{ for } t \geq 0$$

Definition

$\gamma > 0$  is the so called **acceleration factor** and is a constant.

## Expressing the AFT assumption

The AFT assumption can be expressed

- ▶ in terms of survival function as seen before:

$$S_{NS}(t) = S_S(\gamma t)$$

- ▶ in terms of random variables for survival time:

$$\gamma T_{NS} = T_S,$$

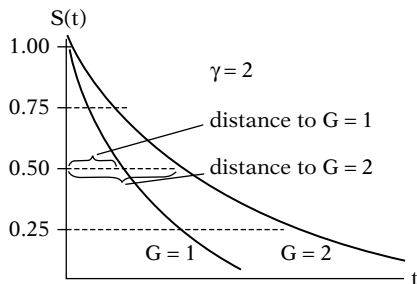
where  $T_{NS}$  is a random variable following some distribution representing the survival time for nonsmokers and  $T_S$  the analogous one for smokers.

## Acceleration factor

The acceleration factor allows to evaluate the effect of predictor variables on the survival time.

### Acceleration factor

The acceleration factor is a ratio of time-quantiles corresponding to any fixed value of  $S(t)$ .



## Interpretation of the Acceleration Factor

Assuming the event to occur is negative for an individual, comparing two groups (levels of covariates) leads to the following general interpretation:

$$\gamma > 1 \quad \Rightarrow \quad \text{exposure benefits survival}$$

$$\gamma < 1 \quad \Rightarrow \quad \text{exposure harmful to survival}$$

For the hazard ratio, we have:

$$HR > 1 \quad \Rightarrow \quad \text{exposure harmful to survival}$$

$$HR < 1 \quad \Rightarrow \quad \text{exposure benefits survival}$$

$$\gamma = HR = 1 \quad \Rightarrow \quad \text{no effect from exposure}$$



# Contents

Introduction

## Parametric Model

Distributional Assumption

Weibull Model

Accelerated Failure Time Assumption

**A More General Form of the AFT Model**

Weibull AFT Model

Log-Logistic Model

Other Parametric Models

The Parametric Likelihood

Frailty Models

Summary

## General Form of AFT Model

Consider an AFT model with one predictor  $X$ . The model can be expressed on the log scale as

$$\log(T) = \alpha_0 + \alpha_1 X + \epsilon,$$

where  $\epsilon$  is a random error following some distribution.

$T$	$\log(T)$
Exponential	Extreme value
Weibull	Extreme value
Log-logistic	Logistic
Lognormal	Normal

## General Form of AFT Model

Some distributions (e.g. Weibull) have an additional parameter  $\sigma$ , which scales  $\epsilon$ .

$$\log(T) = \alpha_0 + \alpha_1 X + \sigma \epsilon$$

Here (and in R) the model is parametrized using  $\sigma = \frac{1}{p}$ :

$$\log(T) = \alpha_0 + \alpha_1 X + \frac{1}{p} \epsilon$$

## General Form of AFT Model

The model in terms of the survival time  $T$  is

$$T = \exp\left(\alpha_0 + \alpha_1 X + \frac{1}{\rho}\epsilon\right) = \exp(\alpha_0) \cdot \exp(\alpha_1 X) \cdot \exp\left(\frac{1}{\rho}\epsilon\right)$$

### Remark

AFT model is multiplicative in terms of  $T$  and additive in terms of  $\log(T)$ .

# Contents

## Introduction

## Parametric Model

Distributional Assumption

Weibull Model

Accelerated Failure Time Assumption

A More General Form of the AFT Model

**Weibull AFT Model**

Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# AFT Model

- ▶ We use again the remission data.
- ▶ We consider the variable  $TRT$  as only predictor ( $TRT = 1$  and  $TRT = 0$ )

## AFT Model Assumption

The ratio of time-quantile is constant ( $\gamma$ ) for all fixed values  $S(t) = q$ .

## Expression for time-quantiles

- ▶ Solve for  $t$  in terms of  $S(t)$
- ▶ Scale  $t$  in terms of predictors

## Weibull AFT Model

Recall:  $S(t) = \exp(-\lambda t^p)$

1. Solving for  $t$  gives:

$$\begin{aligned} S(t) = \exp(-\lambda t^p) &\Leftrightarrow -\log(S(t)) = \lambda t^p \\ &\Leftrightarrow t = (-\log(S(t)))^{1/p} \frac{1}{\lambda^{1/p}} \end{aligned}$$

2. Reparameterizing  $\frac{1}{\lambda^{1/p}} = \exp(\alpha_0 + \alpha_1 TRT)$  yields

$$t = (-\log(S(t)))^{1/p} \exp(\alpha_0 + \alpha_1 TRT).$$

(TRT used to scale time to any fixed value of  $S(t)$ )

3. In terms of any fixed probability  $S(t) = q$  we get

$$t = (-\log(q))^{1/p} \exp(\alpha_0 + \alpha_1 TRT).$$

## Weibull AFT Model: Acceleration Factor

The acceleration factor for a fixed value of  $S(t) = q$  is calculated as follows:

Level of covariates:  $TRT = 1$  and  $TRT = 0$

Then the acceleration factor  $\gamma$  is

$$\begin{aligned}\gamma &= \gamma(TRT = 1 \text{ vs. } TRT = 0) \\ &= \frac{(-\log(q))^{1/p} \exp(\alpha_0 + \alpha_1)}{(-\log(q))^{1/p} \exp(\alpha_0)} \\ &= \exp(\alpha_1).\end{aligned}$$

Note: As in the PH form of the model, this result depends on  $p$  having the same value for  $TRT = 1$  and  $TRT = 0$ .



## R Code and R Output

```
> weibull.aft <- survreg(Surv(Survt,status) ~ TRT,dist='weibull')  
> summary(weibull.aft)
```

Call:

```
survreg(formula = Surv(Survt, status) ~ TRT, dist = "weibull")
```

	Value	Std. Error	z	p
(Intercept)	2.248	0.166	13.55	8.30e-42
TRT	1.267	0.311	4.08	4.51e-05
Log(scale)	-0.312	0.147	-2.12	3.43e-02

Scale= 0.732

Weibull distribution

Loglik(model)= -106.6    Loglik(intercept only)= -116.4

Chisq= 19.65 on 1 degrees of freedom, p= 9.3e-06

Number of Newton-Raphson Iterations: 5

n= 42

## R Code and R Output: Acceleration Factor

The estimated acceleration factor  $\hat{\gamma}$  comparing the treatment group to the placebo group ( $TRT = 1$  vs.  $TRT = 0$ ) is now:

$$\hat{\gamma} = \exp(\hat{\alpha}_1)$$

```
> exp(weibull.aft$coefficient[2])  
      TRT  
3.551374
```

### Interpretation

The survival time for the treatment group ( $TRT = 1$ ) is increased by a factor of 3.55 compared to the placebo group ( $TRT = 0$ )  
 $\Rightarrow$  Treatment is positive.

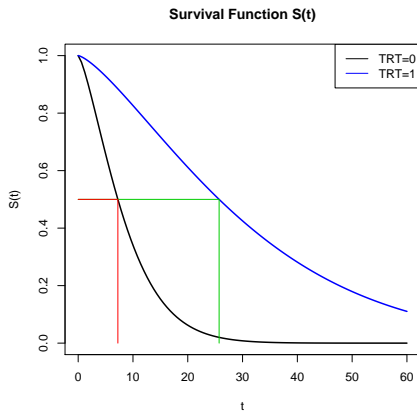
# Survival functions

Computing time-quantiles, for example the median:

$$S(t) = 0.5 \Rightarrow \hat{t}_m = (-\log(0.5))^{1/\hat{p}} \cdot \exp(\hat{\alpha}_0 + \hat{\alpha}_1 TRT)$$

Estimated survival times for the median  $S(t) = 0.5$ :

```
> median <- predict(weibull.aft,  
+                   newdata=list(TRT=c(0,1)),  
+                   type='quantile',p=0.5)  
> median  
      1      2  
7.242697 25.721526  
> median[2]/median[1]  
      2  
3.551374
```



## Relation between Weibull AFT and PH coefficients

$$\begin{aligned}\text{► AFT: } \quad \frac{1}{\lambda^{1/p}} &= \exp(\alpha_0 + \alpha_1 TRT) \\ \Leftrightarrow (1/p) \log(\lambda) &= -(\alpha_0 + \alpha_1 TRT) \\ \Leftrightarrow \log(\lambda) &= -p(\alpha_0 + \alpha_1 TRT)\end{aligned}$$

$$\begin{aligned}\text{► PH: } \quad \lambda &= \exp(\beta_0 + \beta_1 TRT) \\ \Leftrightarrow \log(\lambda) &= \beta_0 + \beta_1 TRT\end{aligned}$$

This indicates the following relationship between the coefficients:

$$\beta_j = -\alpha_j p$$

## Exponential PH and AFT Model

We obtained  $\beta_j = -\alpha_j p$  for the Weibull model. In the special case of the exponential model where  $p = 1$  we have

$$\beta_j = -\alpha_j.$$

### Remark

The exponential PH and AFT are in fact **the same model**, except that the parametrization is different.

## Exponential PH and AFT Model

Example:

The estimated values from the exponential example above support this result.

$$\text{Coefficient PH Model:} \quad \hat{\beta}_1 = -1.527$$

$$\text{Coefficient AFT Model:} \quad \hat{\alpha}_1 = 1.527$$

We also have

$$\widehat{HR}(TRT = 1 \text{ vs. } TRT = 0) = \exp(\hat{\beta}_1) = \exp(-\hat{\alpha}_1) = \frac{1}{\hat{\gamma}}.$$

# Property of the Weibull Model

## Proposition

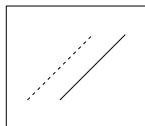
AFT assumption holds  $\Leftrightarrow$  PH assumption holds (given that  $p$  is fixed)

Proof for the considered example ( $TRT = 1$  and  $TRT = 0$ ):

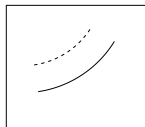
- ▶  $[\Rightarrow]$ :  $\gamma = \exp(\alpha_1)$   
Assume  $\gamma$  is constant  $\Rightarrow \alpha_1$  is constant  
 $HR = \exp(\beta_1) = \exp(-p\alpha_1) \Rightarrow HR$  is constant
- ▶  $[\Leftarrow]$ :  $HR = \exp(\beta_1)$   
Assume  $HR$  is constant  $\Rightarrow \beta_1$  is constant  
 $\gamma = \exp(\alpha_1) = \exp(-\frac{\beta_1}{p}) \Rightarrow \gamma$  is constant

## Possible Plots

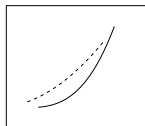
Possible results for plots of  $\log(-\log(\hat{S}(t)))$  against  $\log(t)$ :



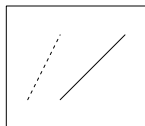
$\Rightarrow$  Weibull (or Exponential if  $p = 1$ ), PH and AFT assumption hold.



$\Rightarrow$  Not Weibull, PH and not AFT.



$\Rightarrow$  Not Weibull, not PH and not AFT.



$\Rightarrow$  Weibull, not PH and not AFT ( $p$  not fixed).



# Contents

## Introduction

## Parametric Model

Distributional Assumption

Weibull Model

Accelerated Failure Time Assumption

A More General Form of the AFT Model

Weibull AFT Model

**Log-Logistic Model**

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Hazard Function of Log-Logistic Model

The log-logistic distribution accommodates an AFT model but not a PH model.

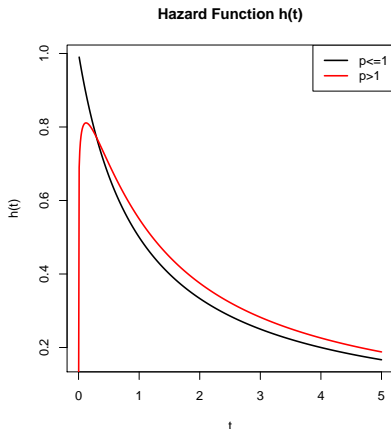
Hazard function is

$$h(t) = \frac{\lambda p t^{p-1}}{1 + \lambda t^p},$$

with  $p > 0$  and  $\lambda > 0$ .

Shape of hazard function:

- ▶  $p \leq 1$ : hazard decreases
- ▶  $p > 1$ : hazard **unimodal**



# PO Assumption

## Definition

In a **proportional odds (PO)** survival model, the odds ratio is constant over time.

- ▶ **Survival odds:** odds of surviving beyond time  $t$

$$\frac{S(t)}{1 - S(t)} = \frac{P(T > t)}{P(T \leq t)}$$

- ▶ **Failure odds:** odds of getting the event by time  $t$

$$\frac{1 - S(t)}{S(t)} = \frac{P(T \leq t)}{P(T > t)}$$

## PO Assumption

The failure odds of the log-logistic survival model are

$$\frac{1 - S(t)}{S(t)} = \frac{\frac{\lambda t^p}{1 + \lambda t^p}}{\frac{1}{1 + \lambda t^p}} = \lambda t^p.$$

The failure odds ratio (OR) for two different groups (1 and 2) is (for  $p$  fixed)

$$OR(1 \text{ vs. } 2) = \frac{\frac{1 - S_1(t)}{S_1(t)}}{\frac{1 - S_2(t)}{S_2(t)}} = \frac{\lambda_1 t^p}{\lambda_2 t^p} = \frac{\lambda_1}{\lambda_2}.$$

Hence, the log-logistic model is a proportional odds (PO) model.

## Graphical Evaluation of Log-Logistic Assumption

The log-failure odds can be written as

$$\log \left( \frac{1 - S(t)}{S(t)} \right) = \log(\lambda t^p) = \log(\lambda) + p \log(t),$$

which is a linear function of  $\log(t)$ .

## Graphical Evaluation of Log-Logistic Assumption

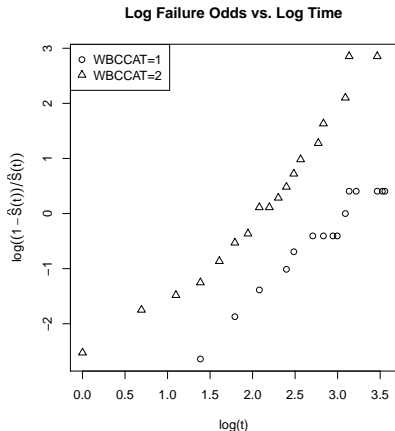
- ▶ Plot  $\log \left( \frac{1 - \hat{S}(t)}{\hat{S}(t)} \right)$  against  $\log(t)$  ( $\hat{S}$  are the KM-survival estimates).
- ▶ If the plot is linear with slope  $p$ , then the survival time follows a log-logistic distribution.

## Log-Logistic Example with the Remission Data

We consider a new categorical variable  $WBCCAT$ :

- ▶  $WBCCAT = 2$  if  $\log WBC \geq 2.5$  (high count)
- ▶  $WBCCAT = 1$  if  $\log WBC < 2.5$  (medium count)

The graphical evaluation of  $WBCCAT = 2$  and  $WBCCAT = 1$ :



$\Rightarrow$  straight lines indicate  
log-logistic distribution

## AFT Log-Logistic Model

AFT log-logistic model with *WBCCAT* as only predictor:

We solve

$$S(t) = \frac{1}{1 + \lambda t^p} = \frac{1}{1 + (\lambda^{\frac{1}{p}} t)^p}$$

for  $t$  and obtain

$$t = \left( \frac{1}{S(t)} - 1 \right)^{\frac{1}{p}} \frac{1}{\lambda^{\frac{1}{p}}}.$$

We reparameterize the factor on the right as

$$\frac{1}{\lambda^{\frac{1}{p}}} = \exp(\alpha_0 + \alpha_1 \text{WBCCAT}).$$

## AFT Log-Logistic Model: Acceleration Factor

We get

$$t = \left( \frac{1}{S(t)} - 1 \right)^{\frac{1}{p}} \exp(\alpha_0 + \alpha_1 WBCCAT).$$

For a fixed probability  $S(t) = q$ , the expression for  $t$  is

$$t = (q^{-1} - 1)^{\frac{1}{p}} \exp(\alpha_0 + \alpha_1 WBCCAT).$$

The acceleration factor  $\gamma$  for  $S(t) = q$  is

$$\begin{aligned} \gamma(WBCCAT = 2 \text{ vs. } WBCCAT = 1) &= \frac{(q^{-1} - 1)^{\frac{1}{p}} \exp(\alpha_0 + 2\alpha_1)}{(q^{-1} - 1)^{\frac{1}{p}} \exp(\alpha_0 + 1\alpha_1)} \\ &= \exp(\alpha_1). \end{aligned}$$



# R Code and R Output

```
> logistic.aft <- survreg(Surv(Survt, status) ~ WBCCAT,  
+ dist='loglogistic',data=remdata)  
> summary(logistic.aft)
```

Call:

```
survreg(formula = Surv(Survt, status) ~ WBCCAT, data = remdata,  
        dist = "loglogistic")
```

	Value	Std. Error	z	p
(Intercept)	4.094	0.586	6.98	2.92e-12
WBCCAT	-0.987	0.337	-2.93	3.40e-03
Log(scale)	-0.564	0.154	-3.67	2.41e-04

Scale= 0.569

Log logistic distribution

Loglik(model)= -111.2    Loglik(intercept only)= -115.4

Chisq= 8.28 on 1 degrees of freedom, p= 0.004

Number of Newton-Raphson Iterations: 4

n= 42

## R Code and R Output: Acceleration Factor

The estimated acceleration factor  $\hat{\gamma}$  comparing  $WBCCAT = 2$  (high count) and  $WBCCAT = 1$  (medium count) is now:

$$\hat{\gamma} = \exp(\hat{\alpha}_1)$$

```
> exp(logistic.aft$coefficient[2])  
WBCCAT  
0.3728214
```

$$\Rightarrow \hat{S}_1(t) = \hat{S}_2(0.37t)$$

( $\hat{S}_i$  is the survival function for  $WBCCAT = i$ ,  $i = 1, 2$ )

### Interpretation

The survival time for the group with high count ( $WBCCAT = 2$ ) is "accelerated" by a factor of 0.37 compared to the group with medium count ( $WBCCAT = 1$ )  $\Rightarrow$  High WBC is negative.

## PO Log-Logistic Model

The proportional odds (PO) form of the log-logistic model can be formulated by reparameterizing  $\lambda$ .

Failure odds:

$$\frac{1 - S(t)}{S(t)} = \frac{\frac{\lambda t^p}{1 + \lambda t^p}}{\frac{1}{1 + \lambda t^p}} = \lambda t^p.$$

Reparameterizing  $\lambda$  gives

$$\lambda = \exp(\beta_0 + \beta_1 WBCCAT).$$

Hence, the failure odds ratio is

$$OR(WBCCAT=2 \text{ vs. } WBCCAT=1) = \frac{t^p \exp(\beta_0 + 2\beta_1)}{t^p \exp(\beta_0 + 1\beta_1)} = \exp(\beta_1).$$

# Comparing AFT and PO Log-Logistic Model

Parameterizations:

- ▶ AFT model:  $\frac{1}{\lambda^{\frac{1}{p}}} = \exp(\alpha_0 + \alpha_1 WBCCAT)$
- ▶ PO model:  $\lambda = \exp(\beta_0 + \beta_1 WBCCAT)$

Hence, we have the relationship

$$\beta_0 = -\alpha_0 p \quad \text{and} \quad \beta_1 = -\alpha_1 p.$$

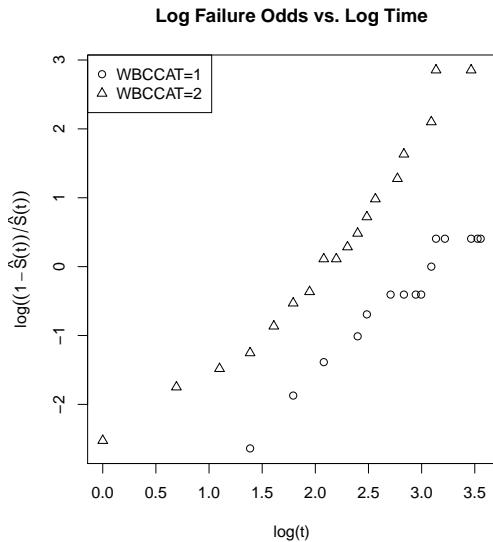
Note: If  $p$  is fixed this leads to: AFT  $\Leftrightarrow$  PO

So we can calculate the estimated OR with the coefficients of the AFT model by:

$$\begin{aligned}\widehat{OR} &= \exp(\hat{\beta}_1) \\ &= \exp(-\hat{\alpha}_1 \hat{p}) = 5.66\end{aligned}$$

```
> alpha1 <- logistic.aft$coefficient[2]
> p <- 1/logistic.aft$scale
> exp(-alpha1*p)
5.662691
```

# Graphical Evaluation



# Graphical Evaluation

## Proposition

1. Straight lines  $\Rightarrow$  Log-logistic
2. Parallel plots and log-logistic  $\Rightarrow$  PO
3. Log-logistic and PO  $\Rightarrow$  AFT

Proof: Consider two groups (1 and 2).

1.  $\log(\text{failure odds}) = \log(\lambda) + p \log(t)$
2. Parallel plots  $\Rightarrow p$  the same for both groups  
 $\Rightarrow OR = \frac{t^p \lambda_1}{t^p \lambda_2} = \frac{\lambda_1}{\lambda_2}$
3. For  $S(t) = q$ , the acceleration factor is

$$\gamma = \frac{(q^{-1} - 1)^{\frac{1}{p}} \lambda_1}{(q^{-1} - 1)^{\frac{1}{p}} \lambda_2} = \frac{\lambda_1}{\lambda_2}.$$

# Contents

## Introduction

## Parametric Model

- Distributional Assumption

- Weibull Model

- Accelerated Failure Time Assumption

- A More General Form of the AFT Model

- Weibull AFT Model

- Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Generalized Gamma Model

The generalized gamma distribution is given by

$$f(t) = \frac{\frac{p}{a^d} t^{d-1} \exp(-(\frac{t}{a})^p)}{\Gamma(d/p)},$$

where

$$\Gamma(z) = \int_0^\infty s^{z-1} e^{-s} ds$$

and  $a > 0$ ,  $d > 0$ ,  $p > 0$ .

- ▶ The three parameters allow great flexibility in the distributions shape.
- ▶ Weibull and lognormal distributions are special cases of the generalized gamma distribution (e.g. setting  $d = p$  gives us the Weibull distribution).



# Lognormal Model

The lognormal distribution is given by

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log(t) - \mu)^2}{2\sigma^2}\right),$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation respectively, of the variable's natural logarithm (by definition, the variable's logarithm is normally distributed).

- ▶ Shape similar to the log-logistic distribution (and yields similar model results)
- ▶ Accommodates an AFT model (as the log-logistic), but is not a proportional odds model (whereas the log-logistic model is a PO model)

# Gompertz Model

- ▶ PH model but not AFT
- ▶ Hazard function (with one predictor (TRT)):

$$h(t) = [\exp(\xi t)] \cdot \exp(\beta_0 + \beta_1 TRT)$$

with parametrically specified baseline hazard  $h_0(t) = \exp(\xi t)$

- ▶  $\xi > 0$ : hazard increases exponentially with  $t$
- ▶  $\xi < 0$ : hazard decreases exponentially with  $t$
- ▶  $\xi = 0$ : constant hazard (exponential model)

# Modeling the Shape Parameter

Many parametric models contain a shape parameter, which is usually considered fixed.

Example:

- ▶ Weibull model

Recall:  $h(t) = \lambda p t^{p-1}$  where  $\lambda = \exp(\beta_0 + \beta_1 TRT)$  and  $p$ , the shape parameter, unaffected by predictors.

- ▶ Alternative Weibull model

Now:  $h(t) = \lambda p t^{p-1}$  where  $\lambda = \exp(\beta_0 + \beta_1 TRT)$  and  $p = \exp(\delta_0 + \delta_1 TRT)$

- ▶ If  $\delta_1 \neq 0$ , the value of  $p$  differs by  $TRT$
- ▶ Not a PH or AFT model if  $\delta_1 \neq 0$ , but still a Weibull model

# Contents

## Introduction

## Parametric Model

Distributional Assumption

Weibull Model

Accelerated Failure Time Assumption

A More General Form of the AFT Model

Weibull AFT Model

Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Parametric Likelihood and Censoring

The likelihood function for a parametric model

- ▶ is a function of the observed data and the unknown parameters of the model.
- ▶ is based on the distribution of the survival time.
- ▶ depends on the censoring of the data.

Right-censored:  $\text{_____} \underset{10}{|} \text{X} \text{ time}$

Left-censored:  $\text{X} \text{_____} \underset{10}{|} \text{ time}$

Interval-censored:  $\text{_____} \underset{8}{|} \text{X} \underset{10}{|} \text{ time}$

## Construction of the Likelihood on an Example

Assume a survival time distribution with probability density function  $f(t)$ .

Subject	Event Time	Likelihood Contribution
Barry	$t = 2$	$f(2)$
Gary	$t > 8$ ( <i>right-censored</i> )	$\int_8^{\infty} f(t)dt$
Harry	$t = 6$	$f(6)$
Carrie	$t < 2$ ( <i>left-censored</i> )	$\int_0^2 f(t)dt$
Larry	$4 < t < 8$ ( <i>interval-censored</i> )	$\int_4^8 f(t)dt$

## Construction of the Likelihood on an Example

The likelihood function  $L$  is the product of each contribution:

$$L = f(2) \cdot \int_8^{\infty} f(t)dt \cdot f(6) \cdot \int_0^2 f(t)dt \cdot \int_4^8 f(t)dt$$

Assumptions for formulating  $L$ :

- ▶ Subjects are independent (product of contributions).
- ▶ No competing risks:  
No competing event prohibits a subject from eventually getting the event of interest.  
Example: Death
- ▶ Follow-up times are continuous without gaps (i.e. subjects do not return into study).

# Maximum likelihood Estimates

The likelihood for  $M$  subjects is

$$L = \prod_{i=1}^M L_i.$$

The maximum likelihood estimates of the parameters are obtained by solving the following system of equations

$$\frac{\partial \log(L)}{\partial \beta_j} = 0, \quad j = 1, 2, \dots, N,$$

where  $N$  is the number of parameters  $\beta_j$ .



## Parametric and Cox likelihood

In a parametric model, the parametric likelihood handles easily right-, left- or interval-censored data.

In a Cox model, the Cox likelihood handles right-censored data, but is not designed to accommodate left- or interval-censored data directly.

Example:

- ▶ Health check for nonsymptomatic outcome every year once
- ▶ If event was detected e.g. at the beginning of the third year, the exact time when the event occurred was between the second and third year
- ▶ Fit a parametric model with the distribution of the outcome denoted by  $f(t)$
- ▶ Each subject's contribution to the likelihood is obtained by integrating  $f(t)$  over the interval in which it had the event.

# Contents

## Introduction

## Parametric Model

- Distributional Assumption

- Weibull Model

- Accelerated Failure Time Assumption

- A More General Form of the AFT Model

- Weibull AFT Model

- Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Frailty

What is frailty?

- ▶ Random component
- ▶ Accounts for variability due to unobserved individual-level factors (unaccounted for by the other predictors)

The **frailty**  $\alpha$  ( $\alpha > 0$ )

- ▶ is an unobserved multiplicative effect on the hazard
- ▶ follows some distribution  $g(\alpha)$  with the mean of  $g(\alpha)$  equal to 1 ( $\mu = 1$ )
- ▶  $\theta = Var(g(\alpha))$ , parameter to be estimated from the data

## Hazard functions, Survival functions and Frailty

Express an individual's hazard function conditional on the frailty as

$$h(t \mid \alpha) = \alpha h(t)$$

This leads to:

$$\begin{aligned} S(t \mid \alpha) &= \exp \left[ - \int_0^t h(u \mid \alpha) du \right] = \exp \left[ - \int_0^t \alpha h(u) du \right] \\ &= \exp \left[ - \int_0^t h(u) du \right]^\alpha = S(t)^\alpha \end{aligned}$$

Suppose  $\alpha > 1$ . Then we get:

- ▶ Increased hazard
- ▶ Decreased survival

And vice versa for  $\alpha < 1$ .

# Survival functions in Frailty Models

Distinguish between

- ▶ the **individual level** or **conditional survival function**  $S(t \mid \alpha)$
- ▶ and the **population level** or **unconditional survival function**  $S_U(t)$ , representing a population average.

Once the frailty distribution  $g(\alpha)$  is chosen we find the unconditional survival function by

$$S_U(t) = \int_0^{\infty} S(t \mid \alpha) g(\alpha) d\alpha$$

Then we can find the corresponding unconditional hazard  $h_U(t)$  using the known relationship between survival and hazard function

$$h_U(t) = \frac{-d[S_U(t)]/dt}{S_U(t)}$$

# Example: Weibull PH model with and without frailty

Vet Lung Cancer Trial

Predictors:

TX (dichotomous: 1 = standard, 2 = test)

PERF (continuous: 0 = worst, 100 = best)

DD (disease duration in months)

AGE (in years)

PRIORTX (dichotomous: 0 = none,  
10 = some)

*Model 1. No Frailty*

Weibull regression (PH form)

Log likelihood = -206.20418

.t	Coef.	Std. Err.	z	p >  z
tx	.137	.181	0.76	0.450
perf	-.034	.005	-6.43	0.000
dd	.003	.007	0.32	0.746
age	-.001	.009	-0.09	0.927
priortx	-.013	.022	-0.57	0.566
_cons	-2.758	.742	-3.72	0.000
/ln_p	-.018	.065	-0.27	0.786
p	.982	.064		
1/p	1.02	.066		

Model 1:

$$h(t) = \lambda p t^{p-1} \text{ where}$$

$$\lambda = \exp(\beta_0 + \beta_1 TX + \beta_2 PERF + \beta_3 DD + \beta_4 AGE + \beta_5 PRIORTX)$$

## Example: Weibull PH model with and without frailty

### Model 2. With Frailty

Weibull regression (PH form)

Gamma frailty

Log likelihood = -200.11338

_t	Coef.	Std. Err.	z	p >  z
tx	.105	.291	0.36	0.719
perf	-.061	.012	-5.00	0.000
dd	-.006	.017	-0.44	0.663
age	-.013	.015	-0.87	0.385
priortx	-.006	.035	-0.18	0.859
_cons	-2.256	1.100	-2.05	0.040
/ln_p	.435	.141	3.09	0.002
/ln_the	-.150	.382	-0.39	0.695
p	1.54	.217		
1/p	.647	.091		
theta	.861	.329		

Likelihood ratio test of theta = 0:

chibar2(01) = 12.18

Prob>=chibar2 = 0.000

Model 2:

$h_j(t \mid \alpha_j) = \alpha_j h(t)$ ,  $j = 1, 2, \dots, n$ , where  $h(t)$  and  $\lambda$  as above,  $\alpha_j$  denoting the frailty for the  $j$ -th subject and where  $\alpha \sim \text{gamma}(\mu = 1, \text{variance} = \theta)$

Note:  $\alpha_j$  not estimable (overparameterization). But the variance of the frailty  $\theta$  is estimated

## Example: Weibull PH model with and without frailty

For Model 1 we get  $\widehat{HR} = \exp(0.137) = 1.15$ .

For Model 2 we get  $\widehat{HR} = \exp(0.105) = 1.11$ .

### Remark

In Model 2 the value we obtained is the estimated hazard ratio for two individuals having the **same** frailty one taking the test and the other taking the standard treatment (and same levels of other predictors).



## Example: Weibull PH model with and without frailty

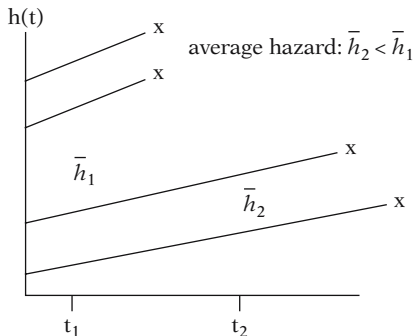
Compare the estimated values for the shape parameter  $p$ :

- ▶ Model 1:  $\hat{p} = 0.982$  ( $\rightarrow$  decreasing hazard)
- ▶ Model 2:  $\hat{p} = 1.54$  ( $\rightarrow$  increasing individual level hazard)

BUT: For frailty models one has to distinguish between the individual level and population level hazard.

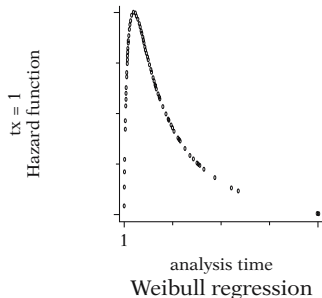
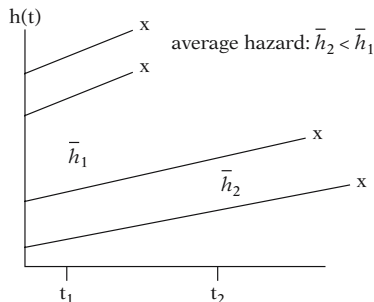
- ▶ Individual level/conditional hazard is estimated to increase
- ▶ Population level/unconditional hazard has an **unimodal** shape (first increasing, then decreasing to 0)

# Conditional and Unconditional Hazards in Frailty Models



- ▶ Hazard for individuals increase
- ▶ Average hazard decreases

# Conditional and Unconditional Hazards in Frailty Models



Population with different levels of frailty

→ "more frail individuals" ( $\alpha > 1$ ) are more likely to get the event earlier

→ "at risk group" has increasing proportion of less frail individuals ( $\alpha < 1$ )

→ decreasing population average hazard  $h_U(t)$

→ **frailty effect**

## Example

Assume plotting the Kaplan-Meier log-log survival estimates for treatment  $TX = 2$  vs.  $TX = 1$  would give us plots starting out parallel but then converge over time.

- ▶ Interpretation 1: Effect of treatment weakens over time  
⇒ PH model not appropriate
- ▶ Interpretation 2: Effect of treatment remains constant over time. Convergence is caused by an unobserved heterogeneity in the population  
⇒ a PH model with frailty would be appropriate

# Contents

## Introduction

## Parametric Model

- Distributional Assumption

- Weibull Model

- Accelerated Failure Time Assumption

- A More General Form of the AFT Model

- Weibull AFT Model

- Log-Logistic Model

## Other Parametric Models

## The Parametric Likelihood

## Frailty Models

## Summary

# Summary

- ▶ Parametric model: assume distribution of survival time
- ▶ PH, AFT and PO (Examples: Weibull and log-logistic models)
- ▶ Parametric likelihood
- ▶ Frailty models: additional variability factor for hazard
- ▶ Distinguish between conditional and unconditional frailty

Thank you for your attention!